

Warszawa, 4 listopada 2025 roku

Prof. dr hab. inż. Jacek Mańdziuk
Wydział Matematyki i Nauk Informatycznych
Politechnika Warszawska

Recenzja rozprawy doktorskiej mgr inż. Macieja Szymkowskiego zatytułowanej
*Uczenie maszynowe oraz sztuczna inteligencja w trimodalnym systemie
biometrycznym bazującym na odcisku palca, jego układzie żył i geometrii*

Recenzja została przygotowana na prośbę Przewodniczącego Rady Naukowej Wydziału Informatyki Politechniki Białostockiej, prof. dr hab. inż. Marka Krętowskiego, wyrażoną w piśmie z dnia 3 lipca 2025 roku.

Tematyka rozprawy

Przedstawiona do recenzji rozprawa dotyczy zagadnień teoretycznych oraz praktycznej implementacji metod biometrycznych, ze szczególnym uwzględnieniem rozwiązań wielomodalnych. Dostęp do różnego rodzaju zasobów czy przestrzeni coraz częściej weryfikowany jest w drodze analizy cech biometrycznych danej osoby. Najpopularniejsze metody weryfikacji dostępu bazują na identyfikacji osoby na podstawie zdjęcia twarzy, odcisku linii papilarnych palca czy skanu tęczówki oka. Wskazane metody bazują na unikalności powyższych cech w ramach całej populacji i – w sensie teoretycznym – zapewniają stuprocentową skuteczność weryfikacji. W praktyce, jednakże, reprezentacja tych cech może nie być doskonała (np. z uwagi na problemy z urządzeniem weryfikującym czy niewłaściwy kąt pomiaru danej cechy). Dodatkowo, istnieje duża grupa metod tzw. *spoofingu*, czyli „podrabiania” cech biometrycznych. W tym kontekście coraz bardziej pożądane stają się metody wielomodalne, tzn. bazujące na więcej niż jednej cesze biometrycznej.

W ramach przeprowadzonych badań Doktorant zaproponował autorski system weryfikacji biometrycznej bazujący na trzech cechach: odcisku palca, układzie żył w palcu i geometrii palca oraz zweryfikował jego skuteczność na grupie 100 osób. Kluczowym elementem zaproponowanego rozwiązania jest urządzenie zbudowane w oparciu o platformę Raspberry Pi 4, pobierające obrazy trzech ww. cech, które są następnie analizowane metodami uczenia maszynowego.

Hipoteza badawcza

W rozprawie Autor stawia i weryfikuje eksperymentalnie następującą tezę badawczą: *„Bazując na minucjach znajdujących się na odcisku palca oraz w układzie żył w palcu, a także geometrii palca, można przeprowadzić dokładne i skuteczne (z progiem minimum 95%, przy bazie użytkowników nie mniejszej niż 100 osób), rozpoznanie tożsamości człowieka przy użyciu systemu multimodalnego bazującego na platformie Raspberry Pi 4”.*

Treść rozprawy

Rozprawa liczy 105 stron, składa się z 7 rozdziałów oraz spisu literatury zawierającego 97 pozycji.

W rozdziale pierwszym, stanowiącym wprowadzenie, Autor zwraca uwagę na praktyczną istotność podjętej tematyki badawczej w kontekście rosnących potrzeb w zakresie skutecznych sposobów ochrony dostępu. Następnie Autor formułuje tezę rozprawy w postaci zacytowanej powyżej, omawia układ pracy oraz przedstawia listę swoich publikacji w obszarze biometrii oraz innych obszarach badań.

Mam dwie uwagi dotyczące tego rozdziału. Po pierwsze, sformułowanie tezy jest moim zdaniem niepoprawne. W rozprawie Autor pokazuje skuteczność swojej metody na poziomie co najmniej 95% dla **konkretnego zbioru 100 osób** wybranych do testów. Ten wynik w żaden sposób nie przekłada się na ogólne stwierdzenie o poziomie skuteczności co najmniej 95% dla **dowolnej grupy złożonej ze 100 osób**, a tym bardziej dla **grup o większej liczebności**. Uwaga druga dotyczy listy publikacji – nie znalazłem w niej wyróżnienia prac, która stanowią bezpośrednią podstawę przedłożonej rozprawy. Takie wskazanie jest oczekiwane z formalnego punktu widzenia, a także bardzo pomocne w ocenie dysertacji.

W rozdziale drugim mgr. Maciej Szymkowski przedstawia bardzo szczegółowy rys historyczny badań biometrycznych, zaczynając od pierwszych zachowanych śladów prawdopodobnego wykorzystania biometrii (ok. 31 tysięcy lat temu), do dnia dzisiejszego, w którym biometria stanowi formalną domenę badań naukowych. Autor dokonuje klasyfikacji cech biometrycznych dzieląc je na fizjologiczne, behawioralne i hybrydowe – w tej ostatniej grupie wyróżniając jako jedyną cechę *głos*. W dalszej części rozdziału przedstawione są skrótkowo główne metody fałszowania cech biometrycznych oraz poruszone są aspekty prawne związane z wykorzystaniem biometrii.

Rozdział napisany jest starannie i nie mam uwag merytorycznych do tej części rozprawy. W sensie językowym moją uwagę zwróciło nadużywanie słów „skorelowany”, „solucja”, „rzeczony” i in., które w języku powszechnym raczej są wykorzystywane stosunkowo rzadko, głównie w celu podkreślenia istotności przekazu. Odnosząc się do kwestii językowych, w dysertacji Autor wielokrotnie błędnie używa słowa „vide” („patrz”) w znaczeniu „czy też”, np. w zdaniu *„Drugim obszarem badawczym jest implementacja algorytmów oraz modeli sztucznej inteligencji vide uczenia maszynowego ...”*. Inny przykład: *„... użytkownicy będą reprezentowani przez zbyt małą ilość skanów (vide nagrań)”*.

Rozdział trzeci omawia stan wiedzy w zakresie multimodalnych systemów biometrycznych, czyli podstawowego obszaru badań Autora zaprezentowanych w rozprawie. Wymienione są możliwe podejścia do zapewnienia różnorodności (multimodalności) próbek, bazujące na wykorzystaniu: (a) różnych cech, (b) różnych urządzeń w procesie akwizycji danych, (c) różnych próbek określonej cechy biometrycznej (np. odciski wielu palców), (d) wielu próbek tej samej cechy (np. wielokrotny odcisk danego palca), (e) wielu aspektów danej cechy (np. minucje odcisku palca w połączeniu z analizą kierunku rozchodzenia się krawędzi). Następnie Autor opisuje kolejno metody biometryczne bazujące na odcisku palca, bazujące na układzie żył w palcu, oraz wykorzystujące geometrię dłoni i palców.

Rozdział kończy przegląd literatury w kontekście systemów multimodalnych oraz dyskusja odnośnie wykorzystania metod sztucznej inteligencji i uczenia maszynowego w biometrii.

Rozdział napisany jest starannie i stanowi dobry przegląd zagadnień oraz metod związanych z wielomodalnymi systemami biometrycznymi. Na str. 47, autor przywołuje jeden z rodzajów sieci neuronowych, tzw. sieci Kołmogorowa-Arnolda, błędnie nazywając je sieciami Kołmogorowa-Antona.

Rozdział czwarty przedstawia autorski system biometrii multimodalnej bazujący na trzech cechach pobieranych z palca: odcisku palca, jego układzie żył oraz geometrii. Na poziomie funkcjonalnym można wyróżnić dwa główne elementy systemu: (1) urządzenie wykonane przez Doktoranta do akwizycji powyższych danych oraz (2) algorytmy uczenia maszynowego umożliwiające ekstrakcję informacji z pobranych cech biometrycznych oraz algorytmy klasyfikacji w oparciu o zebrane próbki danych.

Stanowisko badawcze zostało zbudowane w oparciu o platformę Raspberry Pi, której wybór (spośród kilku rozważanych opcji) Doktorant szczegółowo uzasadnia. Akwizycja próbek polegała na pięciokrotnym poborze danych od każdego ze 100 uczestników eksperymentów, prowadząc do zbioru 15 odczytów per osoba. Grupa osób biorących udział w eksperymencie była zróżnicowana wiekowo (od 22 do 71 lat), 40% stanowiły kobiety.

Pojedynczy eksperyment rozpoznawania użytkownika składał się z następujących kroków: pobrania próbki biometrycznej (obrazu), przetworzenia obrazu (m.in. jego segmentacji), ekstrakcji cech (metodami uczenia maszynowego), podjęcia decyzji (klasyfikacji za pomocą metod uczenia maszynowego). W ramach serii eksperymentów zbadano dwa główne warianty wykorzystania zaproponowanego systemu: (a) niezależna weryfikacja użytkownika w kontekście każdej z trzech uwzględnionych cech biometrycznych, a następnie głosowanie większościowe (w przypadku trzech różnych wskazań odpowiedź systemu była negatywna, tzn. nie wskazywano żadnego użytkownika), (b) dla każdej z trzech cech biometrycznych niezależne pobranie i przetworzenie próbek oraz ekstrakcja cech, konkatenaacja wyekstrahowanych cech do postaci wspólnej reprezentacji, klasyfikacja w oparciu o wytworzone wektory cech.

Mam następujące pytania do omawianego rozdziału:

- (1) Schemat blokowy metody przetwarzania i detekcji cech w obrazie odcisku palca (rys. 4.9) wygląda na standardowy. Czy są w nim jakieś elementy autorskie, tzn. unikalne w kontekście literatury tematu? Analogiczne pytanie odnośnie przetwarzania próbek dotyczących pozostałych dwóch cech.
- (2) Autor definiuje $w_f(T)$ oraz $w_b(T)$ jako „*sumy prawdopodobieństw wartości pixela ...*”. Po czym realizowane jest to sumowanie i jaka jest relacja pomiędzy $w_f(T)$ oraz $w_b(T)$?
- (3) Z czego wynika „przewaga” funkcji Focal Loss (FL) w stosunku do Cross Entropy w omawianym zagadnieniu (str. 67)? Jakie specyficzne własności FL mają tutaj znaczenie?
- (4) Na str. 69 Autor wskazuje, że podczas analizy odcisku palca wzięto pod uwagę dwa rodzaje minucji: zakończenia oraz rozwidlenia. Dlaczego właśnie te dwie formy minucji zostały wybrane i na ile metoda jest wrażliwa na ten wybór?
- (5) Rozdzielczość pobieranych obrazów była stosunkowo mała – 320x240 (str. 74). Na ile metoda jest wrażliwa na dobór rozdzielczości i jak jej zwiększenie / zmniejszenie wpłynęłoby na skuteczność systemu?

- (6) Na str. 76 Autor stwierdza, że „Założono, że suma parametrów (mnożników) nie może przekroczyć wartości jednostkowej”, natomiast z opisu, w szczególności z wzoru (4.11) wynika, że suma ta była dokładnie równa 1.
- (7) Jako klasyfikatora Autor użył sieci neuronowej z 5 warstwami. Jak zdefiniowana była architektura sieci? Dlaczego akurat taką liczbę warstw wybrano i jakie były liczności poszczególnych warstw?

Rozdział piąty przedstawia szczegółowe wyniki eksperymentów. Tabele 5.1-5.6 pokazują kolejno rezultaty różnych ustawień testowych proponowanego systemu identyfikacji biometrycznej. W szczególności, dla wybranej 100-osobowej grupy uczestników eksperymentu, uzyskana skuteczność identyfikacji wynosi 99,2%.

Muszę przyznać, że lektura tego rozdziału wzbudziła u mnie różne wątpliwości. Chętnie usłyszę wyjaśnienia Autora w kontekście poniższych pytań:

- (1) Dlaczego wybrano liczbę 100 uczestników i jaki jest poziom statystycznej istotności zaprezentowanych rezultatów?
- (2) Autor wprowadza podział zbioru na trzy części: treningową, testową i ewaluacyjną. Zakładam, że zgodnie ze standardową nomenklaturą chodzi o zbiory: treningowy, walidacyjny i testowy.
- (3) Wyjaśniając znaczenie wzorów (5.3) i (5.4) Autor stwierdza, że „TP – są to próbki prawidłowo sklasyfikowane (system zwraca dokładnie taką samą informację o tożsamości, jak informacja znajdująca się w bazie danych, FP – są to próbki fałszywie pozytywne, czyli takie w których system przypisuje tożsamość do próbki, pomimo, że użytkownik nie znajduje się w bazie danych, FN – są to próbki fałszywie negatywne – czyli system odrzuca przedstawiony wzorzec biometryczny pomimo tego, że użytkownik znajduje się w bazie danych”. Takie podejście jest niepoprawne metodologicznie i stanowi przykład „porównywania gruszek z jabłkami”. Z jednej strony mamy oczekiwanie prawidłowej identyfikacji konkretnego użytkownika (w TP), a z drugiej w FP i FN mówimy jedynie o przynależności do grupy użytkowników odpowiednio nie znajdujących się i znajdujących się w bazie. Jak zatem zakwalifikujemy sytuację, w której system wskazuje użytkownika A (który jest w bazie) a prawidłowe wskazanie to użytkownik B (także znajdujący się w bazie)?
- (4) Na str. 86 Autor stwierdza, że dla każdej z trzech cech pewność klasyfikacji definiowana jest poprzez normalizację wyjść warstwy wynikowej (funkcja softmax) do przedziału [0,1]. Następnie wyniki te (względem każdej z trzech cech) ważone są odpowiednio wagami 0.4, 0.4, 0.2 i porównywane z progiem na poziomie 0.6. Jaka jest przewaga tego systemu w stosunku do zwykłego głosowania większościowego?

Rozdział szósty poświęcony jest wykorzystaniu technologii haptycznych (dotykowych) w biometrii. W skrócie, zadanie polega na identyfikacji użytkownika na podstawie reakcji na odczucia dotykowe generowane przez odpowiednie urządzenie. Pytanie, które Autor stawia dotyczy możliwości wykorzystania wzorca takiej reakcji do rozpoznania danej osoby.

Autor przedstawił wyniki wstępnych eksperymentów na ograniczonej grupie testowej składającej się z 10 osób, określając je jako obiecujące. System mierzący reakcję emocjonalną twarzy, zmianę rytmu serca oraz ruchy dłonią -- jako odpowiedź na pobudzenie dłoni sygnałem z urządzenia haptycznego, daje „satysfakcjonujące rezultaty” przy niewielkiej liczbie osób w bazie. Przedstawione

badania są na bardzo wstępnym etapie i nie umożliwiają jeszcze wyciągania wiążących wniosków. Autor planuje je kontynuować, poszerzając grupę testowanych osób.

Rozdział siódmy stanowi krótkie podsumowanie dysertacji. Autor streszcza najistotniejsze cechy wprowadzonej przez siebie metody oraz uzyskane rezultaty i płynące z nich wnioski. Wskazuje też możliwe kierunki rozwoju przedstawionych w dysertacji badań. Rozprawę dopełnia spis literatury.

Oryginalny wkład Autora rozprawy

Oryginalny wynik badawczy mgr. Macieja Szymkowskiego dotyczy opracowania, implementacji oraz eksperymentalnej weryfikacji autorskiej metody tri-modalnej weryfikacji biometrycznej w oparciu o minucje znajdujące się na odcisku palca i układzie żył w palcu oraz geometrię palca. Warte podkreślenia jest zbudowanie specjalistycznego narzędzia do akwizycji ww. danych biometrycznych umożliwiającego przeprowadzenie całościowych testów metody.

Przedstawione w rozprawie wyniki badawcze zostały częściowo opisane w publikacjach współautorstwa Doktoranta.

Konkluzja

Przedstawiona do recenzji rozprawa zawiera oryginalne wyniki prac badawczych Doktoranta w obszarze wielomodalnych metod biometrycznych. Rozprawa dotyczy aktualnej tematyki badawczej a jej treść dowodzi wiedzy i umiejętności Autora w podjętym obszarze badań. Tematyka oraz treść rozprawy mieszczą się w obszarze dyscypliny *informatyka techniczna i telekomunikacja*.

Praca napisana jest starannie, a błędy językowe czy interpunkcyjne, które zauważyłem mieszczą się w dopuszczalnym zakresie. Nie dostrzegłem w rozprawie istotnych nieprawidłowości, a wymienione w recenzji uwagi nie podważają mojej ogólnie pozytywnej oceny dysertacji.

Reasumując, **stwierdzam, że rozprawa spełnia wymagania stawiane przez odnośną Ustawę i wnoszę o jej przyjęcie oraz dopuszczenie jej Autora, mgr inż. Macieja Szymkowskiego, do dalszych etapów przewodu doktorskiego.**



