

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

Recenzja rozprawy doktorskiej

Jalil Nourmohammadi Khiarak

Praca pt.

Countermeasure algorithms against subterfuge in mobile biometric systems

1. Tematyka, struktura i kontekst rozprawy

Przedstawiona do recenzji rozprawa doktorska dotyczy zagadnienia biometrycznej weryfikacji użytkownika dokonywanej na podstawie analizy kształtu małżowiny usznej oraz, stanowiącej naturalne uzupełnienie tego wątku, problematyki detekcji prób tzw. ataków prezentacji (ang. Presentation Attack Detection – PAD) polegających na przedstawianiu sensorowi spreparowanych danych. Obydwa przedstawione kierunki prac są osadzone w kontekście procedury uwierzytelniania użytkowników urządzeń mobilnych, mogąc stanowić alternatywę lub uzupełnienie innych powszechnie stosowanych metod biometrycznych, bazujących na analizie obrazu twarzy lub linii papilarnych. Istotnym wyróżnikiem uzasadniającym celowość podjęcia prac w rozważanym obszarze jest uniwersalność implementacji wynikowych metod w środowisku dowolnego współczesnego urządzenia mobilnego (smartfona), z uwagi na nieograniczany dostęp do sensorów rozważanych przez Doktoranta modalności (kamera i ekran dotykowy). Podjęte przez Doktoranta badania mają bardzo aktualny i użyteczny wymiar, a postawione problemy: zarówno poprawna weryfikacja rejestrowanej struktury ucha jak i skuteczna ochrona przed atakami prezentacji, należą do zadań nietrywialnych. W konsekwencji, należy jednoznacznie stwierdzić, że podjęte zagadnienie badawcze odpowiada swoją złożonością i zasadnością wymaganiom stawianym przed rozprawami doktorskimi, a tematyka prac znajduje się w obszarze dyscypliny naukowej informatyka techniczna i telekomunikacja.

Struktura rozprawy jest przejrzysta – w kolejnych rozdziałach Doktorant opisuje prace wykonane w ramach różnych wątków podjętej przez siebie tematyki. Rozprawa rozpoczyna się od prezentacji opracowanej przez Doktoranta bazy danych (część 2). Następnie Doktorant przedstawia zastosowane metody biometrycznej weryfikacji użytkownika na podstawie obrazu ucha i powiązane z tym metody detekcji ataków prezentacji (część 3). Oryginalna propozycja wykorzystania jako biometryki zbioru punktów styku ucha z powierzchnią ekranu smartfona, stanowiącego sensor dotyku jest treścią

czwartej część rozprawy. Dwa ostatnie rozdziały są poświęcone zagadnieniom związanym z weryfikacją tożsamości wykorzystującą dwie modalności powiązane z uchem: obrazów rejestrowanych kamerą smartfona i rozkładów punktów kontaktu ucha z powierzchnią ekranu smartfona. W części 5 pracy została przedstawiona autorska propozycja dwumodalnej metody weryfikacji tożsamości, zaś w rozdziale 6 – metoda dwumodalnej ochrony przed atakami prezentacji.

Przedstawione przez Doktoranta prace są właściwie osadzone w kontekście bieżących prac prowadzonych w obszarze biometrycznej weryfikacji tożsamości na podstawie obrazów ucha i metod detekcji ataków prezentacji na proces akwizycji rozważanej biometryki. Doktorant wykazuje dobrą orientację w aktualnym stanie wiedzy, trafnie identyfikując i korzystając w swoich rozwiązaniach z najnowszych metod i narzędzi algorytmicznych, w tym, zaawansowanych, głębokich architektur neuronowych, jak również poprawnie diagnozując ograniczenia istniejących zasobów danych, prowadzące do opracowanego przez Niego nowego repozytorium danych biometrycznych.

2. Ocena formalnej strony rozprawy

Pomimo nie budzącego zastrzeżeń, klarownego układu rozprawy i czytelnego sposobu prezentacji treści, techniczna strona rozprawy pozostawia sporo do życzenia. Pierwszym, drobnym ale utrudniającym jej śledzenie mankamentem jest użycie szablonu, który w sposób mało przejrzysty uwypukla podział tekstu na poszczególne moduły. Drugorzędny charakter, chociaż również utrudniający jej studiowanie, mają również niedociągnięcia językowe tekstu rozprawy. Język rozprawy nie jest językiem ojczystym Doktoranta, co stanowi oczywistą przeszkodę w klarownym formułowaniu myśli, ale na etapie przygotowania tekstu Doktorant mógł skorzystać z powszechnie dostępnych bezpłatnych narzędzi pozwalających na jego korektę.

Istotniejszym mankamentem formalnej strony pracy są liczne powtórzenia tych samych treści, wynikające ze sposobu przygotowania monografii – składa się ona z fragmentów tekstów publikacji współautorstwa Doktoranta, które jako źródła niezależne zawierały często identyczne lub bardzo zbliżone fragmenty. Całe serie akapitów powielają wcześniej przedstawione informacje dotyczące bazy danych, np. w części 3.3.2, w praktycznie całym wprowadzeniu do części 4, w części 5.4 lub na początku rozdziału 6.4. Również opisy stanu wiedzy w zakresie związanym z proponowanymi metodami się powtarzają (np. 3.2 i 6.1), tak jak i podawane kilkakrotnie wyjaśnienia używanych metryk. Testowaniem cierpliwości Czytelnika jest powtórzenie w rozdziale 5.5.1 i 5.5.2, treści przedstawionych wcześniej w rozdziałach przedstawiających wyniki analiz uzyskiwane osobno dla każdej z dwóch rozważanych modalności – obrazu ucha i punktów kontaktu ucha.

3. Cele i tezy rozprawy

Koncentrując swoje prace badawcze na problematyce biometrycznego uwierzytelniania użytkowników urządzeń mobilnych z wykorzystaniem specyficznej cechy osobniczej, jaką jest morfologia ucha, Doktorant formułuje w rozprawie cztery tezy ('Statements') identyfikujące osiągnięcia pracy i jednocześnie określające szczegółowe kierunki podjętych badań. Pierwsza z nich:

S1: We introduced and investigated on an ear-photo presentation attacks detection and recognition system on mobile device,

informuje o przeprowadzeniu prac nad rozpoznawaniem osób na podstawie obrazów ucha pozyskiwanych kamerą smartfona oraz nad detekcją ataków prezentacji dokonywanych z użyciem podstawionych w postaci obrazów danych. Druga teza:

S2. Ear-touch characteristic is an effective biometric, which is based on mobile devices with multiple touchscreens

dotyczy osiągnięć Doktoranta w zakresie biometrycznej analizy powierzchni kontaktu ucha i ekranu dotykowego smartfona. Trzecia teza:

S3. Ear photo and ear-touch multimodal biometrics improve ear recognition system accuracy in comparison with its single model,

informuje o możliwości zwiększenia poprawności klasyfikacji rozważanej biometryki z użyciem podejścia dwumodalnego, uwzględniającego analizę obrazu ucha i analizę rozkładu punktów kontaktu ucha z powierzchnią sensora. Ostatnia teza:

S4. Ear photo and ear-touch multimodal biometrics improve ear PAD accuracy concerning for to a single model.

sygnalizuje możliwość poprawy skuteczności detekcji ataków prezentacji z wykorzystaniem analizy informacji pochodzących jednocześnie z dwóch rozważanych modalności – obrazu i rozkładu punktów styku. Przedstawione wątki tematyczne w kompletny sposób rysują ramy zrealizowanych przez Doktoranta prac badawczych, uszczegółowionych w dalszej części rozprawy w metodycznie poprawny sposób i poprzedzanych kompetentną prezentacją aktualnego stanu prac dla każdego z rozważanych obszarów.

4. Merytoryczna ocena pracy

Przedstawione w rozprawie efekty prac badawczych Doktoranta mają w mojej ocenie bardzo zróżnicowany poziom naukowy – trzy z przedstawionych rozwiązań ograniczają się do wykorzystania szablonowej metodyki postępowania i nie stanowią żadnego wkładu o charakterze naukowym do dyscypliny, ale dwie to oryginalne, interesujące i wartościowe propozycje o dużym potencjale aplikacyjnym i badawczym. Zdecydowanie najważniejszym osiągnięciem Doktoranta jest nowatorska propozycja biometrycznej weryfikacji tożsamości wykorzystującej jako cechę biometryczną rozkład punktów styku ucha z powierzchnią ekranu dotykowego smartfona. Jest to pomysł zupełnie nowego, naturalnego, wygodnego i nie wymagającego świadomej kooperacji, sposobu uwierzytelnienia użytkownika, który dotychczas nie pojawił się w powszechnym obiegu. Z uwagi na pionierski charakter koncepcji, Doktorant stanął w obliczu konieczności zgromadzenia nieistniejącego w dostępnych zasobach zbioru danych niezbędnych do opracowania odpowiedniego algorytmu. W efekcie przeprowadzonych przez Doktoranta prac powstała unikatowa baza danych (WUT-Ear v1) stanowiąca kompleksowe repozytorium materiału eksperymentalnego wspierającego prace nad biometrią wykorzystującą kształt ucha, obejmująca oprócz rejestracji rozkładów punktów kontaktu ucha z sensorem, również obszerny zestaw obrazów ucha oraz równie obszerny zbiór przykładów ataków na procedurę weryfikacji, zarówno obrazów jak i sztucznie wygenerowanych rozkładów punktów kontaktu ucha z sensorem.

Ponieważ opracowanie bazy danych wymagało zastosowania metodyki badawczej: identyfikacji luk w istniejących repozytoriach, określenia scenariuszy akwizycji i scenariuszy ataków, statystycznej ewaluacji zgromadzonego materiału w celu oceny spełnienia stawianych przed nim wymagań oraz zapewnienia odpowiedniego zróżnicowania materiału eksperymentalnego, ten efekt prac Doktoranta, nie wymieniony jawnie w zbiorze kluczowych elementów dorobku rozprawy, uważam za istotne osiągnięcie mające znamiona naukowe. Jak wspomniałem, najcenniejszym, trzecim komponentem

opracowanej bazy danych, są unikatowe rejestracje położenia punktów styku powierzchni ucha z powierzchnią dotykowego sensora smartfona. Również zaproponowana metodyka ataków prezentacji i zgromadzony w tym celu materiał w czwartym komponencie bazy, są pomysłowe i wartościowe. Dwa pierwsze komponenty, zawierające materiał odnoszący się do rejestracji obrazów ucha są mniej oryginalne, chociaż przedstawione wizualizacje dystrybucji obrazów rzeczywistych i spreparowanych w celu przeprowadzenia ataku prezentacji, w przekonujący sposób uzasadniają wysoką wartość materiału eksperymentalnego przygotowanego przez Doktoranta.

W kolejnych podrozdziałach recenzji ustosunkowuję się do czterech wymienionych we wstępie pracy obszarów dokonań Doktoranta. W odniesieniu do materiału zawartego w kolejnych częściach rozprawy, chcę zwrócić uwagę na mankament metodyki ewaluacji przeprowadzanych eksperymentów, którym jest niekompletne statystyczne podsumowanie, pomijające informację o odchyleniach standardowych (lub przedziałach ufności) uzyskiwanych rezultatów. Przedstawiane na wykresach i w tabelach wartości średnie wyników eksperymentów nie dają pełnego poglądu o stabilności (wrażliwości) stosowanych metod w funkcji zmienności danych używanych do ich testowania, co ma znaczenie z punktu widzenia oceny ich przydatności.

4.1. Weryfikacja tożsamości z użyciem obrazów ucha

Pierwszym wątkiem prac poruszonych w rozprawie jest procedura weryfikacji tożsamości na podstawie obrazów ucha. Prace Doktoranta przedstawione w tej części rozprawy uważam za szablonowe i nie wnoszące żadnych nowych elementów do istniejącego stanu wiedzy. Doktorant używa znanych metodyk ekstrakcji cech: uczenie kontrastowe, stosowane dla budowy głębokich modeli przy skąpych zasobach przykładów treningowych oraz metoda LBP. W pierwszym przypadku, procedura analizy to użycie pretrenowanej sieci syjamskiej jako ekstraktora reprezentacji obrazu testowanego i wzorca reprezentującego weryfikowanego użytkownika. Dla obydwu metod ekstrakcji, wynik analizy jest wyznaczany poprzez ocenę podobieństwa wektorów cech, dokonywaną z użyciem odległości kosinusowej i konfrontowaną z wyznaczoną wcześniej eksperymentalnie wartością progową. Standardowa jest też metodyka ewaluacji działania algorytmu, a z jej wyników nie wynikają żadne zaskakujące wnioski. Na marginesie chcę zauważyć, że przy okazji prezentacji opracowanej przez siebie bazy danych, Doktorant kilkakrotnie podkreśla obecność etykiet płci jako istotny walor bazy danych, ale w przeprowadzonych eksperymentach nie podejmuje ciekawego wątku przetestowania hipotezy o potencjalnych różnicach w strukturze ucha kobiecego i męskiego.

Podobnie jak to miało miejsce w przypadku metody weryfikacji użytkownika, przedstawionej w części 3.3, również metoda detekcji faktu podstawienia w procesie akwizycji danych spreparowanych (PAD) wykorzystuje typowy schemat uczenia transferowego, bez żadnych ciekawych lub oryginalnych pomysłów autorskich. Podejście do budowy algorytmu weryfikacji jest schematyczne i pozbawione wnikliwości – Doktorant nie próbuje optymalizować architektury dodanych przez siebie warstw MLP, stosując arbitralnie przyjęte liczby neuronów w warstwach. Jest to mankamentem, bo MobileNetv2 w swojej bazowej wersji zawiera ok. 3.5 mln. parametrów, zaś przyjęta przez Doktoranta architektura dodanego, gęstego klasyfikatora reprezentacji generowanej przez ekstraktor to sporo ponad 2.5 miliona parametrów (ok. $1280 \times 1024 + 1024 \times 1024 + 1024 \times 512$). W konfrontacji ze stosunkowo skromnym zbiorem uczącym, skala niezbędnych do estymacji parametrów budzi podejrzenie nieuchronnego dopasowania sieci do zbioru danych, pochodzącego mimo starań Doktoranta, z bardzo specyficznego rozkładu prawdopodobieństwa. Brak refleksji nad

relacją liczby parametrów względem ilości dostarczonej informacji uczącej, uważam za uchybienie metodologiczne omawianego wątku pracy. Problem dysproporcji staje się jeszcze bardziej ewidentny w drugim przyjętym przez Doktoranta wariancie, arbitralnego zwiększenia głębokości douczania sieci MobileNetV2. Doktorant sugeruje tu zamrożenie tylko pierwszych 16 warstw, co oznacza dodanie do puli parametrów podlegających uczeniu kolejnych ponad 2.5 miliona elementów. W efekcie problem uczenia przed jakim stawia algorytm może być sprowadzony do zadania: określ ponad 5 mln parametrów na podstawie 8 tysięcy przykładów. Ponieważ w moim przekonaniu architektura sieci jest zbyt złożona w stosunku do posiadanej liczby przykładów, przedstawione próby wyjaśnienia wyników eksperymentów są niewłaściwie ukierunkowane i nie prowadzą do zasadnych wniosków.

4.2. Uwierzytlenie na podstawie rozkładu punktów kontaktu ucha z sensorem

Pierwszym elementem prezentacji opracowanych przez Doktoranta metod biometrycznych, który stanowi oryginalne podejście do rozwiązania złożonego problemu, jest część poświęcona biometrycznej analizie zbiorów punktów styku ucha z powierzchnią sensora – ekranu smartfona. Istotą rozwiązywanego problemu jest znalezienie transformacji najlepiej dopasowującej zarejestrowany zbiór punktów kontaktu do utworzonych w fazie treningu algorytmu wzorców, a następnie, podjęcie decyzji o pozytywnej lub negatywnej weryfikacji próbki. Istnieje kilka metod realizacji tak sformułowanego zadania, o których Doktorant nie wspomina w przeglądzie metod, skupiającym się na kompleksowych algorytmach biometrycznych, a nie na metodach rozwiązania komponentów składowych. Dla porządku więc, należy wskazać, że adekwatnymi metodami mogą tu być tzw. Geometric Hashing, metoda RANSAC czy metoda ICP (Iterative Closest Point). Przedstawione w pracy podejście jest bliskie ostatniej z wymienionych metod, przy czym Doktorant nie przedstawia szczegółów użytego algorytmu wyznaczania parametrów transformacji położeń zbioru dopasowywanych punktów, a jedynie wskazuje na źródło – algorytm opracowany przez Jego Promotora, do którego niestety, nie byłem w stanie dotrzeć (szkoda, że nie pojawił się w załączniku). Autorska modyfikacja tego algorytmu to ekstensywne powtarzanie procedury estymacji macierzy obrotu i wektora translacji dla wszystkich możliwych permutacji zbioru punktów styku. Oprócz metodyki analizy nieznanego rozkładu punktów styku, Doktorant proponuje również metodykę budowy wzorców klas, które są wyznaczone jako regresory zbioru przykładów treningowych klasy. Metoda zaprezentowana w pierwszej części rozdziału 4 jest następnie uogólniona przez Doktoranta na bardziej realistyczny i zarazem trudniejszy przypadek, w którym liczba punktów styku dla różnych rejestracji w zbiorze treningowym jest zmienna, a ich ‘pochodzenie’ podobnie jak poprzednio - nieznanne. W tym miejscu należy zauważyć, że Doktorant nie ułatwia śledzenia swoich wywodów, bo definicję wielkości E , kluczowej dla zrozumienia treści części 4.3.3, zamieszcza dopiero w części 4.3.4. Mimo, że przyjęta przez Doktoranta metoda ma charakter ekstensywny, liczba rozważanych punktów styku jest na tyle mała, że nie stanowi to istotnej bariery obliczeniowej.

Podsumowując ocenę Rozdział 4.3, uważam go zdecydowanie za najbardziej wartościową część pracy, w której Doktorant przedstawia metodę rozwiązania złożonego problemu dopasowania do siebie niekompletnych podzbiorów punktów, pochodzących z tego samego lub z różnych rozkładów i poddawanych rotacjom oraz translacjom z dodatkowym komponentem szumu. Tak sformułowany problem realistycznie odzwierciedla uwarunkowania biometrycznej analizy rozkładów punktów styku ucha z powierzchnią sensora. Metodyka rozwiązania problemu zaproponowana przez

Doktoranta jest poprawna (choć oczywiście, nie jedyna) i oryginalna, a z uwagi na niewielkie liczności rozważanych zbiorów punktów, całkowicie akceptowalna z punktu widzenia złożoności obliczeniowej. Dodatkowo, przydatne z praktycznego punktu widzenia i wykraczające poza standardową ewaluację metody, jest przeprowadzenie analizy poprawności weryfikacji w zależności od liczby posiadanych punktów styku używanych w budowie wzorców.

4.3. Dwumodalna weryfikacja użytkownika

Przedstawione w części 5 pracy dwumodalne podejście do rozpoznawania ucha ma wymiar nowości jedynie z punktu widzenia rodzaju wykorzystywanych źródeł informacji – uwzględnienie obok obrazu, wiedzy o punktach styczności ucha z ekranem dotykowym, wprowadza oryginalny, nowy komponent do arsenału rozważanych dotychczas metod. Na tym jednak oryginalność podejścia się kończy – przyjęta metodyka weryfikacji bazującej na fuzji informacji dokonywanej na etapie podejmowania decyzji jest standardowa, a sposoby dokonywania indywidualnych analiz danych pochodzących z obydwu źródeł w niczym nie wykraczają poza przedstawione wcześniej w literaturze podejścia. Z pewnością na uwagę zasługują raportowane przez Doktoranta efekty zastosowania proponowanego podejścia, stanowiąc potwierdzenie przydatności pomysłu wykorzystania obszarów kontaktu ucha z sensorem dla analizy biometrycznej. Niedosyt pozostawia zaproponowana przez Doktoranta arbitralna metodyka fuzji wyników cząstkowych i brak próby uwzględnienia narzucających się dla stosowanej strategii mechanizmów dywersyfikacji znaczenia informacji pochodzącej z obydwu modalności. Z punktu widzenia badawczego, trudne do akceptacji jest przyjęte z góry założenie, że wynik weryfikacji na podstawie obrazu jest ważniejszy od wyniku uzyskanego dla analizy punktów styku ucha. Oczekiwanym podejściem, które niestety nie zostało podjęte, jest łatwa do realizacji próba oceny jakości danych obrazowych (np. przez prostą ocenę kontrastu lub rozmycia) jako przesłanki do określenia znaczenia tych danych w dokonywanej na etapie podejmowania decyzji syntezie.

4.4. Ochrona przed atakami prezentacji z wykorzystaniem dwóch źródeł informacji

Ostatnia część pracy jest poświęcona problemowi dwumodalnej detekcji ataków prezentacji wykorzystującej osobne analizy danych pochodzących z rozważanych źródeł i fuzję uzyskanych wyników analizy. W przeciwieństwie do metodyki przyjętej dla oceny efektów dwumodalnego rozpoznawania, tym razem Doktorant wzbogaca swoje prace używając dla PAD w odniesieniu do obrazów ucha nowego algorytmu klasyfikacji: transformera (a nie jak, w przypadku PAD dla samych obrazów – sieci konwolucyjnej). Metodą stosowaną w odniesieniu do detekcji ataków prezentacji na podstawie rozmieszczenia punktów styku ucha staje się metoda XGBoost. Ponieważ obydwie metody klasyfikacji są od siebie całkowicie niezależne aż do momentu agregacji wyników końcowych, rodzi się pytanie o zasadność umieszczenia rozważań dotyczących metod w części 6, a nie w częściach 3 i 4, poświęconych zagadnieniom PAD dla każdej z modalności z osobna. Przedstawione w rozdziale wyjaśnienia pozostawiają niedosyt: opis koncepcji transformera DeIT, przedstawiony w p. 6.3.1 jest chaotyczny i zawiera irytujące powtórzenia tych samych treści. W części dotyczącej ewaluacji, Doktorant nie zdradza szczegółów procesu dostrajania transformera (jakie moduły i ile parametrów zostały użyte w douczaniu?).

Rozważenie przez Doktoranta scenariusza z podwójną ścieżką analizy oryginalności danych jest uzasadnione, a uzyskiwana poprawa skuteczności obrony przed takimi atakami stanowi

potwierdzenie przydatności hipotezy sformułowanej przez Doktoranta. Na uznanie zasługuje ostrożność Doktoranta w formułowaniu wniosków wpływających z jego eksperymentów – zdaje sobie sprawę ze szczupłości posiadanego przez siebie zbioru danych, więc przyjmuje prawdopodobne pogorszenie efektów rozpoznawania w konfrontacji z próbkami z innych źródeł. Jednakże z punktu widzenia naukowego, poza samym pomysłem dodania jako źródła informacji o potencjalnym ataku danych z sensora dotyku, pozostała część metodyki nie wykracza poza utarte standardy dziedziny.

5. Uwagi szczegółowe

W pracy znajdują się niejasności wymagające doprecyzowania i usterki o różnym charakterze, które przedstawiam w formie poniższej, być może niekompletnej listy:

22. Nie rozumiem zdania: „*As previously said, we have classified PAD approaches into models inside faces, which include ear-touch, and ear photo biometrics*”

27: Zdanie „*The dataset is specifically designed for research in ear verification and PAD, offering a balanced representation of genders ...*” nie ma pokrycia w przytoczonych danych, z uwagi na proporcje 4:1 między przedstawicielami płci.

29: Niejasne jest znaczenie jednostki ‘unit’ (nie została ona wyjaśniona) stosowanej w opisie statystyk bazy obrazów uszu, zamieszczonych w Tabeli 2-3. Czy przedstawione informacje dotyczą liczby zdjęć przypadających na jednego użytkownika w przetworzonym zbiorze danych?

str. 30: Nie rozumiem zdania: „*The dataset contains a relatively distribution of photos across subjects, which can be useful for analyzing trends and patterns in the data.*”

str. 31: W zdaniu: „*The dimensions of the preprocessed images remained the same as their raw counterparts.*” Autor ma chyba na myśli proporcje (aspect ratio).

str. 31: Jaki ‘model’ został użyty do ekstrakcji wektorów cech, które są przedstawione na rys. 2-5?

35: Nie bardzo widzę potrzebę istnienia tabeli 2-5 – nie ma w niej niczego istotnego ponad informacje podane w tabeli 2-4.

38 Zdanie: „*The average distances for printed photos are considered as genuine images (~15cm)*” miało pewnie informować o tym, że odległość zdjęcia była podobna do odległości autentycznie autoryzowanej osoby.

40: Nie rozumiem fragmentu opisującego zawartość komponentu zawierającego punkty styczności ucha z ekranem: „*there were 92 subjects – 72 men and 20 women. If two ears (left and right ear) were taken for each subject, we would have a total subject of 92.*”

41: Wyjątkowo niezręczne sformułowanie myśli: „*According to our data collection experience, we could record ear-touches with fewer touch points (less than four touch points); hence, we captured ear-touches with at least four touch points.*”

str.43: Wyjątkowo niezręczne sformułowanie myśli: „*To make fake data four hands (Two a man hands and two a woman hands) were used.*”

str.55: Wbrew temu co pisze Autor: „*To train a Siamese network using VGG16 for ear photo recognition, we follow a systematic approach ... used for MobileNetV2, but with certain key differences due to the specific characteristics of the VGG16 architecture.*” – w przedstawionej procedurze nie ma żadnych istotnych różnic.

60: „*Figure 3-5: The flowchart of ear photo PAD using a deep neural network structure*” - nie bardzo wiadomo, co dokładnie przedstawia schemat blokowy: argumentem oceny są jednocześnie wektory cech i augmentowane obrazy?

61: „*The layers highlighted in orange represent the added top layers.*” – nie widzę na rysunku niczego wyróżnionego kolorem pomarańczowym.

str.80: „*... then $D1$ can be reformulated as $D2$ in Eq.6 to ...*” - złe odwołanie, równanie 6 nie zawiera wspomnianych symboli.

str.80: Kolejne błędne odwołanie - „*The definition of E is presented in Eq.7...*” wyrażenie 7 nie ma związku z definicją E (definicja pojawia się w równaniu 9).

str.81: Dla pracy naukowej w obszarze informatyki, dość amatorsko wygląda szacowanie złożoności algorytmu opisywane w sposób prezentowany przez Doktoranta, np. „*... exceeds 10, the computational complexity becomes*

prohibitive due to the factorial growth in the number of permutations ... This exponential increase in permutations would make the algorithm computationally infeasible, ... Analiza złożoności powinna być precyzyjna, czas obliczeń – podany jawnie (liczba 10 na pewno nie jest granicą, powyżej której obliczenia nie są możliwe, więc lepiej było określić próg złożoności przyjętej przez Doktoranta).

str.84: Miara FMR to to samo co FAR, po co więc definicja FAR?

str.85: „*We used single enrollment and multi enrollment ...*” - Czy single-enrollment to budowa wzorca dla pojedynczej rejestracji?

99. Co to jest „*VGG-16 SiamNet fusion model*”. Nie wiadomo, czy jest to jakaś osobna sieć, czy jest to po prostu wyznaczenie ważonej sumy wyników generowanych przez sieć syjamską dla obrazu i przez moduł oceny podobieństwa punktów styku.

103: „*The quality of ... ear spoofing attacks might be because of ...*” - zgubiony wyraz?

136 (Appendix B) Autor powiela błąd przedstawiony w swojej publikacji – zamiast ‘best match of X w.r.t T’, mamy ‘best match of X w.r.t X’

5. Wniosek końcowy

Podsumowując opinię o przedłożonej do recenzji rozprawie doktorskiej, chciałbym stwierdzić, że spełnia ona warunki określone dla rozpraw doktorskich przez odnośne przepisy. W konsekwencji tej konkluzji, **wniosuję o dopuszczenie Autora rozprawy do dalszych etapów przewodu doktorskiego.**