

dr hab. inż. Marcin Kowalski, prof. WAT
Instytut Optoelektroniki
Wojskowa Akademia Techniczna
ul. Gen. S. Kaliskiego 2, Warszawa

Warszawa, 7 lipca 2022

PW WAT Kancelaria
wpłynęło dnia 22.07.2022.
numer

Rada Naukowa Dyscypliny
Informatyka Techniczna i Telekomunikacja
Politechniki Warszawskiej
w Warszawie

RECENZJA

Rozprawy doktorskiej Pani mgr inż. Eweliny Bartuzi-Trokielewicz
pt. "Presentation attack-resistant palm recognition for mobile devices
in unconstrained conditions"

Tematyka pracy dotyczy biometrycznego rozpoznawania osób na podstawie analizy obrazu dłoni zarejestrowanego w warunkach niekontrolowanych na urządzeniach mobilnych oraz metody wykrywania ataków prezentacyjnych dłoni. Dobór tematyki rozprawy jest zasadny, w szczególności z punktu widzenia na zastosowanie modalności, której próbki można pobrać w sposób przystępny i akceptowalny, oraz ze względu na fakt, że opracowana metoda ma w założeniu działać na popularnych urządzeniach mobilnych.

Przedstawiona praca składa się z sześciu rozdziałów, które będą kolejno opiniowane poniżej. Praca została przygotowana w języku angielskim.

Rozdział 1: Wprowadzenie

Autorka pracy podała krótką historię wykorzystania różnych modalności związanych z dłonią w biometrii oraz krótkie uzasadnienie wyboru danej modalności. Autorka pracy twierdzi, że „systemy biometryczne oparte na odcisku dłoni są jednymi z najpopularniejszych”. Jak statystycznie wygląda użycie tego typu systemów w porównaniu do np. modalności dwuwymiarowej twarzy w zakresie światła widzialnego? Jakie są najpopularniejsze aplikacje?

Autorka zamiennie używa określeń „palm recognition” oraz „palmprint recognition”. Ze względu na fakt, że istnieje kilka modalności bazujących na analizie dłoni, nomenklatura w tym zakresie powinna być jednoznaczna w całej pracy.

Rozdział 2: Bazy biometryczne

W drugim rozdziale dysertacji Autorka wskazuje i skrótowo opisuje bazy danych biometrycznych, które zostały wykorzystane w pracy. Pierwsze dwie bazy, Mobibits-HQ oraz Mobibits-VIS są wspólnym dziełem, w którym

Autorka pracy ma wiodącą rolę. Obie bazy podnoszą wartość eksperymentalną niniejszej pracy i stanowią znaczący wkład Autorki w rozwój tej modalności. Szkoda, że metodyka przeprowadzenia eksperymentów, w których starano się uzyskać niekontrolowane warunki nie została przedstawiona w pracy, w szczególności, że Autorka miała w jej opracowaniu znaczący udział.

Rozdział 3: Niezależna od środowiska i dokładna segmentacja dłoni

Niniejszy rozdział dotyczy segmentacji obrazu przedstawiającego wewnętrzną część dłoni. Segmentacja jest wykonywana w celu dalszej analizy obrazu i finalnie, weryfikacji tożsamości. Rozdział składa się z czterech podrozdziałów obejmujących wprowadzenie, przegląd stanu techniki, opis oceny metod referencyjnych oraz proponowaną metodykę segmentacji. Wskazany przegląd stanu techniki w zakresie segmentacji jest skrótowy i ogólnikowy. Dokonano podziału na dwie grupy metod. Pierwsza grupa to najprostsze metody z progiem segmentacji wyznaczonym na podstawie statystycznych wartości obrazu. Druga grupa obejmuje metody wykorzystujące cechy tekstur. Podejścia te wykorzystują deskrytory tekstur, w tym ekstrakcję cech za pomocą banku filtrów. Ta grupa obejmuje również metody segmentacji semantycznej. Autorka zawarła najważniejsze trendy i pozycje, chociaż nie wyjaśnia różnic między metodami.

Eksperymenty z metodami referencyjnymi stanowią znaczącą część niniejszego rozdziału. Autorka zestawiała cztery popularne metody i wykonała testy ewaluacyjne z użyciem określonych baz danych. Niestety w grupie metod referencyjnych nie znalazły się metody segmentacji semantycznej.

W czwartym podrozdziale opisano proponowaną metodykę segmentacji, bazującą na znanych metodach segmentacji semantycznej. Autorka podjęła się żmudnego przygotowania danych do procesu uczenia dla różnych wariantów sieci. Zastosowane w tym zakresie modele sieci neuronowych są znanymi z literatury metodami i równocześnie mogły zostać pokazane jako metody referencyjne. Metody zostały przebadane oraz wyznaczone zostały parametry metryk IoU oraz ESQ. Finalnie wybrano dwie metody, których rozmiary pozwalały na zastosowanie w środowisku urządzeń mobilnych.

W trakcie omawiania eksperymentu z użyciem metod do segmentacji semantycznej wprowadzono pojęcie metryki ESQ, którą można tłumaczyć jako „Jakość segmentacji oceniona przez eksperta (ESQ), procent prawidłowej segmentacji według eksperta”. Niestety w pracy nie zamieszczono sposobu wyznaczania tej metryki, co w zasadzie uniemożliwia odniesienie się do jej skuteczności lub nawet zasadności. Ponadto, Autorka opisuje swoją pracę zwracając się do czytającego przez „my”. Jest to być może niedopatrzenie, które w pracy przewija się w kilku miejscach.

Rozdział 4: Weryfikacja dłoni w środowisku bez ograniczeń

Rozdział dotyczący metody weryfikacji osób na podstawie skanu dłoni podzielony został na 6 podrozdziałów opisujących poszczególne etapy działania metody oraz jej wyniki.

W tej części pracy wykorzystano wybraną uprzednio metodę segmentacji opartą na algorytmie DeepLab3+. Po krótkim wprowadzeniu przedstawiono skrótowo przegląd istniejących rozwiązań z zakresu określania obszaru zainteresowań oraz ekstrakcji cech dystynktywnych. W obu przypadkach zaprezentowane prace nie wyczerpują tematyki natomiast stanowią prawidłowe odniesienie do zakresu istniejących metod.

Następnie zaprezentowane są poszczególne etapy działania metody weryfikacji zaczynając od segmentacji, wyboru obszaru zainteresowań, ekstrakcji cech aż do architektury funkcji decyzyjnej. Autorka zaproponowała zastosowanie syjamskiej sieci neuronowej typu „Consistent Attentive Siamese Network” do uzyskania finalnego wyniku. Wspomniana architektura została dość pobieżnie wyjaśniona. Największy niedosyt informacji dotyczy samego zjawiska „Consistent Attentive”. Nasuwają się pytania, które nie zostały w pracy dostatecznie wyjaśnione:

- jak generowane są attention maps?
- jakie są warunki do generowania attention maps i jakie muszą być ich parametry?
- jak wpływa zmiana sposobu generowania attention maps na wynik porównania dwóch próbek biometrycznych?

Tak postawione pytania są kluczowe do zrozumienia dlaczego taka metoda została wybrana do dalszych badań. W pracy zabrakło porównania ze „standardowymi” sieciami zbudowanymi w architekturze syjamskiej.

Ostatnie dwa podrozdziały pracy dotyczą wyników oraz podsumowania. Przedstawione wyniki obejmują trzy warianty eksperymentów. Z pracy nie wynika, czy Doktorantka zastosowała walidację krzyżową w trakcie swoich badań i jakie są ewentualne wnioski. Z opisu eksperymentów wynika, że do treningu wykorzystane zostały 64 pary próbek, co stanowi relatywnie małą bazę danych. Czy w związku z tym przebadano wpływ ilości danych na uzyskiwane wyniki?

W części odnoszącej się do wyników pracy brakuje odniesienia do zastosowania niniejszej metody na urządzeniach mobilnych, co jest jednym z głównych elementów pracy.

Dodatkowa uwaga odnosi się do nazewnictwa, w którym Autorka niepotrzebnie zamiennie korzysta z wyrażen „palm recognition” oraz „hand recognition”. Nazewnictwo powinno być spójne w całej pracy.

Rozdział 5: Wykrywanie ataków prezentacyjnych

W piątym rozdziale Doktorantka przedstawia pracę w zakresie wykrywania ataków prezentacyjnych na systemy biometrycznego rozpoznawania odcisków dłoni. Rozdział składa się z 5 podrozdziałów. W pierwszych podrozdziałach Autorka wprowadza w tematykę ataków prezentacyjnych opisując różne typy ataków oraz wskazując kierunek swoich badań. Opisanych zostało kilka metod referencyjnych, dla których zaproponowano eksperymenty z trzema typami ataków prezentacyjnych. W trakcie eksperymentów zastosowano trzy typy instrumentów ataku prezentacyjnego, czyli fotografię, wyświetlony obraz oraz próbkę generowaną komputerowo.

Do oceny skuteczności wykrywania ataków zastosowano znane metryki APCER oraz BPCER.

Pewnym mankamentem pracy jest brak wykorzystania rzeczywistych skomplikowanych instrumentów ataku w postaci np. odpowiednio spreparowanych rękawic. Oczywiście, wykonanie takich badań wiązałoby się z relatywnie dużymi kosztami, ale mogłoby urzeczywistnić skuteczność opracowanej metody.

Rozdział zakończony jest krótkim wskazaniem wniosków z pracy.

Rozdział 6: Podsumowanie

Rozdział 6 stanowi podsumowanie pracy doktorskiej. Autorka skrótowo przedstawia wykonaną pracę oraz najważniejsze uzyskane wyniki. W rozdziale pojawiły się drobne błędy edycyjne związane z niepoprawnymi kros-referencjami.

Referencje:

Autorka pracy przytoczyła 112 publikacji, w tym 7 publikacji własnych. Dobór referencji jest rzetelny i aktualny.

Załącznik A: Słownik biometryczny

Załącznik ten zawiera listę podstawowych pojęć związanych z biometrią wraz z wyjaśnieniem. Autorka wskazuje, że niniejsza lista została zaczerpnięta z normy ISO/IEC Information technology – Vocabulary – Part 37: Biometrics standard. Dobór wylistowanych pojęć jest poprawny.

Leovaldo

Załącznik B: Lista publikacji i osiągnięć Autorki

Doktorantka jest autorem 10 publikacji, obejmujących głównie doniesienia konferencyjne. W przypadku 8 publikacji Doktorantka jest pierwszym autorem. Wkład merytoryczny Doktorantki jest dostatecznie opisany w przypadku każdej z pozycji. Atutem jest opublikowanie kilku prac w uznanych międzynarodowych konferencjach.

Załącznik C: Lista konferencji

Załącznik ten zawiera listę konferencji, w których Doktorantka brała aktywny udział i obejmuje 7 pozycji. Pozycje te pokrywają się z przedstawionymi w załączniku B publikacjami.

Załącznik D: Lista grantów i projektów, w których uczestniczył Autor

Załącznik D zawiera listę grantów i projektów badawczych, w których brała udział Doktorantka. Niniejszy załącznik został podzielony na dwie części i obejmuje 3 pozycje, w których Doktorantka odgrywała wiodącą rolę badawczą oraz 5 pozycji, w których Doktorantka pełniła rolę badaczki.

Uwagi krytyczne:

- 1) Skrótowy, mało szczegółowy opis eksperymentów i metod oraz analiza wyników.
- 2) Brak szczegółowego nawiązania do pracy metod w warunkach urządzeń mobilnych.
- 3) Metoda wykrywania ataków prezentacyjnych powinna zostać poddana analizie ablacyjnej. W ramach analizy możliwe było wykazanie stopnia generalizacji opracowanej metody dla ataków nieobecnych w trakcie procesu uczenia.
- 4) Zestaw instrumentów ataku prezentacyjnego jest dość ubogi. Można by się spodziewać zastosowania bardziej skomplikowanych instrumentów, takich jak np. dedykowane, elastyczne rękawice zmieniające geometrię i teksturę dłoni.
- 5) Brak publikacji rezultatów w uznanych czasopismach międzynarodowych – mimo znaczących i nowych rezultatów.

Podsumowując: Niniejsza praca dotyczy rozpoznawania osób na podstawie obrazu wewnętrznej części dłoni. Praca jest napisana dość skrótowo, w szczególności pewien niedosyt można mieć w zakresie opisu eksperymentów i użytych metod analizy danych. Czytając pracę pojawia się mnóstwo pytań, na które odpowiedź wzbogaciłaby wartość badawczą pracy. Uważam jednak pracę za wartościową, zwłaszcza z praktycznego i eksperymentalnego punktu widzenia. Przedstawione rezultaty są wynikiem rzetelnej pracy eksperymentalnej i programowania.

Stwierdzam, że rozprawa pt. „Presentation attack-resistant palm recognition for mobile devices in unconstrained conditions” autorstwa Pani mgr. inż. Eweliny Bartuzi-Trokielewicz **spełnia wszystkie wymagania stawiane rozprawom doktorskim** przez Ustawę o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki z dnia 14 marca 2003 r (Dz.U. 2003 Nr 65 poz. 595). W związku z powyższym stawiam wniosek o przyjęcie przedstawionej pracy, jako rozprawy doktorskiej w dziedzinie nauk technicznych, dyscyplinie Informatyka Techniczna i Telekomunikacja oraz **dopuszczenie** jej Autorki, Pani mgr inż. Eweliny Bartuzi-Trokielewicz **do publicznej obrony**.

Maria Lewalito
07/06/2022