

Podstawy Uczenia Maszynowego - Główne składowe kubków

Kamil Krzempek

1 Zbiór danych

Zbiór danych składa się z 44 zdjęć kubków, z których każdy należy do jednej z 5 kategori - zwykłe, węższe u podstawy (z uchem na lewo lub prawo), matero i przeźroczyste.



Rysunek 1: Przetworzony zbiór danych wraz z przypisanymi klasami

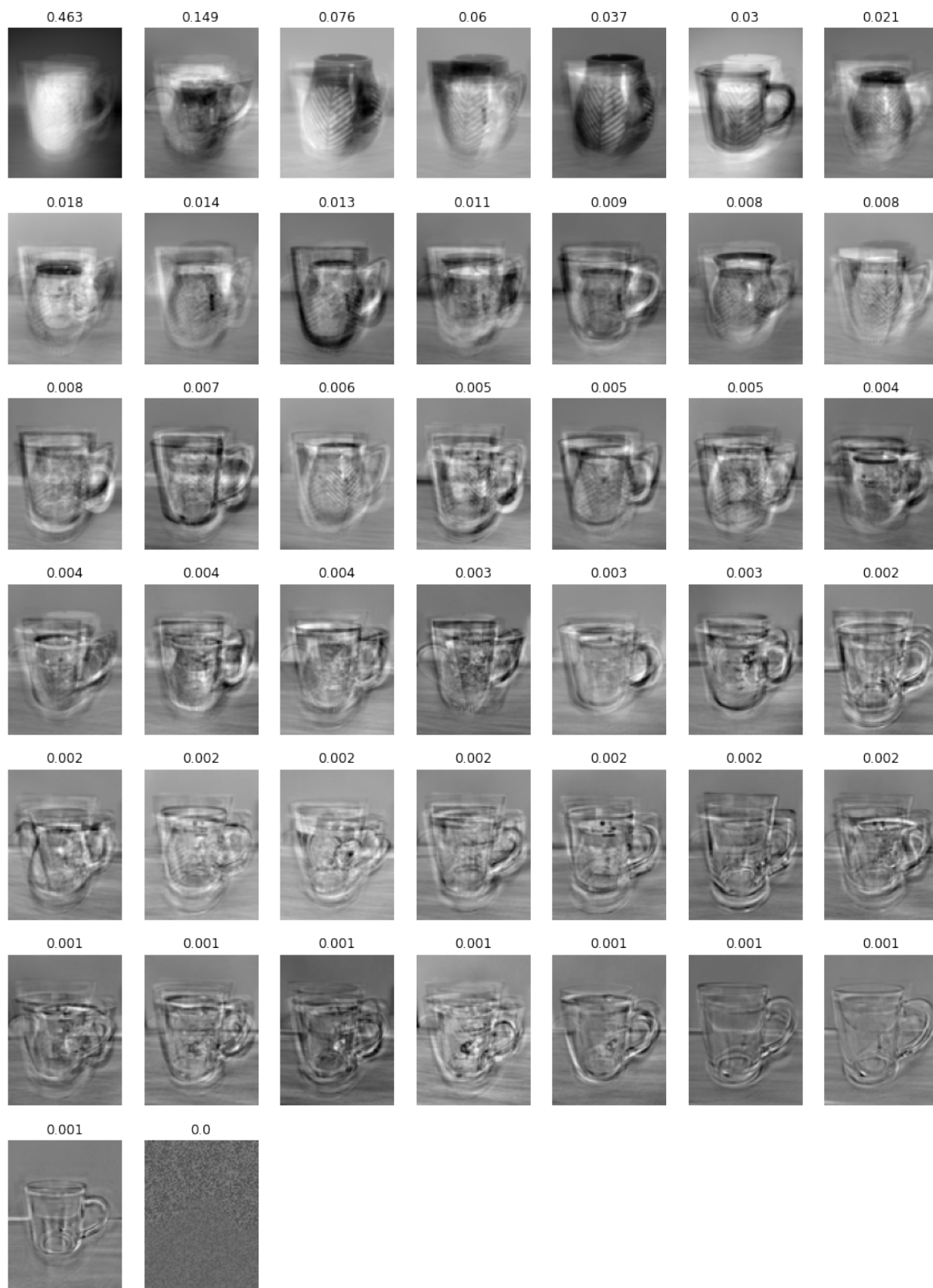
2 Transformacja PCA

2.1 Średnie zdjęcie



Na średnim zdjęciu możemy dostrzec rozmazany kształt kubka. Widać także tło (krawędź między ścianą a półką), które było jednolite dla wszystkich zdjęć.

2.2 Principal components

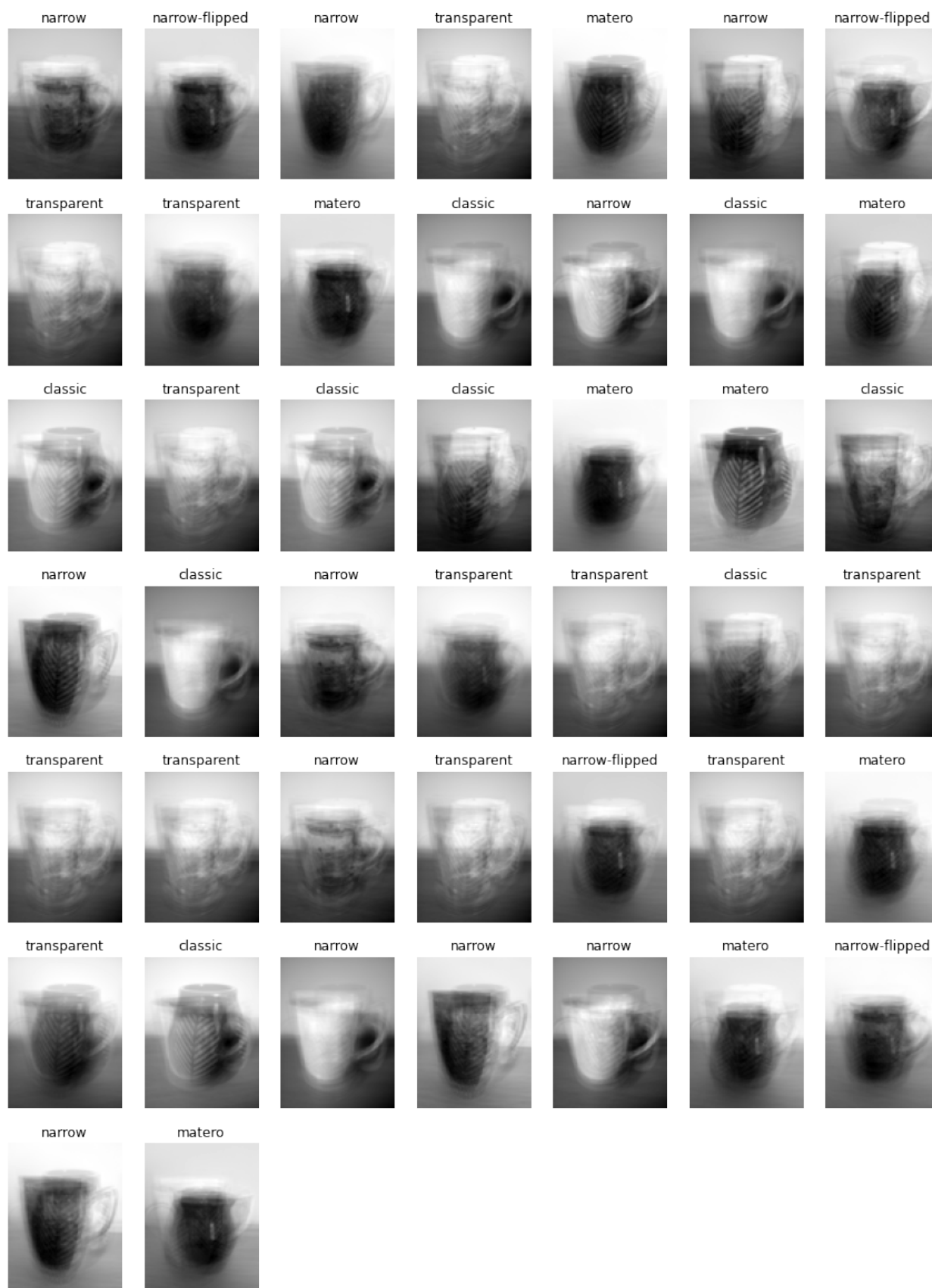


Rysunek 2: Wektory bazowe wraz z powiązanymi wariancjami

3 Redukcja wymiarowości

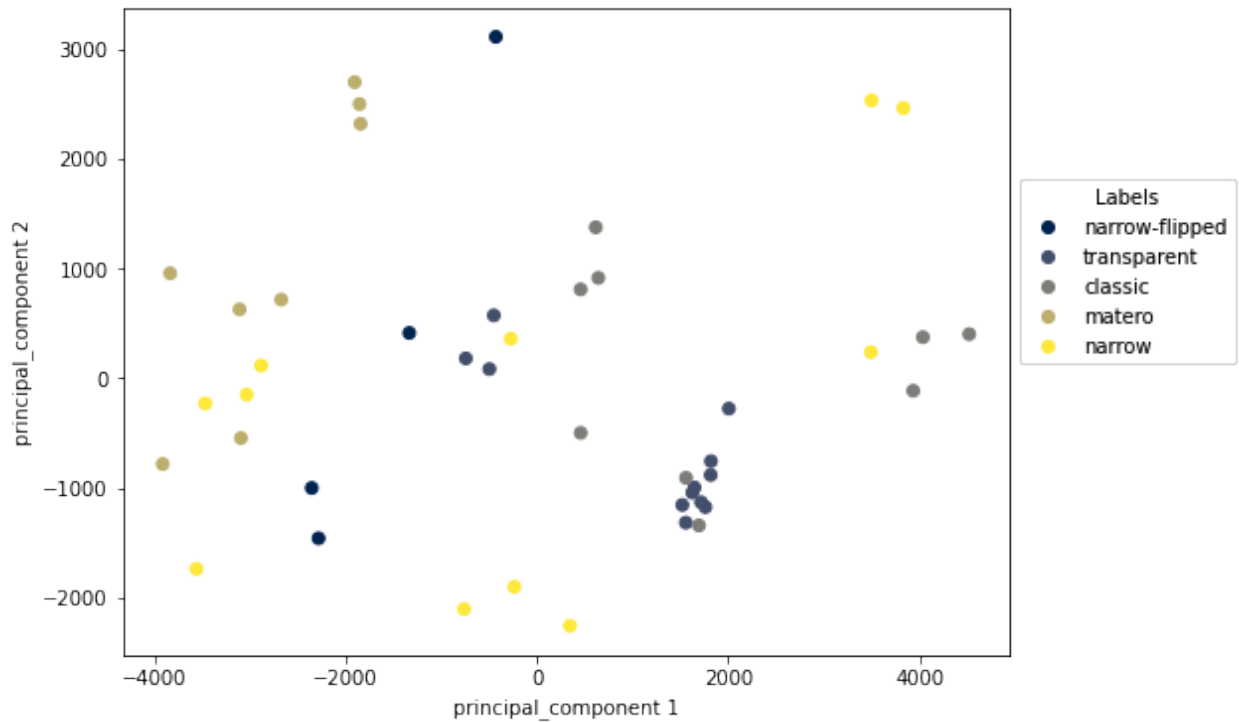


Rysunek 3: Efekt redukcji obserwacji do 16 cech i transformacji z powrotem do pierwotnej bazy



Rysunek 4: Efekt redukcji obserwacji do 4 cech i transformacji z powrotem do pierwotnej bazy

Im więcej głównych składowych tym bardziej były zdjęcia po powrocie do oryginalnej bazy. Przy 4 były bardzo rozmazane, natomiast przy 16 w większości dało się zidentyfikować oryginalne zdjęcie.



Rysunek 5: Obserwacje zredukowane do 2 cech przedstawione na płaszczyźnie

Klasy są rozrzucone (np. narrow) i częściowo przemieszane (np. classic z transparent). Sprawia to, że k-NN mógłby sobie nie poradzić z ich zaklasyfikowaniem.