



Prof. dr hab. inż. Khalid Saeed
Wydział Informatyki
Politechnika Białostocka
ul. Wiejska 45A, 15-351 Białystok
Tel. (+48-85) 746 9196
k.saeed@pb.edu.pl

Rada Naukowa Dyscypliny
INFORMATYKA TECHNICZNA
I TELEKOMUNIKACJA
Sekretariat
Data wpływu... **13.03.2023r.**
Numer.....

Białystok, 6.03.2023 r.

RECENZJA rozprawy doktorskiej

mgr inż. Weroniki Gutfeter

z Wydziału Elektroniki i Technik Informatycznych
Politechniki Warszawskiej

z tytułu *zatytułowanej* "Identyfikacja twarzy na podstawie obrazów
wieloujęciowych z zastosowaniem głębokich sieci agregujących"

Promotor:

Prof. dr hab. inż. Andrzej Pacut
Wydział Elektroniki i Technik Informatycznych
Politechnika Warszawska

Niniejszą recenzję przygotowałem na zlecenie zawarte w piśmie, które otrzymałem od Profesora Jarostawa Arabasa przewodniczącego Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Warszawskiej. Recenzję dokonałem na podstawie dostarczonej mi rozprawy doktorskiej dnia 9.01.2023.

I. Omówienie zawartości rozprawy

Głównym celem pracy doktorskiej mgr inż. Weroniki Gutfeter było zastosowanie metod klasyfikacji obiektów wieloujęciowych do identyfikacji twarzy człowieka jako cechy biometrycznej. Innowacyjne podejścia polegały na opracowaniu algorytmów pozwalających rozwiązać problem identyfikacji twarzy na podstawie zdjęć wieloujęciowych, gdzie obrazy twarzy jednej osoby były zbierane pod różnymi kątami głowy. Doktorantka zaadoptowała metody stosowane dotąd do klasyfikacji obiektów 3D. Zagadnienie to zostało jasno sformułowane przez autorkę pracy, a trzy odpowiednie tezy, które starała się udowodnić, zostały opracowane poprawnie. Rozprawa ma charakter

statystyczno-doświadczalny i zawiera 130 stron. Składa się z pięciu rozdziałów oraz materiałów dodatkowych i bibliografii.

Rozdział pierwszy poświęcony jest przeglądowi wybranych zagadnień związanych z problemem rozpoznawania twarzy. Wybrana literatura do analizy stanu wiedzy jest poprawna, pomimo niedosytu aktualnych pozycji.

W rozdziale drugim autorka przedstawiła metody zastosowania sieci głębokich w identyfikacji twarzy człowieka. Tu szczegółowo pokazane są sposoby generacji wzorców twarzy z użyciem sieci splotowych oraz podstawowy model jednouięciowy będący wyjściowym punktem do prac autorki nad identyfikacją twarzy w systemach obrazów wielouięciowych. W rozdziale tym przedstawiono analizę aktualnego stanu wiedzy dotyczącego metodologii i podejść w kontekście obrazów wielouięciowych. Autorka zbadała i omówiła wpływ pozycji głowy na dokładność systemu identyfikacji twarzy na podstawie modelu jednego ujęcia (single-view model).

W rozdziale trzecim rozprawy autorka przedstawia wybrane podejścia stosowane w identyfikacji obrazów wielouięciowych twarzy. Doktorantka zajmuje się implementacją bardziej skomplikowanych struktur i modułów do agregacji.

W rozdziale czwartym autorka przedstawia szczegółowy opis modeli własnych do agregacji obrazów wielouięciowych twarzy. Doktorantka proponuje własny model agregujący SygnaT wykorzystujący mechanizm *atencji* w strukturze sieci. Mechanizm ten został zaproponowany w pracy „*Attention is all you need*” opublikowanej w Curran Associates w 2017, gdzie autorzy zastosowali go do tłumaczenia języka naturalnego.

Część merytoryczną zamykają wnioski końcowe w rozdziale piątym - *Podsumowanie* oraz Dodatek - *Materiały dodatkowe*. Całość pracy zamyka literatura - *Bibliografia*, która zawiera 73 pozycje, a 3 z nich są współautorstwa doktorantki. Niestety, wśród 73 wybranych pozycji bibliografii, tylko 8 datowanych jest 2020 lub nowsze wydanie.

II. Opinia i ogólna ocena pracy

Rozprawa doktorska mgr inż. Weroniki Gutfeter pt. "*Identyfikacja twarzy na podstawie obrazów wielouięciowych z zastosowaniem głębokich sieci agregujących*" stanowi oryginalne rozwiązanie problemu oraz wkład własny doktorantki w rozwój metod identyfikacji twarzy. Autorka wykonała założone cele pracy, zaprezentowała, rozwiązała problem identyfikacji twarzy na podstawie obrazów wielouięciowych - implementacja algorytmów, rysunki i tabele zostały odpowiednio opracowane. Przykładowe podejścia innych autorów podobnej tematyki zostały wybrane umiejętnie. Doktorantka opracowała systemy identyfikacji twarzy, w których wzorce twarzy są rejestrowane na podstawie zbioru obrazów wielouięciowych i pokazała, że ten system jest lepszy niż system oparty na zdjęciach frontalnych. Opracowała również system identyfikacji

twarzy wykorzystujący modele wieloujęciowe (multi-view) i pokazała, że wskaźniki jakości identyfikacji nie są gorsze niż system oparty o modele jednoujęciowe (single-view). Udowodniła eksperymentalnie i przedstawiła statystycznie, że modele agregujące stosowane dotąd do rozpoznawania obiektów 3D można dostosować do zadania identyfikacji twarzy.

Reasumując należy uznać, że autorka wykazała w swojej rozprawie doktorskiej dobrą znajomość technik identyfikacji twarzy. Z sukcesem zrealizowała postawione zadania pokazując eksperymentalnie wartość proponowanych metod i modeli. Pracę napisała poprawnie, zaś istotne dla tematyki pracy zagadnienia zostały omówione przez autorkę wyczerpująco.

Według mojej oceny uważam, że Pani mgr inż. Weronika Guttfeter osiągnęła wyznaczony cel rozprawy doktorskiej, który wnosi nowe aspekty do nauk technicznych w zakresie informatyki. Wymieniona rozprawa spełnia warunki ustawy o stopniach naukowych.

Uwagi merytoryczne oraz usterki w pisowni pracy

Jak nadmieniałem powyżej, praca napisana jest rzetelnie, ale praktycznie niemożliwe do uniknięcia są drobne usterki merytoryczne lub błędy edycyjne, których przykłady zostały podane poniżej:

Uwagi merytoryczne

- Ogólny stan wiedzy powinien być umieszczony w całości w jednym rozdziale na początku rozprawy.

- Brakuje ogólnego opisu zawartości rozdziałów rozprawy. Doktorantka powinna podać strukturę pracy na początku rozprawy i omówić zawartość każdego rozdziału i pokazać lukę badawczą w literaturze. Autorka taką strukturę pokazała w podsumowaniu. Z drugiej strony, podsumowanie powinno zawierać konkluzje – jakie problemy zostały rozwiązane, opis wyników, a szczególnie na ile jej oryginalne rozwiązania są lepsze od istniejących podejść. Przydatna byłaby analiza porównawcza pod kątem wyników pracy ze stanem wiedzy.

- Rysunek 2.7: W przypadku obrotów o kąty $\pm 60^\circ$, $\pm 30^\circ$, $\pm 15^\circ$ wyniki można uznać za "symetryczne". W przypadku obrotu osób o kąt -90° obserwuje się większy błąd niż w przypadku obrotu o kąt $+90^\circ$. Nie wyjaśniono dlaczego.

- Rysunek 3.9: Doktorantka opisuje: "*Na wykresie 3.9 można zobaczyć że dokładność systemu rośnie wraz ze wzrostem liczby próbek w zestawie, co jest zgodne z intuicją*". Wyraźny skok dokładności widoczny jest w początkowych zmianach liczby odpowiedzi. Dla prawej części wykresu oraz zmiany liczby z 11 do 12 widzimy "nasylenie" oraz marginalne zmiany. Oznacza to, że wniosek autorki nie jest do końca właściwy. Proszę o wyjaśnienie.

- Rysunek 4.3: Rysunek przedstawia miary klasyfikacji, które nie były wcześniej omówione tj. dokładność top-1. Warto to wyjaśnić.

- Tabela 4.4: Przedstawiona tabela budzi pewne wątpliwości - różnice między otrzymanymi wynikami są stosunkowo niewielkie (szczególnie warianty SygnaT (token) oraz SygnaT (avg)). Prosiłbym o wyjaśnienie.

Usterki w pisowni

- Streszczenie, jeśli ma być abstraktem (tak jak dalej podano po angielsku) nie powinno zawierać cytatów.

- Rys. 2.3: Doktorantka nie wyjaśniła „współczynnika dokładności”. Najprawdopodobniej jest to miara skuteczności w procesie nauki, ale powinno się to zdefiniować.

- Rysunki 2.8 oraz 2.9 pokazują metryki z wykorzystaniem miary podobieństwa cosinusowego. Natomiast wskaźnik kolorów dobrany został dość pobieżnie. Przede wszystkim kolory dobierane są w zakresie od 0.0 do 0.9 (górną granicą raczej powinna wynosić 1.0), a liczba przedziałów to 15. W związku z tym podział kolorów nie odpowiada liniowo wartościom. Rysunki zyskałyby na czytelności, gdyby zastosowano 10 kolorów ze skokiem 0.1 lub normalizację koloru do wartości obserwowanych.

- Rysunek 3.1 zastosowanie linii łączących "bloki" w diagramie sugeruje występowanie pewnych zależności (np. kolejności) pomiędzy nimi, ale chyba tak nie jest. Proszę o wyjaśnienie.

- Na stronie 63 czytamy *"Warto jednak zauważyć, że metody tego typu powodują duży narzut na obliczenia i tworzenie bazy wzorców. Wiele z nich jest zoptymalizowanych pod kątem wizualnie akceptowalnych wyników, a nie przenoszenia cech potrzebnych do identyfikacji - następuje wygładzenie obrazu w czasie projekcji i jak zauważają autorzy [10] utrata istotnych cech o wysokiej częstotliwości"*. Jest to bardzo dobry wniosek podsumowujący metodę, ale nie jest we właściwym miejscu. Sugerowałbym przeniesienie go do kolejnego podrozdziału pt. "Podsumowanie eksperymentów z metodami automatycznej korekcji pozy".

III. Merytoryczne osiągnięcia doktorantki oraz jej publikacje

Pani mgr inż. Weronika Gutfeter jest współautorem 2 referatów konferencyjnych oraz 1. artykułu opublikowanego w czasopiśmie IEEE Intelligent Systems (140 punktów na liście Ministerstwa), co świadczy o znaczeniu osiągniętych wyników pracy naukowej doktorantki. Brakuje jednak procentowego udziału doktorantki w najważniejszej publikacji wieloautorskiej. Oświadczenie autorki byłoby całkowicie wystarczające. Na podkreślenie zasługuje również fakt, że doktorantka ma aktywną działalność naukowo-projektową. Wybrane projekty badawcze z udziałem doktorantki jako współwykonawczynie w ramach jej pracy jako asystent naukowej w Instytucie Badawczym NASK to:

- APAKT: automatyczna analiza zagrożeń w sieci, w tym treści pokazujących wykorzystanie seksualne dzieci.

- BLOWIZ: system komputerowy wspierający funkcjonariuszy policji i innych służb mundurowych w identyfikacji osób na podstawie wizerunku utrwalonego na zdjęciu lub materiałach wideo.
- BioWalidator: walidacja fotografii twarzy do dowodów osobistych i innych dokumentów tożsamości.

IV. Wnioski końcowe

Wystawiam ocenę pozytywną rozprawie doktorskiej mgr inż. Weroniki Gutfeter pt. *"Identyfikacja twarzy na podstawie obrazów wieloujęciowych z zastosowaniem głębokich sieci agregujących"* oraz stwierdzam, że praca spełnia wymagania i warunki nakładane przez ustawę o stopniach naukowych. Rozprawa stanowi oryginalne rozwiązanie problemu naukowego oraz osobisty wkład Doktorantki w rozwój metod identyfikacji twarzy człowieka z obrazów wieloujęciowych.

Na tej podstawie wnioskuję o dopuszczenie Autorki wymienionej rozprawy doktorskiej do jej obrony w celu uzyskania stopnia doktora nauk technicznych w dziedzinie nauk inżynieryjno-technicznych w dyscyplinie informatyka techniczna i telekomunikacja.



Khalid Saeed

