Identyfikacja twarzy z wykorzystaniem

Sztucznych Sieci Neuronowych oraz PCA

Michał Pieróg

pierogmichal@gmail.com

Jakub Jaśkowiec

qbajas@gmail.com

Abstrakt— Identyfikacja twarzy jest zadaniem polegającym na weryfikacji tożsamości danej osoby z wykorzystaniem zdjęć jej twarzy. Proces identyfikacji może być użyty w różnych biometrycznych systemach bezpieczeństwa. W niniejszym artykule przedstawiono proces identyfikacji twarzy wykorzystujący sztuczną sieć neuronową typu Feed Forwad oraz statystyczną metodę PCA (Principal Component Analysis). Zdjęcia wejściowe zostały poddane wstępnej filtracji i prostej segmentacji, następnie zredukowano rozmiar danych przy użyciu PCA. Kolejnym krokiem była klasyfikacja przy użyciu sieci neuronowej. Przeprowadzono testy z różnymi parametrami PCA oraz sieci, uzyskując stosunkowo dużą skuteczność identyfikacji dla zbioru testowego, składającego się z 2200 zdjęć w skali szarości.

Słowa kluczowe- identyfikacja twarzy; Sztuczne Sieci Neuronowe; PCA; Back Propagation; Resillent Propagation

# Wstęp

Burzliwy rozwój techniki umożliwił znaczący postęp w dziedzinie bezpieczeństwa. Dlatego coraz większą uwagę naukowców przyciągają biometryczne systemy identyfikacji, szczególnie te oparte na automatycznym rozpoznawaniu twarzy.  Identyfikacja osób na podstawie twarzy jest jedną z ważniejszych gałęzi w przetwarzaniu obrazów. Prawdziwym wyzwaniem jest stworzenie automatycznego systemu, który mógłby dorównać ludzkiej zdolności rozpoznawania twarzy.

Nie poznano dokładnie mechanizmu pozwalającego ludziom w szybki sposób rozpoznawać twarze. Chociaż ludzie radzą sobie bardzo dobrze z rozpoznawaniem znanych twarzy, nie są oni w stanie poradzić sobie z dużymi ilościami twarzy nieznanych. Komputery z ogromną pamięcią i mocą obliczeniową mogłyby przezwyciężyć te ograniczenia.

Głównym problemem związanym z identyfikacją twarzy         jest          sposób jej reprezentacji.  
Istnieją 2 główne metody identyfikacji twarzy. Algorytmy feature-based opierają się na ekstrakcji wektorów cech charakterystycznych części twarzy takich jak oczy, nos i usta wykorzystując zaawansowane modele matematyczne i elastyczne modele kształtu. Najczęściej wykorzystywanymi metodami feature-besed są Elastic Bunch Graph Matching (EBGM) [M1] oraz Active Appearance Model (AAM) [M2].

Metody holistyczne oparte są z kolei na statystyce. Cała twarz rzutowana jest na nową przestrzeń, w której może być opisana wektorem o znacznie mniejszej długości niż w oryginalnej przestrzeni obrazu. Najczęściej wykorzystywanymi metodami holistycznymi w systemach identyfikacji twarzy są Principal Component Analysis (PCA) [M3] oraz Linear Discriminant Analysis (LDA) [M4].

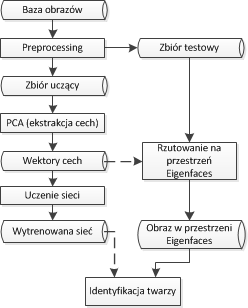
# Proponowane rozwiązanie

W proponowanym systemie skupiono się wyłącznie na etapie identyfikacji twarzy.  Założono, że zarówno zdjęcia znajdujące się w bazie danych jak i zdjęcia do identyfikacji zostały poddane obróbce i posiadają odpowiednie parametry.  Podany proces należy rozpatrywać, jako komponent pewnego, bardziej złożonego systemu rozpoznawania twarzy.

Pierwszy etap identyfikacji polega na wstępnym przetworzeniu obrazów.  Jego celem jest przeskalowanie i przycięcie zdjęć do odpowiednich rozmiarów oraz częściowe usunięcie szumów. Preprocessing bazy danych wykonywany jest w trybie offline, zdjęcia przygotowywane są z wyprzedzeniem, zanim zostaną wykorzystane. Zdjęcia, które będą poddane identyfikacji również muszą przejść wcześniej etap preprocessingu, nie jest on wykonywany w trakcie identyfikacji, gdyż wydłuża czas działania systemu i wprowadza dodatkowe utrudnienia.

Kolejnym krokiem jest przygotowanie danych dla sztucznej sieci neuronowej. Zdjęcia w bazie danych poddawane są analizie PCA, która jest popularną statystyczną metodą wyszukiwania wzorców w wielowymiarowych danych. Redukcja wymiarów przy użyciu PCA jest przeprowadzana w celu zmniejszenia rozmiaru danych wejściowych, usunięcia danych nadmiarowych oraz ekstrakcji cech najlepiej opisujących twarze. Rozmiar wektora wyjściowego może być regulowany, co pozwala na ścisłą kontrolę redukcji danych. Analiza PCA jest zadaniem o dużej złożoności obliczeniowej, pamięciowej i czasowej. Wystarczające jest przeprowadzanie analizy tylko w przypadku zmiany bazy obrazów, dlatego raz obliczone współczynniki PCA są zapisywane w bazie i mogą być wykorzystane w późniejszych etapach identyfikacji, bez konieczności przeprowadzania czasochłonnych obliczeń.

Po wstępnym przetworzeniu danych następuje klasyfikacja twarzy. Wektory własne twarzy, tzw. Eigenfaces stanowią wejście dla sztucznej sieci neuronowej. Sieć jest trenowana z wykorzystaniem algorytmu Back Propagation lub Resillent Propagation. Po fazie treningu można przystąpić do identyfikacji twarzy. Każdy neuron wyjściowy z sieci odpowiada jednej osobie w bazie, wygrywa neuron o najwyższej wartości na wyjściu. Wprowadzono jednak dodatkowy parametr będący progiem. Jeżeli wartość żadnego neuronu nie przekroczy wartości progowej uznaje się, że wektor wejściowy nie opisuje twarzy żadnej osoby znajdującej się w bazie. Przebieg procesu identyfikacji przedstawiono na rys. 1.



Rys. 1 Proces identyfikacji. Linie ciągłe- wejście/wyjście procesu. Linie przerywane- wykorzystanie komponentu w procesie.

## Przetwarzanie wstępne obrazów

Zdjęcia w dostępnej bazie nie mogły być wykorzystane w systemie bez wstępnej obróbki. Były one zaszumione, a twarz nie była wyśrodkowana. Ponadto zdjęcia zawierały duży obszar tła. Czynniki te mogłyby wpłynąć negatywnie na skuteczność systemu, dlatego zdjęcia zostały poddane prostemu przetwarzaniu wstępnemu. W celu zredukowania szumów użyto kilkukrotnej filtracji medianowej, z małą maską. Filtr medianowy bardzo dobrze usuwa zakłócenia typu sól-pieprz oraz w bardzo małym stopniu wpływa negatywnie na krawędzie obrazu powodując tylko minimalne rozmycie obrazu[M5].

Po usunięciu zakłóceń następuje przycięcie obrazu. Zdjęcia przycinane są z góry i z dołu przy użyciu stałego marginesu, dlatego nie jest to rozwiązanie uniwersalne i sprawdza się tylko dla wykorzystanej bazy. Przycięcie boków zdjęcia z wykorzystaniem stałych, ustalonych współrzędnych krawędzi było niemożliwe, gdyż twarze miały różne położenia na zdjęciach. Aby wyciąć poprawnie twarz należy ustalić jej lokalizację. Przy użyciu filtru Prewitt’a[M6] wykryto krawędzie na obrazie, następnie obraz poddano binaryzacji ze stałym progiem oraz operacji morfologicznego zamknięcia[M7], dzięki czemu uzyskano stosunkowo dobry kontur twarzy. Kolejnym krokiem było już wyznaczenie maksymalnej i minimalnej współrzędnej *X* otrzymanego konturu, która posłużyła do przycięcia obrazu. Kolejne kroki preprocessingu przedstawiono na rys. 2.

## Principal Component Analysis

* Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as “3.5-inch disk drive”.
* Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
* Do not mix complete spellings and abbreviations of units: “Wb/m2” or “webers per square meter”, not “webers/m2”. Spell out units when they appear in text: “. . . a few henries”, not “. . . a few H”.
* Use a zero before decimal points: “0.25”, not “.25”. Use “cm3”, not “cc”. (bullet list)

## Sieć neuronowa

Zastosowana została sieć typu feed-forward. W tego typu sieci neurony zgrupowane są w formie warstw. Warstwy otrzymują wejście od warstwy poprzedniej i przekazują swoje wyjście do następnej warstwy. Połączenie pomiędzy neuronami tej samej lub poprzedniej warstwy nie są możliwe [J1].

Dla sieci zostały ustalone następujące parametry:

* jedna warstwa ukryta
* 200 neuronów w warstwie ukrytej
* liczba neuronów w warstwie wejściowej jest zależna od rozmiaru PCA
* liczba neuronów w warstwie wyjściowej jest równa ilości osób, które sieć na zadanie rozpoznać (dla testów przyjęto 30)

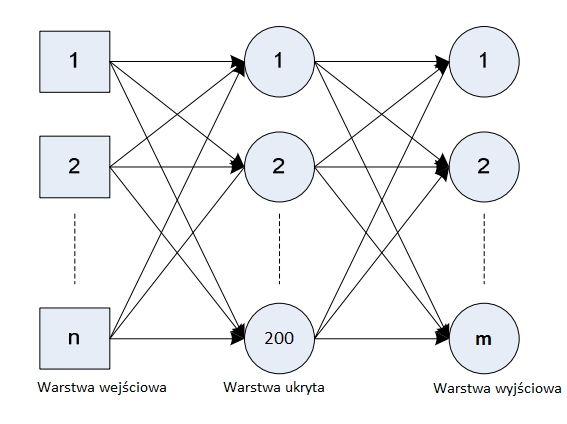
Wyjścia warstwy ukrytej i wyjściowej zostały poddane Sigmoidalnej Funkcji Aktywacji.

## Trening sieci neuronowej

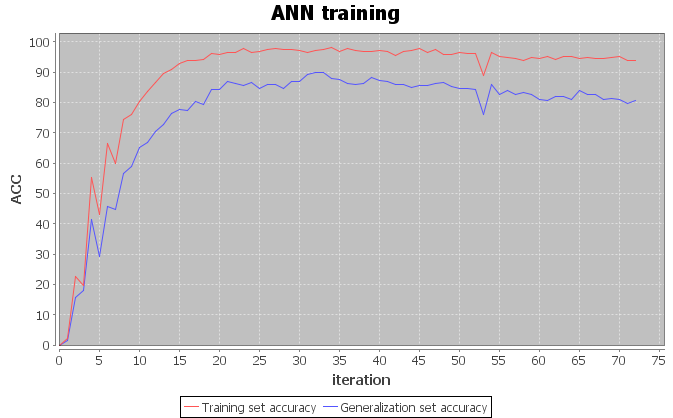
Trening sieci odbywa się przy wykorzystaniu zdjęć z bazy danych.

Sieć można trenować za pomocą dwóch metod: Resilient Propagation oraz Back Propagation. Obie metody są przykładem metod trenowania za pomocą propagacji. Taki trening polega na porównaniu wejścia systemu z pożądanym wyjściem, system jest dopasowywany dopóki różnica pomiędzy dwoma nie jest zminimalizowana. W takim treningu

Rys. 2 Architektura sieci.



błędy propagowane są od warsty wyjściowej do wyjściowej sieci. Propagacja wymaga, aby funkcja aktywacji była różniczkowalna.



## Resilient Propagation

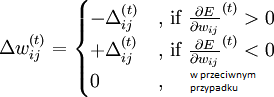
Charakterystyka:

Jest jedną z najlepszych uniwersalnych metod treningu sieci. Nie jest najlepsza w każdym przypadku, ale w większości [J2].

Resilient Propagation zazwyczaj daje lepsze rezultaty niż Back Propagation. Dodatkowo, metoda ta nie wymaga ustalania parametrów, które są konieczne przy Back Propagation. Wysoka skuteczność sieci uzyskiwana jest już przy niewielkiej liczbie iteracji, co obrazuje rysunek nr 3.

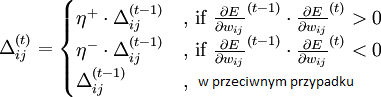
Opis działania:

Działanie tej metody składa się z dwóch kroków. W pierwszym, przeliczamy wagi przy użyciu następującego wzoru:



Znajdujemy zmianę wagi (delta w) w zależności od zmiany wartości (delty).

Po policzeniu wag, ustalamy nową deltę. Uzyskujemy ze wzoru:



gdzie:

|  |  |
| --- | --- |
| Symbol | Znaczenie |
|  | Zmiana wartości dla obecnej iteracji t. |
|  | Zmiana wartości dla iteracji t-1. |
|  | Gradient wagi z i do j dla iteracji t. |
|  | Gradient wagi z i do j dla iteracji t-1. |
| *E* | Błąd treningu |
| *E*(*t* − 1) | Błąd treningu dla iteracji t-1 |
|  | Zmiana wagi z i do j w obecnej iteracji (t). |

Rys. 3 Statystyki treningu dla metody Resilient propagation

## Back Propagation

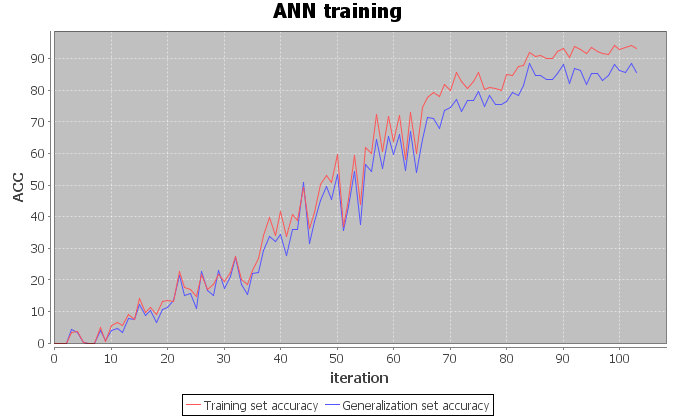
Wymaga ustalenia szybkości uczenia (learning rate) oraz pędu (momentum). Znalezienie optymalnych wartości tych parametrów może być trudne.

Przy użyciu tej metody skuteczność sieci wzrasta systematycznie, jednak wolniej (rys. nr 4).

## Implementacja

Sieć została zaprogramowana w języku programowania Java, przy użyciu biblioteki Encog [J3]. Encog jest zaawansowanym frameworkiem obsługującym sieci

Rys. 4 Statystyki treningu dla metody Back propagation



neuronowe oraz inne algorytmy wykorzystywane w uczeniu maszynowym.

Kod odpowiedzialny za stworzenie sieci wygląda następująco:

*network = new BasicNetwork();*

*network.addLayer(new BasicLayer(null, true, inputs));*

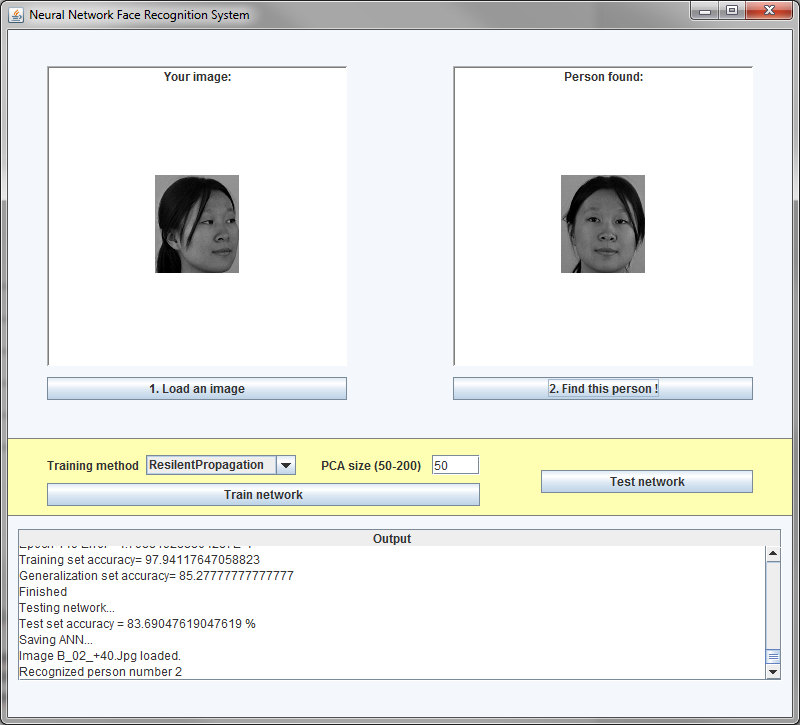
*network.addLayer(new BasicLayer(new ActivationSigmoid(), true, hidden));*

*network.addLayer(new BasicLayer(new ActivationSigmoid(), false, outputs));*

*network.getStructure().finalizeStructure();*

*network.reset();*

## Wyniki testów



Wraz ze zwiększeniem rozmiaru PCA, czas treningu oraz skuteczność sieci wzrasta. Podobne zachowanie obserwujemy zwiększając liczbę neuronów w warstwie ukrytej.

Dodanie dodatkowej warstwy ukrytej nie zwiększyło skuteczności sieci.

Prezentowana tabela pokazuje zależność między rozmiarem PCA oraz ilością neuronów w warstwie ukrytej, a skutecznością sieci dla zbioru testowego. Sieć została trenowana metodą Resilient Propagation, wykonanych zostało 150 iteracji.

|  |  |  |
| --- | --- | --- |
| Ilość neuronów w warstwie ukrytej | Rozmiar PCA | Skuteczność sieci |
| 50 | 20 | 56% |
| 50 | 50 | 68% |
| 50 | 100 | 71% |
| 50 | 200 | 73% |
| 50 | 300 | 78% |
| 200 | 20 | 74% |
| 200 | 50 | 83% |
| 200 | 100 | 85% |
| 200 | 200 | 88% |
| 200 | 300 | 89% |
| 300 | 20 | 84% |
| 300 | 100 | 92% |
| 300 | 300 | 89% |
| 400 | 100 | 88% |

W dalszej części procesu testowania, obrazy testowe zostały poddane modyfikacjom: została zmieniona jasność zdjęć lub wprowadzono zakłócenia. Skuteczność sieci dla takich zdjęć była zdecydowanie niższa.

## Działanie programu

Po treningu sieci, program jest w stanie rozpoznawać twarze osób zarejestrowane pod różnymi kątami.

W pierwszym kroku ładowane jest zdjęcie osoby. Następnie po kliknięciu ‘Find this person’ prezentowane jest zdjęcie znalezionej osoby zarejestrowane pod kątem 0 stopni oraz na konsoli wypisywany jest odpowiedni komunikat.

Rys. 5 Rozpoznanie odoby

1. Alaa Eleyan and Hasan Demirel, “PCA and LDA Based Face Recognition Using Feedforward Neural Network Classifier” Springer-Verlag Berlin Heidelberg 2006
2. http://www.heatonresearch.com/wiki/Resilient\_Propagation
3. http://www.heatonresearch.com/encog