

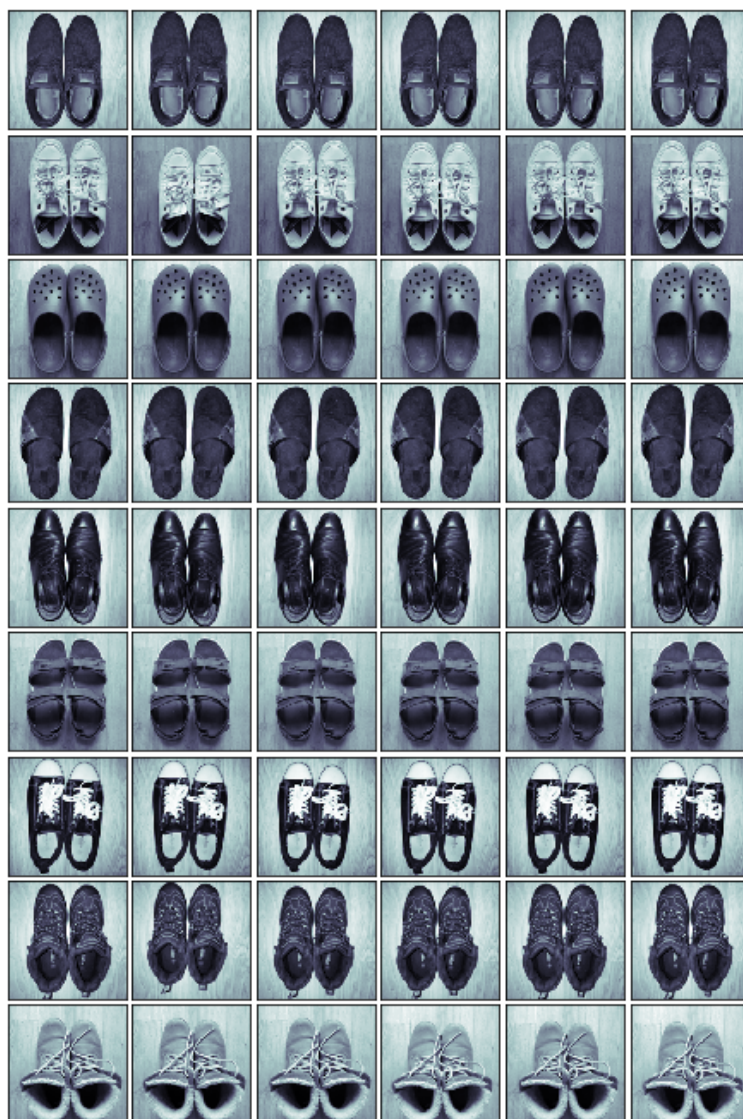
# Podstawy Uczenia Maszynowego

## Laboratorium 3

Maria Polak

### 1. Przygotowanie zbioru

Zbiór danych składa się z 9 różnych par butów. Każda para została sfotografowana 6-krotnie.



*Rys 1. Zbiór danych - buty*

Poszczególne klasy znajdujące się w zbiorze:



*Codzienne*



*Conversy*



*Crocsy*



*Klapki*



*Sandały*



*Pantofle*



*Zimowe*

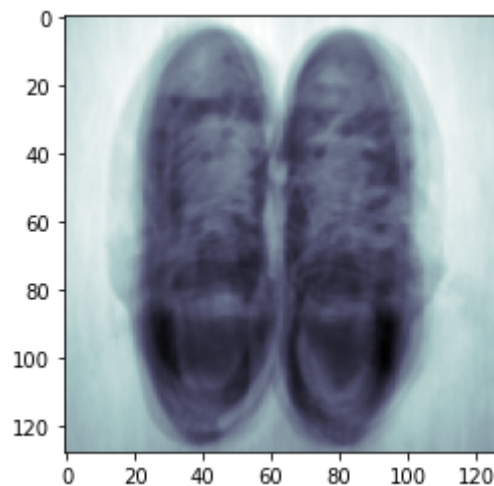


*Zimowe Damskie*



*Trampki*

Średnie zdjęcie, które zostało odjęte celem wycentrowania zbioru wygląda następująco:



**Rys 2. Średni but**

Na średnim zdjęciu widzimy 'poświatę/cień' różnych butów. Zdjęcie na pierwszy rzut oka najbardziej przypomina klasę *Trampki*.

## 2. Wektory bazowe i wariacja

Przygotowany zbiór został przepuszczony przez PCA. Otrzymane wektory bazowe i ich wariancja zostały przedstawione na poniższej grafice:



**Rys 3. Wektory bazowe i ich wariancja**



Widzimy że tylko kilka pierwszych (najważniejszych) wektorów przypomina buty. Każdy kolejny wektor ma coraz mniej szczegółów. Od mniej więcej 13 pozycji widzimy już tylko kontur kształtu, a na kilku ostatnich nie zostało już praktycznie nic.

### 3. Redukcja wymiarowości

Zdjęcia odtworzone z 16 najważniejszych cech prezentują się następująco.



**Rys 4. Buty odtworzone z 16 cech**

Pierwsze 16 cech składa się na 94% wariancji, więc wyniki są stosunkowo trafne. Kolory nie są jednak identyczne z oryginałem oraz wokół butów pojawia się 'poświata' podobna do tej ze średniego zdjęcia.

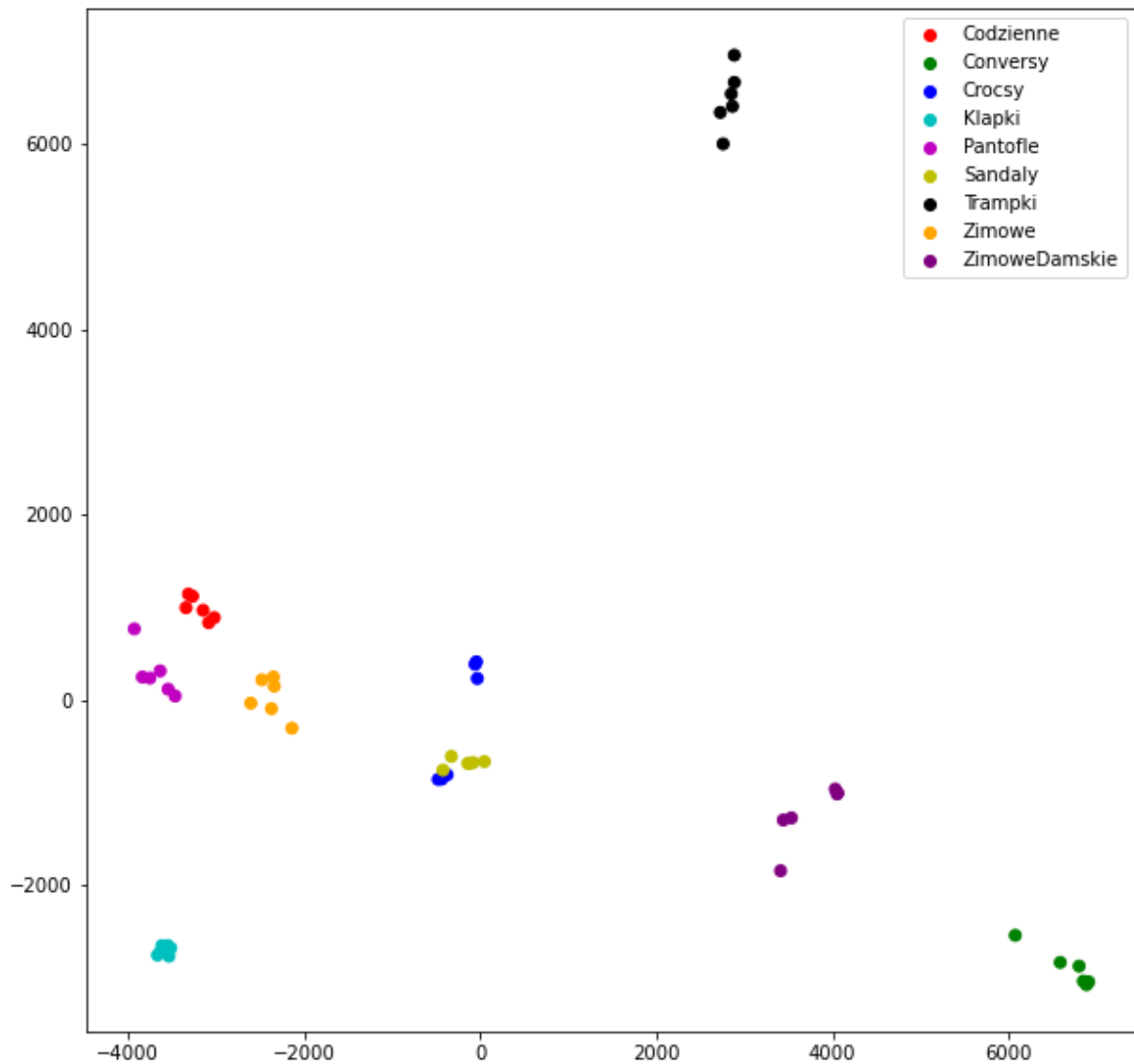
Zdjęcia odtworzone z 4 najważniejszych cech prezentują się następująco.



**Rys 5. Buty odtworzone z 4 cech**

Przy 4 wynikach straciliśmy dużo informacji. Mimo że 4 pierwsze cechy składają się na 70% wariacji to podobne do oryginału pozostały tylko klasy *Trampki*, *Klapki* i *Pantofle*, a reszta klas jest ciężka do rozpoznania. Najbardziej ucierpiała klasa *ZimoweDamskie*, która aktualnie przypomina klasę *Conversy*.

#### 4. Rzut 2D



**Rys 6. Rzut zbioru na 2D**

Klasy są w większości odseparowane, jedynie klasy *crocsy* i *sandaly* lekko na siebie nachodzą. k-NN najprawdopodobniej poradziłby sobie z klasyfikacją tak podanego zbioru całkiem dobrze.