

Warsztaty z Technik Uczenia Maszynowego

Piotr Rowicki, Igor Starega

22 marca 2024

Tabela Zmian

data	zmiana
08.03	dodanie dokumentacji wstępnej
19.03	pełen opis sekcji 2

1 Dokumentacja Wstępna

1.1 Opis problemu

Zagadnieniem którym będziemy zajmować się w naszym projekcie jest znajdowanie punktów charakterystycznych twarzy. Jest to jeden z problemów konkursowych dostępnych na platformie [Kaggle](#). Dokładniej polegać on będzie na znalezieniu pozycji takich części twarzy jak oczy, uszy, nos, usta (wstępnie na zdjęciach, docelowo w czasie rzeczywistym na obrazie z kamery).

1.2 Zbiór danych

Do wytrenowania sieci wykorzystamy zbiór danych udostępniony na wcześniej wspomnianej platformie,

1.3 Użyte narzędzia

Językiem użytym do implementacji będzie Python. Do obróbki danych zostanie wykorzystana biblioteka Pandas, co wynika z natury tego, jak dane zostały nam przekazane. Do implementacji modelu planujemy wykorzystać bibliotekę TensorFlow, a jeżeli pozwoli na to czas, również PyTorch (do implementacji alternatywnego modelu). Do implementacji aplikacji graficznej wykorzystamy również bibliotekę OpenCV, w celu wykrywania twarzy. Decyzję o Frameworku do samego GUI zostawiamy na później

1.4 Podział pracy

1. Piotr Rowicki
 - (a) Analiza obróbka danych
 - (b) możliwa implementacja drugiego modelu
 - (c) wspólna implementacja szaty graficznej
2. Igor Starega

- (a) Implementacja pierwszego modelu
- (b) wspólna implementacja szaty graficznej

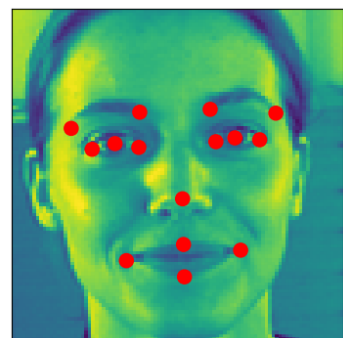
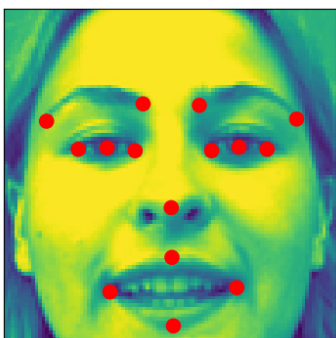
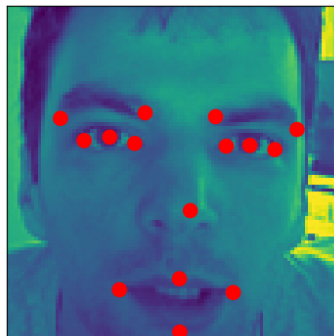
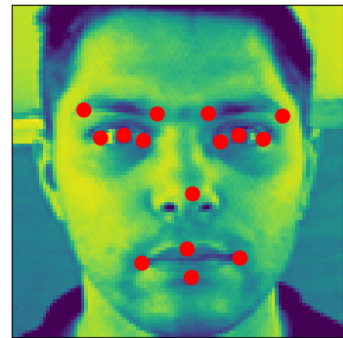
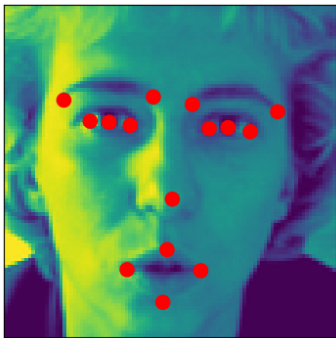
2 Analiza i Obróbka Danych

2.1 Opis zbioru

Dane dostępne są w postaci tabeli zbudowanej z 31 kolumn i 7049 wierszy. Kolumna 31 zawiera ciąg liczb stałoprzecinkowych reprezentujących monochromatyczny obraz twarzy. rozdzielczość tych zdjęć to 96 na 96 pikseli. pozostałe 30 kolumn zawiera współrzędne charakterystycznych cech twarzy (każda cecha ma poświęcone dwie kolumny, na współrzędną x-ową i y-ową). Rozważane cechy to:

1. środek lewego oka
2. zewnętrzny kącik lewego oka
3. wewnętrzny kącik lewego oka
4. środek prawego oka
5. zewnętrzny kącik prawego oka
6. wewnętrzny kącik prawego oka
7. czubek nosa
8. zewnętrzny koniec lewej brwi
9. wewnętrzny koniec lewej brwi
10. zewnętrzny koniec prawej brwi
11. wewnętrzny koniec prawej brwi
12. lewy kącik ust
13. prawy kącik ust
14. środek górnej wargi
15. środek dolnej wargi

kilka wybranych zdjęć, wraz z nałożonymi na nie cechami przedstawione zostały na rysunku 1



Rysunek 1: Zdjęcia przedstawiają kilka wybranych obrazów ze zbioru danych wraz z nałożonymi współzrędnymi kluczowych cech

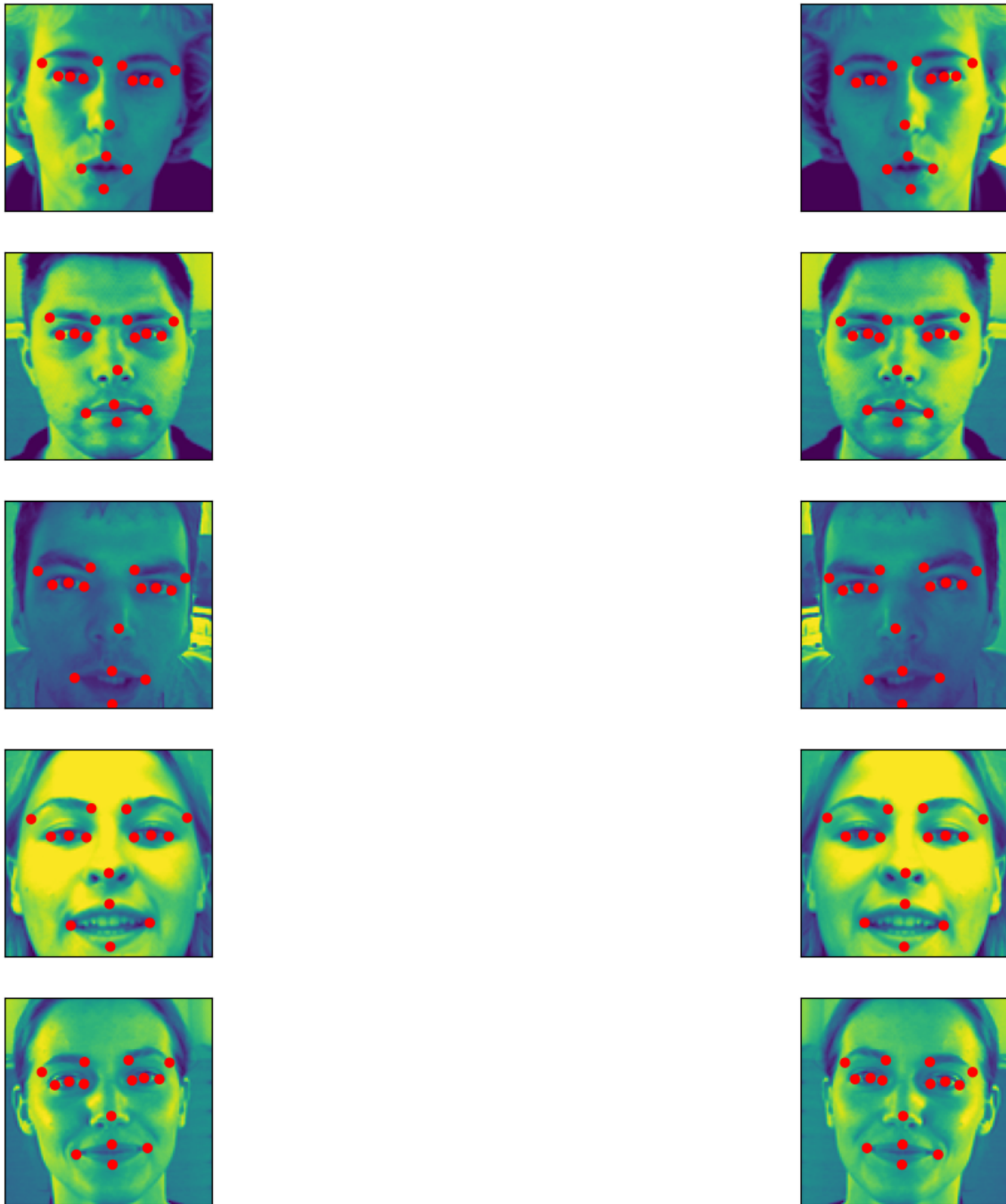
2.2 obróbka danych

2.2.1 problemy ze zbiorem

Niestety w zbiorze danych znajdują się wiele wierszy, które nie mają wszystkich danych. Zdecydowaliśmy się więc pozbyć się tych rekordów. Po tej Operacji pozostało nam 2140 rekordów. Są one gotowe do dostarczenia do modelu, jednak ich ilość jest poniżej oczekiwań.

2.2.2 rozwiązanie problemu

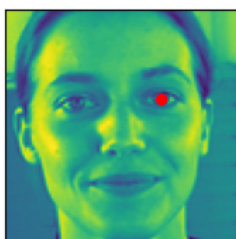
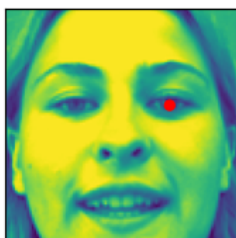
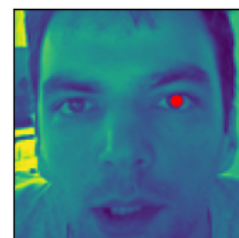
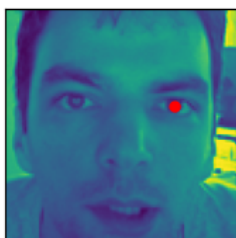
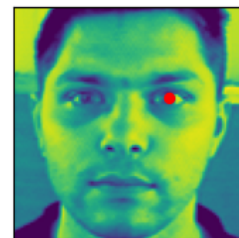
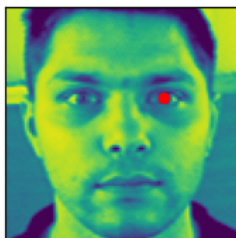
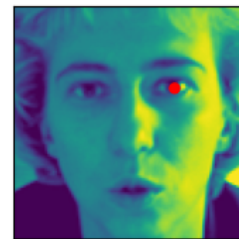
Aby rozwiązać problem małego zbioru danych, postanowiliśmy rozwiązać go, obracając wszystkie obraz i wszystkie cechy lustrzanie odbito względem osi OX . W celu ułatwienia tej operacji, wszystkie współrzędne cech zostały najpierw znormalizowane do przedziału $[-1, 1]$. W wyniku tych operacji otrzymaliśmy nowy zbiór danych. Porównanie odpowiadających sobie zdjęć i cech zostało przedstawione na Rysunku 2.



Rysunek 2: Rysunek przedstawia porównanie zdjęć i cech z wyjściowego zbioru danych z nowym zbiorem otrzymanym poprzez odbicia lustrzane odbicia względem osi OX

Takie operacje rodzą jednak kolejny problem. Mianowicie dla wierszy odbitych, pozycje lewych

cech na zdjęciach, odpowiadają danym opisanym jako prawe cechy i na odwrót. Aby to rozwiązać wystarczy jedynie pozamieniać wartości w odpowiadających kolumnach. Po tej operacji otrzymamy już całkowicie spójne dwa zbiory. dowód tego widać na Rysunku 3



Rysunek 3: Rysunek przedstawia porównania odpowiadających sobie zdjęć z nałożonymi na nie pozycji lewych oczu. jak widać znajdują się one po tej samej stronie niezależnie od tego czy zdjęcie było odbite czy nie

Teraz Wystarczy połączyć oba zbiory w jeden, i otrzymamy zbiór 4280 rekordów gotowych do modelu.