



POLITECHNIKA KRAKOWSKA
WYDZIAŁ INFORMATYKI I TELEKOMUNIKACJI
KIERUNEK: INFORMATYKA



SZEREGI CZASOWE, GIEŁDA I EKONOMIA

PRINCIPAL COMPONENT ANALYSIS

SPRAWOZDANIE Z LABORATORIUM

Maria Guz
Paweł Midura

Zadania zostały zrealizowane za pomocą języka programowania Python oraz dzięki środowisku Jupyter. Wybór tych narzędzi podyktowany był ich przejrzystością, intuicyjnością i prostotą.

Dane

Analizowane dane składają się z 20 zdjęć w rozmiarze 250 na 250 pikseli, z których pierwsze 3 zostały zaprezentowane na grafikach poniżej.



Zdjęcia załadowano z użyciem biblioteki PIL i zmieniono je na zdjęcia czarno białe.



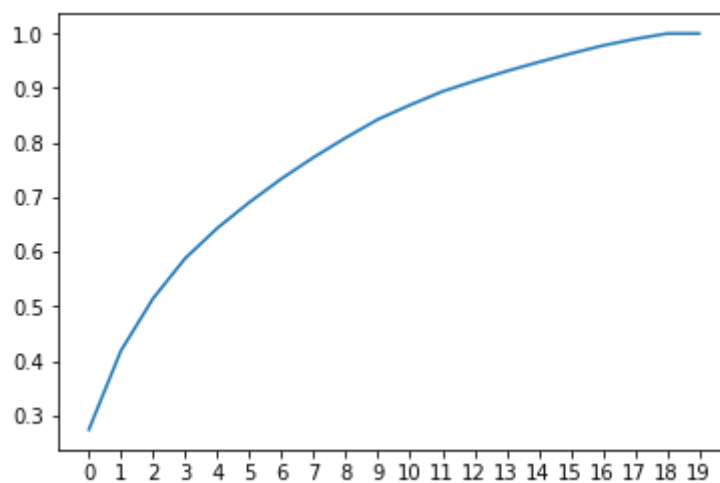
Następnie wykorzystując klasę PCA pakietu sklearn.decomposition przeprowadzono analizę PCA dla zbioru 20 zdjęć czarno białych. Po przeprowadzeniu analizy PCA dla 20 głównych składowych, przedstawiono wartości własne:

```
[ 3.65908230e+04  2.66121333e+04  2.16367269e+04  1.89589621e+04
 1.64612455e+04  1.52301753e+04  1.45542191e+04  1.38251872e+04
 1.32556393e+04  1.28148591e+04  1.13404057e+04  1.09831834e+04
 9.71184483e+03  9.38907446e+03  9.01549644e+03  8.72721924e+03
 8.50390727e+03  7.63999184e+03  7.00708997e+03  6.53971620e-12]
```

Wektory własne przedstawiono niżej:



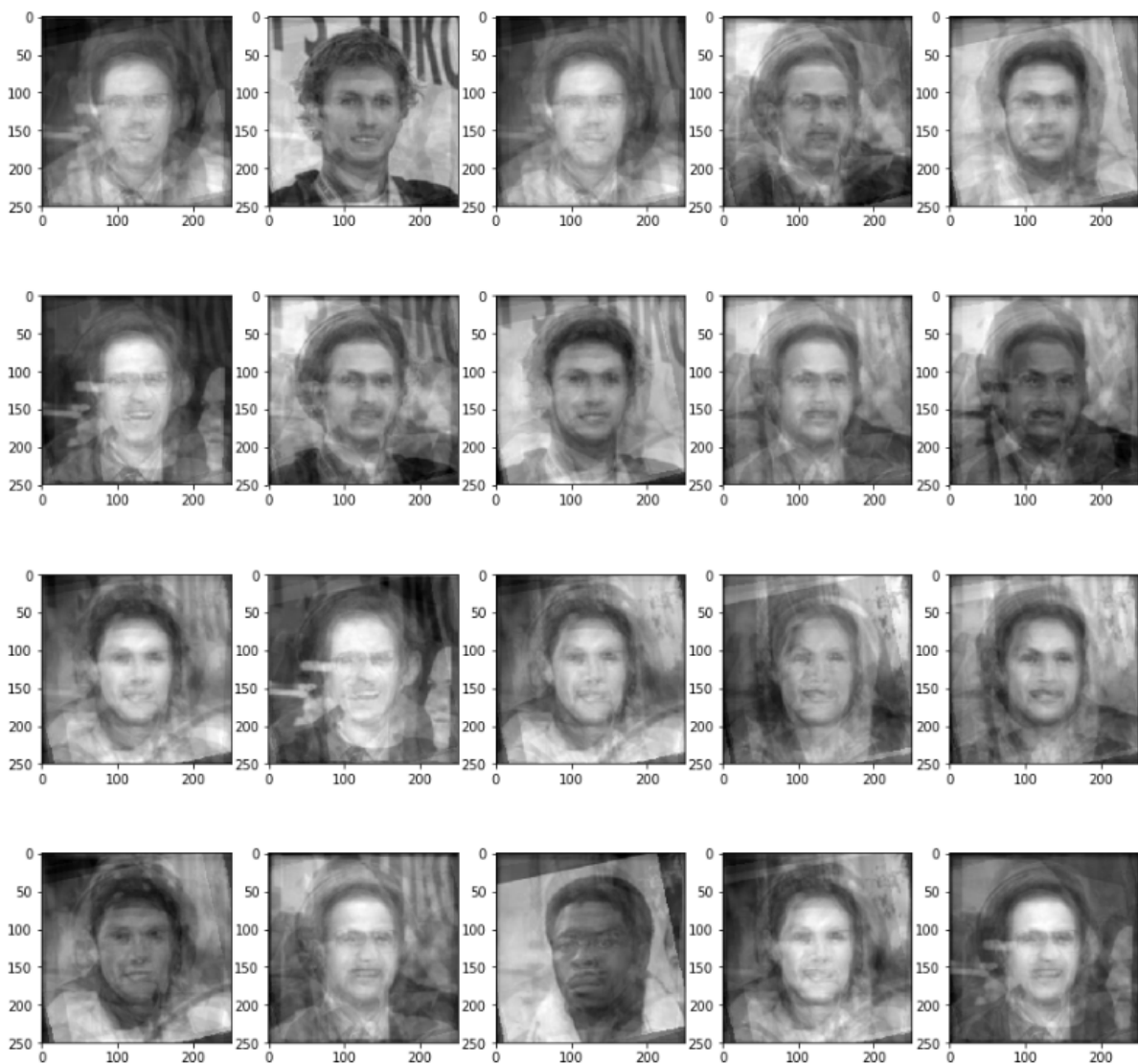
Sprawdzono także procent wyjaśnionej skumulowanej wariancji dla każdej ze składowych.



```
array([ 0.27383225,  0.41867595,  0.51442255,  0.58793643,  0.64335632,
        0.69079691,  0.73411986,  0.77321135,  0.80914834,  0.84273508,
        0.8690376 ,  0.89370917,  0.91299968,  0.93102927,  0.94765266,
        0.96322995,  0.97802026,  0.98995811,  1.          ,  1.
      ])
```

Następnie na podstawie różnej ilości składowych głównych próbowano odtworzyć obrazy ze zbioru oryginalnego. Rezultaty przedstawiono poniżej.

Components = 4



Components = 7



Components = 10



Components =15



Components = 20



Wnioski

W przypadku stosowania większej ilości składowych otrzymywano lepsze rezultaty przy rekonstrukcji obrazów. Jednak niektóre obrazy potrzebowały mniej składowych aby je wyraźnie zrekonstruować, przykładami takich obrazów są obraz drugi oraz obraz osiemnasty. Do rekonstrukcji obrazu, która była rozpoznawalna wystarczyły 4 składowe w przypadku obrazu numer 2 oraz 7 składowych dla obrazu numer 18. Może to być spowodowane tym, że twarze znajdujące się na obrazach są bardziej generyczne od pozostałych. To znaczy, mają w sobie najwięcej cech wspólnych wszystkich obrazów w zbiorze.