**RADA NAUKOWA DYSCYPLINY**

**INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA POLITECHNIKI WARSZAWSKIEJ**

zaprasza na

OBRONĘ ROZPRAWY DOKTORSKIEJ

**mgr inż. Eweliny Bartuzi-Trokielewicz**

która odbędzie się w dniu **17 listopada 2022 roku o godzinie 10.00** w trybie hybrydowym

Temat rozprawy doktorskiej:

***„Presentation attack-resistant palm recognition for mobile devices in unconstrained conditions”***

Promotor: prof. dr hab. inż. Andrzej Pacut - Politechnika Warszawska

Recenzenci: dr hab. inż. Marcin Kowalski – Wojskowa Akademia Techniczna w Warszawie

prof. dr hab. inż. Khalid Saeed – Politechnika Białostocka

prof. dr hab. inż. Krzysztof Ślot – Politechnika Łódzka

* Obrona odbędzie się w Sali nr 116 Gmachu Elektroniki Politechniki Warszawskiej oraz na platformie MS Teams

**Osoby zainteresowane uczestnictwem w obronie w trybie zdalnym** proszone są o zgłoszenie chęci uczestnictwa w formie elektronicznej na adres sekretarza komisji: dr hab. inż. Piotra Gawrysiaka, prof. uczelni – email : piotr.gawrysiak@pw.edu.pl **do dnia 15.11.2022 do godz.15.00.**

Z rozprawą doktorską i recenzjami można zapoznać się w Czytelni Biblioteki Głównej Politechniki Warszawskiej, Warszawa, Plac Politechniki 1.

Streszczenie rozprawy doktorskiej i recenzje są zamieszczone na stronie internetowej: https://bip.pw.edu.pl/Postepowania-w-sprawie-nadania-stopnia-naukowego/Doktoraty/Wszczete-po-30-kwietnia-2019-r/Dyscyplina-informatyka-techniczna-i-telekomunikacja-dziedzina-nauk-inzynieryjno-technicznych/mgr-inz.-Ewelina-Bartuzi-Trokielewicz.

Przewodniczący Rady Naukowej Dyscypliny
Informatyka Techniczna i Telekomunikacja
Politechniki Warszawskiej
***dr hab. inż. Jarosław Arabas***

**Streszczenie**

Cechy dłoni są wykorzystywane w systemach biometrycznych już od początku XX wieku. Początkowo analizowano kształt dłoni i cechy geometryczne, następnie poszerzono jej zastosowanie o odciski palców, wzory naczyń krwionośnych, a także cechy tekstury. Biometria dłoni jest wygodna dla użytkownika, szybka i cechuje się dużą skutecznością rozpoznawania w warunkach pomiaru w warunkach kontrolowanych. W niniejszej rozprawie doktorskiej zidentyfikowano cztery duże problemy badawcze związane z przeniesieniem metod biometrycznych dłoni na urządzenia mobilne i zaproponowano sposoby ich rozwiązania.

W ramach pierwszego problemu badawczego wykonano analizę niezawodności istniejących metod rozpoznawania cech dłoni dla danych pobieranych w warunkach niekontrolowanych, ze szczególnym uwzględnieniem danych z urządzeń mobilnych. Przeprowadzono ocenę uzyskanych błędów rozpoznawania biometrycznego i ich przyczyn. Wyniki eksperymentów wykazały dwie główne przyczyny znacznego spadku jakości rozpoznawania cech dłoni. Pierwszą z nich był niepoprawny przebieg procesu segmentacji dłoni na obrazie, związany z warunkami pomiarowymi, nierównomiernym oświetleniem, trudnym tłem: różnicowanym lub w odcieniach skóry. Podczas eksperymentu obejmującego ręcznej korekcji obszaru maski dłoni, nie uzyskano znacznej poprawy wyników weryfikacji tożsamości. Drugą wykrytą przyczyną błędów są nieliniowe zmiany w teksturze dłoni, związane ze swobodną prezentacją dłoni, różnym ustawieniem dłoni w przestrzeni, ustawieniem i położeniem palców względem siebie, a także napięciem mięśniowym.

W pracy zaproponowano metodę segmentacji dłoni, niwelującą problem określania obszaru dłoni na obrazie. Podejście wykorzystuje głęboką splotową sieć neuronową, która skutecznie wykrywa dłoń na zdjęciach wykonanych w świetle widzialnym, oraz podczerwonym, w różnych warunkach pomiarowych, a także na skomplikowanym tle, bądź tle w odcieniach skóry. Średni wskaźnik skuteczności segmentacji dłoni dla różnych baz danych, IoU (ang. *intersection over union*) wyniósł ponad 98%.

Opracowano także metodę ekstrakcji cech osobniczych, niewrażliwą na nieliniowe zmiany tekstury dłoni. Metoda ta wykorzystuje syjamskie splotowe sieci neuronowe do wyznaczania podobieństwa między cechami tekstury dłoni. Do zwiększenia

jakości weryfikacji wykorzystano mechanizm uczenia z uwagą (ang. *attention learning mechanism*). To podejście pozwoliło na nawet 10-krotne zredukowanie błędu zrównoważonego EER: dla łatwiejszych baz zdjęć dłoni z 6.62% do 0.07%, a dla trudniejszych z 9.28% do 2.92%.

Ostatnim elementem pracy było opracowanie modułu zwiększającego bezpieczeństwo zaproponowanego systemu biometrycznego poprzez wykrywanie prezentacji fałszywych danych (ang. *presentation attacks detection*). System uwzględnia odpieranie trzech popularnych typów ataków w postaci wydruków zdjęć, prezentacji danych na ekranie oraz danych wygenerowanych sieciami neuronowymi z wykorzystaniem techniki przeniesienia stylu (ang. *style transfer*). Zaproponowana metoda wykrywania ataków pozwoliła na 99% poprawnej klasyfikacji fałszywych danych.

Autorka ma nadzieję, iż zaprezentowane w niniejszej rozprawie doktorskiej metody poprawy jakości działania systemów biometrycznych opartych na cechach dłoni w warunkach pomiaru niekontrolowanego będą stanowić wartościowy wkład w dalszy rozwój biometrii, inspiracją do stosowania podobnych podejść w innych gałęziach biometrii, a także ułatwi komercyjne wdrożenia biometrii dłoni.

**Słowa kluczowe:** *biometria, rozpoznawanie cech dłoni, rozpoznawanie w warunkach niekontrolowanych, biometria na urządzenia mobilne, bezpieczeństwo biometrii, detekcja autentyczności charakterystyki biometrycznej, uczenie maszynowe, ekstrakcja cech, klasyfikacja cech, porównywanie cech*

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



Warszawa, 7 lipca 2022

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 deskryptory 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 zestawiła 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 wyrażeń „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.

Kowalski

**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 Doktoranta 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 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**.

Marcin Kowalski
07/06/2022

prof. dr hab. inż. Khalid Saeed
Wydział Informatyki
Politechnika Białostocka
ul. Wiejska 45A, 15-351 Białystok

Tel. (+48-85) 746 9196
Fax: (+48-85) 746 9057
k.saeed@pb.edu.pl

Białystok, 24.08.2022 r.

---

RECENZJA rozprawy doktorskiej

**mgr inż. Eweliny Bartuzi-Trokielewicz**

z Wydziału Elektroniki i Technik Informacyjnych
Politechniki Warszawskiej

*zatytułowanej*

"Presentation attack-resistant palm recognition for mobile devices in unconstrained conditions"

Promotor:

Profesor dr hab. inż. Andrzej Pacut
Wydział Elektroniki i Technik Informacyjnych
Politechnika Warszawska

*Niniejszą recenzję przygotowałem na zlecenie zawarte w piśmie z dnia 24.06.2021 (otrzymane dnia 1.07.2022), które otrzymałem od Profesora Jarosława Arabasa Przewodniczącego Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Warszawskiej na podstawie uchwały Rady, podjętej dnia 24.05.2022 r.*

---

## I. Zawartość rozprawy

Praca doktorska mgr inż. Eweliny Bartuzi-Trokielewicz jest poświęcona zagadnieniu rozpoznawania cech dłoni na urządzeniach mobilnych w warunkach niekontrolowanych. Autorka przedstawia kilka nowych użytecznych metod i algorytmów dotyczących analizy cech dłoni, jej segmentacji i ekstrakcji cech oraz nowego podejścia zwiększającego bezpieczeństwo systemu biometrycznego w oparciu o wykrywanie fałszywych danych. To zagadnienie i inne cele rozprawy zostały przeanalizowane i przedstawione w rozprawie wraz z eksperymentami.

Rozprawa napisana jest w języku angielskim. Zawiera 99 stron tekstu, rysunków, tabel i ilustracji. Składa się z sześciu rozdziałów, bibliografii oraz czterech dodatków.

Rozdział 1. to „Introduction", gdzie autorka przedstawia wprowadzenie do tematyki rozprawy – biometrii obrazu dłoni (od charakterystyki dłoni do jej wstępnego przetwarzania i ekstrakcji jej cech) poprzez badanie sposobu jej ochrony przed atakiem prezentacji. W tym rozdziale autorka wymienia swoje cztery twierdzenia jako tezy pracy do udowodnienia w swojej rozprawie - główne cele to pokazanie sposobów i podejść autorki do poprawienia metod rozpoznawania odcisków dłoni przy różnych warunkach.

W rozdziale 2. „Biometric databases" autorka umieściła opis trzech kategorii baz danych dla dwóch rodzajów danych: obrazy otrzymane z urządzeń mobilnych oraz dodatkowe obrazy zbierane przy różnych warunkach otoczenia. Tytuł rozdziału sugeruje bazy z różnymi cechami biometrycznymi, ale dotyczy głównie dłoni oraz dwóch innych cech - tęczówki i palca.

W rozdziale 3. „Environment-invariant and accurate palm segmentation" opisano autorską metodę segmentacji obrazu dłoni opartą o głębokie splotowe (konwolucyjne) sieci neuronowe (DCNN - *Deep Convolutional Neural Network*). Metoda ta jest stosowana niezależnie od warunków otoczenia, takich jak światło, typ obrazu lub tło.

Rozdział 4. „Palm verification in unconstrained environment" skupia się na metodach weryfikacji dłoni przy nieograniczonych warunkach otoczenia. Rozdział ten zawiera bardzo ciekawy sposób prezentacji wyników przy wykorzystaniu mapy termicznej dłoni oraz wizualizacji wag dla mechanizmu atencji. Uwidoczniono obszary zainteresowania dłoni, które niosą najważniejsze informacje dla klasyfikatorów.

Rozdział 5. „Presentation attack detection" jest ostatnim rozdziałem merytorycznym i jest poświęcony metodom detekcji ataku prezentacji. Autorka przedstawiła biometryczne metody bezpieczeństwa i stan wiedzy metod PAD (*Presentation Attack Detection*) oraz zaproponowaną własną metodę wykrywania takiego ataku. Ten rozdział jest raczej skromny. Tematyka „*Presentation attack*" jest zawarta w tytule rozprawy i stanowi jeden z jej najważniejszych aspektów. Uważam, że rozdział o ataku prezentacji powinien być umieszczony wśród pierwszych rozdziałów, a tutaj w zamian powinien być rozdział o 'Presentation attack-**resistance recognition**'. Niemniej, zawartość tego rozdziału jest istotna.

Praca kończy się wnioskami, rozdział 6. „Summary", w których autorka konkluduje rozwiązanie powierzonych jej zadań oraz osiągnięcie wyznaczonych celów. Podaje również znakomite wyniki swojej pracy w liczbach.

Całość pracy kończy bibliografia „Bibliography", która zawiera 112 pozycji wybranych referatów i artykułów z literatury światowej pokazującej stan wiedzy i odzwierciedlającej szeroką wiedzę autorki. Zabrakło jednak jednej, moim zdaniem, istotnej pozycji o Presentation Attack: „*Ctirad Sousedik and Christoph Busch (2014) - Presentation attack detection methods for fingerprint recognition systems:*

*a survey. IET Biometrics, vol. 3, issue 4, pp. 219-233*). Chociaż praca ta jest o metodach wykrywania ataku prezentacji na przykładzie odcisków palca, to posiada ważną informację o ataku na prezentację cech biometrycznych i jest jedną z pierwszych prac przeglądowych w tej dziedzinie.

Dodatkowo, autorka opracowała cztery załączniki, w których umieściła ważne informacje: W "Appendix A" umieszczono ważne słownictwo i terminologie biometryczne - podano ich definicje lub wyjaśnienia, co znacznie ułatwia czytelnikom zrozumienie tekstu w rozprawie. Niestety, numeracja podrozdziałów nie jest prawidłowa, gdyż podana jako należąca do rozdz. 6. (6.1, 6.2, … .). "Appendix B" to lista publikacji autorki. "Appednix C" zawiera aktywność konferencyjną autorki, gdzie wygłaszała referaty na konferencjach międzynarodowych. W ostatnim dodatku, "Appendix D", podano listę grantów (jest ich 8), w których doktorantka brała udział jako wykonawca, a w trzech z nich jako główny wykonawca. Świadczy to, iż autorka jest dobrym badaczem naukowym z uznaniem otoczenia naukowego.

## II. Opinia o rozprawie doktorskiej

Rozprawę doktorską pani Eweliny Bartuzi-Trokielewicz ocenię w dwóch płaszczyznach: technicznej i merytorycznej oraz klarowności i czytelności rozprawy. Usterki redakcyjne dotyczące klarowności pisowni oraz edycyjne będą umieszczone w załączniku.

### *A. Techniczne brzmienie i merytoryczna kompletność rozprawy*

Autorka wykazała w swojej pracy dobrą znajomość zagadnień popularno-naukowych. Przedstawiła uzyskane przez siebie wyniki w sposób przekonywujący. Cytowana literatura jest prawidłowo dobrana. Według mnie praca jest interesująca, a naukowe osiągnięcia merytoryczne doktorantki są znaczne. Rozprawa jest wzbogacona wieloma przykładami osiągnięć autorki. Bardzo dobre wprowadzenie do tematyki przetwarzania obrazu dłoni jako cechy biometrycznej. Wszystkie aspekty dotyczące tej cechy oraz etapów jej rozpoznawania w różnych warunkach są szczegółowo podyskutowane z odpowiednimi standardami i wymaganymi rysunkami. Autorka poprawnie przedstawiła i wyjaśniła definicje ROI i FIDO. Wskazywała na brak definicji FIDO dla biometrii dłoni i zaproponowała 3 przykładowe stopnie oparte na pokazaniu wydruku, zdjęć na wyświetlaczu oraz wygenerowaniu obrazu przez sieć neuronową. Wskazuje to na wytrwałość i konsekwencje w prowadzonych badaniach. Doktorantka bardzo wyraźnie zaprezentowała problematyczność wykorzystanych baz, co wskazuje na znajomość trudności w analizie poszczególnych korpusów.

Głównymi rezultatami pracy badawczej doktorantki są:

- Wykonanie analizy niezawodności istniejących metod rozpoznawania dłoni przy nieograniczonych warunkach otoczenia.

- Zaprezentowanie metody segmentacji dłoni, która wykorzystuje głęboką splotową sieć neuronową i skutecznie wykrywa obraz dłoni przy różnych warunkach.

- Opracowanie metody ekstrakcji cech dłoni niewrażliwej na nieliniowość zmian tekstury.

- Opracowanie nowego podejścia zwiększającego bezpieczeństwa systemu biometrycznego w oparciu o wykrywanie fałszywych danych.

Dokonując recenzji rozprawy doktorskiej trzeba również zwrócić uwagę na jej słabe strony. Należą do nich, moim zdaniem, następujące punkty:

- Dobry system biometryczny zapewniający wysoki poziom bezpieczeństwa cechuje się tym, że jest odporny na przeróbki i przebudowę cechy biometrycznej (ang. *alteration*). Czujniki (jako czytniki, sensory) nie powinny w ogóle przyjąć fałszywych próbek. Ataki na cechy biometryczne są przeróżne, owszem, wiele z nich było wymienionych w rozprawie, jednakże zabrakło tak ważnych fałszerstw jako atak na prezentację, jak na przykład, cięcie lub transplantacja dłoni lub też zastosowania chemii w celu likwidacji wzorców odcisków dłoni. Niestety, zabrakło dyskusji na ten temat – w jakim stopniu system autorki wykrywa te przeróbki i jak je rezystuje?

- Żywotność (ang. *liveness*): Badanie żywotności cechy biometrycznej przez czujnik systemu biometrycznego jest jego drugą ważną charakterystką. Nie zauważyłem poświęcenia dodatkowego podrozdziału dla tego zjawiska. Autorka cytowała kilka pozycji literatury zajmującej się problematyką żywotności cech biometrycznych. Podała to w podrozdziale 5.2, kiedy odniosła się do stanu wiedzy o metodach PAD na przykładach. Jednak, uważam, że powinna być zawarta szersza dyskusja, a w szczególności, jak system autorki traktuje takie zjawiska.

***B. Klarowność i czytelność rozprawy***

Rozprawa napisana w języku angielskim, czyta się ją dobrze, chociaż nie brakuje błędów gramatycznych i edytorskich. Pewnym usprawiedliwieniem jest to, że nie jest to język ojczysty doktorantki. Mnie się wydaje, że autorka chciała dodać rozprawie charakter światowy pisząc w języku kongresowym, ażeby inni badacze naukowi mogli zapoznać się z jej nowymi osiągnięciami. Algorytmy, twierdzenia, wymagane rysunki i tabele są prawidłowo opracowane. Aktualny stan wiedzy dotyczącej tematyki biometrii dłoni został przedstawiony poprawnie. Istotne dla tematyki pracy zagadnienia omówiono czytelnie.

Mam wrażenie, że autorka nie ustrzegła się pewnej liczby nieścisłości mogących mieć wpływ na zrozumienie tekstu rozprawy. Niektórych szczegółów nie przedstawiono standardowo, podczas, gdy inne zagadnienia można było sformułować trochę inaczej lub w ogóle nie umieszczać w pracy. Oto moje redakcyjne uwagi:

- W opisie bazy danych (rozdział 2.2, str. 24) brakuje informacji, w jaki sposób kodowane były niestandardowe kanały RGB. Baza THID zawiera kanał podczerwieni często kodowany jako UINT16, a korpus CASIA posiada kanał 8-bitowy (dane pomiarowe matrycy CMOS). Informacja o zakresie danych liczbowych pozwoliłaby rozwiać wątpliwości co do normalizacji danych, co z kolei ma znaczący wpływ na pracę sztucznych sieci neuronowych.

- Tabela 2.1 nie jest spójna. Występują w niej zarówno informacje o rozmiarze w postaci rozdzielczości, jak i te, podane w konkretnych wymiarach. Nie jest jednak do końca wiadomo, czy w pierwszych czterech bazach zdjęcia są w postaci kwadratów czy prostokątów. Powiązane jest to z możliwością walidacji rożnych typów sieci neuronowych realizujących mechanizm segmentacji - czy każdy z obrazów ma wymagany wymiar.

- W podrozdziale 2.4., według autorki, jednym z elementów wykonanej pracy, jest rozszerzenie metody segmentacji dłoni w celu ujęcia innych cech i udowodnienia multimodalnego charakteru jej metody. Jednak rozprawa nie zawiera żadnej informacji o multimodalności jej podejścia. Z tego samego powodu, nie widzę sensu dodawania informacji o bazach odcisków palca oraz tęczówki oka.

- W podrozdziale 3.4.1 przedstawiono skrócone opisy dotyczące architektury sieci neuronowych CNN wykorzystanych w procesie segmentacji. Zbrakło jednak krótkiego zestawienia czy modele te wykorzystywały dane wejściowe o stałej (tj. tej samej) wielkości. Czy dane wejściowe były przetwarzane inaczej dla każdej z sieci (skalowanie, okno przesuwne)? W jaki sposób przetworzono dane pochodzące z matrycy CMOS (8 bitowa skala szarości) z korpusu CASIA? Czy wykorzystywano zbiór walidacyjny do doboru hiperparametrów np. liczby epok? Jest wyraźny brak szczegółów technicznych.

- W podrozdziale 3.4.3. autorka decyduje się na zastosowanie technik augmentacji (… *translation, rotation, scaling, adding noise, reflation, changing brightness, contrast, saturation and hue*). Przy czym brak jest jakiejkolwiek informacji dlaczego zdecydowano się na takie modyfikacje, czy wynikają one z literatury czy zaczerpnięte zostały z dziedzin pokrewnych?

- Autorka napisała w podrozdziale 4.5.3: “The experiments were carried out with a ten-fold subject-disjoint division of the data into training and test sets in the ratio of 80:20.”. Nie jest to typowo stosowany sposób walidacji i dlatego wymaga dopowiedzenia.

- W podrozdział 4.5.4. pomimo, że doktorantka prezentuje wagi map w sposób graficzny dla mechanizmu atencji, brak jest zwięzłego podsumowania. Z prezentowanych grafik wynika, że w zdecydowanej większości przypadków wysokie wagi koncentrowały się w obszarach dłoni, pomijając palce. Brak jest wnioskowania w tekście.

- Rysunek 5.1: brak źródła danych, w jaki sposób wykresy zostały wygenerowane.

- Na stronie 64 w 5.3. jako *Proposed type of attacks* zaproponowano 3 przykłady stopni FIDO dla analizy zdjęć dłoni. Przy czym dużą część rozdziału poświęcono na sztuczne modelowanie obrazu przez sieci neuronowe. W dużym stopniu pominięto aspekty techniczne, jak np. współpracę z rejestratorami obrazu. Rozprawa zyskałaby na jakości przy odwołaniu się do literatury przeglądowej. Jako przykład wskazałem pracę opublikowaną w IET Biometrics i podałem w opisie bibliografii.

Powyżej wymienione uwagi mają charakter dyskusyjny i nie obniżają wartości rozprawy, jednakże chciałbym, żeby autorka ustosunkowała się do nich na obronie.

## III. Merytoryczne osiągnięcia doktorantki

Mgr inż. Ewelina Bartuzi-Trokielewicz osiągnęła cele pracy i udowodniła wyznaczone tezy rozprawy doktorskiej, która wnosi nowe aspekty do nauk technicznych w zakresie informatyki.

Pani Ewelina jest współautorką 10. recenzowanych referatów i artykułów opublikowanych w międzynarodowych czasopismach i konferencjach. Wszystkie znajdują się na liście Ministerstwa Edukacji i Nauki. Są to 2 artykuły w recenzowanych czasopismach, jeden rozdział w książce wydawnictwa Springer Nature oraz 7 referatów konferencyjnych. Autorka ma znaczny udział w projektach naukowych (osiem grantów naukowo-badawczych). Świadczy to o wysokim znaczeniu osiągniętych wyników jej pracy naukowej w dziedzinie biometrii. Oznacza to, iż problematyka rozprawy wpisuje się w bieżący trend zagadnień z tej właśnie dziedziny.

## IV. Wnioski końcowe

Wystawiam pozytywną ocenę rozprawie doktorskiej mgr inż. Eweliny Bartuzi-Trokielewicz pt. „*Presentation attack-resistant palm recognition for mobile devices in unconstrained conditions*". Stwierdzam, że praca spełnia wymagania i warunki nakładane przez ustawę o stopniach naukowych i wnoszę bez zastrzeżeń o dopuszczenie doktorantki do obrony pracy w celu uzyskania stopnia doktora nauk technicznych w dziedzinie nauk inżynieryjno-technicznych w dyscyplinie informatyka techniczna i telekomunikacja.

Khalid Saeed

## Załącznik do opinii

### Drobne usterki

1. Strona 12 – Appendix A w spisie treści: numeracja podrozdziałów jest podana jako należąca do rozdziału 6.
2. Podrozdział 1.2.3 wymaga cytowania.
3. Strona 18 – drugi paragraf wymaga cytatu.
4. "2. Biometric databases", zbyt ogólny tytuł rozdziału. Lepiej byłoby podać precyzyjnie, np. "Databases for selected biometric features".
5. Warto podać cytat o *Grayworld algorithm* (str. 34).
6. Rysunek 3.1 – brak podpisanych osi.
7. W podrozdziale 3.4.2 autorka pisze: „*For each palm image, one mask, the most accurate from all four methods, was subjectively selected by an expert (here: the Author)"*. Nie jest do końca klarowne kim jest ekspert/autor. Czy jest to autor rozprawy, publikacji czy bazy danych?
8. Strona 49 – pojedyncze referencje [46, 47, 48, 49, 50, 51, 52] zamiast ogólnego standardu [46-52].
9. Rysunek 4.6 – brak legendy.

10. Praca posiada błędy gramatyczn, literówki. Poniżej kilka takich przykładów:
    - "in a form of the following" (str. 18)
    - Datum liczba pojedyncza, a mnoga data. Chociaż stosuje się data w liczbie pojedynczej, ale trzeba trzymać jeden standard. Autorka raz pisze "data were" a drugi raz "data was"
    - The Table 3.1 (str. 37)
    - The Table 3.2 (str. 41)

Łódź, 30.08.2022

Prof. dr hab. inż. **Krzysztof Ślot**
Instytut Informatyki Stosowanej
Politechnika Łódzka

Recenzja rozprawy doktorskiej Pani mgr inż.

**Eweliny Bartuzi-Trokielewicz**

pt.

# Presentation attack-resistant palm recognition for mobile devices in unconstrained conditions

## 1. Tematyka i cele rozprawy

Przedstawiona do recenzji rozprawa doktorska dotyczy tematyki biometrycznej ochrony dostępu do zasobów aplikacji działających na urządzeniach mobilnych, gdzie źródłem informacji biometrycznej są obrazy linii papilarnych spodu dłoni. Prace nad rozwojem metod rozpoznawania biometrycznego, które są dedykowane implementacjom w urządzeniach mobilnych, są niezwykle **aktualne** z uwagi na dążenie do zapewnienia wysokiego poziomu bezpieczeństwa korzystania z powszechnie dostępnych usług, wymagających uwierzytelniania użytkowników. Założenie nienadzorowanego kontekstu akwizycji danych biometrycznych stanowi oczywiste wyzwanie dla algorytmów analizy biometrycznej z uwagi na konieczność zastosowania skutecznych metod ekstrakcji informacji biometrycznej z obrazów o potencjalnie bogatej treści, nieoptymalnej prezentacji biometryk oraz możliwości dokonywania prób oszukania systemu analizy przez podstawienie spreparowanych danych, nazywane w literaturze ‘atakami prezentacji’. W konsekwencji, podjęty w rozprawie obszar tematyczny jest **trudny**, a sformułowanie oryginalnych metod pozwalających na uzyskanie poprawy procesu analizy biometrycznej może stanowić istotne osiągnięcie w obszarze dyscypliny naukowej **informatyka techniczna i telekomunikacja**, w której prowadzony jest przewód Doktorantki.

Wybór biometrii dłoni jako podstawy rozpoznawania osób stanowi uzasadnioną alternatywę dla częściej stosowanego podejścia, wykorzystującego biometrię twarzy. Na marginesie warto stwierdzić, że podjęte przez Doktorantkę prace mają w części charakter ogólniejszy niż sugerowany tytułem rozprawy, wykraczając poza obszar identyfikacji użytkowników urządzeń mobilnych – detekcja w obrazach obszarów zawierających linie papilarnych dłoni, a następnie, identyfikacja na tej podstawie osób, ma istotne znaczenie w kryminalistyce jako narzędzie automatyzacji analizy materiału dowodowego.

Celem prac Doktorantki było opracowanie funkcjonalnego, kompleksowego algorytmu analizy obrazów dłoni prezentowanych kamerze urządzenia mobilnego, obejmującego trzy komponenty: ekstrakcję obszaru

spodu dłoni, przeprowadzenie procedury rozpoznawania na bazie informacji zawartej w wydzielonym obszarze oraz zapewnienie dodatkowej detekcji prób ataków prezentacji (ang. Presentation Attack Detection – PAD). W obszarze każdego z wymienionych komponentów procedury, Doktorantka podjęła próby sformułowania oryginalnych metod pozwalających na uzyskanie lepszych wyników odpowiednich analiz: zwiększenia dokładności wydzielania obszaru zainteresowania, zwiększenia poprawności rozpoznawania oraz zwiększenia skuteczności detekcji ataków prezentacji.

## 2. Struktura i tezy rozprawy

Rozprawa została napisana w języku angielskim, który w większości oceniam za poprawny, umożliwiający bezproblemowe śledzenie przekazywanych treści. Kompozycja pracy nie budzi zastrzeżeń – po prezentacji kontekstu prac, identyfikacji istniejących wyzwań i problemów oraz określeniu głównego i szczegółowych celów podjętych prac, przedstawionej we Wprowadzeniu, Doktorantka w Rozdziale 2 obszernie omawia zbiory danych, używane w pracach jako podstawa do tworzenia i testowania algorytmów, a następnie przechodzi do prezentacji swoich pomysłów i wyników ich ewaluacji. W kolejnych trzech rozdziałach opisane są badania Doktorantki dotyczące trzech wymienionych wcześniej aspektów biometrycznej analizy obrazów dłoni:

- poszukiwania metod ekstrakcji obszaru dłoni z obrazów pozyskiwanych kamerą urządzenia mobilnego, bez wprowadzania ograniczeń na rodzaj tła
- poszukiwania metod reprezentacji tekstury linii papilarnych spodu dłoni, pozwalających na uzyskanie wysokiej poprawności rozpoznawania
- opracowania skutecznych metod detekcji ataków prezentacji, a więc ataków na procedurę przedstawiania systemowi analizy biometrycznej spreparowanych danych wejściowych

Kompozycja każdego z modułów nie budzi zastrzeżeń – Doktorantka poprzedza prezentację swoich pomysłów przeglądem stanu wiedzy w podejmowanym obszarze tematycznym, uzupełnionym oceną wybranych, uznanych przez Nią jako reprezentatywne metod referencyjnych, po czym próbuje (ze zróżnicowanym stopniem szczegółowości) wyjaśnić istotę oraz szczegóły proponowanych przez nią metod, poddawanych w ostatniej części ewaluacji i konfrontacji z istniejącymi podejściami.

Doktorantka formułuje cztery stwierdzenia (‘statements’), wykazanie słuszności których jest przedmiotem jej prac:

1. Palm recognition in unconstrained conditions is more difficult than in laboratory, and the recognition accuracy of existing algorithms is insufficient.
2. The proposed method of hand segmentation allows for generating precise, consistent binary mask predictions for different conditions of data acquisition.
3. The proposed palmprint feature extraction method, based on the Siamese neural network, increases the accuracy of recognition in unconstrained conditions.

4. Presentation attack detection method, which was proposed in this doctoral study is resistant to common types of presentation attack instruments.

## 3. Merytoryczna ocena pracy

Prace Doktorantki mają bardzo silny wymiar praktyczny: opracowane przez Nią procedury są z powodzeniem implementowane i dają bardzo dobre wyniki w konfrontacji z rzeczywistymi danymi pozyskiwanym z użyciem urządzeń mobilnych, co bez wątpienia zasługuje na uznanie. Jednakże, w odniesieniu do wartości naukowej przedstawianych przez Nią koncepcji, moja opinia jest zróżnicowana – część z zaprezentowanych przez Doktorantkę pomysłów jest oryginalna i wartościowa, natomiast część nie stanowi w mojej opinii zauważalnego wkładu do dziedziny i wymaga przeprowadzenia dodatkowych, obszernych prac w celu potwierdzenia słuszności formułowanych tez. Struktura dalszej części oceny odpowiada trzem tematycznie odrębnym wątkom przedstawionym w rozprawie, wyrażonym postawionymi przez Doktorantkę tezami.

### 3.1. Segmentacja obszaru dłoni

Prace Doktorantki ukierunkowane na wykazanie słuszności drugiej z przedstawionych tez rozprawy, poświęcone poszukiwaniu nowych metod ekstrakcji obszaru dłoni w obrazach rejestrowanych kamerą urządzeń mobilnych, a więc w obrazach charakteryzujących się dowolnie skomplikowanym tłem, zostały zamieszczone w Rozdziale 3 rozprawy. Autorka rozpoczyna prezentację treści przeglądem podstawowych metod stosowanych w segmentacji obrazów: punktowych - bazujących na analizie jasności lub koloru oraz obszarowych – wykorzystujących deskryptory statystyczne tekstury obrazu. W przedstawionym przeglądzie, pojawia się jedynie wzmianka o segmentacji semantycznej, dokonywanej z użyciem głębokich sieci neuronowych – Autorka zamieszcza referencje do trzech pozycji [39-41] opisujących wykorzystanie sieci konwolucyjnych do segmentacji obszaru dłoni. Zaprezentowany przegląd stanu wiedzy w obszarze segmentacji semantycznej wykorzystującej sieci głębokie nie jest kompletny i pomija publikacje prezentujące podobną jak przyjęta przez Nią metodykę analizy (np. wykorzystanie architektury RefineNet do segmentacji obrazów dłoni fotografowanych na dowolnym tle bez narzucania ograniczeń co do sposobu prezentacji dłoni [1], czy sieci U-Net do realizacji tego samego zadania [2]). Co więcej, segmentacja obszaru dłoni nie jest wcale jedynym pomysłem na precyzyjne określenie regionu zainteresowania, niezbędnego dla potrzeb biometrycznej analizy linii papilarnych. Doskonale rokującą alternatywą dla realizacji tego zadania jest wykorzystanie metod bazujących na identyfikacji lokalnych punktów charakterystycznych. Sztandarowym przykładem skutecznej detekcji punktów charakterystycznych dłoni, pozwalające również na wyznaczenie obszaru spodu dłoni dla potrzeb analizy biometrycznej jest, wykorzystująca głębokie klasyfikatory neuronowe, metoda zaimplementowana w powszechnie stosowanymi obecnie pakiecie MediaPipe [3]. Pominięcie alternatywnych pomysłów na realizację kompletnej ścieżki rozpoznawania biometrycznego, bazującego na analizie linii papilarnych spodu dłoni, zubaża tło prezentowanych rozważań.

Istotą podejścia zaproponowanego przez Doktorantkę do realizacji zadania ekstrakcji obszaru dłoni z obrazów jest półautomatyczna procedura przygotowania zbioru danych uczących, w oparciu o które trenowana ma być docelowa, głęboka sieć neuronowa o architekturze adekwatnej dla zadania segmentacji semantycznej. Celem Doktorantki jest uzyskanie zbioru dokładnych masek, stanowiących oczekiwane

wyniki segmentacji. Dla osiągnięcia tego celu wykorzystuje wyniki segmentacji obrazów uzyskiwane przez cztery różne pretrenowane głębokie sieci neuronowe (FCN, DeepLab, SegNet i U-Net), z których wybiera subiektywnie najlepsze jako materiał do ponownego treningu wybranych sieci 'bazowych' (DeepLabv3 i SegNet). Następnie, nauczony na 'dobrych' danych klasyfikator jest aplikowany do segmentacji obrazów, które wcześniej sprawiały problemy, a uzyskane wyniki są poddawane manualnym korektom, z wykorzystaniem napisanego przez Doktorantkę narzędzia do edycji obrazów. Prawdopodobnie, jest to robione po to, by wytworzyć dodatkowy podzbiór treningowy obrazów 'trudnych', na którym douczane są docelowe klasyfikatory, ale w tekście nie znalazłem takiej informacji, więc jest to tylko moja spekulacja. Przedstawiona przez Doktorantkę procedura pozwala na zwiększenie skuteczności uczenia poprzez zwiększenie liczebności zbioru treningowego, chociaż osiągnięcie takiego rezultatu wymaga zaangażowania eksperta w procesie subiektywnej oceny poprawności segmentacji oraz korygowania błędów segmentacji.

Ocena przydatności zaproponowanego przez Doktorantkę podejścia jest dokonywana w drodze weryfikacji eksperymentalnej i konfrontacji z konkurencyjnymi podejściami. Jako referencję dla opracowanej przez siebie metody Doktorantka przyjmuje trzy metody segmentacji punktowej i czwartą, stanowiącą kombinację metody punktowej i obszarowej, używającej cech lokalnych wyznaczonych na podstawie wyników filtracji filtrami Gabora. Niestety, wbrew temu co twierdzi Doktorantka, żadna z tych metod nie zasługuje na miano metody 'state of the art', więc sformułowane na podstawie porównania efektów segmentacji, mocno tryumfalne wnioski o supremacji przedstawionej przez Nią metody nad aktualnie istniejącymi rozwiązaniami są nieuprawnione. Najskuteczniejsze obecnie metody ekstrakcji dłoni w warunkach dowolności tła, dowolności ułożenia dłoni i dowolności warunków akwizycji, wykorzystują głębokie architektury do segmentacji semantycznej, podobne do używanych przez Doktorantkę, jako komponenty składowe przeprowadzanej przez Nią procedury. Przekonywującym potwierdzeniem korzyści płynących z zastosowania Jej procedury byłaby konfrontacja uzyskanych za jej pomocą wyników segmentacji z wynikami osiąganymi przez te elementarne architektury segmentacji semantycznej, a nie z wynikami uzyskiwanymi dla metod segmentacji punktowej i obszarowej. Niestety nie wiadomo, jak można by powiązać ze sobą informacje zawarte w Tabeli 3.2 (segmentacja w sieciach składowych, trenowanych bez dodatkowych zabiegów) i Tabeli 3.4 (segmentacje w sieci wytrenowanej metodą zaproponowaną przez Doktorantkę).

Doktorantka do oceny wyników eksperymentów używa, obok powszechnie stosowanej metryki ilościowej 'IoU', przede wszystkim subiektywnych miar 'poprawności segmentacji', bazujących na opinii eksperta. Taki sposób podsumowania wyników daje jedynie poglądowe oszacowanie efektów analizy i nie może stanowić obiektywnej podstawy do formułowania zbyt daleko idących wniosków. Zamiast stosowania subiektywnej miary porównawczej jakości wydzielania obszaru dłoni, Doktorantka mogła dokonać pośredniej, ale obiektywnej oceny efektów segmentacji, porównując wyniki pełnej biometrycznej analizy obrazu. Możliwym scenariuszem takiego postępowania byłoby przeprowadzenie najpierw segmentacji obrazu dłoni za pomocą metody Doktorantki i metod referencyjnych, a następnie, zastosowanie identycznego algorytmu dalszej analizy, obejmującego wybór obszaru zainteresowania i klasyfikację (a więc metodykę opisaną w Rozdziale 4). Jeżeli poprawna segmentacja dłoni jest rzeczywiście istotnym czynnikiem dla powodzenia biometrycznej analizy dłoni, uzyskiwane wyniki klasyfikacji dawałyby ilościową podstawę oceny korzyści użycia Jej podejścia w porównaniu do stosowania prostych metod segmentacji semantycznej.

Doktorantka proponuje wykorzystanie zastosowanego przez siebie schematu budowy zbioru treningowego do realizacji zadań segmentacji obrazów dwóch innych modalności biometrycznych: wydzielania tęczówki oka i wydzielania obszaru palca. Niestety, wątki te są przedstawione w pracy w sposób niezwykle skąpy, co z jednej strony jest zrozumiałe, bowiem odbiegają od tematyki biometrycznej analizy dłoni, z drugiej jednak strony, wprowadzają sporo zamieszania i niejasności. Przykładowo, zupełnie nie rozumiem informacji zawartej w Tabelach 3.4 i 3.5, odnoszącej się do 'trzeciego' scenariusza eksperymentalnego (zgodnie z wyjaśnieniem, ten scenariusz to 'rozszerzenie' wariantu 'drugiego' o dwie dodatkowe modalności biometryczne, co już wymaga szerszego komentarza, bo nie bardzo wiem, jak segmentacja tęczówki może być utrudniana przez obecność '47 kategorii tekstur tła'). Co oznaczają dane przedstawione w ostatniej kolumnie obydwu Tabel? Czy zaproponowana przez Doktorantkę procedura segmentacji obrazów dłoni działa lepiej, gdy sieć trenuje się również na obrazach oka? Taki wniosek wydaje się być absurdalny, ale jest on uprawniony na podstawie przedstawionego tekstu. Czy może wyniki z ostatniej kolumny to jakaś agregacja różnych eksperymentów, więc nie można ich porównywać z danymi z wcześniejszych kolumn?

Podsumowując ocenę merytoryczną pierwszego z zaprezentowanych w rozprawie wątków, chcę stwierdzić, że Doktorantka sformułowała oryginalną, półautomatyczną koncepcję zwiększenia poprawności realizacji segmentacji semantycznej. Szkoda, że ewaluacja metody została dokonana w sposób niekompletny, co utrudnia ocenę jej znaczenia z punktu widzenia wkładu do rozważanej dyscypliny naukowej.

### 3.2. Identyfikacja biometryczna na podstawie linii papilarnych dłoni

Zrealizowane przez Doktorantkę prace nad biometryczną analizą linii papilarnych dłoni, opisane w Rozdziale 4 rozprawy, stanowią najciekawszy w mojej opinii fragment pracy, odnoszący się przede wszystkim do trzeciej ze sformułowanych we wstępie pracy tez (teza pierwsza zawiera dość oczywiste przypuszczenie o zwiększeniu stopnia trudności rozpoznawania w warunkach braku możliwości narzucenia ograniczeń na proces akwizycji obrazu). Doktorantka prezentuje dwuelementową procedurę, której pierwszym etapem jest autorska metoda ekstrakcji obszaru zainteresowania, a drugim – wyznaczenie deskryptora tego obszaru i jego klasyfikacja. O ile zaproponowana przez Doktorantkę procedura ekstrakcji obszaru analizy, mimo wykazanej w eksperymentach skuteczności, ma czysto inżynierski charakter, o tyle zaproponowana, oryginalna metoda klasyfikacji udowadnia Jej doskonałą orientację w obszarze zaawansowanych metod uczenia maszynowego, kreatywność i intuicję badawczą.

Doktorantka trafnie identyfikuje dwa podstawowe źródła problemów rozpoznawania linii papilarnych dłoni: nieliniowe deformacje treści obrazu (zależną od napięcia mięśni, zmienną strukturę wewnętrznej powierzchni dłoni) oraz niewielką liczbę dysponowanych dla klas przykładów, stanowiącą fundamentalny problem dla opracowania skutecznych metod uczenia głębokiego. Aby zmierzyć się z pierwszym problemem, Doktorantka decyduje się na wybór metody analizy, pozwalającej na zapewnienie elastyczności wyboru obszarów istotnych dla podejmowania decyzji, oferowanej w przypadku głębokich sieci neuronowych przez mechanizm skupiania uwagi ('attention'). Możliwość kontekstowej dywersyfikacji znaczenia różnych fragmentów informacji podawanej na wejście sieci stanowi, jak wykazuje Doktorantka, skuteczny pomysł radzenia sobie z analizą struktur podlegających nieliniowym deformacjom. W przypadku drugiego z przedstawionych problemów – praktycznego braku możliwości 'wyuczenia' standardowego klasyfikatora głębokiego w obliczu posiadania skrajnie ograniczonych liczebności danych reprezentujących dane klasy, uwaga Doktorantki skupia się na architekturze syjamskiej, stanowiącej rozwiązanie dedykowane dla budowy dyskryminatywnej reprezentacji klas w

warunkach rozważanego ograniczenia liczby dostępnych przykładów. Efektem obydwu pomysłów jest zaproponowana przez Doktorantkę architektura głęboka, trenowana z użyciem dwuskładnikowej funkcji celu, maksymalizującej, poprzez zastosowanie schematu uczenia sieci syjamskiej na parach próbek zgodnych i różnych, dyskryminatywność reprezentacji obszaru zainteresowania i jednocześnie wymuszającej dla przetwarzanych par obrazów zgodność przestrzennego rozkładu informacji uznawanej za kluczową w podejmowaniu decyzji.

Przeprowadzona przez Doktorantkę weryfikacja eksperymentalna metody nie budzi zastrzeżeń, a uzyskane wyniki konfrontowane są z rezultatami rozpoznawania uzyskiwanymi za pomocą dobrze zidentyfikowanych metod referencyjnych, stanowiących aktualnie najbardziej skuteczne podejścia do realizacji rozważanego problemu.

Przedstawiona przez Doktorantkę prezentacja prac nad rozpoznawaniem linii papilarnych jest w większości przejrzysta i ciekawa, niestety, zawiera również liczne niedopowiedzenia. Doktorantka pozostawia domyślności czytelnika ogólną strukturę algorytmu identyfikacji dłoni: pisze o tym, jak wytrenować model wyznaczający dyskryminatywną reprezentację linii papilarnych dłoni, jednak nie informuje, jak przeprowadza analizę w wytrenowanej sieci. Jeżeli dobrze rozumiem, podstawą klasyfikacji obrazu jest wektor cech wyznaczany przez wytrenowaną sieć, który jest następnie poddawany porównaniu (nie wiadomo według jakiej zasady) z wektorami reprezentującymi klasy (nie wiadomo jak określonymi – czy z każdym z wektorów zbioru treningowego, czy klasa ma jednego reprezentanta). Podobnie, bardzo wartościowa analiza możliwości redukcji rozmiaru obrazów, jakie mają być poddawane analizie, pozwalająca na uproszczenie architektury sieci używanej do realizacji zadania, jest komunikowana w zdawkowy sposób. Doktorantka jako kryterium używa współczynnika korelacji wzajemnej, szkoda że nie podaje precyzyjnie, jak go definiuje (obrazy mają różne rozmiary). Co więcej, jeżeli celem analizy jest zachowanie jak największej ilości informacji oryginalnej, dlaczego nie używa jako kryterium informacji wzajemnej? Wreszcie, na podsumowującym analizy rysunku 4.6 Doktorantka nie uznaje za stosowne podać, w jakich jednostkach mierzy rozmiar obrazu (jeśli są to piksele, to oznacza, że mozaika o rozmiarze 7 x 7 punktów zapewnia informację wystarczającą dla przeprowadzenia klasyfikacji dla obrazów z dwóch baz danych, w co szczerze wątpię).

Podsumowując prace Doktorantki w obszarze poszukiwania nowych metod biometrycznej identyfikacji linii papilarnych spodu dłoni chcę jednoznacznie stwierdzić, że są one wartościowe i oryginalne, stanowiąc w mojej opinii zauważalny wkład naukowy do rozważanej dziedziny.

### 3.3. Detekcja ataków prezentacji

Przedstawione w Rozdziale 5 efekty prac Doktorantki w obszarze detekcji ataków prezentacji na biometrię linii papilarnych spodu dłoni uważam za słabszą część rozprawy, z uwagi na niezwykle zdawkową komunikację prezentowanych treści oraz powierzchowność merytorycznej analizy podnoszonych problemów i proponowanych rozwiązań.

Ogólna struktura prezentacji materiału jest przejrzysta – Doktorantka najpierw przedstawia problem, kompetentnie informuje o przyjętej taksonomii ataków prezentacji i stanie wiedzy w zakresie detekcji ataków prezentacji, następnie omawia rozważone przez Nią metody ataków oraz sposoby ochrony przed atakami, podsumowując tekst wynikami weryfikacji przeprowadzonych eksperymentów. Niestety, na

każdym z etapów prezentacji, pojawia się deficyt wyjaśnień szczegółowych, utrudniający lub uniemożliwiający śledzenie przedstawianych treści.

Pierwszym obszarem prac Doktorantki była generacja syntetycznych obrazów dłoni, które mają stanowić najbardziej ‘wyrafinowaną’ formę ataków prezentacji na system biometrycznej analizy dłoni. Niestety, nie rozumiem (nie zostało to wyjaśnione) dlaczego wygenerowane sztucznie obrazy dłoni z fałszywą, transferowaną teksturą linii papilarnych, miałyby stanowić większe wyzwanie dla algorytmu detekcji ataku prezentacji niż zdjęcie (aby utworzyć obraz syntetyczny według metodyki przyjętej przez Doktorantkę, konieczne jest posiadanie zdjęcia ‘atakowanej’ dłoni). Ponieważ prezentacja wygenerowanego obrazu jest dokonywana z użyciem wyświetlacza czy monitora, wydaje się że skuteczność proponowanego ataku, nie powinna różnić się od ataku przedstawienia zdjęcia dłoni.

W swoich pracach Doktorantka wyróżnia trzy metody generacji sztucznych obrazów – dwie z nich są gotowymi implementacjami, pozyskanymi z Internetu, zaś trzecią wskazuje Ona jako metodę własną. Niestety, nie wyjaśnia na czym polega autorski wymiar metody – opis przedstawiony w części 5.3.3 nie prezentuje żadnego oryginalnego pomysłu, a jedynie przytacza dwa wzory i ilustracje pochodzące ze źródeł. Co więcej, w prezentowanym opisie jest wiele nieścisłości. Wzór (5.1), wbrew zapowiedzi: ‘The content of the image ... can be described by the formula:’ nie kwantyfikuje ‘zawartości obrazu’, ale wyraża miarę błędu oceniającą różnicę między wynikami przetwarzania obrazów: oryginalnego i wygenerowanego, dla określonej warstwy sieci. Doktorantka przytacza wyrażenie (5.1) jako komponent funkcji celu stosowany do oceny podobieństwa treści, tymczasem w materiale źródłowym jest on wskazany jako komponent używany do oceny rekonstrukcji komponentu ‘stylu’ (liczba ‘4’ w mianowniku wzoru (5.1) wynika z iloczynowego charakteru komponentów budujących wyrażenie oceniające styl).

W części dotyczącej detekcji ataków prezentacji, Doktorantka informuje o dwóch zastosowanych przez Nią, różnych podejściach: analizie statystycznej lokalnej tekstury obrazu z użyciem trzech różnych deskryptorów (LBP, BSIF i częstotliwościowej) oraz analizie z użyciem klasyfikatorów konwolucyjnych. Moim podstawowym zastrzeżeniem wobec przedstawionej przez Doktorantkę metodyki detekcji ataków prezentacji jest całkowite pominięcie kwestii kompresji obrazów i ich ewentualnego wpływu na uzyskiwane wyniki detekcji. Kompresja wprowadza do mikrostruktury obrazu powtarzalne artefakty, które być może, stanowią kluczowy czynnik w podejmowaniu decyzji o prawdziwości lub fałszywości prezentowanego obrazu. Dlatego też, przeprowadzając ataki prezentacji, należałoby co najmniej podzielić dysponowane obrazy na dwie kategorie: poddawanych kompresji stratnej (najpowszechniejsza to oczywiście format ‘jpg’), należących do wskazanej przez Doktorantkę kategorii najmniej zaawansowanych sposobów ataku, oraz niepoddawanych kompresji lub poddawanych kompresji bezstratnej, co odpowiada bardziej zaawansowanemu atakowi. Pomijając wątek wpływu kompresji, Doktorantka naraża się na wyciąganie z przeprowadzanych przez siebie eksperymentów fałszywych wniosków – na przykład, o wysokiej skuteczności ataków przeprowadzanych z użyciem obrazów generowanych metodą transferu stylu, wynikających być może z faktu braku artefaktów kompresji metodą jpg.

Przedstawiona przez Doktorantkę weryfikacja eksperymentalna sprawdzanych przez Nią metod detekcji mogłaby być obszerniejsza – jako miarę oceny przedstawia współczynnik fałszywego odrzucenia próbki dla jednego (wybranego arbitralnie?) poziomu współczynnika fałszywej akceptacji (10%). Bazując na

uzyskanych wynikach, przedstawia w Podsumowaniu pracy zupełnie nieuprawnioną, w świetle przytoczonych rezultatów, informację o osiągnięciu za pomocą swojej metody poprawności detekcji ataku prezentacji na poziomie 99%. Nie mam pojęcia, jak Doktorantka policzyła ten wskaźnik, chyba że dla przedstawionego przypadku współczynnik fałszywej akceptacji był bliski zeru (dlaczego więc nie przedstawiła ilościowego podsumowania tak znakomitych efektów w postaci macierzy pomyłek?).

Dokonania Doktorantki w obszarze rozwoju metod detekcji ataku prezentacji, mają w mojej ocenie, charakter implementacyjny – z powodzeniem wdraża metody generacji skomplikowanych treści oraz różne algorytmy binarnej klasyfikacji obrazów. W przedstawionym materiale nie znajduję jednak wystarczająco jasno opisanych pomysłów, które mógłbym uznać za naukowo znaczące z perspektywy rozwoju metod biometrycznej analizy danych.

## 4. Uwagi dodatkowe

W przedstawionej rozprawie znajduje się niewielka liczba dostrzeżonych przeze mnie usterek technicznych. Pierwsza z nich pojawia się w zdaniu definiującym cel pracy:

*The aim of this doctoral study is to investigate factors that may influence existing palmprint recognition algorithms and to propose methods taking these changes into consideration and enabling correct recognition, including presentation attack detection.*

gdzie prawdopodobnie słowo ‘factors’ zostało omyłkowo zastąpione słowem ‘changes’. W prezentacji stanu wiedzy Doktorantka wprowadza taksonomię metod segmentacji, wyróżniając dwie kategorie metod ‘klasycznych’: ‘bazujące na progu’ i ‘bazujące na teksturze’, co jest niepotrzebną próbą redefinicji istniejącej od dawna, powszechnie przyjętej systematyzacji podejść do segmentacji na metody ‘punktowe’ i ‘obszarowe’ (‘point-wise, region-wise’).

W prezentowanych przez Doktorantkę wzorach 3.2-3.4 brakuje indeksu G przy wartości średniej. W opisie metody segmentacji semantycznej: *„fully convolutional network (FCN) [27], that uses blocks of convolution and maxpooling layers to first decompress an image to 1/32th”*, Doktorantka omyłkowo używa słowa decompress zamiast compress. Ostatnie dwa akapity podsumowania Rozdziału 3 są zupełnie niepotrzebne – pierwszy powiela informacje ze wstępu, drugi – powiela informacje z wcześniejszej części podsumowania. Zamieszczona na stronie 49 referencja [53] jest wskazana jako źródło informacji o metodzie ekstrakcji obszaru zainteresowania dłoni, tymczasem dotyczy rozpoznawania tęczówki. Wreszcie, wydaje się, że wektory Mm* na rys. 4.9 są przedstawione w odwrotnej orientacji (rozumiem, że wartości przypisane wierszowi to maksymalna wartość mapy ‘atencji’, co jest sprzeczne z zawartością rysunku).

## 5. Wniosek końcowy

W podsumowaniu niniejszej recenzji chciałbym stwierdzić, że przedstawiona praca zawiera oryginalne i wartościowe koncepcje, stanowiące zauważalny wkład do dziedziny biometrycznej analizy danych, pozyskiwanych w realistycznych warunkach z użyciem urządzeń mobilnych. Zaproponowane i zweryfikowane przez Doktorantkę metody analizy obrazów, ukierunkowane na rozpoznawanie osób na

podstawie struktury linii papilarnych spodu dłoni, mają silny wymiar praktyczny i mogą stanowić elementy przydatne dla tworzenia zaawansowanych systemów biometrycznych.

Konkludując recenzję, chciałbym stwierdzić, że rozprawa doktorska Pani magister inżynier Eweliny Bartuzi-Trokielewicz pt. „Presentation attack-resistant palm recognition for mobile devices in unconstrained conditions" **spełnia** w moim przekonaniu wymagania określone w odnośnej ustawie o stopniach i tytule naukowym i tym samym **wnioskuję o jej dopuszczenie do publicznej obrony**.

Literatura

[1] A. Urooj, A. Borji, Analysis of hand segmentation in the wild, in: IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 4710–4719.

[2] ] W. Wang, Y.u. Kaicheng, J. Hugonot, Pascal Fua and Mathieu Salzmann, Recurrent U-Net for Resource-Constrained Segmentation, ICCV, 2019

[3] https://google.github.io/mediapipe/solutions/hands