

POLITECHNIKA ŁÓDZKA
KATEDRA MIKROELEKTRONIKI I TECHNIK INFORMATYCZNYCH

MICHAŁ WŁODARCZYK
ROZPRAWA DOKTORSKA (STRESZCZENIE)

**Wydajne metody analizy biometrycznej dla
niekooperacyjnych scenariuszy uwierzytelniania osób**

Promotor:
prof. dr hab. inż. Andrzej NAPIERAŁSKI
Promotor pomocniczy:
dr inż. Kamil GRABOWSKI

Łódź, 2018

Spis treści

1	Wprowadzenie	10
1.1	Motywacja	10
1.2	Cel i zakres pracy	11
1.3	Tezy	12
2	System COMPACT	13
2.1	Wprowadzenie	13
2.2	Konstrukcja systemu	14
2.3	Przetwarzanie obrazów	16
2.4	Podsumowanie	17
3	Baza danych COMPACT	18
3.1	Wprowadzenie	18
3.2	Opis bazy danych	19
3.3	Podsumowanie	21
4	Rozpoznawanie osób w oparciu o cechy twarzy	22
4.1	Wprowadzenie	22
4.2	Opis wykorzystanych algorytmów	22
4.3	Eksperymenty	25
4.4	Podsumowanie	25
5	Estymacja pozy głowy	27
5.1	Wprowadzenie	27
5.2	Opis zaproponowanej metody	28
5.3	Eksperymenty	29
5.4	Podsumowanie	30
6	Rozpoznawanie osób w oparciu o okolicę oka	32
6.1	Wprowadzenie	32
6.2	Opis zaproponowanej metody	32
6.3	Eksperymenty	34
6.4	Podsumowanie	37
7	Podsumowanie	38

Rozdział 1

Wprowadzenie

1.1 Motywacja

Na przestrzeni ostatnich lat obserwuje się, że liczba ataków terrorystycznych czy innych nielegalnych działań drastycznie wzrosła na świecie. Instytucje publiczne i podmioty prywatne inwestują dużą ilość pieniędzy w technologię, która pozwoli zastąpić tradycyjne systemy ochrony i zwiększyć ogólne poczucie bezpieczeństwa społeczeństw, szczególnie w zatłoczonych miejscach. Z tego powodu, można zaobserwować, że znacząco wzrosła liczba kamer monitorujących przestrzeń publiczną. Przykładowo, w samej Wielkiej Brytanii, jest to obecnie około 6 milionów urządzeń [1]. Jednakże, pomimo powszechnej opinii, aktualne rozwiązania nie pozwalają na identyfikację konkretnych osób, ograniczając się jedynie do wykrywania niebezpiecznych zachowań, takich jak pozostawiony bagaż czy bójka. Z tego powodu systemy te nie są użyteczne w niektórych zastosowaniach, na przykład podczas weryfikacji osób wchodzących na stadion czy lotnisko [2].

Jednocześnie, systemy uwierzytelniania biometrycznego są uznawane za jedno z ważniejszych osiągnięć naszych czasów. Wraz ze wzrostem zapotrzebowania na zautomatyzowany dostęp do dóbr i usług, zagadnienia związane z biometrią zyskały w ostatnim czasie ogromne znaczenie [3]. Coraz więcej instytucji, jak na przykład banki, pozwalają na autoryzację w oparciu o cechy twarzy czy odcisk palca. Niemniej jednak, dostępne aktualnie rozwiązania wciąż wymagają czynnego udziału podmiotu w procesie uwierzytelniania. Przykładowo, w systemach rozpoznawania opartych o tęczówkę, konieczne jest, aby osoby identyfikowane znajdowały się blisko kamery i patrzyły w nią przez kilka sekund, aż do pozyskania odpowiedniej jakości zdjęć. W konsekwencji, rozwiązania tego typu nie mogą być wykorzystane w sytuacjach, w których zakłada się brak lub znikomą kooperację osoby z systemem.

W związku z tym, można zauważyć, że wciąż oczekuje się innowacji w zakresie automatyzacji procesów uwierzytelniania. Odpowiednie rozwiązania muszą uwzględniać zmiany wielu parametrów, takich jak: zmienne warunki świetlne, zmiany perspektywy czy próby oszustwa. Identyfikacja biometryczna w warunkach niekooperacyjnych stała się więc jedną z najbardziej rozwijanych gałęzi nauki oraz obszarem aktywności wielu ośrodków badawczych.

1.2 Cel i zakres pracy

Poniższa praca doktorska poświęcona jest zagadnieniom związanym z identyfikacją biometryczną dla niekooperacyjnych scenariuszy uwierzytelniania. Przeprowadzone badania zostały wykonane w ramach projektu LIDER/027/591/L-4/12/NCBR/2013, zatytułowanego "Biometryczny system uwierzytelniania pozytywnego dla scenariuszy niekooperacyjnych" (COMPACT). Prace badawcze realizowane były pod kierunkiem dra inż. Kamila Grabowskiego. Oprócz autora, zespół badawczy składał się z następujących osób: dr inż. Wojciech Sankowski oraz mgr inż. Damian Kacperski. Celem projektu było zaprojektowanie biometrycznego systemu uwierzytelniania pozytywnego o wysokiej przepustowości niewymagającego od analizowanego podmiotu znaczącej kooperacji. Główne zadania projektu COMPACT mogą zostać podsumowane w następujący sposób:

- Zaprojektowanie stanowiska akwizycji danych, dzięki któremu możliwa będzie rejestracja wysokiej jakości obrazów od osób w ruchu oraz z różnej odległości. Akwizycja obrazów uznawana jest za jeden z kluczowych problemów, który musi zostać rozwiązany, aby dokonać skutecznej identyfikacji w warunkach niekontrolowanych;
- Opracowanie algorytmów detekcji oraz śledzenia osób, dostosowanych do warunków niekontrolowanych. Precyzyjna analiza sceny ma kluczowe znaczenie dla kolejnych etapów przetwarzania. Pozwala na predykcję ruchu śledzonej osoby oraz przewidywanie jej położenia w przyszłości, co umożliwi dokładne ustalenie pozycji kamery wąskiego pola widzenia, w celu pobrania wysokiej jakości zdjęć biometrycznych;
- Opracowanie algorytmów rozpoznawania biometrycznego, dostosowanych do warunków niekontrolowanych. Przedstawione rozwiązanie musi uwzględniać zmiany wielu parametrów, takich jak: zmiany perspektywy, rozmycie ruchu, różna odległość osób od kamery, zasłonięcia czy zmiany mimiki.

Zgodnie z przyjętym podziałem prac, obszar aktywności autora skupiał się na zagadnieniach związanych z automatyzacją procesów rozpoznawania osób. W tym samym czasie, mgr inż. Damian Kacperski poświęcił się pracy nad technikami detekcji i śledzenia osób. Cały zespół badawczy zaangażowany był w konstrukcję bramki biometrycznej i stanowiska akwizycji danych. W tym zakresie, autor był odpowiedzialny za wszystkie aspekty związane z inżynierią oprogramowania (projekt architektury oprogramowanie, implementacje prototypów). W rezultacie, wkład autora może zostać scharakteryzowany w następujący sposób:

- Zaprojektowanie architektury oprogramowania dla systemu biometrycznego średniej skali. Opracowanie rozwiązania oparte jest na zestawie niezależnych usług oraz serwisów, które zapewniają wysoki poziom skalowalności oraz izolacji błędów. Wydajność obliczeniowa uzyskanego rozwiązania została przetestowana oraz pozytywnie oceniona;
- Opracowanie przeglądu wiedzy dotyczącego dostępnych systemów rozpoznawania osób w warunkach niekooperacyjnych. Dodatkowo, autor scharakteryzował i porównał biometryki, które mogą zostać wykorzystane w takich scenariuszach. W rezultacie, określone zostały ograniczenia aktualnych rozwiązań oraz możliwości ich rozwoju;

- Zebranie bazy danych pozwalającej na rozwój algorytmów rozpoznawania osób w warunkach niekooperacyjnych. Przygotowana baza danych zawiera automatycznie pobrane zdjęcia osób przechodzących przez bramkę biometryczną. Zebrane zdjęcia zostały wykonane w zakresie bliskiej podczerwieni oraz zawierają charakterystyczne dla warunków niekooperacyjnych zniekształcenia: rozmycie ruchu, zmienną mimikę, zakrycia twarzy czy różną odległość podmiotu od kamery;
- Opracowanie przeglądu wiedzy dotyczącego najbardziej istotnych prac badawczych związanych z rozpoznawaniem osób w oparciu o cechy twarzy w warunkach niekooperacyjnych. Na podstawie przeanalizowanej literatury, autor zaimplementował oraz porównał skuteczność następujących algorytmów: ThreePatchLBP, HighDimLBP oraz ResNet. Badania zostały zrealizowane na zebranej wcześniej bazie danych;
- Opracowanie algorytmu do filtrowania bazy danych wzorców w celu ograniczenia liczby porównań i zmniejszenia czasu potrzebnego na identyfikację wybranego podmiotu. W celu pokazania poprawności i użyteczności zaproponowanego rozwiązania, autor porównał skuteczność rozpoznawania osób w oparciu o cechy twarzy z wykorzystaniem otrzymanego algorytmu z metodą, w której baza wzorców nie była w żadnej sposób przetwarzana;
- Opracowanie algorytmów rozpoznawania osób w oparciu o okolicę oka (ang. *periocular recognition*), z wykorzystaniem sieci głębokiego uczenia. W celu pozyskania danych uczących, autor przetworzył oraz połączył ze sobą wiele dostępnych publicznie baz danych. Skuteczność otrzymanego rozwiązania została porównana ze znanymi w literaturze metodami klasycznymi;
- Określenie skuteczności rozpoznawania osób w oparciu o okolicę oka (ang. *periocular recognition*) w warunkach niekooperacyjnych. Otrzymane rezultaty zostały porównane ze skutecznością algorytmów rozpoznawania osób w oparciu o cechy twarzy. Na tej podstawie możliwe było ilościowe określenie różnic w skuteczności i przydatności biometryk okolicy oka oraz twarzy;

1.3 Tezy

Teza pierwsza:

Zastosowanie estymacji pozy głowy w biometrycznych systemach uwierzytelniania pozwala na ograniczenie liczby zbioru wzorców użytych do identyfikacji badanej osoby bez znaczącej utraty skuteczności biometrycznego systemu uwierzytelniania.

Teza druga:

Dla biometrycznych systemów uwierzytelniania, w których dąży się do minimalizacji błędu fałszywego odrzucenia (maksymalizacja wygody), algorytm oparty o sieć neuronową głębokiego uczenia daje lepsze wyniki niż dotychczas znane metody identyfikacji osób w oparciu o okolicę oka.

Rozdział 2

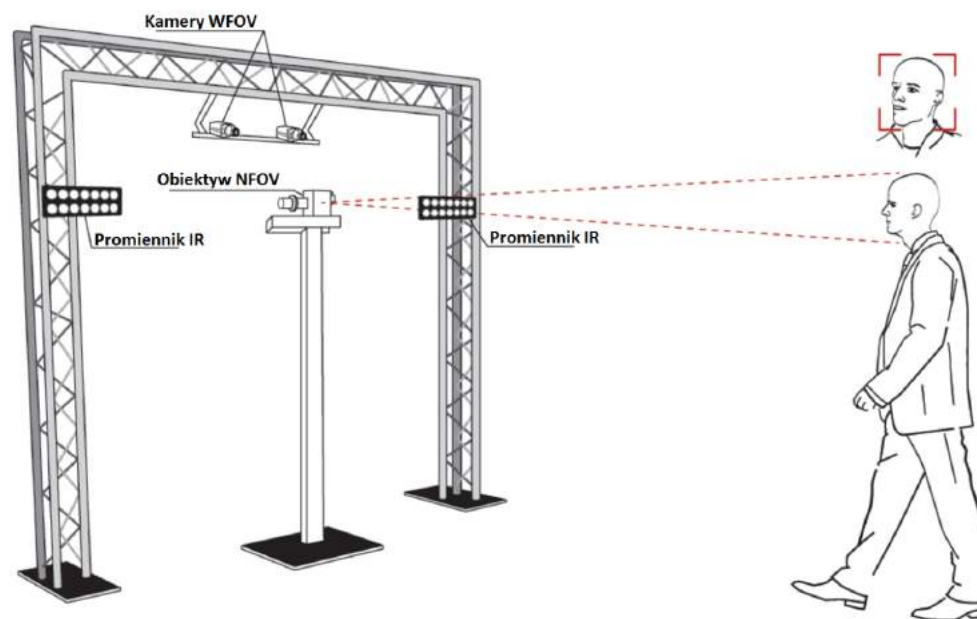
System COMPACT

2.1 Wprowadzenie

Jak wspomniano w Rozdziale 1, opracowanie systemu biometrycznego, który byłby w stanie uwierzytelnić osobę bez jej znaczącej kooperacji, stało się głównym obszarem aktywności wielu badaczy oraz ośrodków akademickich. Większość dostępnych implementacji, jak na przykład [4], [5] lub [6], opiera akwizycję zdjęć biometrycznych na kamerach Pan-Tilt-Zoom (PTZ). Okazuje się jednak, że otrzymywane stopy błędów uwierzytelniania wciąż są zbyt wysokie, aby można było stosować tego typu rozwiązania dla dużych populacji. W związku z tym, w ramach prac badawczych prowadzonych w Katedrze Mikroelektroniki i Technik Informatycznych Politechniki Łódzkiej, zaproponowany został system COMPACT jako alternatywne podejście do omawianych problemów. Przed przystąpieniem do projektowania, zespół postanowił:

- Skoncentrować się warunkach panujących we wnętrzach budynków (korytarze, wejścia do budynków);
- Wykorzystać stereowizję w celu pobierania informacji o lokalizacji osób w scenie trójwymiarowej;
- Zastosować obrazowanie w bliskiej podczerwieni w celu pobierania danych biometrycznych;
- Wykorzystać cechy twarzy oraz okolicy oka (ang. *periocular recognition*) do przeprowadzenia procesu identyfikacji osób;
- Wykorzystać metody estymacji pozycji głowy w celu filtrowania bazy danych zdjęć rejestracyjnych.

Z wymienionych powyżej powodów, system COMPACT został oparty na dwóch kamerach szerokiego pola widzenia (WFOV) oraz jednej kierowanej kamerze wąskiego pola widzenia (NFOV). Kamery WFOV pracują w konfiguracji systemu stereowizyjnego. Obserwują i analizują całą scenę w celu lokalizacji oraz śledzenia położenia osób, które mają zostać zidentyfikowane. W momencie, w którym system uzna, że możliwe jest pobranie cech biometrycznych (na przykład twarz danej osoby jest widoczna), kamera NFOV jest nakierowywana na konkretny obszar sceny w celu wykonania zdjęcia wysokiej rozdzielczości. Następnie, moduły biometryczne wykonują próbę identyfikacji. Koncepcja działania systemu COMPACT została przedstawiona na Rys. 2.1.



Rysunek 2.1: Schemat działania systemu COMPACT

2.2 Konstrukcja systemu

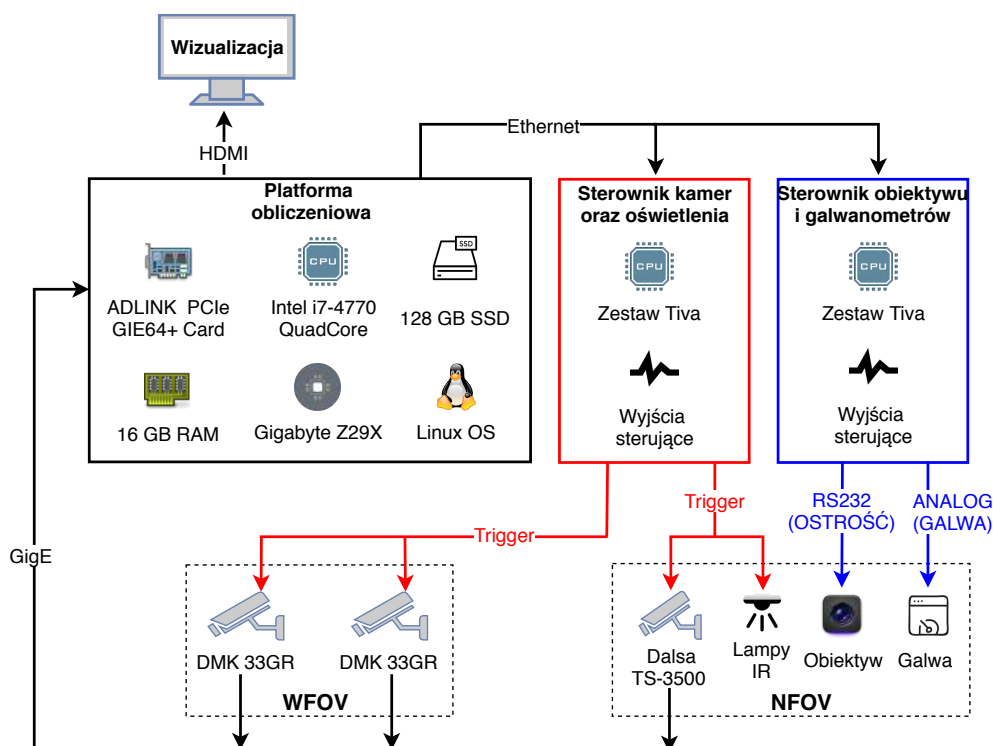
Jak pokazano na Rys. 2.2, system COMPACT składa się z następujących elementów:

- Platforma obliczeniowa;
- System wizyjny WFOV;
- System wizyjny NFOV z systemem światła bliskiej podczerwieni;
- Sterownik kamer oraz światła bliskiej podczerwieni;
- Sterownik obiektywu NFOV oraz galwanometrów.

Platforma obliczeniowa została zrealizowana przy pomocy komputera PC wyposażonego w procesor Intel Core i7-4770 Quad-Core, 16 GB pamięci RAM oraz dysk twardy SSD o pojemności 128 GB. Głównym celem platformy obliczeniowej jest akwizycja obrazów, analiza danych oraz sterowanie modułami zewnętrznymi. Ponadto, platforma obliczeniowa obsługuje dodatkowe aplikacje, które służą do zarządzania i monitorowania pracy systemu. Architektura oprogramowania oraz sposób komunikacji pomiędzy poszczególnymi modułami został opracowany przez autora.

System wizyjny WFOV składa się z dwóch monochromatycznych kamer Imaging Source DKM33GR0134. Każda z kamer pobiera obrazy w rozdzielczości 1280 x 960 pikseli. Kamery wyposażone są w obiektywy H614-MQ wyprodukowane przez firmę Pentax oraz są zamontowane na dedykowanym uchwycie. Oprogramowanie do kalibracji oraz rekonstrukcji współrzędnych w scenie trójwymiarowej zostało zrealizowane przez dra inż. Wojciecha Sankowskiego.

System wizyjny NFOV został oparty na dedykowanym rozwiązaniu zaproponowanym w ramach projektu COMPACT. Składa się z kamery wyposażonej w dedykowany obiektyw z wysuniętą źrenicą, napęd ostrości oraz silniki galwanometryczne. Dzięki temu, możliwe jest nakierowanie urządzenia



Rysunek 2.2: Komponenty sprzętowe systemu COMPACT

na wybrany obszar sceny, tak jak w kamerach PTZ. Jednakże, w tym przypadku ruch jest dużo szybszy i pozbawiony niepożądanych wibracji. Akwizycja danych zrealizowana z wykorzystaniem kamery Teledyne DALSA TS-C3500, która pobiera obrazy w rozdzielczości 3520 x 220 pikseli. Podczas akwizycji, załączane są lampy bliskiej podczerwieni, w celu zapewnienia stabilnych warunków oświetleniowych. Koncepcja obiektywu NFOV oraz założenia konstrukcyjne zostały zaproponowane przez dra inż. Kamila Grabowskiego. Obiektyw został wykonany przez prof. Tadeusza Kryszczyńskiego oraz dra Tomasz Kozłowski z Instytutu Optyki Stosowanej w Warszawie. Mgr inż. Damian Kacperski odpowiedzialny był za integrację poszczególnych modułów oraz oprogramowanie służące do sterowania napędem ostrości oraz silnikami galwanometrycznymi.

W celu wyzwania kamer oraz lamp bliskiej podczerwieni, wykonany został dedykowany sterownik zewnętrzny. Oparty jest on na zestawie *Tiva Connected Launchpad Evaluation Kit*. Problem synchronizacji został zrealizowany biorąc pod uwagę następujące założenia: zapewnienie synchronizacji pomiędzy kamerami WFOV oraz zapewnienie synchronizacji pomiędzy systemem wizyjnym NFOV i WFOV. W rezultacie, w tym samym czasie może pracować tylko jeden zestaw kamer: WFOV lub NFOV, natomiast lampy bliskiej podczerwieni są aktywowane tylko na czas akwizycji obrazów biometrycznych. Sterownik podłączony jest do platformy obliczeniowej przy pomocy sieci Ethernet. Dzięki temu, możliwa jest zmiana parametrów kamer oraz lamp bliskiej podczerwieni w czasie rzeczywistym. Przykładowo, moc światła może być zmniejszona w zależności od odległości osoby od kamery. Podobne rozwiązanie zostało wykorzystane w celu sterowania kamerą NFOV. W tym przypadku, należało zapewnić kontrolę nad mechanizmem ostrości oraz możliwość zmiany położenia zwierciadeł obiektywu. Oba sterowniki zostały wykonane oraz oprogramowane przez mgr inż. Damiana Kacperskiego.

Rys. 2.3 przedstawia prototyp systemu COMPACT.



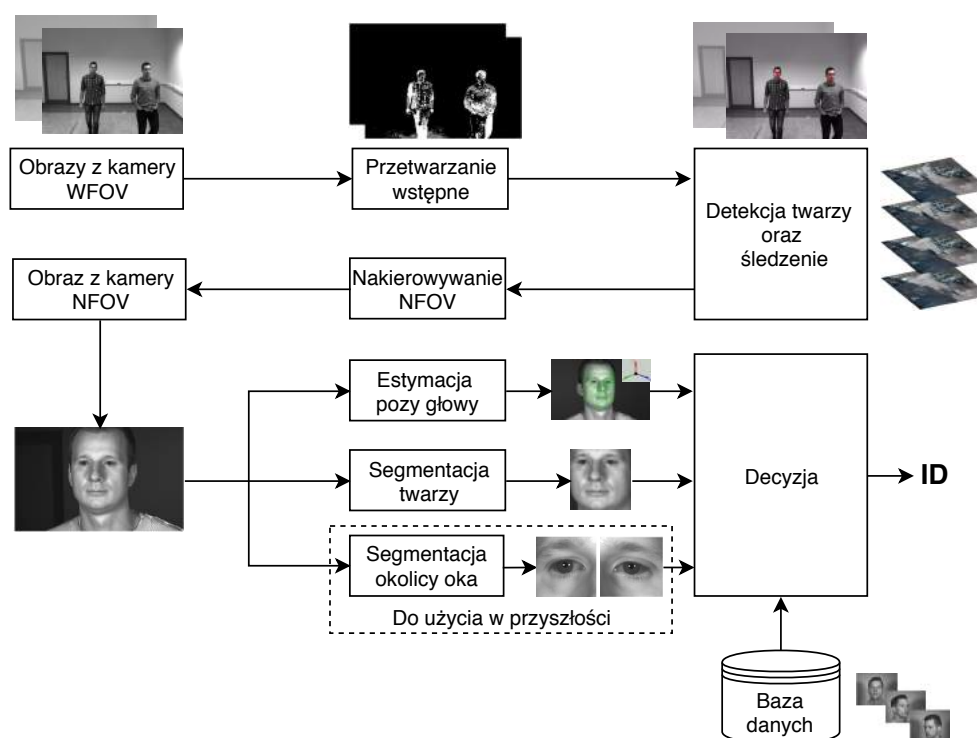
Rysunek 2.3: Zrealizowany prototyp systemu COMPACT

2.3 Przetwarzanie obrazów

Rys. 2.4 przedstawia najważniejsze etapy przetwarzania obrazów w systemie COMPACT. Ogólnie rzecz biorąc, przetwarzanie rozpoczyna się w momencie odebrania obrazów z kamer WFOV. W tym czasie, scena jest analizowana w celu znalezienia położenia osób, które mają zostać zidentyfikowane. Po określeniu lokalizacji twarzy wszystkich podmiotów na każdym z obrazów stereowizyjnych, obliczane jest ich położenie w przestrzeni trójwymiarowej. Następnie, sekwencje pozycji 3D są zapisywane i wykorzystywane przez moduł śledzenia, w którym osoby są oznaczane i grupowane (*Osoba A, B, ...*). W ostatnim etapie, kamera NFOV jest nakierowywana na konkretny obszar sceny w celu pobrania wysokiej rozdzielczości zdjęcia biometrycznego. Systemy wizyjne WFOV i NFOV są zsynchronizowane i skalibrowane ze sobą w taki sposób, że możliwa jest wymiana współrzędnych sceny pomiędzy nimi. Moduły detekcji oraz śledzenia osób zostały zrealizowane przez mgra inż. Damiana Kacperski i są opisane w jego pracy doktorskiej [7].

W momencie, gdy dostępny jest wysokiej rozdzielczości obraz twarzy danej osoby, przeprowadzana jest próba identyfikacji biometrycznej. Moduły odpowiedzialne za rozpoznawanie osób w oparciu o cechy twarzy, rozpoznawanie osób w oparciu o okolicę oka (ang. *periocular recognition*) oraz estymację pozycji głowy zostały opracowane przez autora. Algorytmy identyfikacji zostały wybrane na podstawie szeregu badań wykonanych przez autora oraz zrealizowanych z wykorzystaniem bazy danych zebranej przy pomocy systemu COMPACT. Opis bazy danych znajduje się w Rozdziale 3, natomiast opisy eksperymentów i uzyskane wyniki są zaprezentowane w Rozdziałach 4, 5 oraz 6. W aktualnej

implementacji, rozpoznawanie osób oparte jest o cechy twarzy, jako że uzyskane wyniki pokazały, że ta biometryka pozwala na uzyskanie o rząd wielkości lepszych wyników w porównaniu do okolicy oka, jak zaprezentowano w Sekcji 6.3. Wykorzystany algorytm, scharakteryzowany w Sekcji 4.2, oparty jest o sieć głębokiego uczenia ResNet. Jednocześnie, w celu ograniczenia liczby porównań w bazie danych zdjęć rejestracyjnych, zbiór ten jest filtrowany przy zastosowaniu algorytmów estymacji pozycji głowy. Dzięki temu możliwe jest zmniejszenie czasu potrzebnego na identyfikację pojedynczej osoby, co pokazano w Sekcji 5.3. Jeżeli chodzi o rozpoznawanie osób w oparciu o okolicę oka, to okazało się, że ta biometryka może być użyteczna w analizowanych warunkach, na przykład w sytuacjach, gdy twarz jest zasłonięta. Jednakże, prace nad automatycznym wykrywaniem zasłonień wykraczają poza projekt COMPACT i są uważane jako dalszy element rozwoju systemu. Z tego powodu, ten moduł nie został wykorzystany w obecnej implementacji systemu.



Rysunek 2.4: Schemat blokowy przetwarzania obrazów w systemie COMPACT

2.4 Podsumowanie

System COMPACT został zaproponowany jako nowatorskie oraz alternatywne podejście do rozpoznawania osób w warunkach niekooperacyjnych. Opracowana konstrukcja łączy specjalistyczną wiedzę z zakresu wielu dziedzin, takich jak: elektronika, informatyka czy fizyka optyczna. Zaprezentowane rozwiązanie jest skalowane i uniwersalne. Kluczową nowością jest dedykowany system wizyjny NFOV, który jest w stanie śledzić twarze osób w czasie rzeczywistym. Eliminuje ograniczenia kamer PTZ, szeroko stosowanych w innych rozwiązaniach, takich jak [4], [5] czy [6]. W konsekwencji, możliwe jest pobieranie wysokiej rozdzielczości zdjęć od osób w ruchu oraz przeprowadzenie procesu uwierzytelniania podczas swobodnych przejść przez bramkę biometryczną.

Rozdział 3

Baza danych COMPACT

3.1 Wprowadzenie

Biometryczne bazy danych są jednym z ważniejszych komponentów, które pozwalają na obiektywne porównanie dostępnych algorytmów identyfikacji osób. W momencie, gdy pojawiają się nowe, trudniejsze zbiory, przykuwa to uwagę badaczy, którzy w konsekwencji opracowują nowe metody, żeby uzyskać lepsze wyniki. Jako przykład można podać bazę danych Labeled Faces in The Wild (LFW) [8], która okazała się kluczowa przy rozwoju algorytmów rozpoznawania twarzy w warunkach niekooperacyjnych. Na przestrzeni lat 2007 - 2017, skuteczność dostępnych metod wzrosła z 87% [9] do 98% [10].

Okazuje się jednak, że LFW i inne popularne bazy danych, nie były pobrane w pełni automatycznych warunkach, przez co zdjęcia w nich zawarte nie zawierają niektórych zniekształceń, takich jak rozmycie obrazu czy źle ustawiona ostrość. Z tego powodu, przy wykorzystaniu systemu COMPACT, autor zebrał nową bazę danych, która jest dostępna publicznie dla innych ośrodków badawczych. Akwizycja zdjęć odbyła w semestrze zimowym roku akademickiego 2015/2016. Każda osoba dostarczyła pisemną zgodę, że jej zdjęcia mogą być wykorzystywane w celach naukowych. W porównaniu do znanych już zbiorów, baza COMPACT posiada następujący wkład:

- Obrazy są pobrane w wysokiej rozdzielczości (3520 x 2200 pikseli) przy wykorzystaniu światła bliskiej podczerwieni;
- Obrazy są pobrane w automatyczny sposób, w odległości od 1 do 4 metrów, podczas swobodnych przejść osób przez bramkę biometryczną;
- Obrazy zawierają wiele cech biometrycznych: twarz, okolica oka, ucho. Dzięki temu możliwe jest opracowanie algorytmów fuzji wielu biometryk. Jest to aktualnie jeden z dominujących trendów badawczych związanych z systemami niekooperacyjnymi;
- Dla każdego nagrania, poza zdjęciami biometrycznymi, dostępne są również dane wykorzystywane do realizacji detekcji i śledzenia osób na scenie. Przykładowo, zdjęcia te można wykorzystać do opracowania algorytmów predykcji ruchu;

- Zebrane dane podzielone są na cztery grupy, w zależności od poziomu trudności. Najprostsza grupa zawiera zdjęcia osób patrzących na wprost, podczas gdy najtrudniejsze zdjęcia wprowadzają takie zniekształcenia jak zmienna mimika czy zasłonięcia twarzy. Lista zniekształceń występujących na zebranych zdjęciach przedstawiona jest na Rys. 3.1.



Rysunek 3.1: Lista zniekształceń występujących na zdjęciach dostępnych w bazie danych COMPACT

3.2 Opis bazy danych

Baza danych COMPACT została zebrana wewnątrz jednego z pomieszczeń Katedry Mikroelektroniki i Technik Informatycznych Politechniki Łódzkiej. Akwizycja danych składała się z następujących etapów: wypełnianie ankiety osobowej, pobieranie zdjęć rejestracyjnych oraz akwizycja zdjęć weryfikacyjnych. W celu zapewnienia wiarygodności wyników, zdjęcia rejestracyjne i weryfikacyjne były pobierane w innym czasie, z przerwą trwającą przynajmniej 7 dni. Tabela 3.1 przedstawia najważniejsze parametry bazy danych COMPACT.

Tabela 3.1: Parametry bazy danych COMPACT

Widmo światła	Bliska podczerwień
Liczba osób	108
Liczba zdjęć rejestracyjnych	12312
Liczba zdjęć weryfikacyjnych	31078
Średnia liczba zdjęć weryfikacyjnych na osobę	287
Rozmiar zdjęć	2560x2048 px (rejestracyjne) 3520x2200 px (weryfikacyjne)

Przed przystąpieniem do akwizycji, każda osoba musiała wypełnić kwestionariusz osobowy, który zawierał dziewięć pytań wielokrotnego wyboru, takich jak: płeć, kolor oczu czy kolor włosów. Lista pytań w ankiecie została opracowana na podstawie pracy Tome *et al.* [11]. Celem kwestionariusza było pobranie danych, które mogą być wykorzystane podczas przetwarzania wstępnego, przykładowo w celu przefiltrowania bazy danych zdjęć rejestracyjnych. Tabela 3.2 zawiera listę wszystkich pytań wraz z możliwymi odpowiedziami.

Dane rejestracyjne to wysokiej rozdzielczości zdjęcia twarzy pobrane przy pomocy dedykowanej platformy obrotowej, opisanej w [12]. Pojedyncze przejście składa się z $19 * 3 = 57$ obrazów, które odpowiadają trzem pozycjom wertykalnym (spojrzenie w górę, spojrzenie przed siebie, spojrzenie w dół), pobranym w zakresie kątów od -45° do 45° , z krokiem 5° . Dla każdej osoby pobrano dwa niezależne przejścia, tak, aby zredukować ilość błędów, wynikających z mrugania czy ruchów głowy. Rys. 3.2 przedstawia przykładowe zdjęcia rejestracyjne.

Tabela 3.2: Lista pytań zawartych w ankiecie bazy danych COMPACT

Cecha	Odpowiedzi
Wiek	N
Wzrost	N
Płeć	Mężczyzna, kobieta
Kolor włosów	Łysy, Blond, Brunet, Czarny
Rodzaj włosów	Łysy, Krótkie, Kręcone, Faliste, Proste
Długość włosów	Łysy, krótkie, Średnie, Długie
Zarost	Brak, Krótka, Długa
Kolor oczu	Niebieski, Zielony, Brązowy
Okulary	Tak, Nie



Rysunek 3.2: Przykładowe zdjęcia rejestracyjne zawarte w bazie danych COMPACT

Dane weryfikacyjnej zostały pobrane w sposób automatyczny, podczas swobodnych przejść osób przez bramkę biometryczną COMPACT (patrz Rozdział 2). Rys. 3.3 przedstawia przykładowe zdjęcia weryfikacyjne. Składają się z dwóch podzbiorów: wysokiej rozdzielczości zdjęć twarzy wykonanych przez kamerę NFOV oraz sekwencji video zarejestrowanych przez kamery WFOV. Ponieważ nie wszystkie obrazy zawierały użyteczne informacje (na przykład niewidoczna twarz z powodu błędnej detekcji), dostępne zdjęcia zostały ręcznie przejrane i przefiltrowane. W celu zapewnienia odpowiednich poziomów trudności, autor zaproponował cztery scenariusze podczas nagrywania przejść. Pierwsze dwa scenariusze odbywały się w warunkach kontrolowanych, natomiast dwa pozostałe w warunkach niekontrolowanych. Osoby, które nosiły okulary, były poproszone o ich zdjęcie dla pierwszych trzech scenariuszy. Każdy scenariusz został powtórzony dwukrotnie, co w sumie daje 8 nagrań dla każdej osoby. Zaproponowane zostały następujące scenariusze:

- **Scenariusz 1:** Osoba przechodzi przez bramkę patrząc na wprost. Pozwala to na pobranie obrazów frontalnych, niezawierających żadnych zniekształceń. Takie dane są użyteczne podczas implementacji algorytmów oraz służą do wyznaczenia wyników referencyjnych;

- **Scenariusz 2:** Osoba przechodzi przez bramkę patrząc w lewo (pierwsze przejście) lub w prawo (drugie przejście). Takie dane mogą zostać wykorzystane do badań określających wpływ pozycji głowy na skuteczność algorytmów identyfikacji;
- **Scenariusz 3:** Osoba przechodzi przez bramkę rozglądając się w sposób losowy, zarówno w płaszczyźnie poziomej jak i pionowej. Takie dane reprezentują warunki niekooperacyjne, zawierając zarówno zdjęcia frontalne jak i profilowe;
- **Scenariusz 4:** Osoba przechodzi przez bramkę rozglądając się w sposób losowy, podobnie jak w scenariuszu trzecim. Dodatkowo, osoby były proszone o wykonanie gestu, który spowodowałby zasłonięcie części twarzy, na przykład poprawa fryzury podczas przejścia. Takie dane reprezentują najbardziej rzeczywiste zdjęcia niekooperacyjne, ze wszystkimi możliwymi zniekształceniami.



Rysunek 3.3: Przykładowe zdjęcia weryfikacyjne zawarte w bazie danych COMPACT (lewa kamera WFOV, prawa kamera WFOV, kamera NFOV)

3.3 Podsumowanie

Baza danych COMPACT została opracowana w celu rozwoju biometrycznych systemów uwierzytelniania działających w warunkach niekooperacyjnych. Zawiera zdjęcia 108 osób oraz składa się z trzech elementów: ankiety osobowej, zdjęć rejestracyjnych twarzy oraz zdjęć weryfikacyjnych. Ankieta zawiera dziewięć pytań wielokrotnego wyboru, takich jak: wzrost, płeć czy rodzaj włosów. Zdjęcia rejestracyjne zostały zebrane na platformie obrotowej, dzięki czemu dostępne są ujęcia twarzy z różnych kątów wertykalnych i horyzontalnych. W sumie dla każdej osoby zebrano 114 zdjęć. Zdjęcia weryfikacyjne zawierają swobodne przejścia osób przez bramkę biometryczną COMPACT. Dane zostały następnie pogrupowane w cztery kategorie, w zależności od stopnia trudności zdjęć. Najważniejszą różnicą w porównaniu do innych, podobnych baz danych jest fakt, że zdjęcia weryfikacyjne zostały zebrane w pełni automatyczny sposób. Dzięki temu na zdjęciach dostępne są zniekształcenia odzwierciedlające rzeczywiste warunki pracy systemów niekooperacyjnych, takie jak brak ostrości lub rozmycie spowodowane ruchem osoby.

Rozdział 4

Rozpoznawanie osób w oparciu o cechy twarzy

4.1 Wprowadzenie

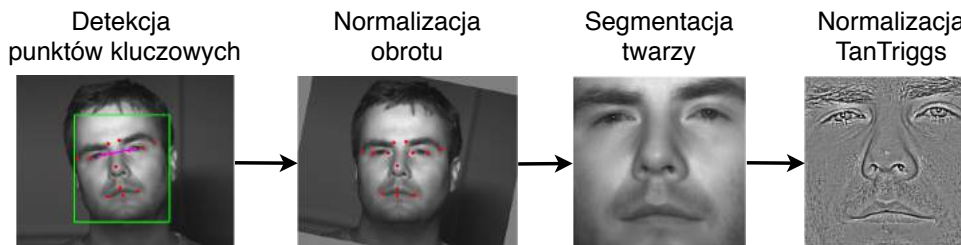
Algorytmy rozpoznawania osób w oparciu o cechy twarzy znajdują szerokie zastosowanie w systemach uwierzytelniania biometrycznego, które działają w warunkach niekooperacyjnych. Dzieje się tak głównie ze względu na fakt, że pozyskiwanie obrazów twarzy jest stosunkowo łatwe, nieinwazyjne oraz nie wymaga interakcji pomiędzy systemem a osobą, która ma zostać zidentyfikowana. W ciągu ostatnich lat, ze względu na rosnące zapotrzebowania na tego typu rozwiązania, nastąpił znaczący postęp w zakresie technik rozpoznawania twarzy. Przykładowo, skuteczność dostępnych algorytmów wzrosła z 87% [9] w 2009 roku do 98% w 2017 roku [10]. Jednakże, okazuje się, że dostępne zbiory danych testowych, które wykorzystuje się przy opracowywaniu nowych algorytmów (na przykład baza danych LFW [8]), nie odzwierciedlają w pełni rzeczywistych warunków pracy systemów nieoperacyjnych. Dostępne tam zdjęcia nie zawierają takich zniekształceń jak rozmycie obrazu czy źle ustawiona ostrość. Z tego powodu uważa się, że brak odpowiednich baz danych testowych jest jednym z głównych ograniczeń dalszego rozwoju algorytmów rozpoznawania twarzy, przeznaczonych do pracy w warunkach niekooperacyjnych.

Z tego powodu, w ramach projektu COMPACT została zebrana nowa baza danych, która zawiera swobodne przejścia osób przez bramkę biometryczną (patrz Rozdział 3). W celu ukazania jej użyteczności oraz dostarczenia referencyjnych rezultatów dla innych badaczy, autor wykonał badania pokazujące skuteczność wybranych algorytmów służących do rozpoznawania osób w oparciu o cechy twarzy na zebranych zbiorze. Ponadto, przeprowadzone eksperymenty pozwoliły określić skuteczność identyfikacji biometrycznej w systemie COMPACT, ponieważ twarz jest główną biometryką wykorzystywaną w procesie uwierzytelniania.

4.2 Opis wykorzystanych algorytmów

Aby zwiększyć skuteczność algorytmów rozpoznawania, dane wejściowe są normalizowane w procesie wstępnej obróbki. W tym celu zastosowana została następująca procedura. W pierwszej kolejności, przy

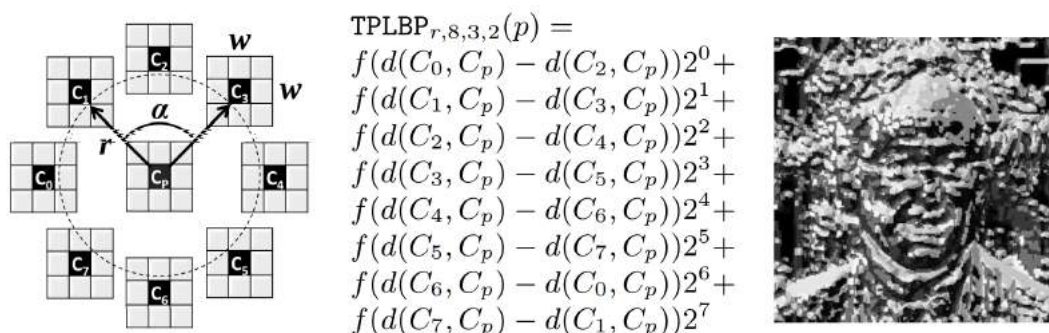
użyciu punktów kluczowych twarzy (na przykład położenie nosa lub ust) [13], obraz wejściowy jest obracany w taki sposób, aby środki oczu były umieszczone na tej samej wysokości. Następnie, twarz jest wycinana i skalowana do rozmiarów 150 x 150 pikseli. W ostatnim etapie, następuje wyrównanie iluminacji poprzez zastosowanie algorytmu TanTriggs, opisanego w [14]. Rys. 4.1 przedstawia schemat blokowy zaprezentowanej procedury.



Rysunek 4.1: Schemat blokowy procedury służącej do normalizacji obrazów twarzy

W celu przedstawienia szerokiej analizy, autor zaimplementował trzy algorytmy służące do rozpoznawania osób w oparciu o cechy twarzy: ThreePatchLBP [15], HighDimLBP [16] and ResNet [17]. Wybór tych metod nie był przypadkowy, każdy algorytm opiera się na innym sposobie działania. Poniżej znajduje się krótka charakterystyka zastosowanych algorytmów.

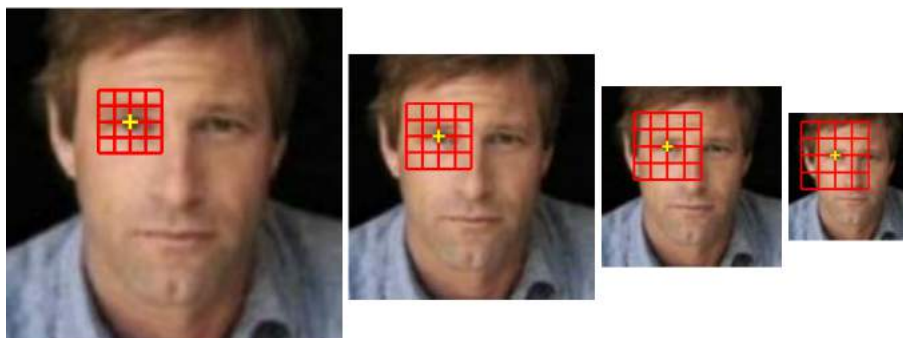
Algorytm ThreePatchLBP (TPLBP) jest wariacją algorytmu Local Binary Pattern (LBP). Zasada działania algorytmu TPLBP opiera się na analizie otoczenia piksela w postaci okręgu o zadanym promieniu i stałej liczbie punktów. Następnie, otoczenie to jest binaryzowane lokalnie z progiem w postaci elementu środkowego. Zostało to zaprezentowane na Rys. 4.2. W kolejnym kroku, zakodowany w ten sposób obraz dzielony jest na wiele małych regionów. W tym celu, wykorzystuje się siatkę o rozmiarze 10 na 10. Następnie, dla każdego otrzymanego w ten sposób obszaru, obliczany jest histogram, który mierzy częstotliwość występowania poszczególnych kodów binarnych. W ostatnim etapie, otrzymane histogramy łączone są w jeden wektor cech. W celu uzyskania współczynnika podobieństwa, otrzymane wektory porównywane są przy pomocy dystansu Battacharyya [18].



Rysunek 4.2: Lewe zdjęcie: procedura pozyskiwania kodu TPLBP dla wybranego piksela. Środkowe zdjęcie: przykładowy kod TPLBP obliczony dla $S = 8$, $w = 3$ oraz $\alpha = 2$. Prawe zdjęcie: przykładowy obraz zakodowany przy pomocy algorytmu TPLBP. Zaadaptowano z [15]

Algorytm HighDimLBP [16] również jest wariacją algorytmu LBP. W tym przypadku, wektor cech konstruowany jest poprzez analizę obszarów zlokalizowanych wokół punktów kluczowych twarzy (kąćki ust, kąćki oczu). Procedura jest następująca. Punkty kluczowe lokalizowane są przy pomocy algorytmu zaprezentowanego w [13]. Kolejno, dla każdego punktu kluczowego, definiowana jest siatka

o rozmiarze 4 na 4. Następnie, na każdym elemencie siatki wykonywany jest deskryptor LBP. W rezultacie otrzymujemy lokalny wektor cech, unikalny dla każdego punktu. W kolejnych krokach, obraz wejściowy jest skalowany w dół i cały proces jest powtarzany, przy zachowaniu takiego samego rozmiaru siatki. Opisana procedura została przedstawiona na Rys. 4.3. W ostatnim etapie, wszystkie lokalne wektory cech łączone są w jeden wektor opisujący całą twarz. Początkowa skala wykorzystywana jest do opisu lokalnych informacji wokół każdego z punktów kluczowych, podczas gdy kolejne zmniejszenia obrazu pozwalają na pobranie informacji dotyczących cech globalnych, takich jak kształt twarzy. W celu uzyskania współczynnika podobieństwa, otrzymane wektory cech porównywane są przy pomocy dystansu Battacharyya [18], tak jak w przypadku algorytmu TPLBP.



Rysunek 4.3: Lokalne wektory cech pobrane dla przykładowego punktu kluczowego twarzy. Zaadaptowano z [16]

Algorytm ResNet to wariacja sieci neuronowej głębokiego uczenia ResNet-34 [17]. Zawiera 29 warstw splotowych, a jej implementacja pochodzi z biblioteki dlib [10]. Architektura wykorzystywanej sieci została podsumowana w Tabeli 4.1. Model sieci został nauczony przy użyciu 3 milionów zdjęć twarzy. Baza ucząca została skonstruowana w wyniku połączenia ze sobą wielu publicznych zbiorów, takich jak FaceScrub, VGG oraz dużej liczby zdjęć pobranych z Internetu. Dla każdej twarzy, wykorzystywana sieć zwraca wektor cech składający się ze 128 elementów. Następnie, otrzymane w ten sposób wektory są ze sobą porównywane przy użyciu dystansu Euklidesowego [19].

Tabela 4.1: Architektura sieci ResNet-29

	Oryginalny ResNet-34	Zmodyfikowany ResNet-29
Conv-1	7x7, 64, stride 2	7x7, 32, stride 2
Conv-2	3x3 max pool, stride 2	
	3x $\begin{pmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{pmatrix}$	3x $\begin{pmatrix} 3 \times 3, 32 \\ 3 \times 3, 32 \\ 3 \times 3, 32 \end{pmatrix}$
Conv-3	4x $\begin{pmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{pmatrix}$	3x $\begin{pmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \\ 3 \times 3, 64 \end{pmatrix}$
Conv-4	6x $\begin{pmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{pmatrix}$	2x $\begin{pmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{pmatrix}$
Conv-5	3x $\begin{pmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{pmatrix}$	2x $\begin{pmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{pmatrix}$
	average pool, 1000-d fc	average pool, 128-d fc

4.3 Eksperymenty

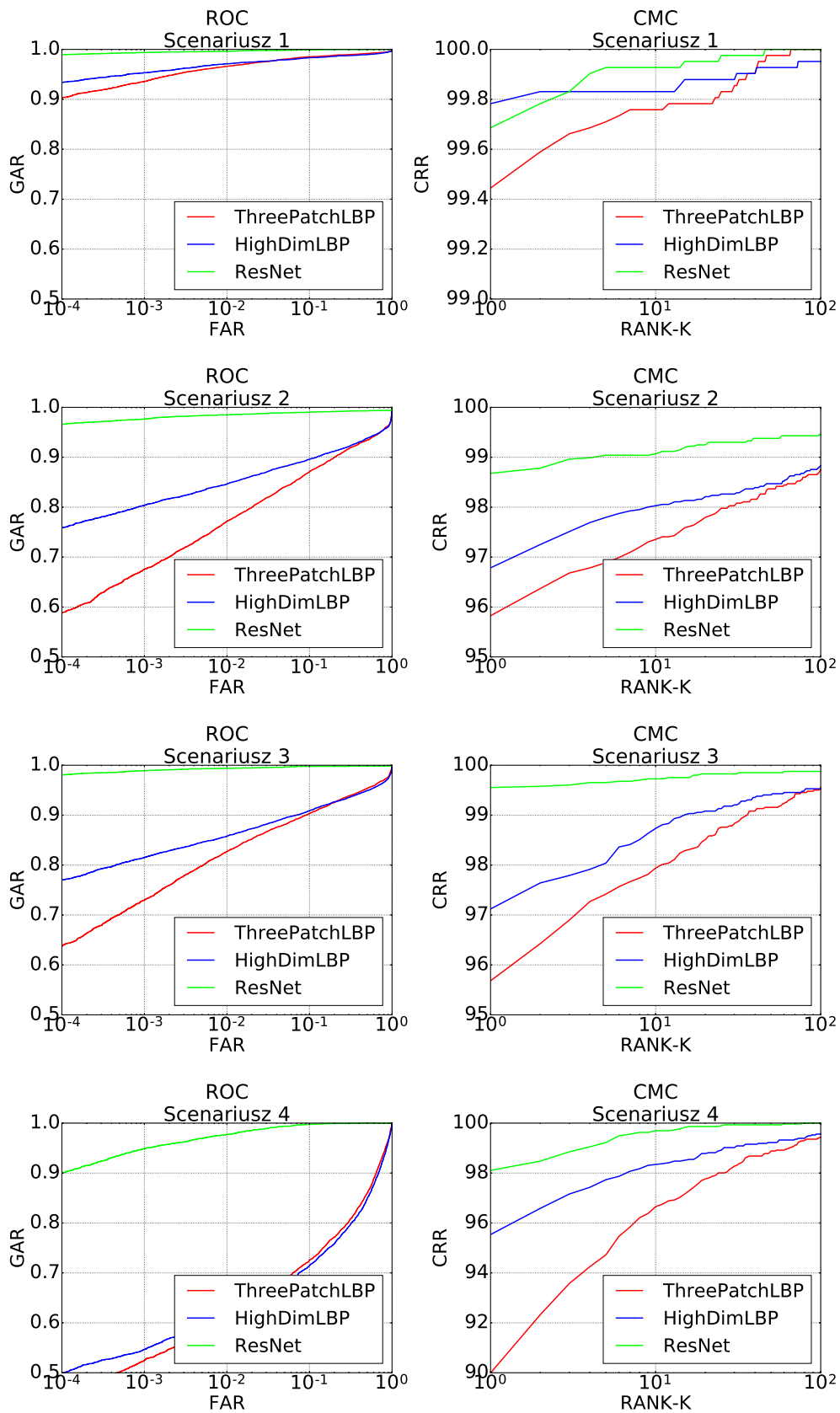
Skuteczność algorytmów rozpoznawania osób w oparciu o cechy twarzy została przebadana dla każdego scenariusza dostępnego w bazie danych COMPACT. Otrzymane wyniki są podsumowane w Tabeli 4.2. Rys. 4.4 przedstawia uzyskane krzywe ROC oraz CMC.

Tabela 4.2: Skuteczność algorytmów rozpoznawania twarzy na bazie danych COMPACT

Scenariusz 1			
	ThreePatchLBP	HighDimLBP	ResNet
EER, %	2.54 ± 0.25	2.43 ± 0.12	0.44 ± 0.08
AUC, %	99.06 ± 0.16	98.91 ± 0.13	99.96 ± 0.01
RANK-1, %	99.20 ± 0.14	99.39 ± 0.47	99.67 ± 0.08
Scenariusz 2			
	ThreePatchLBP	HighDimLBP	ResNet
EER, %	12.07 ± 0.74	10.33 ± 0.92	1.41 ± 0.65
AUC, %	92.53 ± 0.89	93.32 ± 0.85	99.25 ± 0.61
RANK-1, %	95.48 ± 0.82	95.88 ± 0.97	98.61 ± 0.67
Scenariusz 3			
	ThreePatchLBP	HighDimLBP	ResNet
EER, %	9.67 ± 0.72	9.21 ± 0.72	0.62 ± 0.20
AUC, %	94.64 ± 0.59	94.56 ± 0.62	99.83 ± 0.12
RANK-1, %	95.20 ± 0.75	96.47 ± 0.86	99.56 ± 0.14
Scenariusz 4			
	ThreePatchLBP	HighDimLBP	ResNet
EER, %	21.76 ± 1.84	22.67 ± 1.42	1.58 ± 0.28
AUC, %	85.44 ± 1.29	84.41 ± 1.14	99.88 ± 0.02
RANK-1, %	89.03 ± 1.42	93.69 ± 2.15	98.06 ± 0.64

4.4 Podsumowanie

W celu ukazania użyteczności bazy danych COMPACT oraz dostarczenia referencyjnych rezultatów dla innych badaczy, autor wykonał badania określające skuteczność znanych w literaturze algorytmów rozpoznawania osób w oparciu o cechy twarzy. Przeprowadzone eksperymenty pozwoliły na sformułowanie następujących wniosków. Po pierwsze, można zauważyć, że skuteczność rozpoznawania maleje dla kolejnych scenariuszy dostępnych w bazie danych COMPACT. Dla każdej metody, najlepsze rezultaty otrzymano dla scenariusza pierwszego, który zawiera głównie zdjęcia frontalne. Najgorsza skuteczność zaraportowana jest dla scenariusza czwartego, który najbardziej oddaje rzeczywiste warunki pracy systemów niekooperacyjnych. Oznacza to, że baza danych COMPACT została zaprojektowana poprawnie i spełnia swoje wymagania. Po drugie, uzyskane wyniki wskazują, że sieci głębokiego uczenia pozwalają na uzyskanie znacząco lepszej skuteczności rozpoznawania niż metody klasyczne. Można uznać, że jednym z ograniczeń rozwoju systemów uwierzytelniania jest brak dostępu do dużych zestawów danych uczących. W końcu, przedstawione wyniki pozwalają na określenie efektywności procesu uwierzytelniania w systemie COMPACT. Wykorzystując algorytm ResNet, możliwa jest identyfikacja osób przechodzących przez bramkę biometryczną ze skutecznością równą około 98% (parametr RANK-1). Oznacza to, że opracowany prototyp systemu został wykonany poprawnie.



Rysunek 4.4: Krzywe ROC i CMC dla algorytmów rozpoznawania twarzy na bazie danych COMPACT

Rozdział 5

Estymacja pozy głowy

5.1 Wprowadzenie

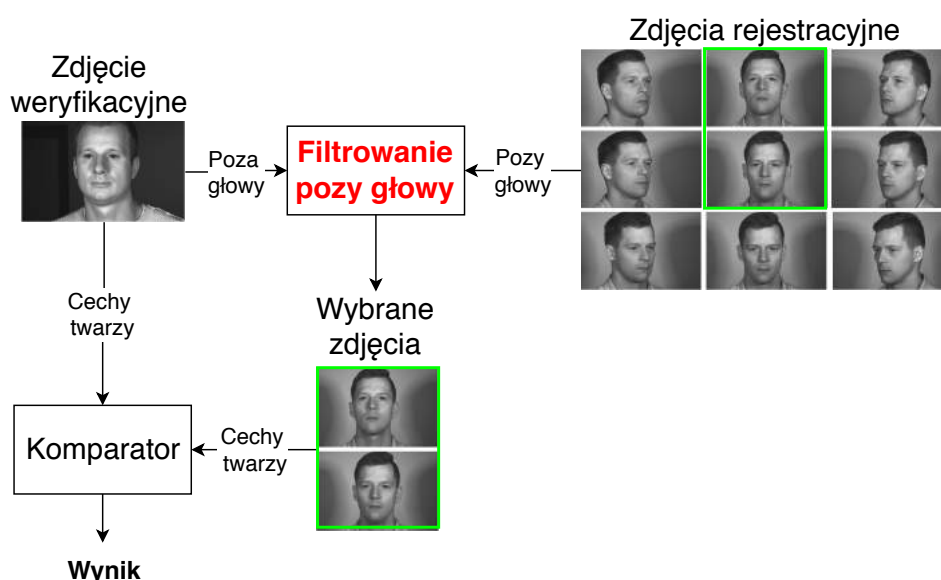
Istnieje wiele zastosowań, gdzie estymacja pozy głowy odgrywa ogromne znaczenie. Przykładowo, techniki tego typu są szeroko wykorzystywane w systemach śledzenia [20], systemach wspomagania kierowcy [21] czy programach służących do interakcji człowiek-komputer [22]. W większości tych zastosowań, dokładność modułu odpowiadającego za estymację pozy głowy jest kluczowa dla dalszych etapów przetwarzania. W konsekwencji, algorytmy estymacji pozy głowy muszą być odporne na zmiany takich parametrów jak: zniekształcenia kamery, zmienna mimika twarzy czy zakrycia twarzy (na przykład ze względu na obecność okularów).

Jednym z innych zastosowań, gdzie estymacja pozy głowy odgrywa istotną rolę i ma znaczący wpływ na skuteczność przetwarzania, są systemy uwierzytelniania biometrycznego, które działają w warunkach niekooperacyjnych. Jako, że osoba nie bierze czynnego udziału w procesie identyfikacji, zmiany pozy głowy są jednym z głównych czynników wpływających na skuteczność rozpoznawania w takich scenariuszach. Aby rozwiązać ten problem, powszechne jest, że dane rejestracyjne składają się z wielu zdjęć poszczególnej osoby, pobranych pod różnymi kątami. Takie rozwiązanie zostało zastosowane również w systemie COMPACT (patrz Rozdział 3). Zdjęcia weryfikacyjne porównywane są wtedy ze wszystkimi zdjęciami rejestracyjnymi i ostateczny wynik obliczany jest przy wykorzystaniu technik fuzji. Jednakże, takie podejście powoduje relatywnie duży narzut obliczeniowy, ze względu na dużą liczbę porównań potrzebnych do wykonania podczas pojedynczej próby identyfikacji. W związku z tym, autor zaproponował algorytm do filtrowania bazy danych zdjęć rejestracyjnych w celu ograniczenia liczby porównań i zmniejszenia czasu potrzebnego na identyfikację wybranej osoby. Optymalizacja polega na wykorzystaniu algorytmów estymacji pozy głowy do ograniczenia liczby wzorców tylko do tych, dla których poza głową jest podobna do tej, która jest widoczna na badanej próbce weryfikacyjnej. W celu pokazania poprawności i użyteczności zaproponowanego rozwiązania, autor porównał skuteczność rozpoznawania osób w oparciu o cechy twarzy z wykorzystaniem opracowanego algorytmu filtrowania z metodą, w której baza wzorców nie była w żadnej sposób przetwarzana (*brute force*).

5.2 Opis zaproponowanej metody

Rys. 5.1 przedstawia schemat blokowy zaproponowanej metody do filtrowania bazy danych wzorców. Realizacja rozpoznawania twarzy została oparta na algorytmie ResNet, opisanym w Rozdziale 4. Estymacja pozy głowy została z kolei zrealizowana przy pomocy algorytmu 3DMM (3D Morphable Model), wykorzystując implementację dostępną w bibliotece *eos* [23]. Zaproponowana metoda składa się z następujących etapów:

- Dla każdej próbki rejestracyjnej i weryfikacyjnej obliczany jest kąt odchylenia i pochylenia głowy;
- Zdjęcia rejestracyjne są ograniczane tylko do tych, na których poza głowy jest podobna do tej, która jest na aktualnie analizowanym zdjęciu weryfikacyjnym;
- Zdjęcie weryfikacyjne jest porównywane z każdym pozostawionym zdjęciem rejestracyjnym, każde porównanie zwraca wartość określającą procent podobieństwa. Ostateczny wynik wybierany jest jako maksymalna wartość podobieństwa.



Rysunek 5.1: Schemat blokowy zaproponowanej metody do filtrowania bazy danych wzorców

Filtrowanie zdjęć rejestracyjnych odbywa się w następujący sposób. Dla każdej próbki weryfikacyjnej określone jest, czy obliczona poza głowy jest zaufana czy też nie. Załóżmy, że α_y oraz α_p przedstawiają odpowiednio kąt odchylenia i pochylenia głowy. Estymacja pozy głowy uważana jest za zaufaną, jeżeli wartość bezwzględna obliczonego kąta α_y oraz kąta α_p jest mniejsza lub równa wartości progowej T_y oraz T_p . Zostało to zilustrowane na Równaniu (5.1). W przeciwnym wypadku, estymacja uważana jest za niezaufaną. Jak zauważono w [24], dokładność algorytmów estymacji pozy głowy maleje wraz ze wzrostem kąta odchylenia lub nachylenia głowy. Z tego powodu, w opracowanej metodzie, wartości progów T_y oraz T_p ustalono na 40° .

$$estymacja\ zaufana = \begin{cases} \text{prawda} : & |\alpha_y| \leq T_y \wedge |\alpha_p| \leq T_p \\ \text{fałsz} : & \text{w przeciwnym razie} \end{cases} \quad (5.1)$$

W sytuacji, gdy estymacja pozy głowy uznawana jest jako niezauważalna, wykorzystywana jest metoda *brute force*, w której próbka weryfikacyjna porównywana jest ze wszystkimi zdjęciami rejestracyjnymi. W przeciwnym wypadku, zbiór rejestracyjny ograniczany jest tylko do tych zdjęć, na których pozy głowy są podobne. Załóżmy, że $\Delta\alpha_y$ przedstawia różnicę pomiędzy kątami odchylenia dla przykładowej próbki weryfikacyjnej i rejestracyjnej. Analogicznie, $\Delta\alpha_p$ opisuje różnicę pomiędzy kątami pochylenia. Poza głowy ze zdjęcia rejestracyjnego uznawana jest za podobną w sytuacji, gdy wartość bezwzględna parametrów $\Delta\alpha_y$ oraz $\Delta\alpha_p$ jest mniejsza lub równa wartości progowej t_y oraz t_p . Zostało to zilustrowane na Równaniu (5.2). Kąt przechyłu nie jest brany pod uwagę, ponieważ obrazy twarzy są normalizowane w taki sposób, że środki oczu znajdują się na tej samej wysokości (patrz opis procedury normalizacji w Sekcji 4.2).

$$\text{podobna poza} = \begin{cases} \text{prawda} : & |\Delta\alpha_y| \leq t_y \wedge |\Delta\alpha_p| \leq t_p \\ \text{fałsz} : & \text{w przeciwnym razie} \end{cases} \quad (5.2)$$

5.3 Eksperymenty

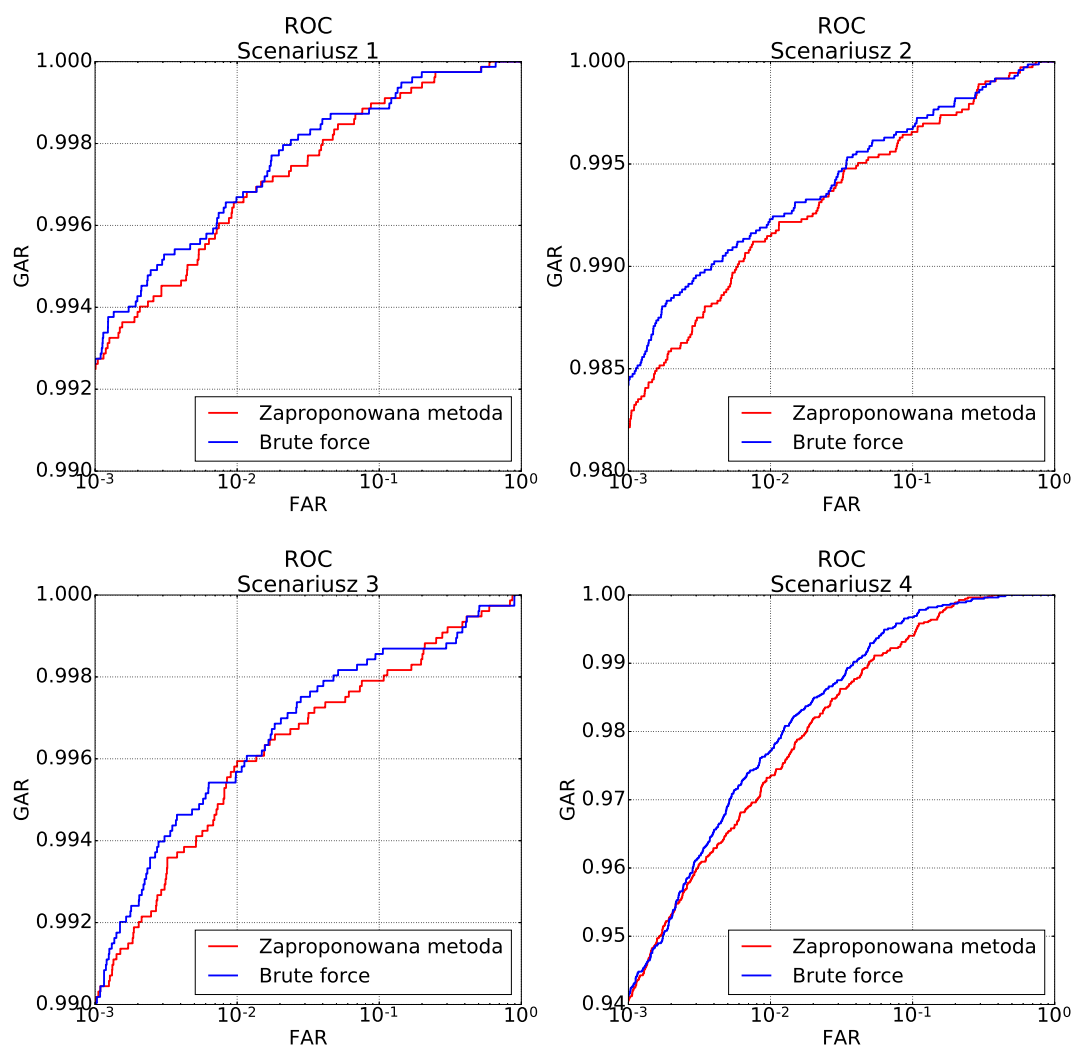
W celu pokazania poprawności i użyteczności zaproponowanego rozwiązania, autor porównał skuteczność rozpoznawania osób w oparciu o cechy twarzy z wykorzystaniem otrzymanego algorytmu z metodą, w której baza wzorców nie była w żaden sposób przetwarzana. Progi t_y oraz t_p zostały ustawione odpowiednio na 20° oraz 15° . Eksperymenty zostały przeprowadzone osobno dla każdego scenariusza dostępnego w bazie danych COMPACT. Otrzymane wyniki są podsumowane w Tabeli 5.1. Tabela 5.2 prezentuje średnią liczbę porównań wykonanych przy analizie pojedynczej próbki. Rys. 4.4 przedstawia natomiast uzyskane krzywe ROC oraz CMC.

Tabela 5.1: Porównanie skuteczności rozpoznawania twarzy z wykorzystaniem zaprezentowanego algorytmu filtrowania z metodą, w której baza wzorców nie była przetwarzana (*brute force*)

Scenariusz 1		
	Zaproponowana metoda	Brute force
EER, %	0.50 ± 0.04	0.46 ± 0.06
AUC, %	99.96 ± 0.01	99.96 ± 0.01
RANK-1, %	99.75 ± 0.05	99.72 ± 0.05
Scenariusz 2		
	Zaproponowana metoda	Brute force
EER, %	0.88 ± 0.13	0.82 ± 0.11
AUC, %	99.87 ± 0.04	99.88 ± 0.03
RANK-1, %	99.01 ± 0.11	99.18 ± 0.12
Scenariusz 3		
	Zaproponowana metoda	Brute force
EER, %	0.58 ± 0.07	0.52 ± 0.06
AUC, %	99.92 ± 0.03	99.92 ± 0.04
RANK-1, %	99.56 ± 0.07	99.58 ± 0.07
Scenariusz 4		
	Zaproponowana metoda	Brute force
EER, %	1.89 ± 0.22	1.68 ± 0.16
AUC, %	99.82 ± 0.03	99.86 ± 0.03
RANK-1, %	97.65 ± 0.49	97.87 ± 0.38

Tabela 5.2: Średnia liczba porównań wykonanych przy analizie pojedynczej próbki weryfikacyjnej dla algorytmu z wykorzystaniem filtrowania zdjęć rejestracyjnych oraz metodą (*brute force*)

	Scenariusz 1	Scenariusz 2	Scenariusz 3	Scenariusz 4
Brute force	57	57	57	57
Zaproponowana metoda	12	30	15	14
Zysk z optymalizacji	79%	47%	74%	75%



Rysunek 5.2: Porównanie skuteczności rozpoznawania twarzy z wykorzystaniem zaprezentowanego algorytmu z metodą, w której baza wzorców nie była przetwarzana (*brute force*)

5.4 Podsumowanie

W systemach działających w warunkach zmniejszonej kooperacji (na przykład system COMPACT), powszechne jest, że dane rejestracyjne często składają się z wielu zdjęć poszczególnej osoby, pobranych pod różnymi kątami. Z tego powodu, aby skutecznie przeprowadzić proces identyfikacji, musi zostać wykonana duża liczba porównań. W celu optymalizacji, autor zaproponował wykorzystanie algorytmów estymacji pozy głowy do ograniczenia liczby wzorców tylko do tych, dla których poza głową jest podobna do tej, która jest widoczna na badanej próbce weryfikacyjnej. Opracowane rozwiązanie okazało się poprawne i skuteczne. Dla każdego przeprowadzonego eksperymentu, liczba porównań została

znacząco zmniejszona, natomiast uzyskane stopy błędów pozostały na tym samym poziomie. Oznacza to, że system może zignorować dużą część zdjęć rejestracyjnych i mimo to przeprowadzić pozytywną próbę uwierzytelnienia. **Otrzymane wyniki i przedstawione wnioski dowodzą tezy pierwszej.** Należy również wspomnieć, że optymalne progi filtrowania zależą od dokładności algorytmu estymacji pozy głowy oraz tego, jak gęsto próbkowane są zdjęcia rejestracyjne. Z tego powodu, dobranie odpowiednich wartości powinno odbywać się indywidualnie, w zależności od implementacji użytych algorytmów oraz przypadku użycia. Wadą zaprezentowanego algorytmu jest to, że baza danych zdjęć rejestracyjnych musi być skonstruowana w ściśle określony sposób i zawierać dużą liczbę zdjęć dla każdej z dostępnych osób.

Rozdział 6

Rozpoznawanie osób w oparciu o okolicę oka

6.1 Wprowadzenie

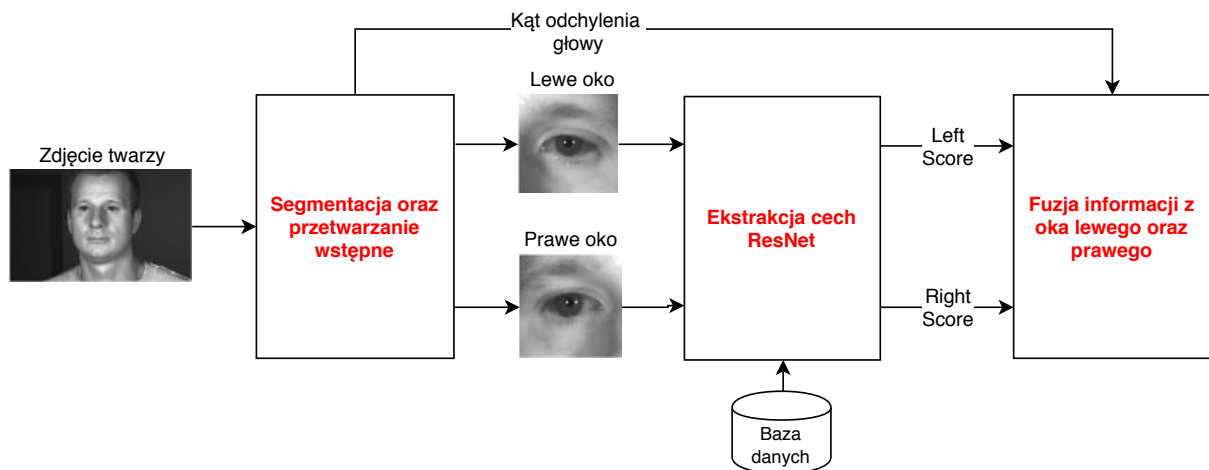
W ostatnich latach, duża liczba badaczy i ośrodków naukowych poświęciła się pracy nad technikami rozpoznawania osób w oparciu o okolicę oka (*ang. periocular recognition*). Informacja biometryczna pobierana jest z w tym przypadku z tekstury skóry oraz kształtu brwi, rzęs czy powiek. Rozpoznawanie osób w oparciu o okolicę oka jest przydatne w sytuacjach, w których obraz tęczówki jest niedokładny lub twarz jest zasłonięta, na przykład podczas demonstracji lub choroby. Dodatkowo, uważa się, że obszar okolicy oka wydaje się być bardziej odporny na takie czynniki jak efekt starzenia [25] czy zmiana perspektywy [26].

Mimo rosnącego zainteresowania, większość algorytmów służących do rozpoznawania osób w oparciu o okolicę oka opiera się na relatywnie prostych technikach holistycznych. Z drugiej strony, kiedy analizuje się postęp metod rozpoznawania osób w oparciu o cechy twarzy, można zauważyć, że najlepsze wyniki otrzymywane są przez algorytmu bazujące na sieciach głębokiego uczenia. Z tego powodu, autor zaproponował nową metodę rozpoznawania osób w oparciu o okolicę oka, bazującą na sieci neuronowej ResNet. Skuteczność opracowanego rozwiązania została następnie porównania ze znanymi w literaturze metodami klasycznymi. Aby zapewnić wiarygodność wyników, eksperymenty zostały przeprowadzone na publicznych bazach danych: FOCS [27] oraz UbiPR [28]. Dodatkowo, autor zweryfikował skuteczność zaproponowanego modelu na bazie danych COMPACT oraz porównał otrzymane wyniki do skuteczności algorytmów rozpoznawania osób w oparciu o cechy twarzy. Przeprowadzone badania pozwoliły na ilościowe określenie różnicy w skuteczności obu biometryk oraz określeniu, czy jest możliwe wykorzystanie obszaru okolicy oka w warunkach niekooperacyjnych.

6.2 Opis zaproponowanej metody

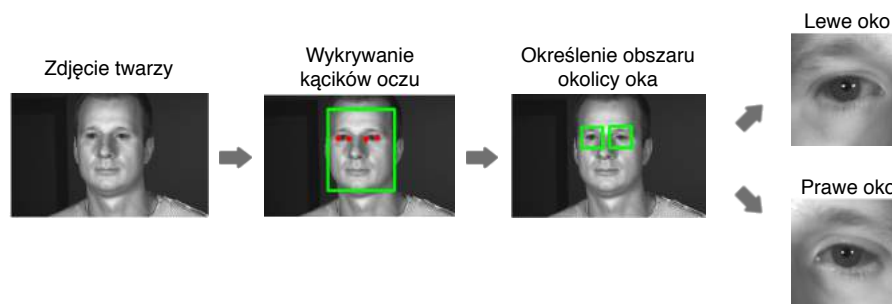
Rys. 6.1 przedstawia schemat blokowy zaproponowanego algorytmu służącego do rozpoznawania osób w oparciu o okolicę oka. Przedstawiona metoda składa się z następujących elementów:

- Dla oka lewego oraz prawego, obszary wokół oka wycinane są ze zdjęcia twarzy na bazie położenia kąćków oczu. Otrzymane fragmenty są następnie normalizowane i skalowane do stałej wielkości. Dodatkowo, dla każdego obrazu wejściowego, obliczany jest kąt odchylenia głowy;
- Ekstrakcja cech jest realizowana przy pomocy sieci neuronowej ResNet. W celu określenia podobieństwa pomiędzy dwoma wektorami cech, wykorzystywany jest dystans Euklidesowy [19];
- Wartości podobieństwa wynikające z porównania oka lewego oraz prawego łączone są ze sobą przy wykorzystaniu technik fuzji klasyfikacyjnej. W efekcie, możliwe jest otrzymanie pojedynczego wskaźnika określającego podobieństwo dwóch próbek.



Rysunek 6.1: Schemat blokowy procedury rozpoznawania osób w oparciu o okolice oka

Segmentacja obszaru okolicy oka ze zdjęcia twarzy realizowana jest przy pomocy następującej procedury. W pierwszym kroku, autor znajduje kąćki oczu, wykorzystując do tego położenie punktów kluczowych twarzy [13]. Następnie, obliczane jest położenie środka oka. Na bazie tego punktu, wybierany jest obszar o rozmiarze $2w \times 2w$, gdzie w to szerokość oka. Taka normalizacja zapewnia, że zawsze wycinany jest taki sam fragment wokół oka, niezależnie od odległości osoby od kamery. W ostatnim etapie, uzyskany obraz okolicy oka skalowany jest do rozmiarów 128 na 128 pikseli. Schemat blokowy opisanej procedury przedstawiony jest na Rys. 6.2. Dodatkowo, dla każdego zdjęcia obliczany jest kąt odchylenia głowy, który służy jako dodatkowa informacja wykorzystywana w momencie fuzji cech z oka lewego i prawego. Estymacja pozy głowy odbywa się z wykorzystaniem algorytmu opisanego w Rozdziale 5.



Rysunek 6.2: Schemat blokowy algorytmu segmentacji obszaru okolicy oka ze zdjęć twarzy

W celu ekstrakcji cech, autor wykorzystał sieć neuronową głębokiego uczenia. Aby otrzymać odpowiednio skuteczny model potrzebne było posiadanie relatywnie dużego zbioru danych uczących. Jednakże, dostępne publiczne bazy danych zawierające zdjęcia okolicy oka są relatywnie małe, szczególnie w porównaniu do baz danych twarzy. Z tego powodu, autor połączył ze sobą wiele zbiorów, aby otrzymać reprezentacyjny zbiór uczący. Wykorzystane zostały następujące bazy danych: PUT [29], FOCS [27], UBIPr [28], CrossEye [30], SCFace [31], CASIA V4 [32], FaceScrub [33], GBU [34]. W efekcie, otrzymano bazę danych, która składała się z 1856 osób oraz 146504 zdjęć, ze średnią liczbą 79 zdjęć na osobę. Otrzymana baza danych jest podsumowana w Tabeli 6.1, gdzie czynniki zniekształceń oznaczone są jako: E (mimika), O (zasłonięcia), L (warunki świetlne), P (poza) oraz D (odległość). Dodatkowo, ponieważ sieci neuronowe nie powinny być uczone z wykorzystaniem podobnie znormalizowanych zdjęć, autor przeprowadził augmentację danych wejściowych, wykorzystując operację losowego kadrowania oraz wypełniania. Implementacja sieci neuronowej została oparta o architekturę ResNet-29 [35], tę samą, która została zastosowana do rozwiązania problemu rozpoznawania twarzy (patrz Sekcja 4.2 oraz Tabela 4.1).

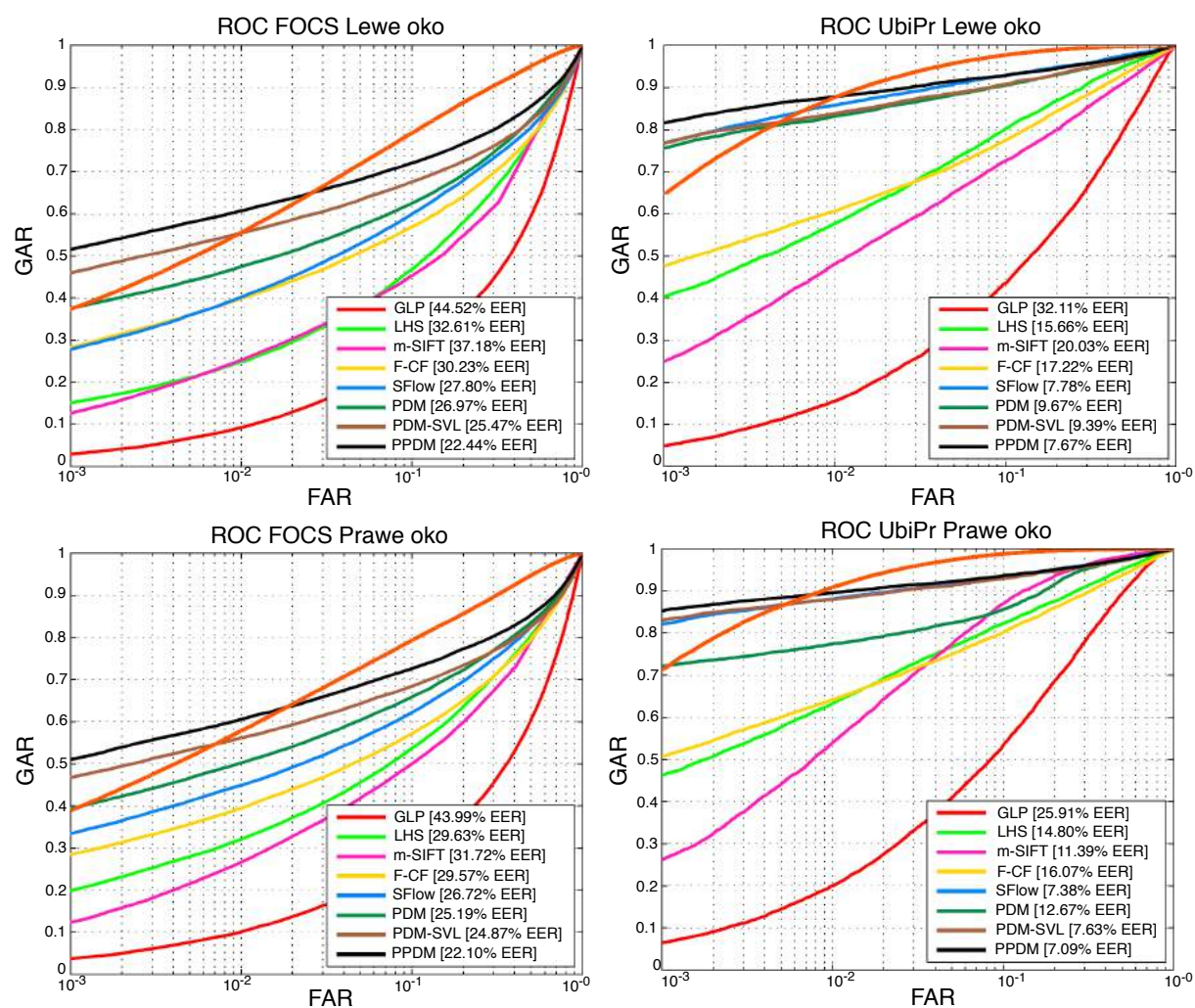
Tabela 6.1: Bazy danych wykorzystane do uczenia sieci neuronowej

Baza danych	Liczba osób	Liczba zdjęć	Źródło światła	Zniekształcenia
PUT	100	9971	Widzialne	E, P
FOCS	136	9581	Bliska podczerwień	D, L, O, P
UBIPr	261	10950	Widzialne	D, L, O, P
CrossEye	120	1920	Widzialne oraz bliska podczerwień	L
SCFace	130	1170	Widzialne	E, L, P
CASIA V4	142	2567	Bliska podczerwień	E
FaceScrub	530	106863	Widzialne	D, E, L, P, O
GBU	437	3482	Widzialne	D, E, L
ŁĄCZNIE	1856	146504	Widzialne oraz bliska podczerwień	E, O, L, P, D

W celu otrzymania pojedynczej wartości określającej stopień podobieństwa pomiędzy dwiema próbkami, autor wykorzystał informację o kącie odchylenia głowy. W warunkach niekooperacyjnych, osoba może zostać sfotografowana z dowolnej perspektywy, przez co jedno oko może być w wybranych przypadkach w ogóle niewidoczne. Potrzebna jest zatem dodatkowa informacja określająca użyteczność zarówno oka lewego, jak i prawego. Taka fuzja może być postrzegana jako problem kombinacyjny lub klasyfikacyjny [36]. Na bazie przeprowadzonych analiz i eksperymentów, autor wykorzystał podejście klasyfikacyjne bazujące na algorytmie Support Vector Machines (SVM).

6.3 Eksperymenty

Skuteczność zaproponowanego algorytmu została porównana z znanymi w literaturze metodami: LBP-GIST (GLBP) [37], LBP-HOG-SIFT (LHS) [38], Modified SIFT (m-SIFT) [39], Full Correlation (F-CF) [26], SIFT-Flow (SFlow) [40] oraz trzema wersjami algorytmu Probabilistic Deformation Models (PDM, PDM-SVL, PPDM) [26]. Porównanie zostało przeprowadzone na publicznych bazach danych: UbiPR oraz FOCS. Dla celu tego eksperymentu, zbiory te zostały wykluczone z procedury uczenia sieci. Skuteczność referencyjnych algorytmów pochodzi z [26]. Orzymane wyniki są podsumowane w Tabelach 6.2 oraz 6.3. Rys. 6.3 przedstawia uzyskane krzywe ROC.



Rysunek 6.3: Skuteczność algorytmów rozpoznawania osób w oparciu o okolice oka na bazach danych UbiPR oraz FOCS. Kolor pomarańczowy przedstawia zaproponowany algorytm ResNet

Tabela 6.2: Skuteczność algorytmów rozpoznawania osób w oparciu o okolice oka na bazie danych UbiPR

	Lewe oko		Prawe oko	
	EER, %	RANK-1, %	EER, %	RANK-1, %
GLBP	32.11	14.99	25.91	20.48
LHS	15.66	71.87	14.80	78.74
m-SIFT	20.03	58.14	11.39	70.15
F-CF	17.22	71.13	16.07	74.18
SFlow	7.78	91.67	7.38	94.67
PDM	9.67	89.77	12.67	93.38
PDM-SVL	9.39	90.39	7.63	93.57
PPDM	7.67	91.97	7.09	93.77
ResNet	4.86	94.32	3.83	95.94

W drugiej części badań, autor zweryfikował możliwość zastosowania opracowanego modelu w warunkach niekooperacyjnych. W tym celu, wykorzystana została baza danych COMPACT. Otrzymane wyniki zostały następnie porównane z wynikami uzyskanymi przy wykorzystaniu algorytmów rozpoznawania osób w oparciu o cechy twarzy (patrz Rozdział 4 oraz Sekcja 4.3). Eksperymenty

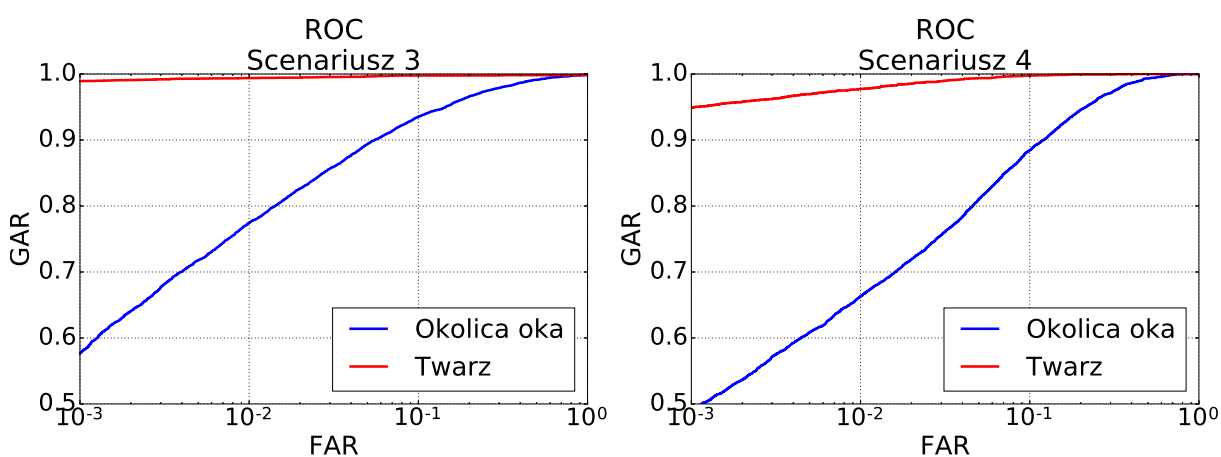
Tabela 6.3: Skuteczność algorytmów rozpoznawania osób w oparciu o okolicę oka na bazie danych FOCS

	Lewe oko		Prawe oko	
	EER, %	RANK-1, %	EER, %	RANK-1, %
GLBP	44.52	24.59	43.99	29.01
LHS	32.61	73.58	29.63	77.71
m-SIFT	37.18	66.41	31.72	67.42
F-CF	30.23	87.02	29.57	86.35
SFlow	27.80	89.88	26.72	88.96
PDM	26.97	94.39	25.19	92.15
PDM-SVL	25.47	94.20	24.87	91.77
PPDM	22.44	94.56	22.10	92.44
ResNet	17.04	94.28	17.62	91.52

przeprowadzono dla scenariusza trzeciego oraz czwartego. Otrzymane wyniki podsumowane są w Tabeli 6.4, podczas gdy Rys. 6.4 przedstawia uzyskane krzywe ROC.

Tabela 6.4: Porównanie skuteczności rozpoznawania osób w oparciu o cechy twarzy oraz okolicę oka na bazie danych COMPACT

Scenariusz 3		
	Twarz	Okolica oka
EER, %	0.67+-0.18	7.9+-0.7
AUC, %	99.80+-0.14	97.5+-0.3
RANK-1, %	99.53+-0.11	82.3+-1.7
Scenariusz 4		
	Twarz	Okolica oka
EER, %	1.65+-0.21	10.8+-0.8
AUC, %	99.87+-0.03	96.3+-0.4
RANK-1, %	97.96+-0.45	72.1+-3.4



Rysunek 6.4: Porównanie skuteczności rozpoznawania osób w oparciu o cechy twarzy oraz okolicę oka na bazie danych COMPACT

6.4 Podsumowanie

Rozpoznawanie osób w oparciu o okolicę oka jest coraz częściej wykorzystywane w biometrycznych systemach uwierzytelniania, które działają w warunkach niekooperacyjnych. Dzieje się tak ze względu na fakt, że zawarta wokół oka informacja jest niezwykle użyteczna w sytuacjach, w których obraz tęczówki jest niedokładny lub twarz jest zasłonięta, na przykład podczas demonstracji lub choroby. Autor zaproponował nowy algorytm ekstrakcji cech, bazujący na sieci głębokiego uczenia opartej na architekturze ResNet. Przygotowany model ResNet okazał się wydajny i odporny na takie zniekształcenia jak: różne oświetlenie, różnorodna mimika czy różne pozy głowy. Skuteczność opracowanego algorytmu została porównana z dziewięcioma powszechnie znanymi w literaturze metodami. Okazuje się, że w systemach, gdzie wygoda jest głównym wymaganiem (na przykład weryfikacja osób przychodzących do pracy), zaproponowane rozwiązanie jest znacząco lepsze od innych algorytmów. Dla systemów o wymaganym wysokim bezpieczeństwie, trzy spośród badanych metod zaprezentowały wyższą skuteczność. **Otrzymane wyniki i przedstawione wnioski dowodzą tezy drugiej.** Dodatkowo, opracowany algorytm został wykorzystany do określenia skuteczności rozpoznawania osób w oparciu o okolicę oka w scenariuszach o zmniejszonej kooperacji. W tym celu wykorzystano bazę danych COMPACT. Uzyskane stopy błędów okazały się większe o rząd wielkości w porównaniu do biometryki twarzy. Oznacza to, że okolica oka nie może być wykorzystana jako pojedyncza cecha w takich warunkach. Mimo to, pozostaje wciąż przydatna jako dodatkowa informacja, na przykład służąca do wstępnej filtracji bazy danych.

Rozdział 7

Podsumowanie

Poniższa praca doktorska poświęcona jest zagadnieniom związanym z identyfikacją biometryczną dla niekooperacyjnych scenariuszy uwierzytelniania. Przeprowadzone badania zostały wykonane w ramach projektu LIDER/027/591/L-4/12/NCBR/2013, zatytułowanego "Biometryczny system uwierzytelniania pozytywnego dla scenariuszy niekooperacyjnych" (COMPACT). Prace badawcze realizowane były pod kierunkiem dra inż. Kamila Grabowskiego. Oprócz autora, zespół badawczy składał się z następujących osób: dr inż. Wojciech Sankowski oraz mgr inż. Damian Kacperski. Prace autora w głównej mierze skupione były wokół części biometrycznej systemu.

W pierwszym etapie prac skonstruowany został prototyp systemu COMPACT. Zgodnie z przyjętym podziałem prac, autor zaprojektował strukturę oprogramowania. Zaproponowane rozwiązanie zostało oparte o architekturę mikroservisów. Pozwoliło to na odseparowanie poszczególnych modułów oraz uzyskanie dużej skalowalności rozwiązania. W kolejnych krokach, autor zaimplementował oprogramowanie odpowiedzialne za akwizycję obrazu z kamer oraz poświęcił się implementacji aplikacji, której celem była synchronizacja wszystkich modułów oraz kontrolowanie przepływu informacji w systemie. W międzyczasie autor aktywnie uczestniczył również w pracach projektowych związanych z systemami wizyjnymi (na przykład wybór kamer czy testy dostępnych urządzeń).

W drugim etapie prac, autor wykorzystał stanowisko badawcze COMPACT do zebrania bazy danych osób przechodzących przez bramkę biometryczną. Zebrany zbiór jest publicznie dostępny i składa się ze 108 osób. Dla każdego podmiotu, zarejestrowane zostały następujące dane: kwestionariusz osobowy, zdjęcia rejestracyjne dla różnych kątów odchylenia głowy oraz zdjęcia weryfikacyjne. W porównaniu do istniejących baz danych, ma on następujące zalety. Po pierwsze, obrazy weryfikacyjne pobrane są w dużej rozdzielczości (3520 x 2200 px) oraz przy użyciu światła bliskiej podczerwieni. Po drugie, zdjęcia weryfikacyjne zostały pobrane w ruchu i z różnych odległości, dzięki czemu zawierają takie zniekształcenia jak rozmycie czy nieprawidłowe ustawienie ostrości kamery. Po trzecie, na zdjęciach dostępnych jest wiele cech biometrycznych, takich jak twarz, okolica oka czy ucho.

Następnie, w celu ukazania użyteczności bazy danych COMPACT oraz dostarczenia referencyjnych rezultatów dla innych badaczy, autor wykonał badania określające skuteczność znanych w literaturze algorytmów rozpoznawania twarzy. W tym celu wykorzystano następujące metody: ThreePatchLBP, HighDimLBP oraz sieć głębokiego uczenia opartą o architekturę ResNet. Przeprowadzone eksperymenty pozwoliły na sformułowanie następujących wniosków. Po pierwsza, baza danych COMPACT okazała

się wymagającym zbiorem, który może być z powodzeniem wykorzystywany do projektowania i optymalizacji algorytmów służących do identyfikacji biometrycznej w scenariuszach o zmniejszonej kooperacji. Po drugie, uzyskane wyniki wskazują, że sieci głębokiego uczenia pozwalają na uzyskanie znacząco lepszej skuteczności identyfikacji niż metody klasyczne. Można uznać, że jednym z ograniczeń rozwoju systemów uwierzytelniania jest brak dostępu do dużych, realistycznych zestawów danych uczących.

W kolejnym etapie prac, autor zaproponował algorytm do filtrowania bazy danych wzorców w celu ograniczenia liczby porównań i zmniejszenia czasu potrzebnego na identyfikację wybranego podmiotu. W systemach działających w warunkach zmniejszonej kooperacji (na przykład system COMPACT), powszechne jest, że dane rejestracyjne często składają się z wielu zdjęć poszczególnej osoby, pobranych pod różnymi kątami. Z tego powodu, aby skutecznie przeprowadzić proces identyfikacji, musi zostać wykonana duża liczba porównań. W celu optymalizacji, autor zaproponował wykorzystanie algorytmów estymacji pozycji głowy do ograniczenia liczby wzorców tylko do tych, dla których poza głową jest podobna do tej, która jest widoczna na badanej próbce weryfikacyjnej. Opracowane rozwiązanie okazało się poprawne i skuteczne. Dla każdego przeprowadzonego eksperymentu, liczba porównań została znacząco zmniejszona, natomiast uzyskane stopy błędów pozostały na tym samym poziomie. Wadą zaprezentowanego algorytmu jest to, że baza danych wzorców musi być skonstruowana w ściśle określony sposób i zawierać dużą liczbę zdjęć dla każdej z zarejestrowanych osób.

W ostatnim etapie prac, autor skupił się na algorytmach identyfikacji biometrycznej opartych o okolicę oka (*ang. periocular recognition*). Ta cecha jest uznawana za szczególnie użyteczną w scenariuszach niekooperacyjnych, przykładowo gdy tęczówka jest niedostępna lub gdy twarz jest częściowo zasłonięta, na przykład podczas demonstracji. Autor zaproponował nowy algorytm ekstrakcji cech, bazujący na sieci głębokiego uczenia opartej na architekturze ResNet. Przygotowany model ResNet okazał się wydajny i odporny na takie zniekształcenia jak: różne oświetlenie, różnorodna mimika czy różne pozycje głowy. Skuteczność opracowanego algorytmu została porównana z dziewięcioma powszechnie znanymi w literaturze metodami, a przeprowadzone eksperymenty pozwoliły autorowi obronić drugą tezę. W systemach, gdzie wygoda jest głównym wymaganiem (na przykład weryfikacja osób przychodzących do pracy), zaproponowane rozwiązanie okazało się znacząco lepsze od innych algorytmów. Dla systemów o wymaganym wysokim bezpieczeństwie, trzy spośród badanych metod zaprezentowały wyższą skuteczność. Dodatkowo, opracowany algorytm został wykorzystany do określenia skuteczności rozpoznawania osób w oparciu o okolicę oka w scenariuszach o zmniejszonej kooperacji. W tym celu wykorzystano bazę danych COMPACT. Uzyskane stopy błędów okazały się większe o rząd wielkości w porównaniu do biometriki twarzy.

Podsumowując, w zaprezentowanej pracy doktorskiej przedstawiony został szereg zagadnień i problemów, które muszą zostać przeanalizowane i rozwiązane podczas projektowania biometrycznego systemu uwierzytelniania działającego w warunkach o zmniejszonej kooperacji. Pomimo tego, że na świecie znane są różne implementacje tego typu systemów, żadne z dostępnych rozwiązań nie jest w stanie działać wystarczająco skutecznie, by mogło być wdrożone w aplikacji o dużej skali. Autor ma nadzieję, że zaprezentowane badania stanowią krok naprzód w stronę konstrukcji tego typu rozwiązań.

Bibliografia

- [1] David Barret. One surveillance camera for every 11 people in Britain, says CCTV survey. *The Telegraph*, 2014.
- [2] J. Klontz and A. Jain. A case study of automated face recognition: The Boston marathon bombings suspects. *IEEE Computer* 46, 2013.
- [3] Tractica. Biometrics market forecasts. <https://www.tractica.com/research/biometrics-market-forecasts>, 2017.
- [4] F. W. Wheeler, R. L. Weiss, and P. H. Tu. Face recognition at a distance system for surveillance applications. *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pp. 1-8, 2010.
- [5] U. Park, H. C. Choi, A. K. Jain, and S. W. Lee. Face tracking and recognition at a distance: A coaxial and concentric ptz camera system. *IEEE Transactions on Information Forensics and Security*, pp. 1665-1677, 2013.
- [6] J. C. Neves, J. C. Moreno, S. Barra, and H. Proença. Acquiring high-resolution face images in outdoor environments: A master-slave calibration algorithm. *2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pp. 1-8, 2015.
- [7] Zaawansowane metody akwizycji próbek biometrycznych dla niekooperacyjnych systemów uwierzytelniania. *PhD Thesis from Technical University of Lodz*, 2018.
- [8] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. *Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition*, 2008.
- [9] M. Guillaumin, J. Verbeek, and C. Schmid. Is that you? metric learning approaches for face identification. *IEEE International Conference on Computer Vision*, 2009.
- [10] D. King. dlib - open source machine learning library. 2017.
- [11] P. Tome, J. Fierrez, R. Vera-Rodriguez, and M. S. Nixon. Soft biometrics and their application in person recognition at a distance. *IEEE Transactions on Information Forensics and Security*, 2014.
- [12] D. Kacperski, P. Krotewicz, M. Włodarczyk, and K. Grabowski. Pose-oriented face images acquisition platform. *2016 MIXDES - 23rd International Conference Mixed Design of Integrated Circuits and Systems*, 2016.

- [13] Vahid Kazemi and Josephine Sullivan. One millisecond face alignment with an ensemble of regression trees. *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014.
- [14] X. Tan and B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Transactions on Image Processing*, 2010.
- [15] L. Wolf, T. Hassner, and Y. Taigman. Descriptor based methods in the wild. *Real-Life Images workshop at the European Conference on Computer Vision (ECCV)*, 2008.
- [16] D. Chen, X. Cao, F. Wen, and J. Sun. Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification. *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 2013.
- [17] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, 2015.
- [18] Anil Kumar Bhattacharyya. On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society*, 1943.
- [19] Michel Marie Deza Elena Deza. Encyclopedia of distances. *Springer*, 2009.
- [20] Gangqiang Zhao, Ling Chen, Jie Song, and Gencai Chen. Large head movement tracking using sift-based registration. *Proceedings of the 15th ACM International Conference on Multimedia*, 2007.
- [21] Simon Baker, Iain Matthews, Jing Xiao, Ralph Gross, Takahiro Ishikawa, and Takeo Kanade. Real-time non-rigid driver head tracking for driver mental state estimation. *Tech. report CMU-RI-TR-04-10*, 2004.
- [22] Y. Fu and T. S. Huang. hmouse: Head tracking driven virtual computer mouse. *Applications of Computer Vision, 2007. WACV '07. IEEE Workshop on*, 2007.
- [23] P. Huber, G. Hu, R. Tena, P. Mortazavian, W. P. Koppen, W. Christmas, M. Ratsch, and J. Kittler. A multiresolution 3d morphable face model and fitting framework. *International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 2016.
- [24] M. Włodarczyk, D. Kacperski, P. Krotewicz, and K. Grabowski. Evaluation of head pose estimation methods for a non-cooperative biometric system. *2016 MIXDES - 23rd International Conference Mixed Design of Integrated Circuits and Systems*, 2016.
- [25] F. Juefei-Xu, Khoa Luu, M. Savvides, T. D. Bui, and C. Y. Suen. Investigating age invariant face recognition based on periocular biometrics. *2011 International Joint Conference on Biometrics (IJCB)*, 2011.
- [26] J. M. Smereka, V. N. Boddeti, and B. V. K. Vijaya Kumar. Probabilistic deformation models for challenging periocular image verification. *IEEE Transactions on Information Forensics and Security*, 2015.

- [27] Raghavender Jillela, Arun A. Ross, Vishnu Naresh Boddeti, B. V. K. Vijaya Kumar, Xiaofei Hu, Robert Plemmons, and Paúl Pauca. Iris segmentation for challenging periocular images. *Handbook of Iris Recognition*, 2016.
- [28] C. N. Padole and H. Proenca. Periocular recognition: Analysis of performance degradation factors. *2012 5th IAPR International Conference on Biometrics (ICB)*, 2012.
- [29] A. Kasinski, A. Florek, and A. Schmidt. The put face database. *Image Processing Communications*, 2008.
- [30] A. Sequeira, L. Chen, P. Wild, J. Ferryman, F. Alonso-Fernandez, K. B. Raja, R. Raghavendra, C. Busch, and J. Bigun. Cross-eyed - cross-spectral iris and periocular recognition database and competition. *2016 International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2016.
- [31] Mislav Grgic, Kresimir Delac, and Sonja Grgic. Siface — surveillance cameras face database. *Multimedia Tools Appl.*, 2011.
- [32] Casia databases. <http://biometrics.idealtest.org/>, 2017.
- [33] H.W. Ng and S. Winkler. A data-driven approach to cleaning large face datasets. *Conference on Image Processing (ICIP)*, 2014.
- [34] P. J. Phillips, J. R. Beveridge, B. A. Draper, G. Givens, A. J. O’Toole, D. S. Bolme, J. Dunlop, Y. M. Lui, H. Sahibzada, and S. Weimer. An introduction to the good, the bad, and the ugly face recognition challenge problem. *Face and Gesture 2011*, 2011.
- [35] D. King. Max-margin object detection. *Computing Research Repository (CoRR)*, 2015.
- [36] Anil Jain, Karthik Nandakumar, and Arun Ross. Score normalization in multimodal biometric systems. *Pattern Recogn.*, 2005.
- [37] S. Bharadwaj, H. S. Bhatt, M. Vatsa, and R. Singh. Periocular biometrics: When iris recognition fails. *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, 2010.
- [38] U. Park, R. R. Jillela, A. Ross, and A. K. Jain. Periocular biometrics in the visible spectrum. *IEEE Transactions on Information Forensics and Security*, 2011.
- [39] A. Ross, R. Jillela, J. M. Smereka, V. N. Boddeti, B. V. K. V. Kumar, R. Barnard, X. Hu, P. Pauca, and R. Plemmons. Matching highly non-ideal ocular images: An information fusion approach. *2012 5th IAPR International Conference on Biometrics (ICB)*, 2012.
- [40] Ce Liu, Jenny Yuen, and Antonio Torralba. Sift flow: Dense correspondence across scenes and its applications. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2011.