

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

11. 01. 2026 r.

Opinia o rozprawie doktorskiej **Tomasza Walczyny**
pt.: "Aktywna ochrona tożsamości w obrazach twarzy z wykorzystaniem odpornego na
manipulacje deepfake znakowania wodnego"
przygotowanej w Wojskowej Akademii Technicznej
pod kier. dr hab. inż. Zbigniewa Piotrowskiego, prof. WAT

Opinię tę opracowałem na zlecenie Przewodniczącego Rady Dyscypliny Naukowej
„Informatyka techniczna i telekomunikacja” na podstawie uchwały nr 86/RDN ITiT/2025.

Postępowanie toczy się na podstawie przepisów Ustawy z dn. 20 lipca 2018 r. Prawo o
szkolnictwie wyższym i nauce (Dz. U. z 2022 r. poz. 574 z późniejszymi zmianami).

Rozprawa doktorska ma formę przewodnika po zbiorze dołączonych publikacji (kompilacji
11 prac naukowych), uzupełnionego przewodnikiem narracyjnym, w którym doktorant prezentuje
skróty publikacji oraz dyskusję ich wyników w kontekście realizacji tezy i celów badawczych.

Doktorant Tomasz Walczyna jest pierwszym autorem większości załączonych prac (m.in.
przeeglądów literatury dotyczących konkwersji głosu i deepfake'ów opartych na podmianie
wizerunków twarzy, modelu Fast Fake, analizy regiony zainteresowania, wpływu deepfake na
znakowanie wodne oraz kluczowej metody RE-Mark), co świadczy o jego dominującym udziale
w koncepcji, metodologii, implementacji, eksperymentach i redakcji manuskryptów. W pracach
współautorskich (np. z promotorem dr hab. inż. Zbigniewem Piotrowskim lub innymi
współpracownikami) jego wkład obejmuje zazwyczaj główne elementy badawcze, takie jak
projektowanie architektur neuronowych, trening modeli oraz analiza wyników. Większościowy
udział doktoranta jest potwierdzony w dołączonych oświadczeniach współautorów publikacji,

Publikacje ukazały się w czasopismach open access, głównie wydawnictwa MDPI (Applied
Sciences, Sensors) i na konferencjach, a jedna z nich (RE-Mark) została zgłoszona do bardziej
renomowanego czasopisma IEEE Access (stan na początek 2026 r.), co potwierdza odpowiedni
poziom naukowy osiągnięć doktoranta i ich wkład w dziedzinę bezpieczeństwa multimedialnego
oraz przetwarzania obrazów z wykorzystaniem sieci neuronowych.

Układ mojej opinii ma formę odpowiedzi na typowe pytania stawiane recenzentom prac
kwalifikacyjnych.

1. Jakie zagadnienie naukowe/badawcze jest rozpatrywane w pracy (cel i teza rozprawy) i
czy zostało ono dostatecznie sformułowane przez autora

W rozprawie doktorskiej Tomasza Walczyny rozpatrywane jest zagadnienie naukowe związane z aktywną ochroną tożsamości w obrazach twarzy przed zagrożeniami wynikającymi z technologii deepfake. Autor położył szczególny nacisk na metody znakowania wodnego (watermarking) odpornego na manipulacje typu podmiany twarzy (face-swap). Głównym celem pracy jest wypełnienie luki badawczej polegającej na braku narzędzi, które nie tylko wykrywają manipulacje, ale także umożliwiają aktywne odzyskanie utraconej tożsamości po podmianie twarzy. Teza rozprawy brzmi: „Modele sieci neuronowych umożliwiają ukrycie reprezentacji wizerunku twarzy w obrazie oraz jego rekonstrukcję po manipulacjach deepfake”. Do jej weryfikacji autor wyznaczył cztery cele badawcze (C1–C4), obejmujące analizę schematów manipulacji, opracowanie sterowalnych generatorów deepfake, badanie metod ukrywania danych oraz projektowanie metody watermarkingu z rekonstrukcją (RE-Mark), wsparte założeniami szczegółowo opisanymi w pracy.

Teza i cele zostały dostatecznie jasno sformułowane przez autora. Są ostre, mierzalne i osadzone w kontekście literatury, z jasnym podziałem na założenia do weryfikacji empirycznej. Struktura wprowadzenia logicznie prowadzi od motywacji do hipotez, co ułatwia zrozumienie i ocenę wkładu naukowego.

2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł, w tym, literatury światowej, stanu wiedzy

W rozprawie doktorskiej analiza źródeł, w tym literatury światowej i stanu wiedzy, została przeprowadzona w sposób właściwy i systematyczny. Autor w rozdziale 2 („Kontekst”) prezentuje kompleksowy przegląd kluczowych zagadnień, takich jak ochrona tożsamości, generatywne sieci neuronowe (np. GAN, VAE, modele dyfuzyjne), manipulacja tożsamością oraz metody ochrony przed deepfake, w tym znakowanie wodne. Omówienie opiera się na międzynarodowych źródłach. Autor zacytował prace takich autorów jak Goodfellow (GANs), Kingma (VAE) czy Dhariwal (modele dyfuzyjne), co odzwierciedla aktualny stan wiedzy do 2025 roku. Dodatkowo, załączone publikacje (sekcja 8) to przeglądy literatury, np. „Overview of Voice Conversion Methods Based on Deep Learning” i „Quick Overview of Face Swap Deep Fakes”, które pogłębiają analizę, klasyfikując metody i wskazując luki badawcze. Bibliografia (ponad 40 pozycji) obejmuje renomowane źródła z arXiv, ACM i IEEE, reprezentujące globalny dorobek. Brakuje jedynie szerszego omówienia najnowszych przeglądów z 2024–2025 (np. „A Survey on Proactive Deepfake Defense” z ACM), ale całość jest spójna, krytyczna i prawidłowo osadza tezę w kontekście naukowym.

3. Czy autor rozwiązał postawione zagadnienia; czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione, czy wykazał umiejętność poprawnego przedstawienia uzyskanych przez siebie wyników?

Autor w rozprawie doktorskiej w sposób zadowalający rozwiązał postawione zagadnienia badawcze. Zastosował metody adekwatne do sformułowanego problemu i w większości przypadków uzasadnił przyjęte założenia. Umiejętność przedstawienia wyników również należy ocenić pozytywnie, choć z pewnymi zastrzeżeniami dotyczącymi szczegółowości i

powtarzalności. Autor zrealizował wszystkie cztery cele badawcze (C1–C4):

- przeprowadził analizę schematów manipulacji tożsamością (głównie podmiana, częściowo również konwersja głosu),
- opracował i zaadaptował własne, kontrolowane generatory deepfake,
- zbadał odporność różnych technik ukrywania danych (w tym zero-bit i bitowych) na procesy podmiany twarzy,
- zaprojektował i zweryfikował nową metodę RE-Mark dla potrzeb rekonstrukcyjnego znakowania wodnego umożliwiającego odzyskanie oryginalnego wyglądu twarzy po manipulacji.

W rozdziale 7 autor dokonał oceny realizacji celów i tezy. Wnioski są spójne z wynikami i nie wykazują inklinacji do nadmiernego nadinterpretowania uzyskanych efektów.

Zastosowane metody są w pełni adekwatne do problemu i odpowiadają aktualnemu stanowi wiedzy w 2024–2025 roku w zakresie:

- Autoenkoderów + ArcFace do tworzenia kompaktowej reprezentacji tożsamości,
- sieci GAN, U-Net i DenseNet do ukrywania i ekstrakcji znaku wodnego,
- własnych, stopniowalnych generatorów obrazu twarzy do podmiany (w tym Fast-Fake i modyfikacje z rozmytym Gaussem), posiadające cenne zalety w badaniach nad odpornością,
- testowanie na kilku popularnych modelach face-swap (SimSwap, FaceShifter, HifiFace i in.) oraz na różnych poziomach siły manipulacji,
- użyto metryk zarówno percepcyjnych (LPIPS, SSIM, FID), jak i specyficznych dla tożsamości (ArcFace cosine similarity, verification accuracy).

4. Oryginalność rozprawy, jej mocne strony, pozycja rozprawy w stosunku do stanu wiedzy, przydatność rozprawy dla nauk inżyniersko-technicznych, w szczególności wkład do dyscypliny Informatyka techniczna i telekomunikacja.

Mocne strony pracy to: (i) kompleksowe podejście, tzn. od analizy manipulacji, przez własne generatory deepfake o sterowanej sile, po empiryczną weryfikację; (ii) praktyczna użyteczność RE-Mark (przezroczystość wizualna, odporność na silne degradacje, metryki PSNR/SSIM/FID potwierdzające wysoką jakość rekonstrukcji); (iii) publikacje w punktowanych czasopismach (np. Sensors, Applied Sciences) i preprinty (TechRxiv), w tym analiza wpływu podmiany twarzy na znakowanie wodne w oparciu o świadomy wybór regionu zainteresowania.

W stosunku do stanu wiedzy (2025–2026) rozprawa plasuje się korzystnie: większość prac (np. SepMark, DiffMark, LampMark) koncentruje się dotąd na detekcji lub śladach źródłowych, a nie na aktywnej rekonstrukcji tożsamości. RE-Mark wypełnia lukę wskazaną w przeglądach (np. „A Survey on Proactive Deepfake Defense”, 2024–2025), gdzie brakuje metod umożliwiających odzyskanie autentycznego wizerunku post-factum.

Szczególnie mocną stroną przyjętej przez autora metodologii jest analiza oparta na świadomym wyborze regionu zainteresowania (różne maski: cała twarz, tylko centralna

część, okolice oczu itp.), co potwierdza hipotezę, że różne obszary twarzy mają odmienną semantykę i różną podatność na degradację.

Założenia są w większości dobrze uzasadnione literaturą i własnymi eksperymentami wstępnymi, m.in.:

- możliwość przenoszenia nadmiarowej informacji semantycznej w znaku wodnym (oparte na pracach z rekonstrukcyjnym znakowaniem wodnym i głęboką steganografią),
- użyteczność ArcFace jako odpornego embeddingu tożsamościowego,
- założenie, że znak wodny może przetrwać silne manipulacje oparte na podmianie twarzy dzięki redundancji i uczeniu end-to-end.

Jedno założenie budzi najwięcej wątpliwości i jest najsłabiej uzasadnione. Chodzi o możliwość praktycznego, masowego nakładania znaku wodnego na wszystkie zdjęcia twarzy w świecie rzeczywistym przed pojawieniem się deepfake'ów. Autor wspomina o tym ograniczeniu, ale nie rozwija tematu dystrybucji i egzekwowania takiego rozwiązania.

Wyniki są przedstawione klarownie, z dużą liczbą tabel i wykresów porównawczych. Autor konsekwentnie pokazuje:

- porównania wizualne (przed/po rekonstrukcji),
- wartości metryk w tabelach,
- wykresy zależności jakości rekonstrukcji od siły manipulacji,
- analizy statystyczne (w tym testy istotności w rozdz. 6.3).

Największe zastrzeżenie (słaba strona pracy) dotyczy powtarzalności: w pracy brakuje dokładnego opisu hiperparametrów, seedów, podziału zbiorów treningowych/ewaluacyjnych oraz publicznego udostępnienia kodu i modeli (co w dziedzinie deepfake jest obecnie oczekiwanym standardem). Bez tego weryfikacja wyników przez niezależnego badacza będzie znacząco utrudniona.

5. Uwagi polemiczne

Rozprawa doktorska Tomasza Walczyny, choć wykazuje wysoki poziom oryginalności i solidne podstawy metodologiczne, budzi kilka uwag polemicznych, wynikających z analizy w kontekście aktualnego stanu wiedzy (stan na początek 2026 r.).

Po pierwsze: twierdzenie o pionierskim charakterze metody RE-Mark jako pierwszej umożliwiającej rekonstrukcję tożsamości po podmianie twarzy wymaga doprecyzowania. Autor podkreśla, że RE-Mark wypełnia lukę, pozwalając na aktywne odzyskanie oryginalnego wyglądu twarzy z pojedynczego obrazu bez zewnętrznych referencji, w odróżnieniu od metod detekcyjnych lub śledzących źródło (np. Proactive Deepfake Defence via Identity Watermarking z 2023, czy Robust Identity Perceptual Watermark z 2024–2025). Jednak w literaturze z 2024–2025 pojawiają się zbliżone koncepcje rekonstrukcyjne, np. DFREC (DeepFake Identity Recovery Based on Identity-aware Masked Autoencoder, arXiv 2024), które explicite dążą do odzyskania pary źródłowej i docelowej twarzy z deepfake w celu śledzenia tożsamości, oraz inne prace nad „reconstructive forensics” w przeglądach proactive

defense (np. Enhancing Deepfake Detection: Proactive Forensics Techniques Using Digital Watermarking, 2025). Choć RE-Mark różni się szczegółami (zero-bit, symmetric U-Net z attention, fokus na semantycznej redundancji), brak szerszego porównania z tymi nowszymi podejściami mógłby osłabić argument o unikatowości wkładu.

Po drugie: ograniczona generalizacja i odporność na nieznane ataki. Eksperymenty potwierdzają skuteczność RE-Mark na wybranych generatorach deepfake i degradacjach, ale testy skupiają się głównie na znanych modelach (np. z InsightFace). W kontekście szybkiego rozwoju generatorów (np. modele dyfuzyjne), brak walidacji na najnowszych manipulacjach (np. zaawansowane StyleGAN-based lub podmianie twarzy opartej na dyfuzji) rodzi wątpliwość co do długoterminowej odporności. Podobnie, choć autor bada wpływ klasycznych degradacji, nie uwzględnia specyficznych ataków na neuronowe znakowanie wodne (np. regeneration attacks z użyciem VAE/diffusion do usunięcia/rekonstrukcji, opisane w pracach 2024–2025).

Po trzecie: struktura rozprawy jako kompilacji publikacji z założenia niesie ryzyko fragmentaryczności. Choć pozwala to na prezentację ewolucji badań, brakuje głębszej syntezy ograniczeń poszczególnych komponentów (np. sterowalnych generatorów czy analizy w oparciu o świadomy wybór regionu zainteresowania) w kontekście finalnej metody RE-Mark. To utrudniło recenzentowi ocenę spójności całego wkładu.

Podsumowując, uwagi te nie podważają wartości pracy, bowiem RE-Mark stanowi istotny krok w kierunku rekonstrukcyjnego znakowania wodnego, Jednak wskazują one na potrzebę szerszego benchmarkingu z najnowszymi konkurentami oraz dyskusji potencjalnych słabości w obliczu ewoluujących zagrożeń deepfake.

6. Uwagi szczegółowe

Nr	Lokalizacja w rozprawie	Uwaga szczegółowa
1	Strona 5–7 (Streszczenia), Rozdział 1 (Wprowadzenie), Rozdział 6 oraz publikacja 8.11 (RE- Mark, strony ~164–183)	Twierdzenie o pionierskim charakterze metody RE-Mark jako pierwszej umożliwiającej rekonstrukcję oryginalnej tożsamości z pojedynczego obrazu po podmianie twarzy wymaga doprecyzowania. Metoda jest innowacyjna (zero-bit watermark z semantyczną redundancją, symmetric U-Net z attention), jednak w literaturze z końca 2024–2025 pojawia się zbliżona praca DFREC (DeepFake Identity Recovery Based on Identity-aware Masked Autoencoder, arXiv 2412.07260), która również rekonstruuje obie tożsamości (source i target) bezpośrednio z deepfake, choć w sposób pasywny (bez pre-embedding watermark). Brak szerszego porównania z nowszymi metodami rekonstrukcyjnymi osłabia to argumentację o unikalności.
2	Rozdział 6 (Metoda RE-	Eksperymenty potwierdzają wysoką odporność RE-

Nr	Lokalizacja w rozprawie	Uwaga szczegółowa
	Mark), publikacja 8.11 (strony ~164–183), Rozdział 6.3 (Istotność statystyczna)	Mark na wybrane manipulacje i degradacje, ale testy ograniczają się głównie do znanych generatorów podmiany twarzy (np. InsightFace). Brak walidacji na najnowszych modelach opartych na dyfuzji (2025+) oraz specyficznych atakach na neuronowe znakowanie wodne (np. regeneration attacks via VAE/diffusion) rodzi wątpliwości co do generalizacji i długoterminowej odporności w szybko ewoluującym krajobrazie deepfake.
3	Cała struktura rozprawy (Rozdziały 3–6 jako skróty publikacji, Sekcja 8 z pełnymi publikacjami)	Brak głębszej syntezy w części narracyjnej (np. jak komponenty z wcześniejszych publikacji – sterowalne generatory czy region-aware analiza – integrują się z finalną metodą RE-Mark) utrudnia ocenę spójności całego wkładu naukowego. Przegląd literatury jest solidny i obejmuje kluczowe prace do ~2024, ale brakuje omówienia najnowszych przeglądów proactive defense (np. „A Survey on Proactive Deepfake Defense: Disruption and Watermarking”, ACM 2025) oraz metod rekonstrukcyjnych (np. DFREC 2024–2025), co mogłoby lepiej osadzić RE-Mark w aktualnym stanie wiedzy na 2025–2026.
4	Rozdział 2 (Kontekst), Bibliografia (strony 184–187)	
5	Rozdział 7 (Zakończenie), podrozdział 7.3 (Ograniczenia i kierunki dalszych prac)	Autor trafnie wskazuje ograniczenia (np. zależność od jakości embeddingów ArcFace), ale nie rozważa potencjalnych ataków przeciwstawnych na sam proces embedding/extraction watermark (np. adversarial examples w U-Net), co jest istotne w kontekście bezpieczeństwa metod neuronowych.

7. Podsumowanie

Rozprawa doktorska kpt. mgr. inż. Tomasza Walczyny pt. „Aktywna ochrona tożsamości w obrazach twarzy z wykorzystaniem odpornego na manipulacje deepfake znakowania wodnego” stanowi wartościowy i oryginalny wkład w dyscyplinę Informatyka techniczna i telekomunikacja, szczególnie w obszarze bezpieczeństwa informacji i przetwarzania multimedialnego. Autor precyzyjnie sformułował tezę oraz cele badawcze, które w pełni zrealizował, proponując nowatorską metodę RE-Mark. Metoda ta stanowi neuronowe rekonstrukcyjne znakowanie wodne umożliwiające odzyskanie oryginalnego wyglądu twarzy po manipulacji związanej z podmianą twarzy w pojedynczym obrazie, przy zachowaniu wysokiej jakości wizualnej.

Praca wyróżnia się solidną analizą stanu wiedzy, odpowiednim doбором metod (architektury U-Net z attention, embeddingami ArcFace, własnymi sterowanymi generatorami deepfake) oraz rzetelną prezentacją wyników, wspartych metrykami ilościowymi i testami

statystycznymi. W kontekście szybko rozwijającego się dziedziny aktywnej ochrony przed deepfake'ami (stan na początek 2026 r.), RE-Mark plasuje się w czołówce metod rekonstrukcyjnych, wypełniając lukę między detekcją a aktywnym przywracaniem autentyczności, co ma bezpośrednie zastosowanie w cyberbezpieczeństwie, w dochodzeniach i w systemach biometrycznych.

Zgłoszone uwagi polemiczne i szczegółowe (m.in. potrzeba szerszego porównania z najnowszymi metodami rekonstrukcyjnymi, takimi jak DFREC z 2024–2025, oraz walidacji na atakach opartych na dyfuzji) nie podważają wysokiej jakości rozprawy, lecz wskazują kierunki dalszego rozwoju. Rozprawa spełnia wszelkie kryteria naukowe i metodyczne wymagane dla stopnia doktora nauk inżynieryjno-technicznych w dyscyplinie Informatyka techniczna i telekomunikacja. Wniosek o jej wyróżnienie byłby uzasadniony, gdyby dotychczas artykuły ukazały w czasopismach niezaliczanych do tzw. „drapieżnych” (wydawnictwo MDPI). W przypadku, jeżeli przed terminem posiedzenia komisji doktorskiej ukazały się kolejne publikacje w wydawnictwach o niekwestionowanej renomie, będę gotów do zgłoszenia wniosku o wyróżnienie tej rozprawy doktorskiej.

Wniosek

Rozprawa pana mgr inż. Tomasza Walczyny została zrealizowana w sposób odzwierciedlający wymagane kwalifikacje jej Autora, wystarczający nakład pracy badawczej, implementacyjnej i eksperymentalnej. W mojej opinii treść rozprawy mgr inż. Tomasza Walczyny spełnia wymogi Prawa o Szkolnictwie Wyższym i Nauce, z dnia 20 lipca 2018 r. (Dz. U. z 2022 r. poz. 574 z późniejszymi zmianami), stawiane kandydatom do stopnia naukowego doktora.