

Katedra Biosensorów i Przetwarzania Sygnałów Biomedycznych

Wydział Inżynierii Biomedycznej

POLITECHNIKA ŚLĄSKA



Uczenie Maszynowe

Projekt

Rozpoznawanie twarzy jako narzędzie uwierzytelnienia biometrycznego

Suchanek Kamil, Piecuch Edyta

Kierunek studiów: *Inżynieria Biomedyczna*

Specjalność: *Przetwarzanie i Analiza Informacji Biomedycznej*

Prowadzący projekt: Mgr inż. Konrad Duraj

ZABRZE – 2020

Spis treści

Cel i założenia	3
Lokalne wzorce binarne	4
Histogram zorientowanych gradientów	4
Maszyna wektorów wspierających	5
Realizacja zadania	7
Przygotowanie zdjęć twarzy	7
Ekstrakcja cech twarzy	8
Skuteczność rozwiązania	8
Specyfikacja oprogramowania	12
Funkcja Get_snapshot	13
Funkcja Cut_the_ROI	14
Funkcja detect_face	14
Funkcja create_training_data_set	16
Funkcja test_images	17
Instrukcja obsługi	18
Przygotowanie danych uczących i testowych	18
Przygotowanie klasyfikatora	18
Zastosowanie klasyfikatora	19
Podsumowanie	21

1. Cel i założenia

Celem projektu jest zaprojektowanie procesu uwierzytelnienia biometrycznego na podstawie obrazu twarzy. W ramach projektu zakłada się oprogramowanie przy pomocy dowolnych technologii realizacji następujących zadań:

- ❖ wykstrahowanie danych biometrycznych z obrazów,
Wykrycie twarzy w analizowanych obrazach przy pomocy algorytmu LBP oraz wyznaczenie ROI w celu dalszej analizy.
- ❖ zastosowanie maszyny wektorów wspierających w celu rozróżnienia konkretnych twarzy,
- ❖ przetestowanie rozwiązania i ocena skuteczności identyfikacji.

Zakłada się zastosowanie zaprojektowanego rozwiązania i ocenę jego skuteczności zgodnie z ogólnie przyjętymi metodami oraz ewentualną próbę poprawy jakości uzyskanych wyników.

2. Lokalne wzorce binarne



3. Histogram zorientowanych gradientów

Histogramy zorientowanych gradientów HOG są deskryptorami obrazu, które opisują kształt i pomagają w znalezieniu w obrazie konkretnego obiektu. Działają na zasadzie zliczania występowania gradientów, które występują w tej samej orientacji przestrzennej (pod określonym kątem), w konkretnym, dokładnie określonym fragmencie obrazu. W równomiernie rozmieszczonych fragmentach obrazu są liczone gradienty, w przeciwieństwie do innych deskryptorów. W celu poprawienia skuteczności wykrywania obiektów stosuje się lokalną normalizację kontrastu w tych regionach, które się na siebie nakładają.

Poprzez podzielenie obrazu na małe fragmenty (komórki), uzyskiwane są deskryptory HOG, a następnie dla każdej komórki obliczany jest histogram występowania orientacji krawędzi. Połączenie tych histogramów, obliczonych dla wszystkich komórek jest deskryptorem obrazu HOG.

Obraz, który zawiera poszukiwany obiekt, po zastosowaniu na nim dozwolonej transformacji nadal będzie zawierał cechy deskryptora HOG, pozwalające na wskazanie obecności poszukiwanego obiektu. Deskryptor HOG jest odporny na wszystkie transformacje geometryczne, poza rotacją.

Po fazie obliczenia gradientów dla każdej komórki tworzony jest histogram, który przedstawia rozkład orientacji krawędzi (rozkład tych gradientów) we wszystkich pikselach w komórce. Następnie komórki obrazu są grupowane w większe bloki, w celu przeprowadzenia operacji normalizacji kontrastu.

Na rysunku 1.1 przedstawiono realizację HOG (jedna komórka na blok, 1616 pikseli w komórce, 8 kanałów) dla przykładowego obrazu.



Rys. 1.1. Histogram zorientowanych gradientów: a) obraz wejściowy, b) histogram zorientowanych gradientów HOG dla obrazu wejściowego



4. Maszyna wektorów wspierających

Rozpoznanie twarzy dzięki maszynie wektorów wspierających rodzi wyzwanie dla określenia funkcji reprezentujących poszczególne twarze. Problem leży w mnogości wariantów wizualnej prezentacji zależnych od punktu widzenia, oświetlenia i mimiki.

Metody geometryczne oparte na cechach wykrywa rysy twarzy, takie jak oczy, nos, usta i podbródek. Jako deskryptory twarzy używane są odległości, kąty i powierzchnie. To podejście umożliwia łatwą redukcję danych i jest niewrażliwe na zmiany oświetlenia i punktu widzenia. Niestety są one dość niepewne [1].

Metody szablonowe i neuronowe zwykle działają bezpośrednio na macierzach intensywności pikseli reprezentujących twarze. Wdrożenie i początkowa analiza obrazów jest znacznie uproszczona względem metod geometrycznych. Ekstrakcja cech oddana jest w ręce algorytmowi sieci.





“” W podpunkcie 4 (o SVM) poprosiłbym o rozwinięcia zasady działania SVM oraz opis hiperparametrów, których Państwo użyli,”” ///K.D.

5. Realizacja zadania

Wykorzystano środowisko Matlab do wstępnego przetwarzania obrazów, wyszkolenia modelu i akwizycji nowych zdjęć. Do testowania rozwiązania posłużyły również fragmenty publicznej bazy danych BiID.

5.1. Przygotowanie zdjęć twarzy

Zdjęcia zostały podzielone na dwa równoliczne (z drobnymi odstępstwami) zestawy: testowy i treningowy. Każdy z zestawów został rozłożony do osobnego folderu z numerem: s0, s1 itd. Nazwy tych folderów posłużyły jako nazwy klas do rozróżnienia. Program sporządza listę dostępnych folderów obrazów w zadanych folderach zawierających tylko takie. Następnie program przeskakuje po każdym z folderów tożsamości i sporządza kolejną listę, tym razem z obrazami. Każde zdjęcie z listy podlega detekcji twarzy - w przypadku braku wykrycia owej, lokalizacja obrazu jest wskazywana w oknie Command Window, a program bezawaryjnie kontynuuje pracę. Sprawdzana jest również wymiarowość obrazu i w razie konieczności kolorowy obraz jest konwertowany na skalę szarości.

Proces wczytywanie obrazów i pozyskiwania ROI twarzy jest wspólny dla obu procesów. Dzięki informacjom zwrotnym, z folderów można wykluczyć obrazy bez wykrytych twarzy, oraz zwrócić uwagę na te błędnie sklasyfikowane.

Obrazy umieszczone w folderach o nazwach "sN", gdzie N jest unikatowym numerem danej tożsamości, mogą być dowolnie dodawane do folderu z innymi obrazami treningowymi oraz testowymi. Oprogramowanie przyporządkuje temu folderowi oddzielną klasę do rozpoznawania o ile folder będzie wśród danych treningowych, kwestia obecności folderu testowego jest dowolna.

Zdjęcia użytkowe do bieżącej identyfikacji pozyskiwane są z kamery internetowej przy pomocy dodatkowego modułu pomocniczego Matlab do akwizycji obrazów. Następnie program upewnia się czy akceptujemy zdjęcie do identyfikacji i przeprowadza detekcję i rozpoznawanie twarzy.

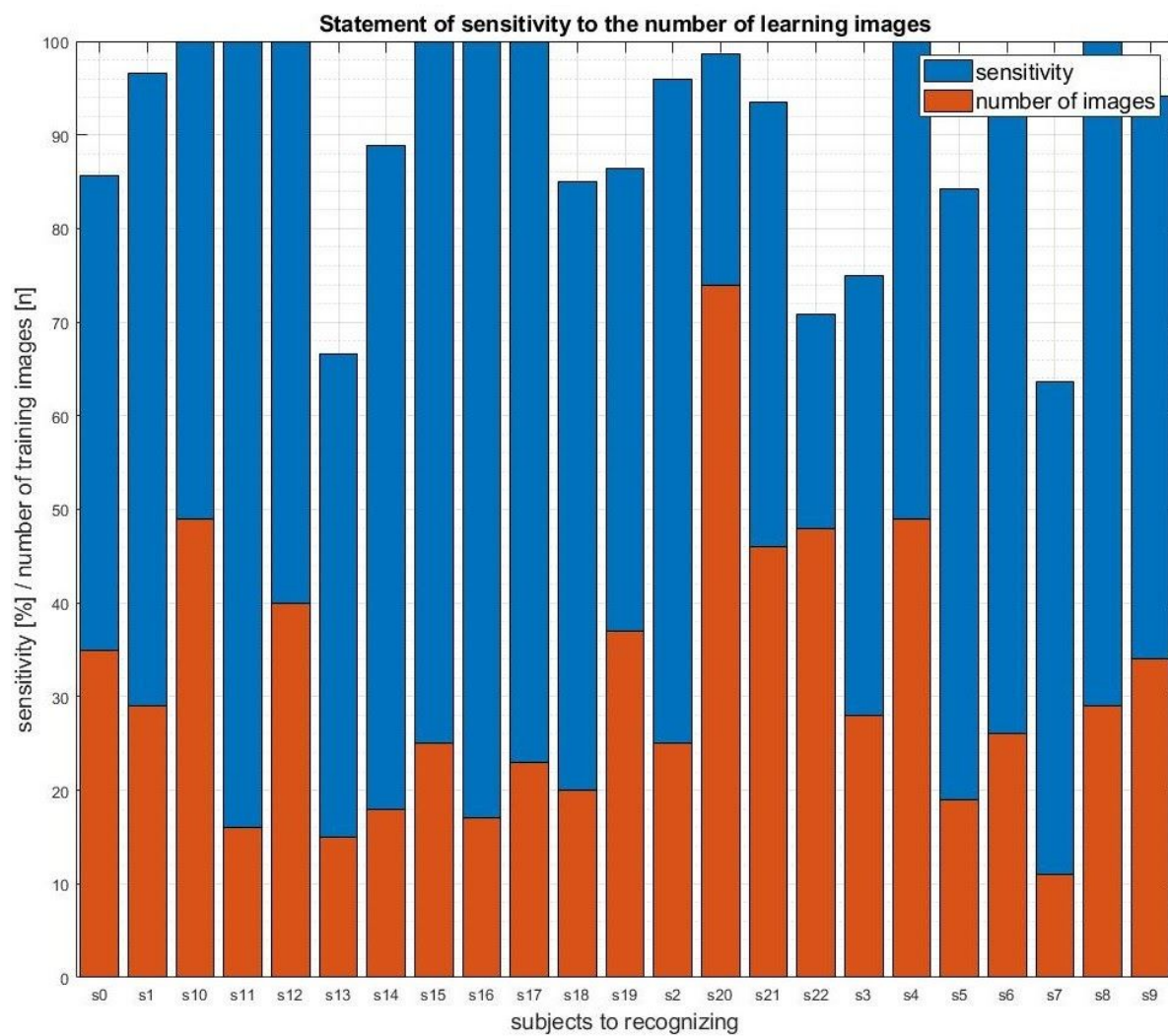
5.2. Ekstrakcja cech twarzy

Kwadratowe ROI obrazów zawierający twarze zostaje wycięty z oryginalnych obrazów i w skali szarości jest skalowany do rozmiaru 255x255 pikseli. W ten sposób przygotowany obraz trafia do algorytmu HOG. Powstały wektor $1 \times M$, przy stałym wymiarze M bierze udział w procesie uczenia modelu SVM oraz rozpoznawania, na już przeszkolonym modelu.

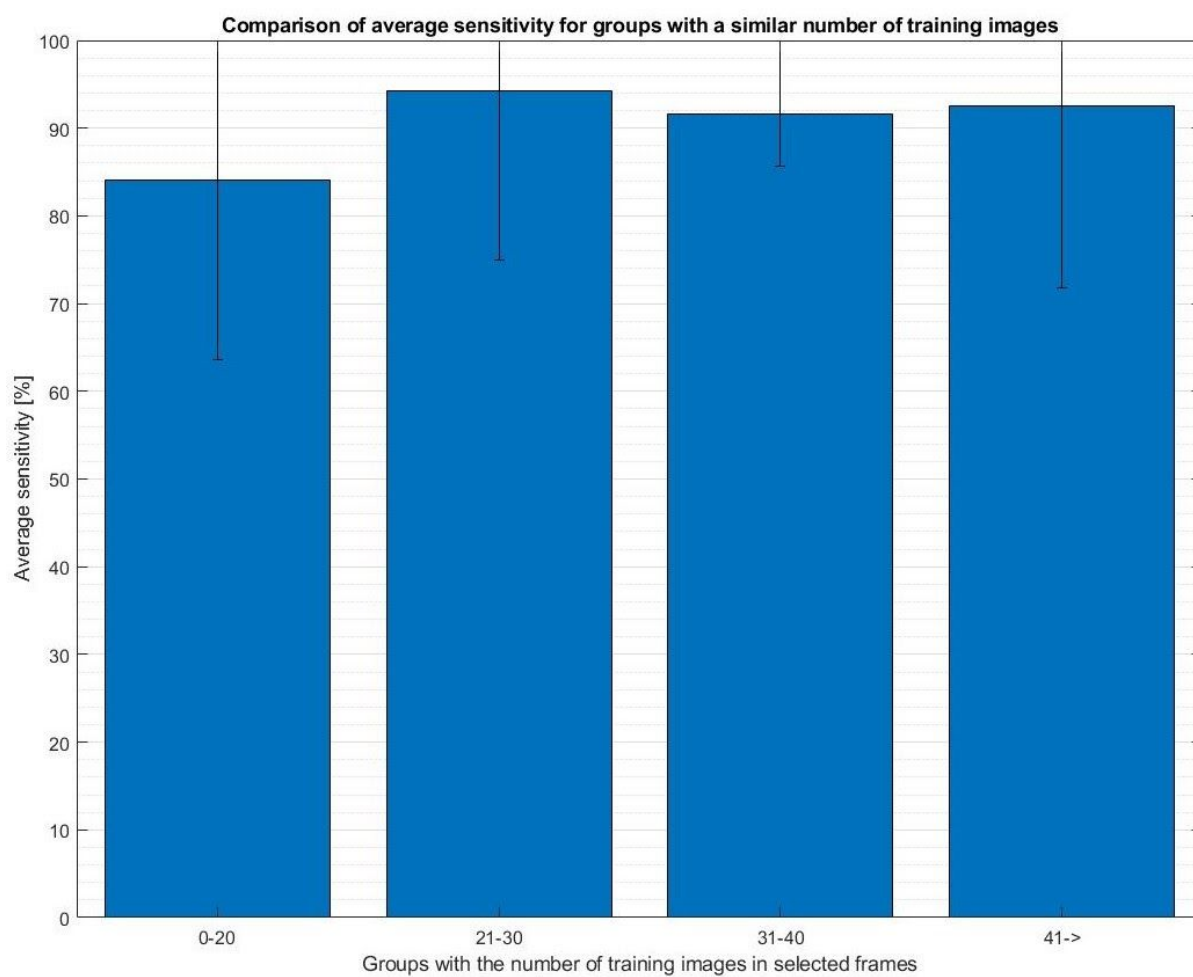
5.3. Skuteczność rozwiązania

Skuteczność metody została sprawdzona przy pomocy podstawowych statystyk i macierzy pomyłek. Wynika z nich (przy uwzględnieniu znajomości zdjęć, ich jakości), że w celu uzyskania zadowalającego poziomu rozpoznawania potrzeba około 30 zdjęć dobrej jakości, tzn. odpowiednio zróżnicowanych i wyraźnych, dających możliwość uchwycenia pewnych zmienności w obrębie danej tożsamości. Większa liczba zdjęć wykonanych bez odpowiedniej dynamiki mogą dawać podobne lub gorsze rezultaty.

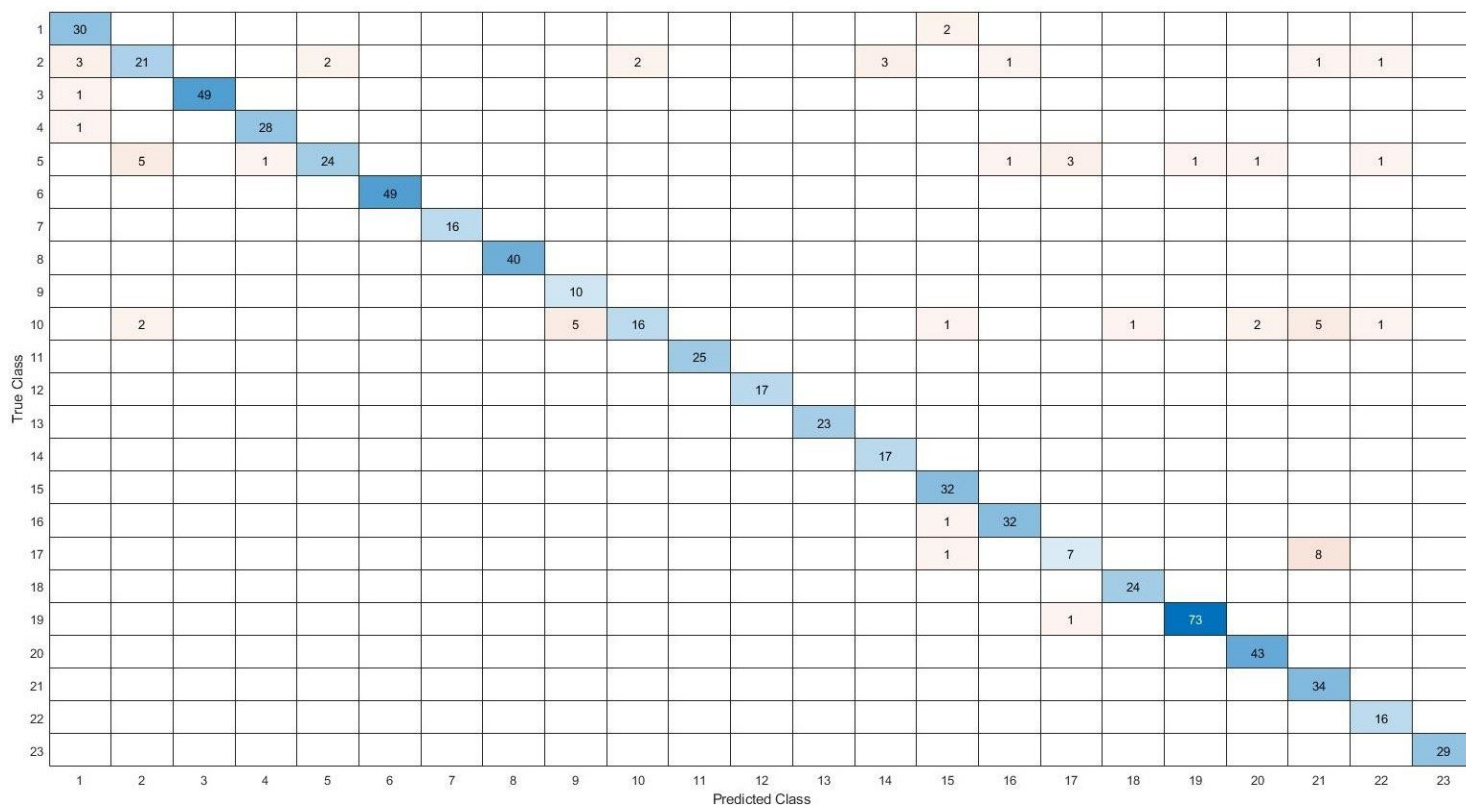
Na wykresie nr 1 przedstawiono zestawienie czułości i ilości obrazów treningowych dla każdej rozpoznawanej osoby. Na wykresie nr 2 przedstawiono uśrednienie tych wartości dla grup o zbliżonej ilości zdjęć wykorzystanych do nauki modelu. Całkowita skuteczność dla wszystkich osób wyniosła 91.8654%. Wykres nr 3 zawiera macierz pomyłek.



Wykres nr 1 - Zestawienie czułości z liczbą obrazów uczących dla każdej klasy.



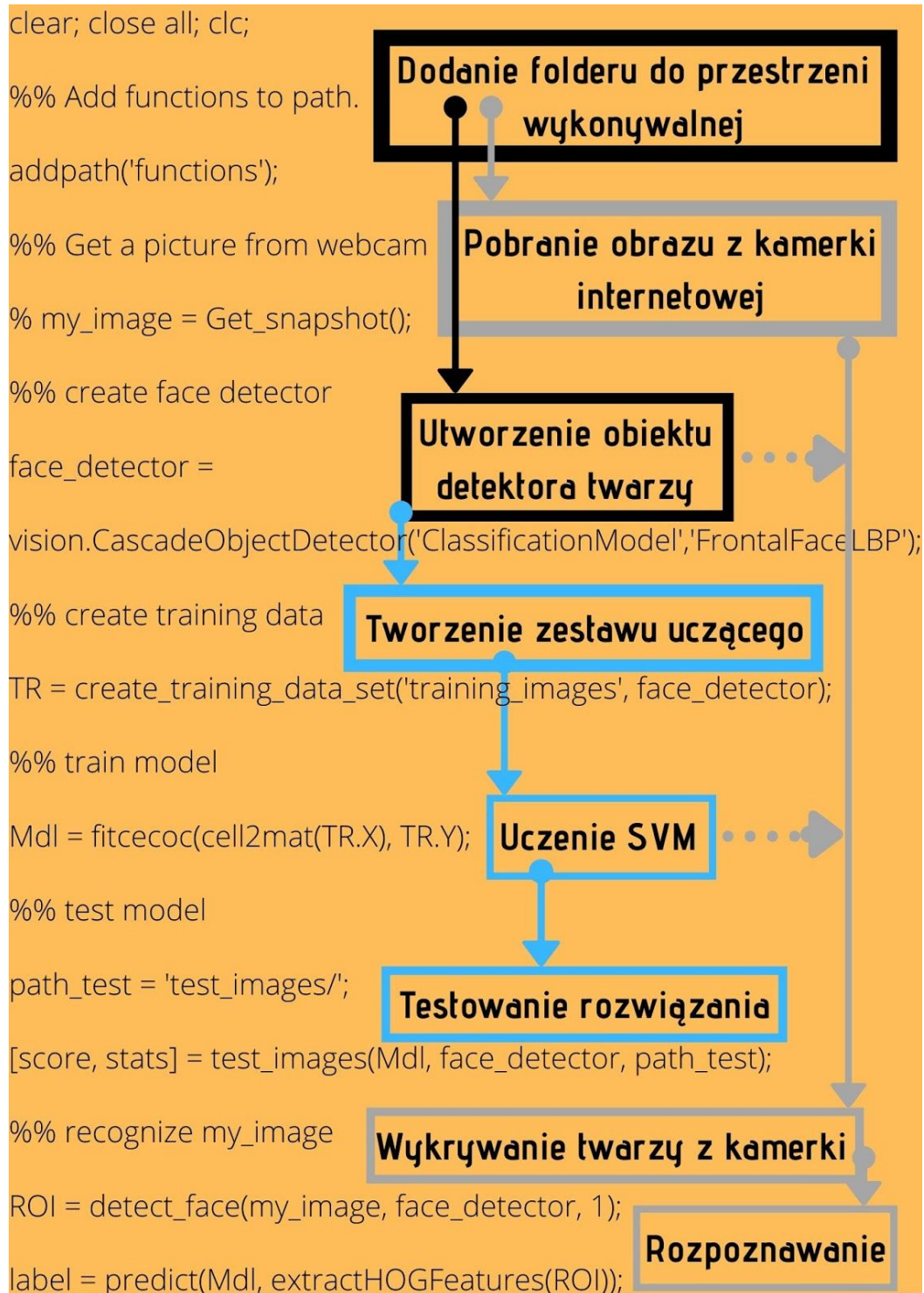
Wykres nr 2 - Zestawienie średnich czułości dla klas o zbliżonej liczbie obrazów uczących.



Wykres nr 3 - Macierz pomyłek.

6. Specyfikacja oprogramowania

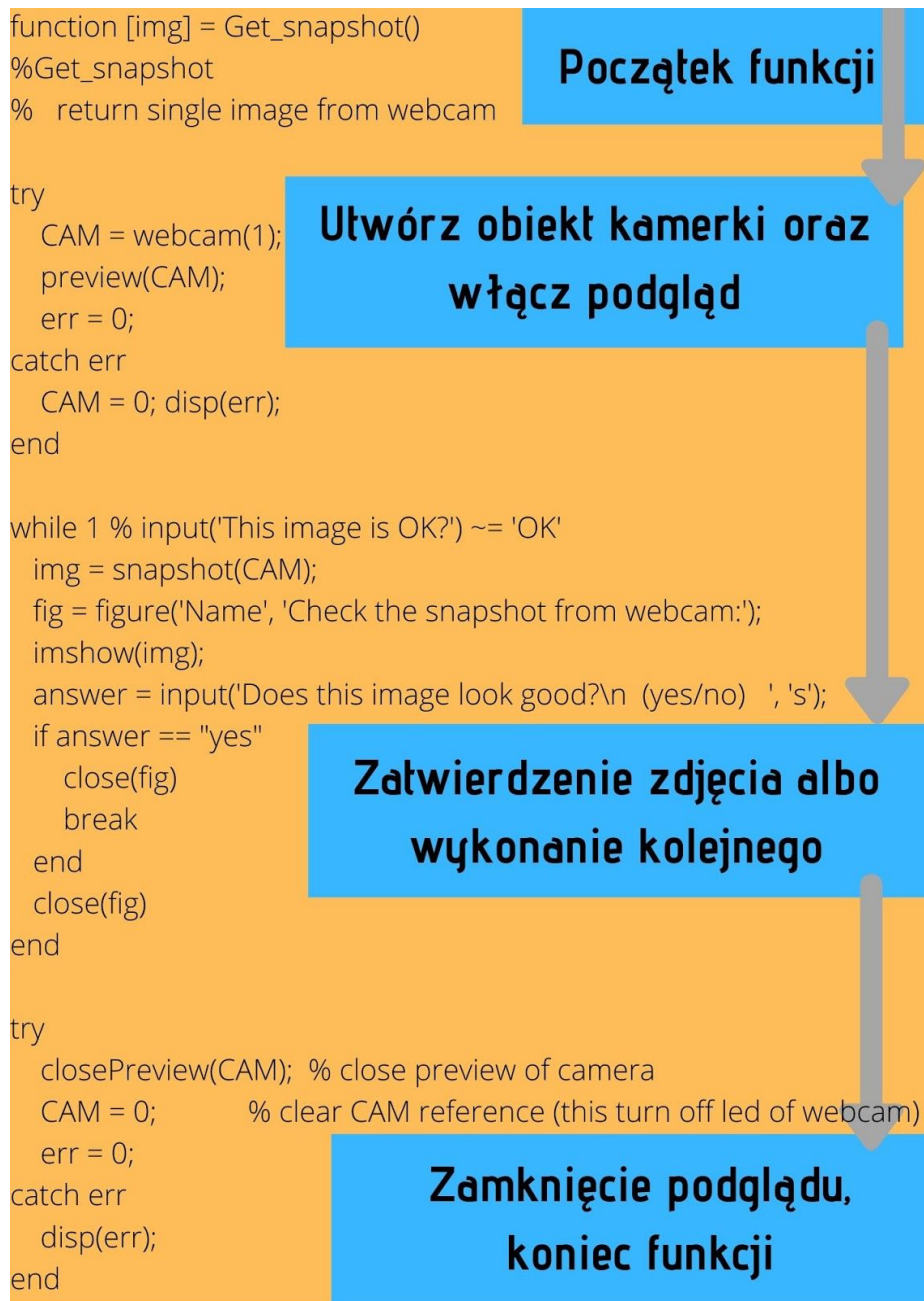
Program główny składa się z następujących poleceń:



Funkcja `addpath()`, dodaje odpowiednią lokalizację z plikami wykonywalnymi, w tym przypadku udostępnia funkcje programu z zewnętrznego folderu. Następnie funkcja

Get_snapshot() wykonuje zdjęcia przy pomocy kamery internetowej i w razie powodzenia zapisuje je do zmiennej my_image. W kolejnym kroku tworzony jest obiekt wstępnie przeszkolonego modelu wykrywania twarzy przy pomocy algorytmu LBP. W celu przeszkolenia modelu tworzony jest zestaw danych uczących przy pomocy funkcji create_training_data_set. Następnie model jest szkolony funkcją fitcecoc() i jest gotowy do testów oraz rozpoznania osoby z obrazu w zmiennej my_image.

6.1. Funkcja Get_snapshot



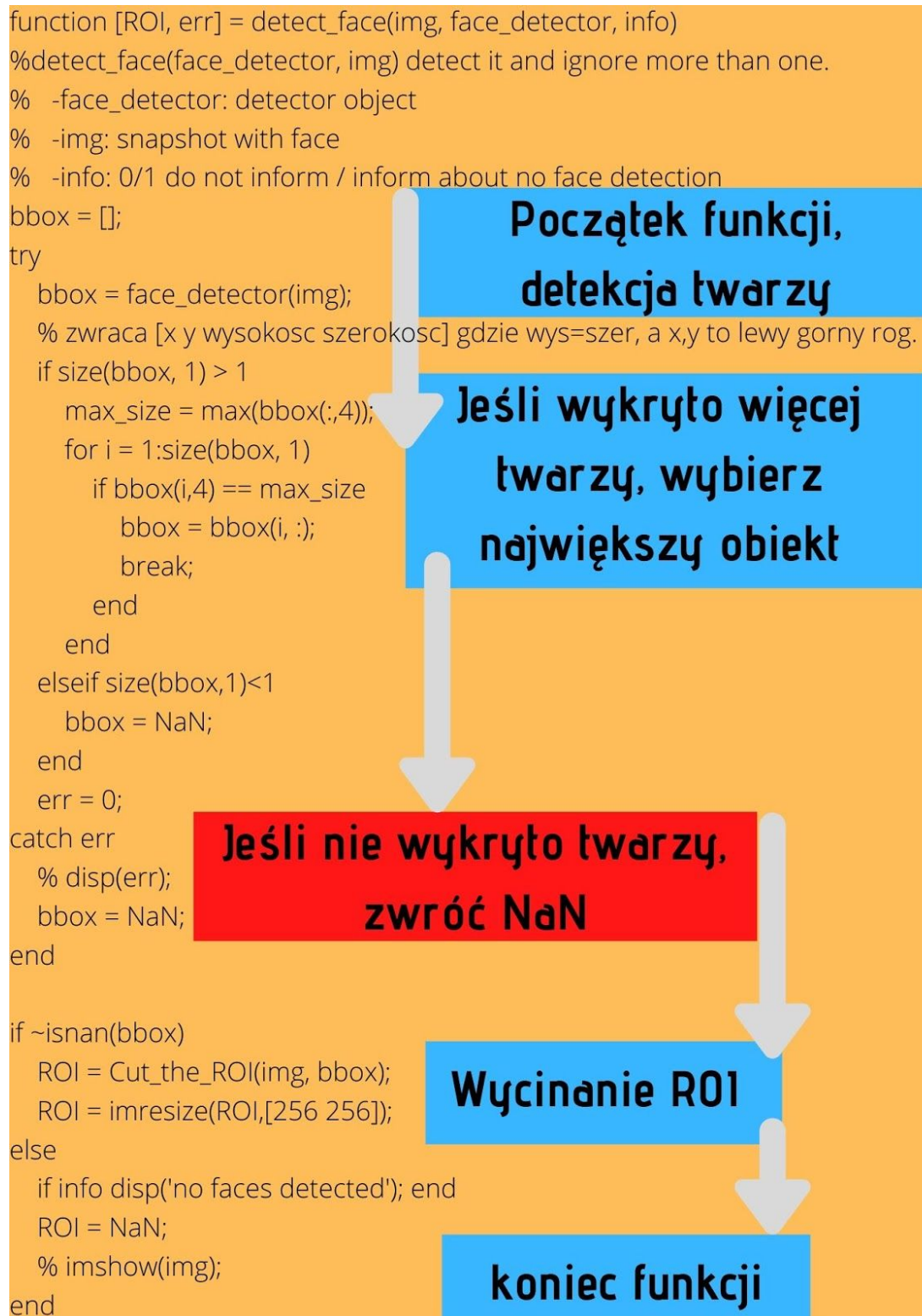
Funkcja ta służy do pobierania obrazu z kamery internetowej. W pierwszej kolejności próbuje uzyskać dostęp do kamery, w przypadku powodzenia wyświetla podgląd i aktualny zrzut do zatwierdzenia. W oknie Command Window należy wpisać yes/no w przypadku odpowiednio zaakceptowania zrzutu w podglądzie albo no w celu wykonania nowego zrzutu. Ostatecznie zatwierdzony zrzut ekranu zapisywany jest do zmiennej i zwracany przez funkcję.

6.2. Funkcja Cut_the_ROI

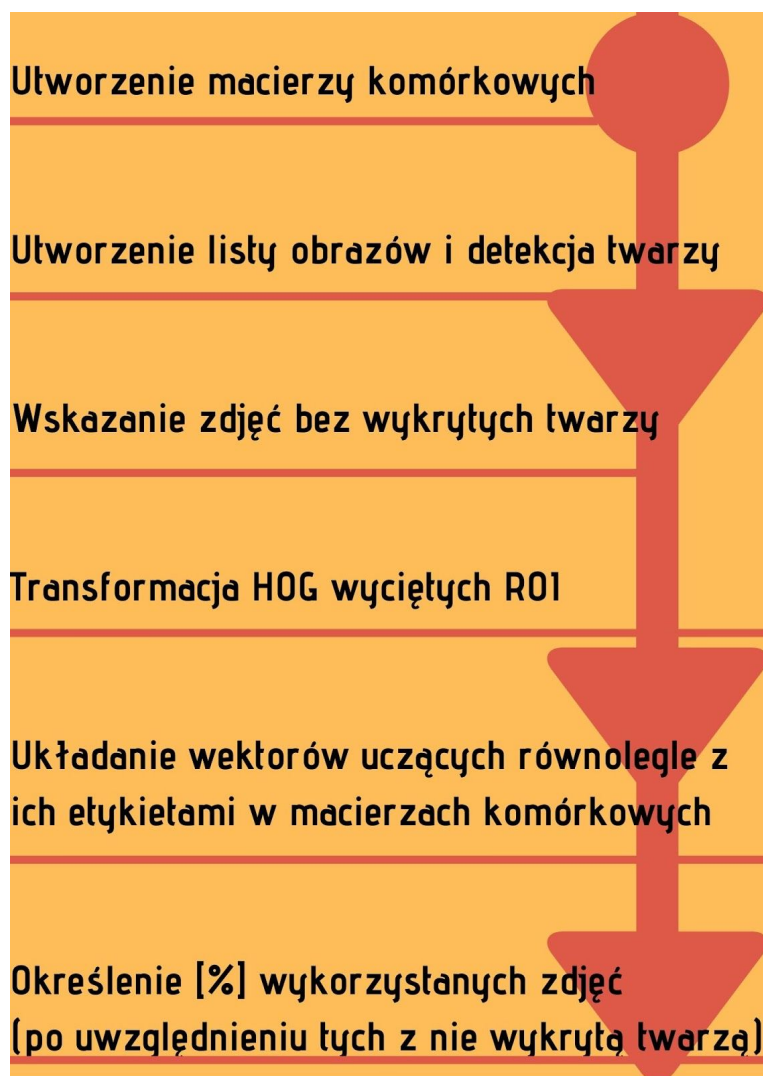
Funkcja odpowiedzialna za wydzielenie ROI z obrazu, na którym wykryjemy twarz. Wykorzystuje ona zmienną bbox z współrzędnymi z metody obiektu vision.CascadeObjectDetector programu MATLAB oraz obraz z stwierdzoną obecnością twarzy. Argumentem wyjściowym jest odpowiednio przycięty obraz.

6.3. Funkcja detect_face

Funkcja kolejno wykrywa twarz przy pomocy obiektu vision i sprawdza, czy nie zostało wykrytych więcej niż jedna twarz. Dla pewnych obrazów to następuje, więc zdecydowano aby wziąć pod uwagę największą zwróconą ramkę, a resztę zignorować. Jest to słuszne, ponieważ twarz w zdjęciach, gdzie ich poszukujemy jest zazwyczaj największym tego rodzaju obiektem, na którym nam zależy, często mniejsze (nawet kilkukrotnie), niewyraźne elementy zostają natomiast błędnie uznane za twarz. W przypadku braku wykrytej twarzy, zamiast przeskalowanego ROI zostanie zwrócony NaN. Przy pomocy argumentu info możemy zdecydować, czy chcemy być informowaniu i takim przypadku.



6.4. Funkcja create_training_data_set



Funkcja służy do generowania struktury danych uczących i raportowania o ilości wykorzystanych zdjęć treningowych, dzięki czemu można w prosty sposób wykluczyć niepotrzebne obrazy.

6.5. Funkcja test_images

Fragment komentarzy funkcji opisujący jej argumenty wejściowe i wyjściowe:

```
function [scores, stats] = test_images(Mdl, face_detector, path_of_test_images)  
%test_images test model  
%input:  
% Mdl: trained model  
% face_detector: trained object model of vision.CascadeClassifierObject  
% path_of_test_images: path with test images  
%return:  
% scores: struct of scores and success  
% .success = {0 1 0} % 0 - failure, 1 - correct recognition  
% .score = {score score score} % score from predict matlab function  
% stats: struct of statistics  
% .TP = {TP for s0, ... for s1, ...}  
% .FN = {numel(images) - TP - cort for s0, ... for s1, ...}  
% .label = {"s0", "s1", ...}  
% .sen = {sensitivity % for s0, ...}  
% .miss = {miss rate % for s0, ...}
```

Funkcja odczytuje wszystkie obrazy testowe, przeprowadza klasyfikację i na bieżąco oblicza statystyki dla każdej z testowanych klas.

7. Instrukcja obsługi

Przygotowane oprogramowanie umożliwia podstawienie własnych danych i przeszkolenie domyślnie skonfigurowanego klasyfikatora SVM.

7.1. Przygotowanie danych uczących i testowych

Dane do nauki i testów należy umieścić w folderach odpowiednio: "training_images" oraz "test_images". W obu folderach, obrazy odpowiadające określonym tożsamościom powinny znajdować się w odpowiadających sobie folderach s0, s1 itd...

7.2. Przygotowanie klasyfikatora

W celu zastosowania skryptu *main.m*, należy uruchomić jego ostatnią wersję, dostępną w czerwcu 2020 roku na repozytorium w serwisie github:

github.com/KamilSuchanek95/Biometria-twarzy-UM-m,

zakładając, że w odpowiednich folderach mamy już przygotowane obrazy do nauki i testów.

W wyniku uruchomienia programu, w oknie Command Window powinny pojawić się następujące komunikaty:

the percentage of images used from s0 is:100.00% (or 35 out of 35)

the percentage of images used from s1 is:100.00% (or 29 out of 29)

...

Oraz podczas procesu testowania:

test image folder: s0

incorrect classification as s3 on image: t0 (14).jpg score: -0.015977

incorrect classification as s3 on image: t0 (16).jpg score: -0.017234

incorrect classification as s4 on image: t0 (17).jpg score: -0.023112

incorrect classification as s3 on image: t0 (20).jpg score: -0.024862

incorrect classification as s1 on image: t0 (7).jpg score: -0.024719

s0 TP = 30 FN = 5 sensitivity = 0.85714 miss rate = 0.14286

=====

...

W wyniku pracy programu powstał przeszkolony obiekt Mdl oraz face_detector widoczne w oknie Workspace. Aby uniknąć ponownych obliczeń można zapisać owe zmienne do pliku poprzez zaznaczenie ich, kliknięcie prawym przyciskiem myszy oraz wybranie poprawnej opcji. Tak zapisane zmienne będzie można załadować przy pomocy kodu (funkcja load) oraz przeciągając owe do okna Workspace albo Command Window w dowolnej chwili.

7.3. Zastosowanie klasyfikatora

W ten sposób przygotowany klasyfikator i detektor można zastosować poprzez załadowanie przygotowanych wcześniej zmiennych oraz uruchomienie kolejno sekcji kodu o nazwach (tylko sekcji, nie całego kodu, chyba że niepotrzebne sekcje zostały własnoręcznie zakomentowane):

- %% Add functions to path
- %% Get a picture from webcam

Po uruchomieniu tej sekcji, pojawi się podgląd na żywo z kamery internetowej oraz dodatkowe okno przedstawiające aktualnie uchwycone zdjęcie, program zapyta w oknie Command Window, czy takie ujęcie nam odpowiada:

Does this image look good?

(yes/no) yes

W tym przypadku wprowadzono słowo "yes" i zatwierdzono przyciskiem ENTER. Można wpisać również "no", lub cokolwiek poza "yes" i po zatwierdzeniu pojawi się nowy podgląd uchwyconego ujęcia (samo zatwierdzenie bez wpisywania czegokolwiek również skutkuje wykonaniem kolejnego zrzutu z kamery).

- %% recognize my_image

Po uruchomieniu tej sekcji kodu program przystąpi do wykrywania twarzy, w przypadku braku wykrytej twarzy zostaniemy o tym poinformowani w oknie

Command Window. W tym przypadku zaobserwujemy również błąd programu ze względu na wykonanie predykcji w następnej linijce, W przypadku wątpliwości czy nasze zdjęcie będzie miało dość dobrą jakość można więc utworzyć nową sekcję dla kolejnej linii kodu. Aby uzyskać informację o wykrytej osobie należy wpisać w okno Command Window "label":

```
>> label  
  
label =  
  
1×1 cell array  
  
{'s6'}
```

W tym przypadku wskazało na tożsamość s6 (Kamil Suchanek, poprawnie). Na własne potrzeby można wprowadzić mapowanie wartości i osób, które umieściliśmy w poszczególnych folderach.

8. Podsumowanie

W ramach realizacji projektu powstał prototyp rozwiązania elementu systemu biometrii twarzy opartego o transformację HOG i SVM. Aktualne repozytorium projektu zawiera skrypty zgodne z funkcjonalnym opisem w tym dokumencie, zawierające drobne zmiany zmiennych i szyku linii kodu w celu poprawy czytelności.

[...]

[...] ""Jak ma się wynik pracy do założeń??""/kamil

[...]

Bibliografia

1. Guodong Guo, Stan Z. Li., Kapluk Chan - Face Recognition by Support Vector Machines - School of Electrical and Electronic Engineering Nanyang Technological University, Singapore.
2. .
3. .