Agregacja wzorców twarzy

mgr inż. Weronika Gutfeter

Agenda

1. Opis wybranego zadania identyfikacji
2. Sieci neuronowe – przykład dwuwymiarowy
3. Badanie metod agregacji wzorców w systemie identyfikacji

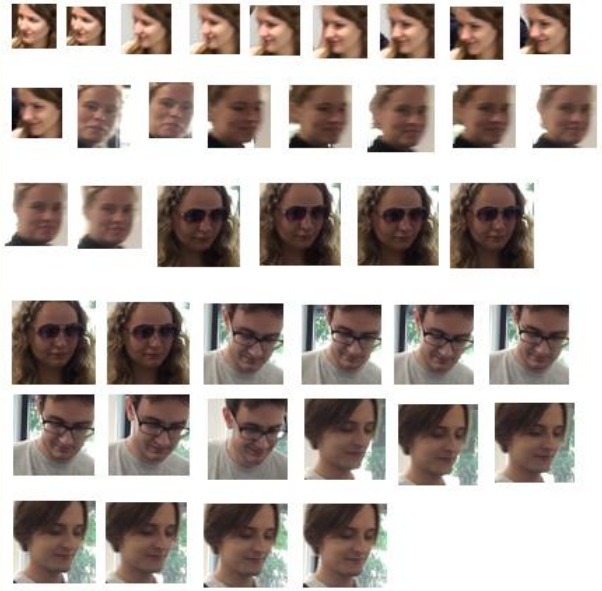
Baza NASK

- Baza zdjęć i nagrań wideo zebrana w NASK'u w ramach projektu BIOWIZ
- Zbiór ma emulować zastosowania policyjne – zdjęcia referencyjne przypominają policyjne zdjęcia sygnalityczne
- Nagrania wideo zostały przetworzone do wycinków zawierających twarze za pomocą pakietu Neurotechnology Sentiveillance
- Zdjęcia referencyjne: 5 osób, 3 sesje, 5 ujęć
- Dystraktory z bazy FEI : 200 klas po 5 ujęć (dodane jako tło do bazy referencyjnej)
- Zdjęcia „operacyjne" 4600 wycinków twarzy z nagrania wideo (średnio 920 obrazków na osobę)

Baza NASK

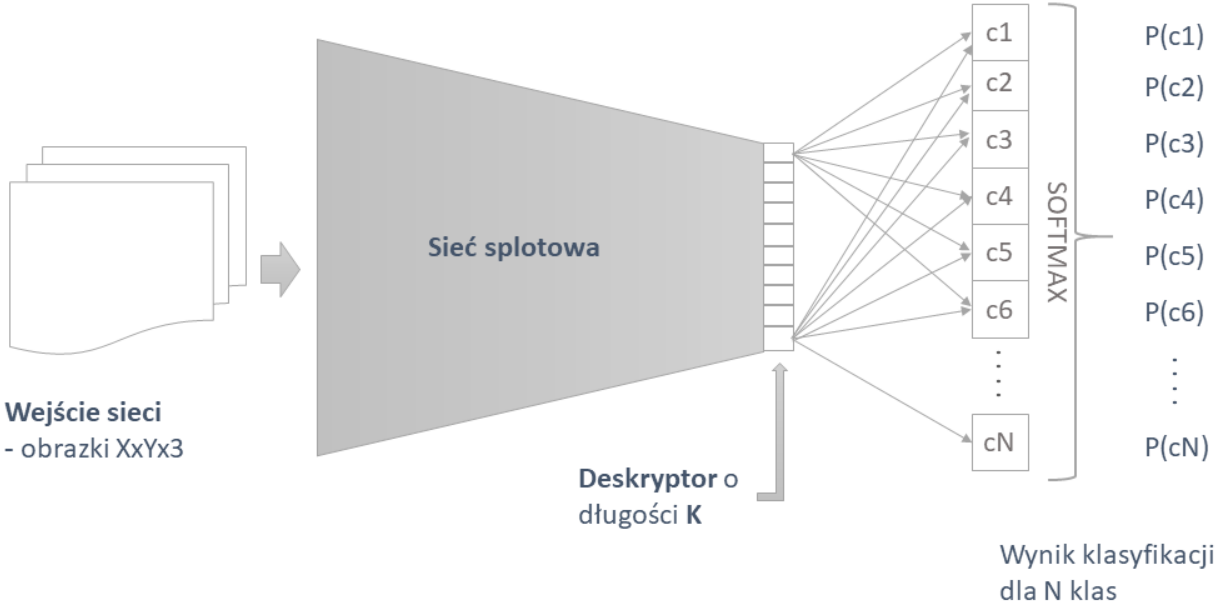
Zdjęcia referencyjne
5 klas/75 próbek

Dystraktory
200 klas/1000 próbek



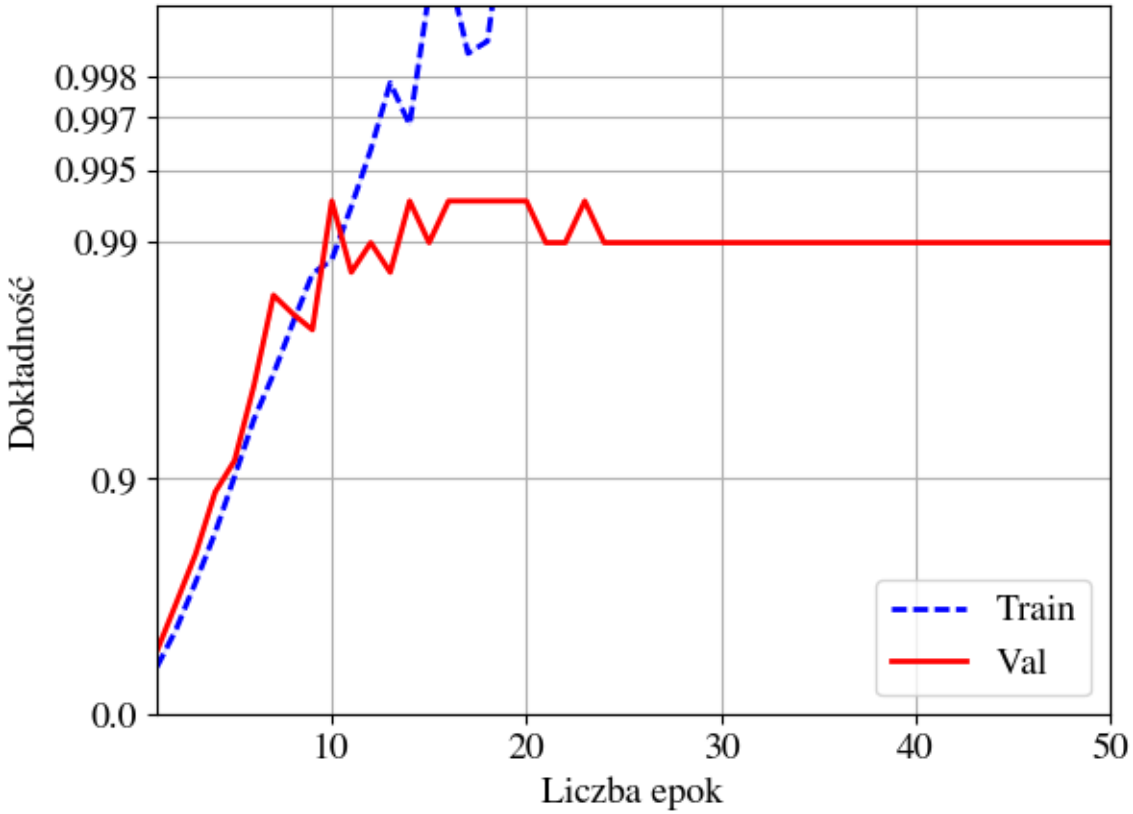
Zdjęcia referencyjne
5 klas/ok. 4600 wycinków z wideo

Schemat modelu opartego o głęboką sieć splotową

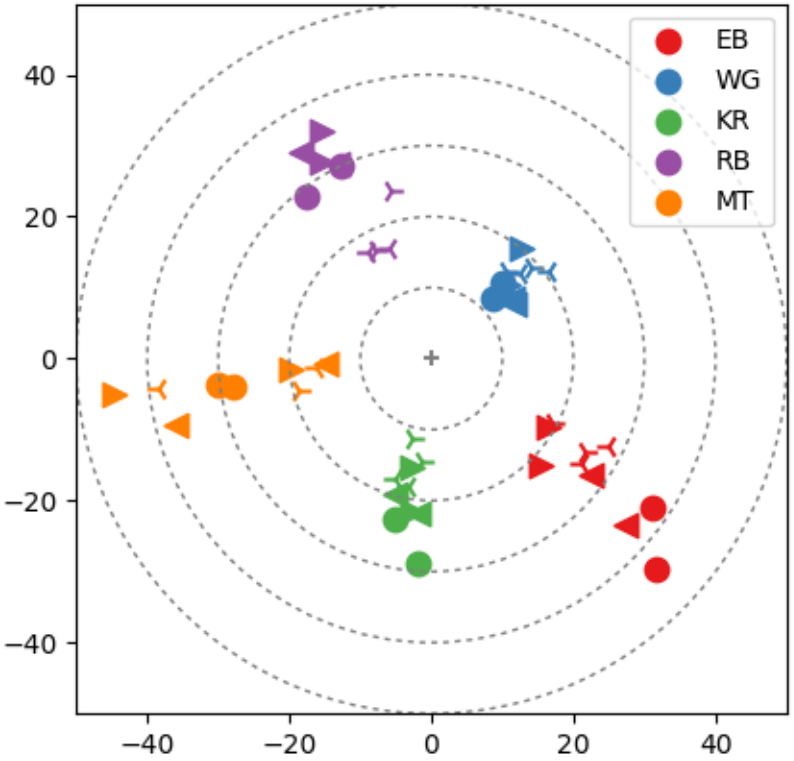


UPROSZCZONY MODEL 2D – DŁUGOŚĆ DESKRYPTORA K=2

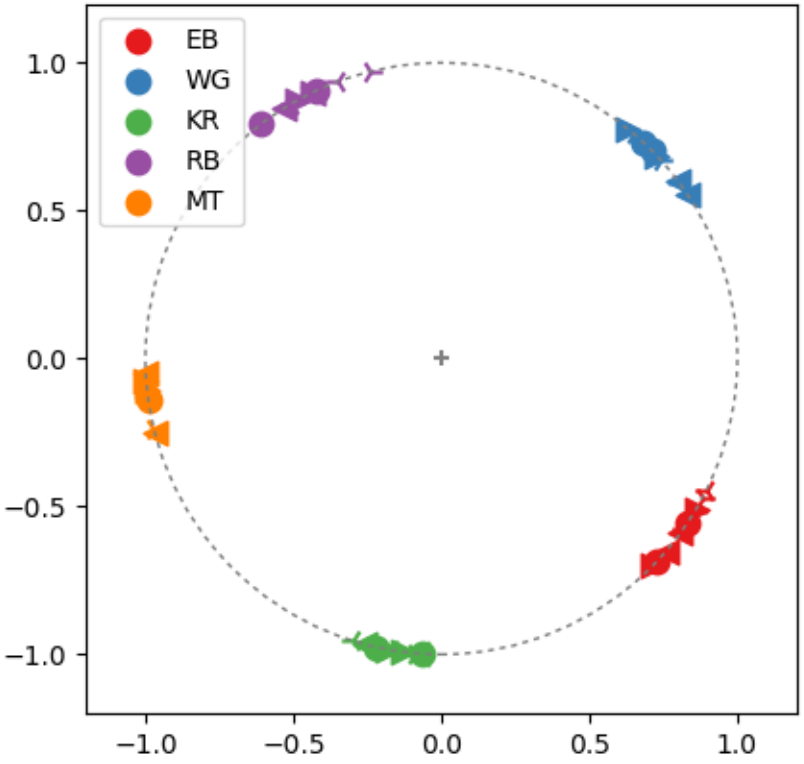
Model 2D– przebieg procesu uczenia



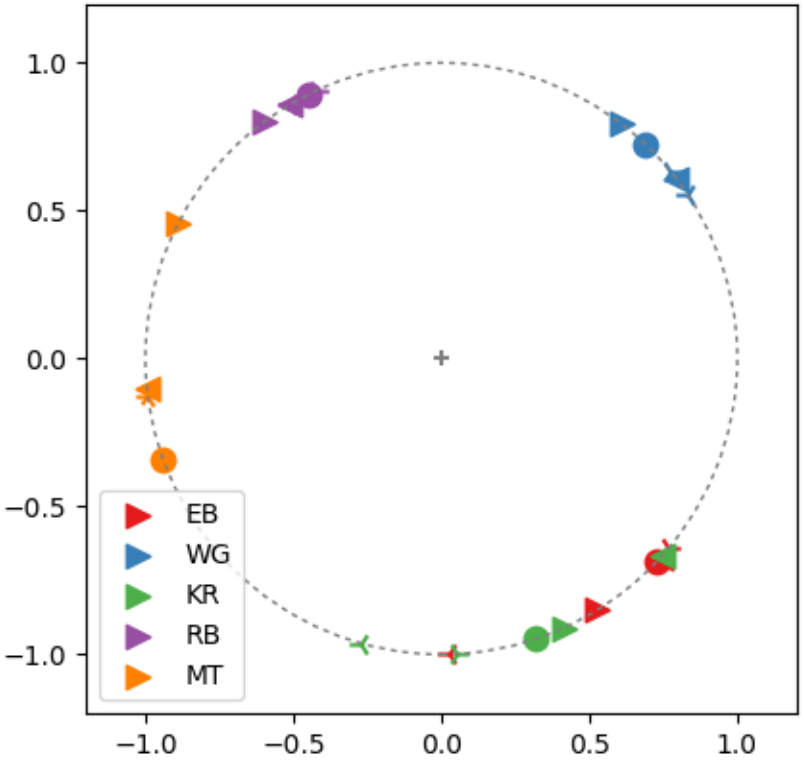
Model 2D– deskryptory dla zbioru **trenujacego**



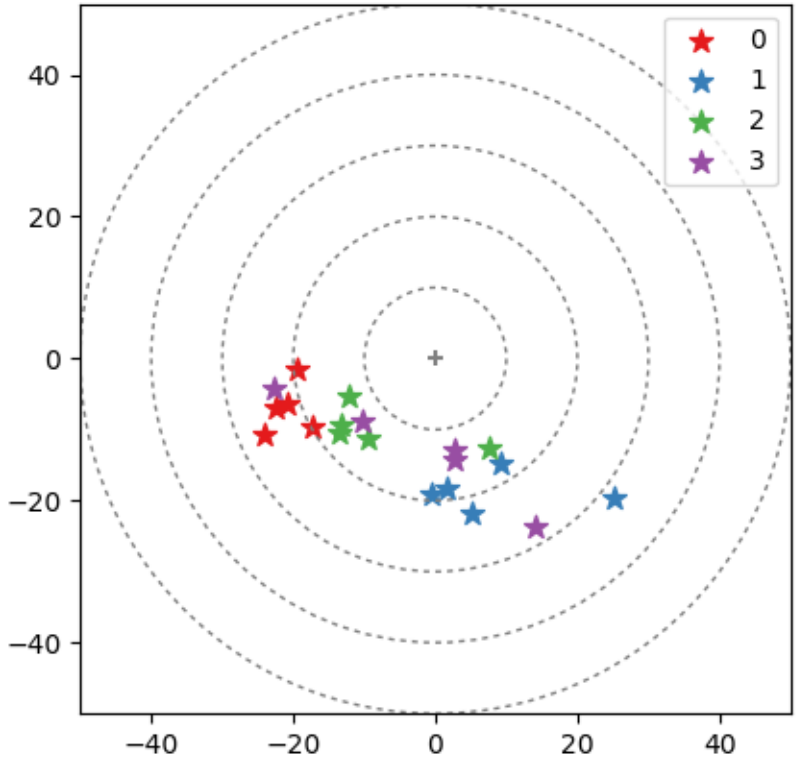
Model 2D – **znormalizowane** deskryptory dla zbioru trenującego



Model 2D – znormalizowe deskryptory dla zbioru **walidacyjnego**

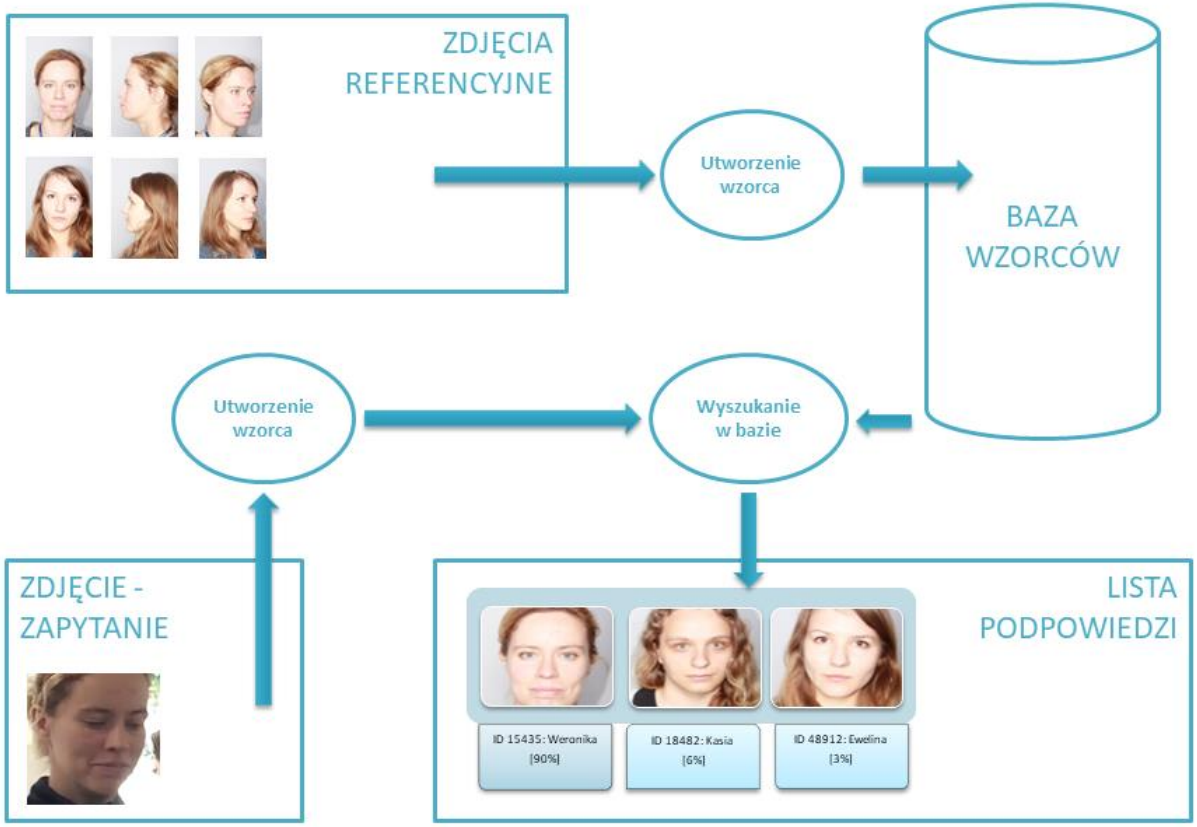


Model 2D– przykłady spoza bazy (inne klasy)



BADANIE NA ZBIORZE NASK Z WYTRENOWANYM DESKRYPTOREM

Budowa systemu identyfikacji



Różne podejścia do zapisywania wzorców w bazie

JEDNA PRÓBKA – JEDEN WZORZEC

- Prostsza implementacja
- Łatwiejsze dodawanie nowych wzorców i skalowalność bazy

JEDNA OSOBA – JEDEN WZORZEC

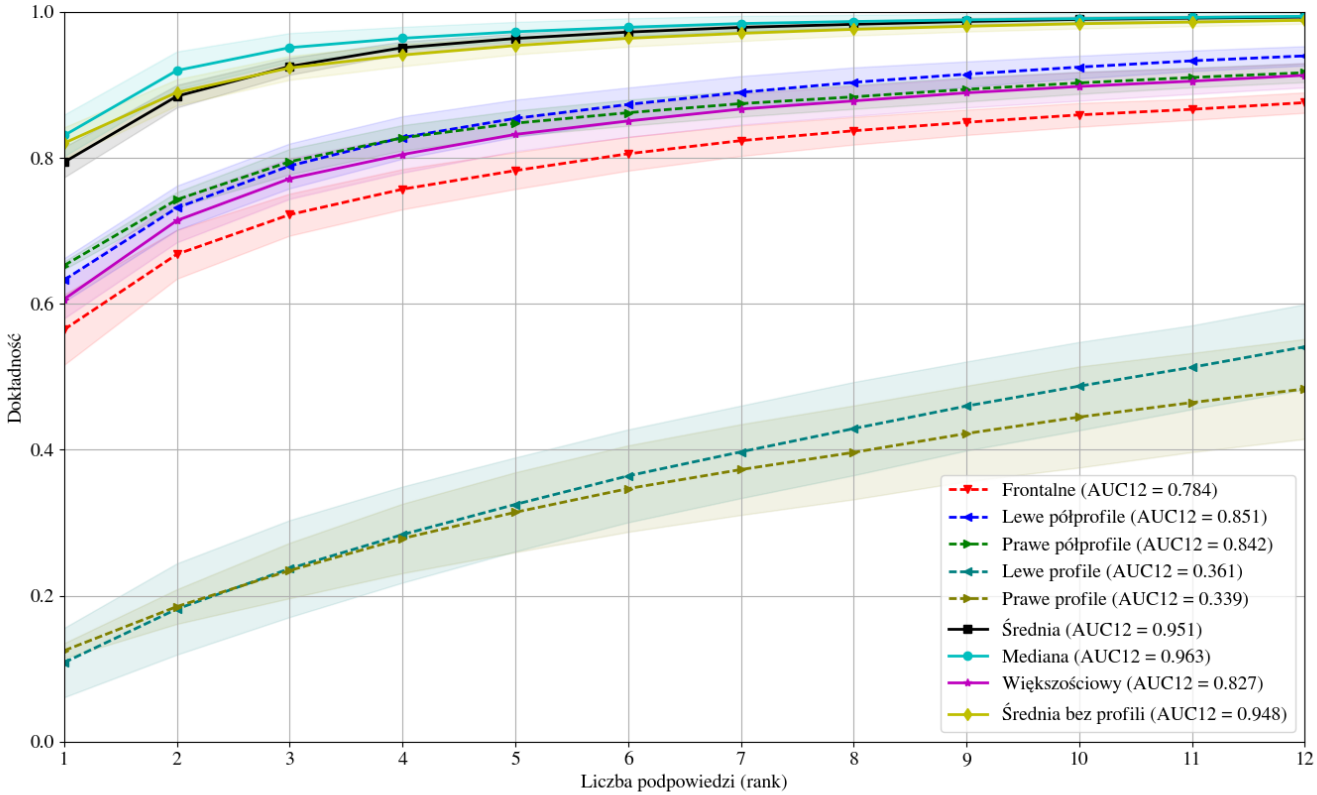
- Krótszy czas wyszukiwania (mniej wzorców do przeszukania)
- Trudniejsze modyfikacje w bazie
- Konieczność zapisywania próbek
- Intuicyjnie złożenie wzorców powinno dać silniejszy wzorzec niż wzorce pojedyncze ???

Klasyfikator użyty w badaniu

- VGG – Face Descriptor (głęboka sieć splotowa o głębokości 16 warstw)
- Klasyfikator nie zna klas użytych w badaniu – porównanie znormalizowanych deskryptorów
- Scenariusz identyfikacji próbek zakładał znalezienie pojedynczej próbki ze zbioru „operacyjnego” z w bazie złożonej ze zdjęć NASK i FEI (razem 205 klas)
- Wyniki przedstawione są za pomocą krzywych CMC

AGREGACJA PO STRONIE BAZY REFERENCYJNEJ

Agregacja zdjęć referencyjnych z różnych ujęć

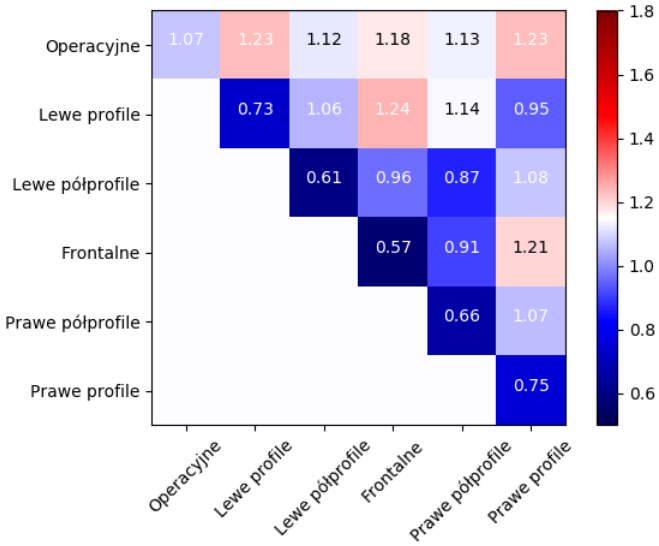


Wnioski

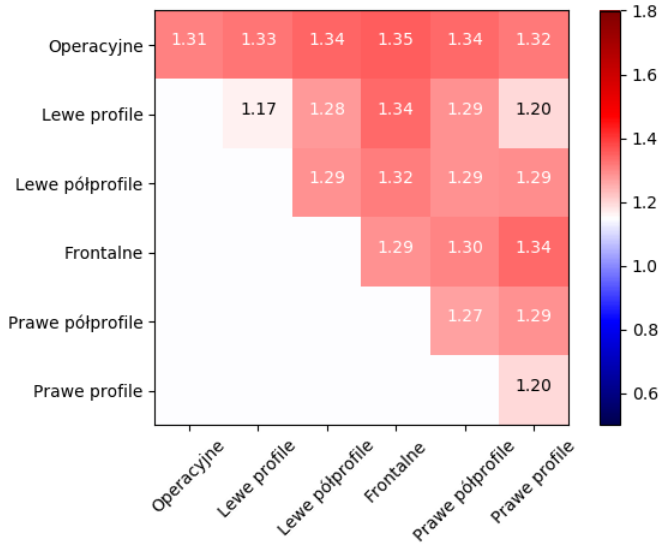
1. Złożenie wzorców referencyjnych z próbek z różnych ujęć polepsza wyniki identyfikacji rank-1 o ok. 47% dla wzorca medianowego i o ok. 40% dla wzorca średniego w stosunku do identyfikacji na zbiorze samych zdjęć frontalnych
2. Pełne profile są stosunkowo najmniej użyteczne w identyfikacji za pomocą tego klasyfikatora
3. Dla zbioru NASK wyszukiwanie próbek wśród zdjęć półprofilowych dało lepsze wyniki niż dla zdjęć frontalnych

Porównywanie ze zdjęciami z różnych ujęć

Średnie odległości między próbkami należącymi do tej samej klasy

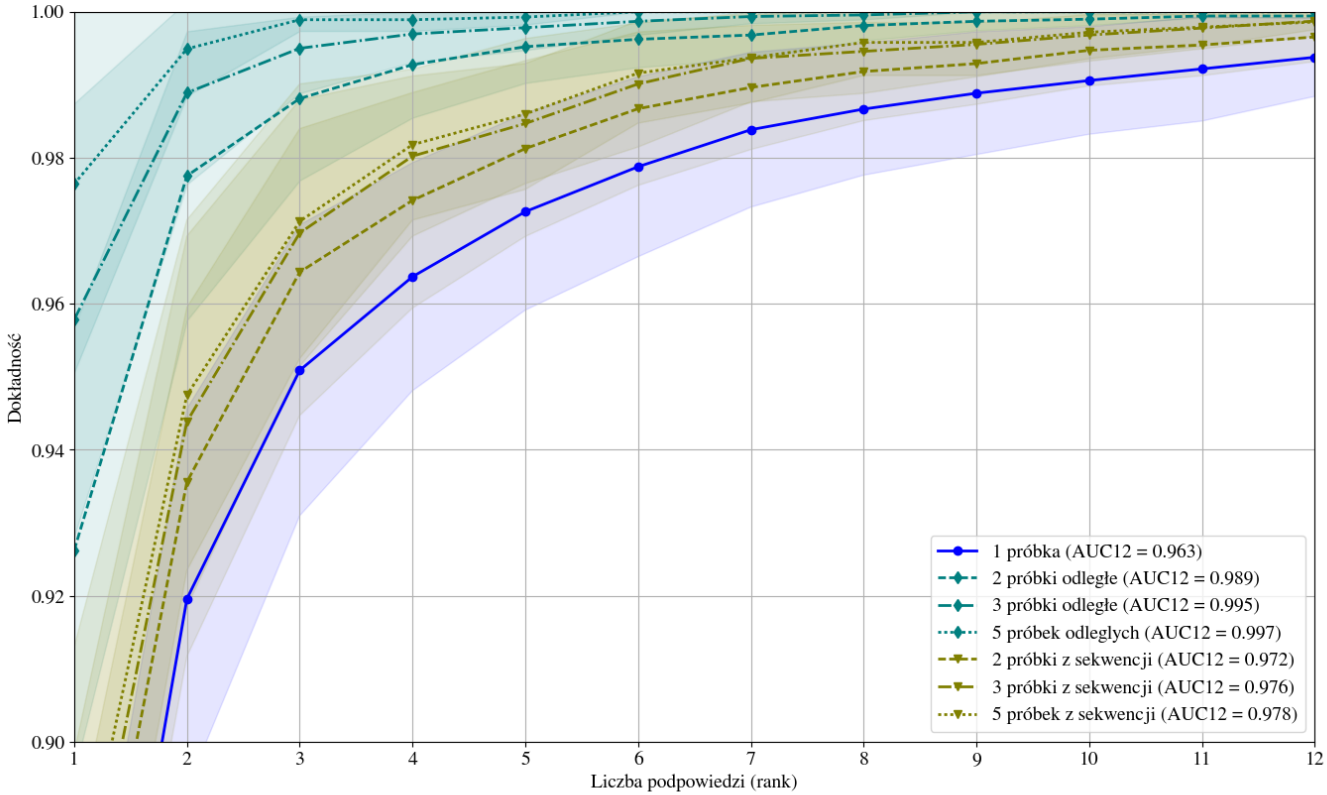


Średnie odległości między próbkami należącymi do różnych klas



AGREGACJA PO STRONIE PRÓBEK OPERACYJNYCH

Agregacja próbek wycinków z wideo



Wnioski

1. Agregacja obrazów z wideo polepsza wyniki identyfikacji:
   1. Złożenie wzorca z 3 próbek odległych dało wynik rank-1 większy o 15,3% niż dla wzorców pojedynczych
   2. Złożenie wzorca z 3 próbek sekwencyjnych dało wynik rank-1 o 4,4% lepszy
2. Wzrost dokładności jest większy w przypadku agregacji próbek niezależnych (odległych)

weronika.gutfeter@gmail.com