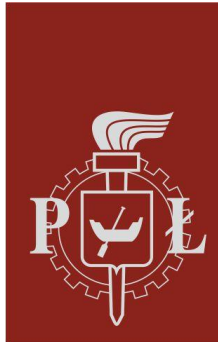


07.10.2018, Łódź



System rozpoznawania osób na podstawie zdjęcia ucha

Dokumentacja



1. Wprowadzenie	3
1.1. Detekcja ucha	4
1.2. Uszy własne	4
1.3. Konwolucyjne sieci neuronowe	6
2. Aplikacja	7
1.1. Uszy własne	7
1.2. Konwolucyjne sieci neuronowe	7



1. Wprowadzenie

Twarz nie jest jedynym elementem ciała, które może posłużyć do identyfikacji osób, mimo że jest jednym z najbardziej oczywistych wyborów, ponieważ używamy jej w codziennym życiu. Niemniej jednak, nie jest to wybór idealny. Znajomość twarzy danej osoby w wieku dziecięcym nie gwarantuje rozpoznania twarzy tej samej osoby w wieku dorosłym, a tym bardziej starczym. Dodatkowo zarost, biżuteria, makijaż i wiele innych elementów wpływają na to czy i z jaką pewnością jesteśmy w stanie określić czyjąś tożsamość. Zależnie od dnia twarz może być całkowicie inna, jest ona dość zmienna, dlatego też zapewne nie jest ona najlepszym sposobem na identyfikację czyjejś tożsamości.

Jak ucho sprawdziłoby się w roli takiego identyfikatora? W porównaniu do twarzy ucho zmienia się w bardzo niewielkim stopniu przez całe życie, zachowując swój kształt. Dodatkowo nie zmienia się ono zależnie od dnia czy nastroju. W porównaniu do twarzy niezależnie od perspektywy wykonania zdjęcia ucho jest stosunkowo podobne – co może nie mieć miejsca w przypadku zdjęcia twarzy lewego i prawego profilu. W pierwszej chwili ucho nie zostałoby wybrane jako identyfikator, gdyż ludzie sami nie są w stanie rozróżnić innych na podstawie wyłącznie ucha. Nie oznacza to, że proces ten jest cięższy, po prostu naturalnym dla człowieka jest skupienie się na twarzy i poprzez przyglądanie się wszelakim twarzom przez kilka, kilkanaście, kilkadziesiąt lat – uczymy się je rozpoznawać i rozróżniać coraz lepiej. Nie jest więc zaskoczeniem, że taki proces identyfikacji mógłby zostać wykonany również dla ucha.

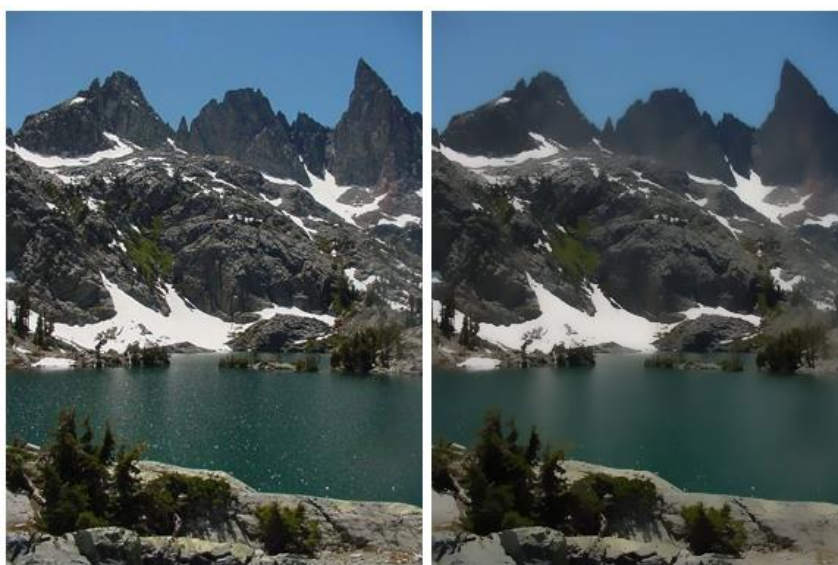
1.1. Detekcja ucha

Detekcja ucha jest analogiczna do detekcji twarzy gdyż tutaj również metoda Viola-Jonesa została użyta. Tak jak detektor mógł być wytrenowany do wykrywania twarzy na zdjęciach – to samo może zostać wykonane dla uszu. Jest jednak pewna znacząca różnica. W przypadku uszu, ucho lewe różni się od ucha prawego i detekcja uszu tylko i wyłącznie za pomocą jednego wytrenowanego modelu nie byłaby możliwa. Z tego też względu stworzone zostały dwa modele do wykrywania poszczególnych uszu.

1.2. Uszy własne

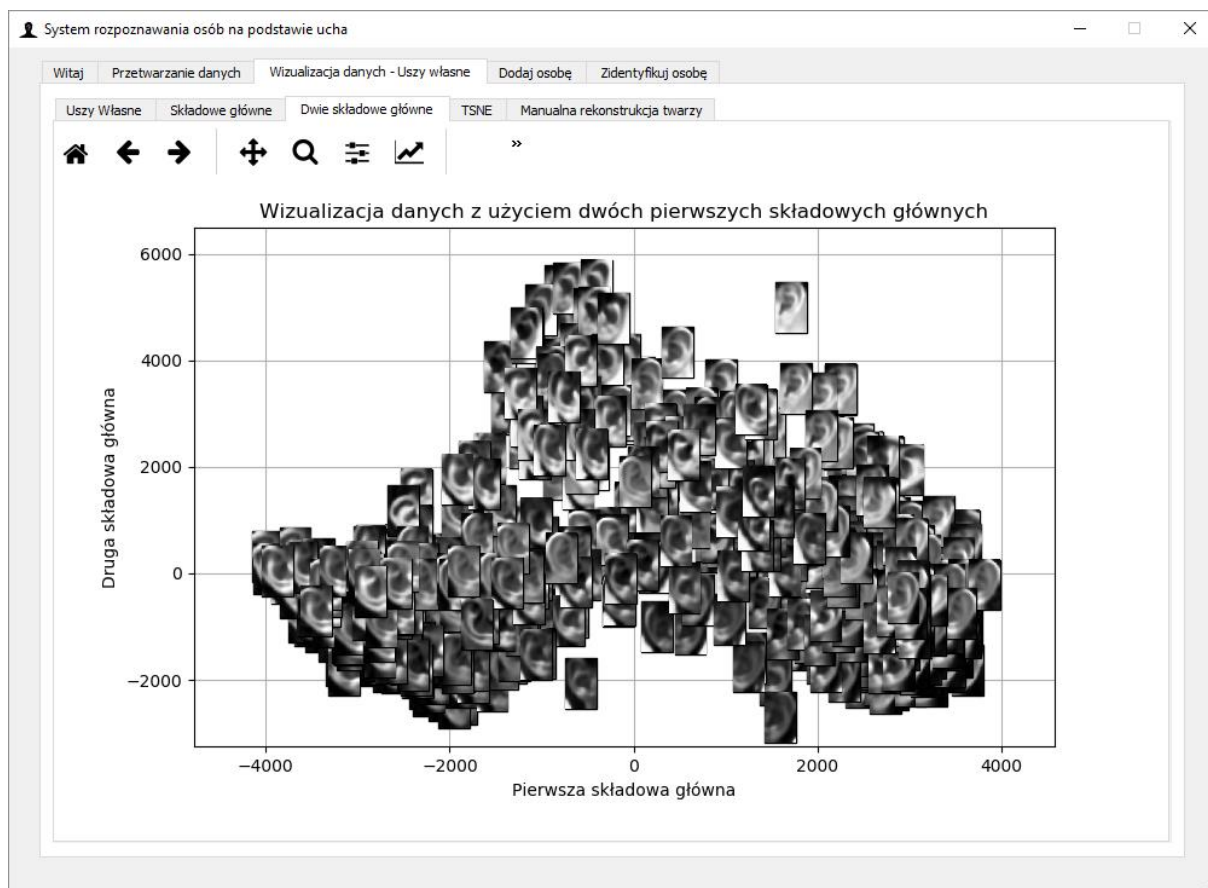
Opisana wcześniej metoda twarzy własnych nie jest oczywiście zarezerwowana wyłącznie dla twarzy. Nie używa ona żadnych konkretnych cech twarzy, nie polega ona również na odległościach geometrycznych czy innych podobnych cechach. Dlatego też metoda ta może zostać wykorzystana do identyfikacji osób na podstawie zdjęcia ucha. Metoda działania nie różni się w żadnym stopniu od tej przedstawionej w twarzach własnych. Zbiorem w bazie danych mogą być zdjęcia ucha lewego jak i prawego jednocześnie.

Jedyną różnicą pomiędzy wcześniej przedstawioną implementacją twarzy własnych jest fakt użycia filtra bilateral na każdym ze zdjęć. Jest to nieliniowy filtr zachowujący krawędzie i redukujący szum w zdjęciach – w efekcie „wygładzający” zdjęcie. Zastępuje on wartość piksela średnią ważoną wartości z pikseli w jego otoczeniu – gdzie wielkość otoczenia jest parametrem definiowanym samodzielnie. Zastosowanie takiego filtra umożliwia wyeksponowanie krawędzi ucha i jednocześnie pominięcie nieistotnych elementów w zdjęciu. Efekt działania takiego filtra jest zaprezentowany na rysunku 1.



Rys 1. Działanie filtra bilateral.

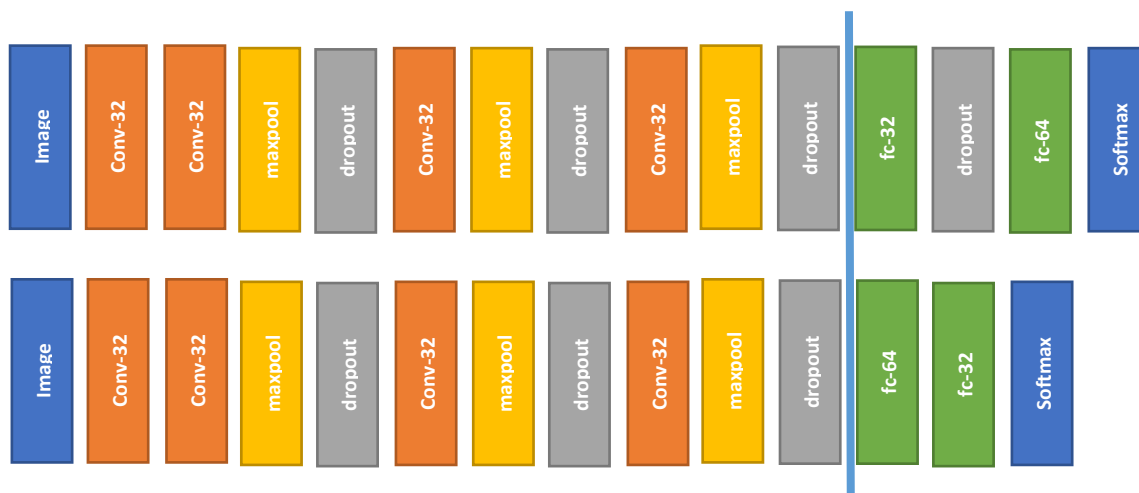
Sama aplikacja zawiera te same metody wizualizujące różne aspekty związane z uszami własnymi. To na co warto zwrócić uwagę to analiza głównych składowych gdzie pierwsza składowa główna odpowiada w jakimś stopniu za orientację ucha zaś drugi za jasność zdjęcia, co można zaobserwować na rysunku 2.



Rys 2. Pierwsze dwie składowe główne.

1.3. Konwolucyjne sieci neuronowe

W porównaniu do wcześniejszego zastosowania sieci – tutaj transfer wiedzy nie był możliwy. Jako że nie ma ogólnodostępnych modeli konwolucyjnych sieci neuronowych stworzonych do identyfikacji uszu model ten musiał zostać stworzony i przetrenowany samodzielnie. Struktura sieci użyta do treningu sieci jest przedstawiona na rysunku 3.



Rys 3. Struktura użytej sieci neuronowej do początkowego treningu (u góry) i do dotrenowania i identyfikacji (u dołu).

Poza opisanym przy twarzach warstwami pojawia się jeszcze warstwa nazywana dropout. Ma ona za zadanie wyłączenie z procesu uczenia z pewnym prawdopodobieństwem poszczególnych neuronów, zapewniając tym samym mniejszą wrażliwość modelu na overfitting, czyli tzw. przeuczenie. Przeuczenie może nastąpić w przypadku małej ilości danych, które będą użyte do trenowania sieci z wielką w porównaniu do danych ilością parametrów. W takim przypadku sieć uczy się odpowiedzi „na pamięć” co nie jest oczekiwanym efektem, ponieważ traci ona umiejętność generalizacji.

W przypadku tego modelu ilość zdjęć wynosiła 721 dla 24 klas (modeli uszu). Po przetrenowaniu górnego modelu z rysunku 3 na takiej ilości danych, został on zapisany. W przypadku identyfikacji ucha model jest wczytywany a jego ostatnie cztery warstwy są zastępowane trzema warstwami, z których dwie są w pełni połączone a ostatnią jest softmax. Następnie sieć jest dotrenowywana na nowo dodanych w aplikacji uszach.

Tak jak w przypadku uszu własnych bipolar filter również został zastosowany do każdego ze zdjęć.



2. Aplikacja

Funkcjonalność aplikacji składa się z dwóch rozdzielnych części, którymi są zaimplementowane różne metody tak samo jak w przypadku rozpoznawania twarzy. Jediną różnicą jest to iż dwie klasy odpowiadają jednej osobie (ucho lewe i prawe).

Sama procedura używania aplikacji również jest identyczna jak w przypadku rozpoznawania twarzy. Z racji trudności otrzymania dużych zbiorów danych uszu dla sieci neuronowej zdjęcia w skali szarości zostały wykorzystane, dlatego też identyfikacja z użyciem konwolucyjnych sieci neuronowych nie przedstawia zdjęć w kolorze. Dodatkowo, początkowa baza danych użyta dla Uszu Własnych jest większa i z tego powodu załadowanie jej jest bardziej czasochłonne. Jest to również używalne przy dodawaniu uszu w Uszach Własnych, gdyż dane w bazie muszą zostać przeliczone ponownie.

Z racji dwóch modeli do wykrywania ucha (lewego i prawego) używanie ich jednocześnie spowodowałoby bardzo wolną pracę aplikacji w trakcie dodawania nowego ucha. Z tego też względu pierwsze sekundy dodawania osoby służą do kalibracji gdzie liczona jest ilość wystąpienia ucha prawego i lewego – bądź w przypadku wystąpienia obu sprawdzana jest ich wielkość i większe z nich jest zaliczane. Na koniec kalibracji jeśli różnica w ilości wystąpienia uszu lewego i prawego jest większa niż 5 ucho z większą ilością wystąpień będzie dodawane i drugi z modeli wykrywania jest wyłączany. Przykładowo 15 razy mogło pojawić się prawe ucho w kadrze, zaś lewe tylko 3 razy – w tym przypadku jedynie prawe ucho będzie poszukiwane po kalibracji.

1.1. Uszy własne

Tak jak wcześniej zostało to wspomniane – procedura i działanie są takie same jak w przypadku rozpoznawania twarzy.

1.2. Konwolucyjne sieci neuronowe

Tak jak w przypadku sieci neuronowych dla twarzy – sieć przestanie się uczyć gdy przez 5 epok precyzja nie urośnie o więcej niż jeden procent. Dodatkowo sieć jednocześnie uczy się uszu lewych i prawych bez rozdzielania zbiorów. Procedura dodawania uszu jest identyczna jak dodawania twarzy, każde ucho to kolejna dodana klasa.