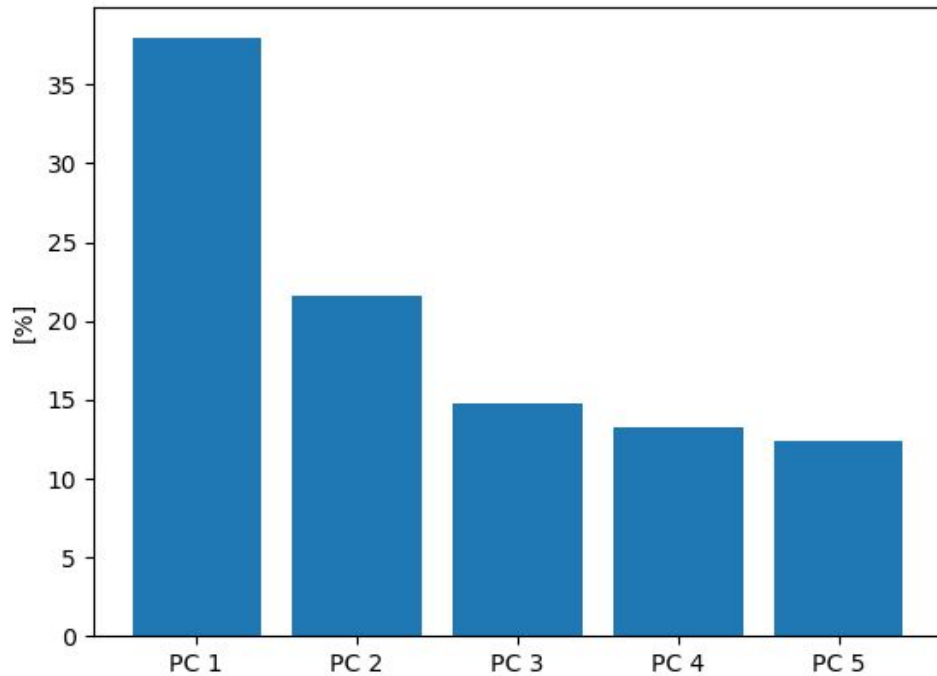


PCA/KPCA sprawozdanie

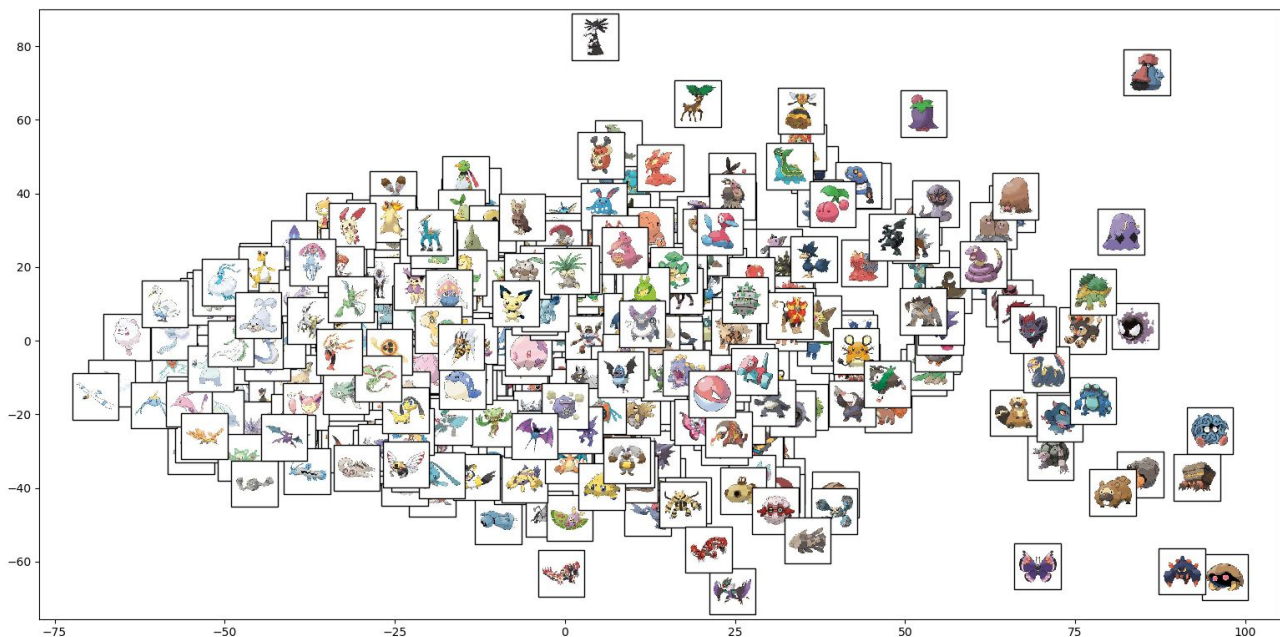
Józef Jasek

Analiza PCA na zbiorze Pokemon

Analizowanym zbiorem jest zbiór 819 obrazów RGB rozmiaru 64x64. Mówimy zatem o przestrzeni $64 * 64 * 3 = 12\,288$ wymiarowej. Po przeprowadzeniu PCA dla 5 składowych otrzymujemy:



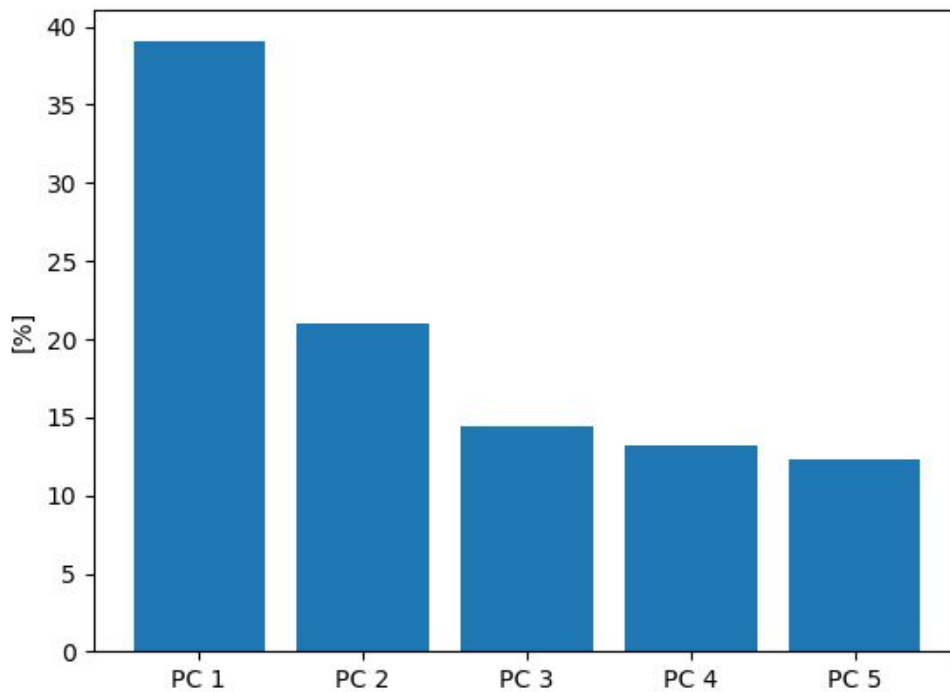
Co dla dwóch wymiarów generuje wykres:



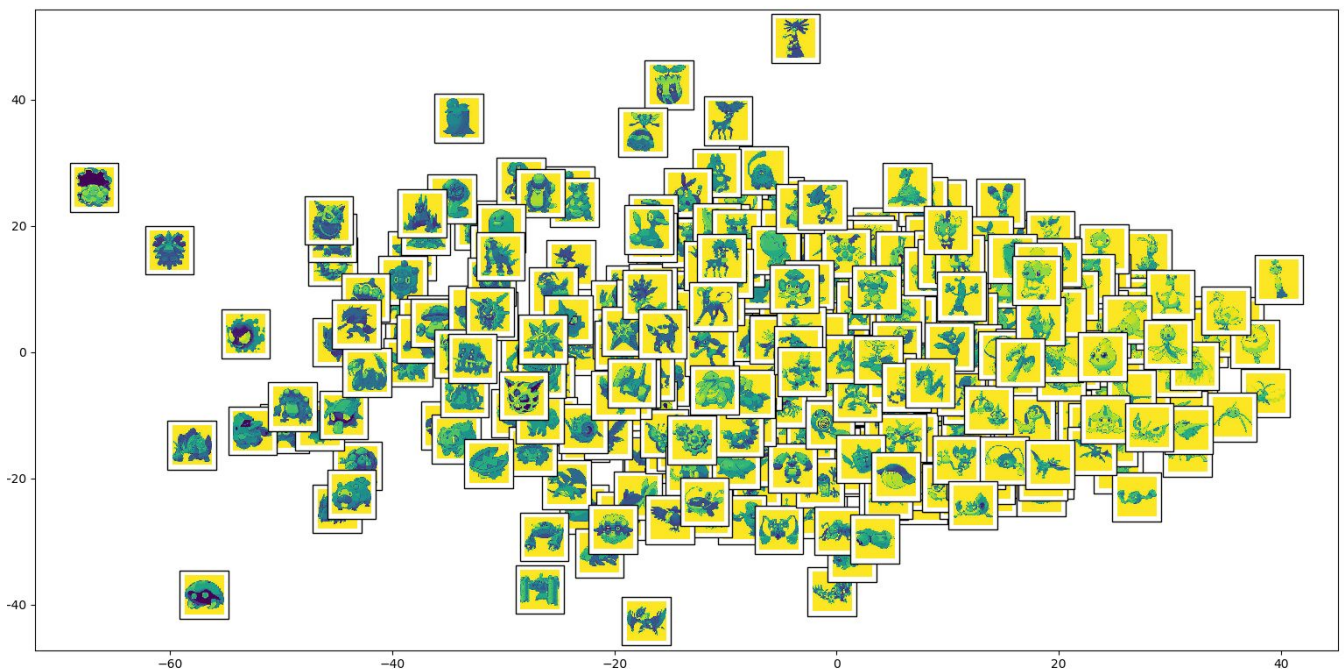
Z pola `explained_variance_` odczytujemy, że biorąc 5 wymiarów zachowamy 57,3% początkowej wariancji. Średnia liczba tych samych sąsiadów w zbiorze początkowym co w tym po przeprowadzeniu PCA dla każdego pokemona wyniosła 1.33089.

Analiza KPCA na zbiorze Pokemon

Tym razem zastosujemy czarno-białe wersje obrazków z wyżej. Otrzymujemy



Widzimy jeszcze większe znaczenie PC1 - i rzeczywiście największy wpływ na rozmieszczenie na wykresie stanowi to jak duży fragment obrazka jest czarny/szary. Tym razem otrzymaliśmy 59,8% zachowanej wariancji. Średnia liczba tych samych sąsiadów dla każdego pokemona wyniosła 1.32601.



Wnioski

W naszym zbiorze z uwagi na duże skomplikowanie elementów użycie większej liczby komponentów PCA jak również zastosowanie KPCA nie wpłynęło istotnie na poprawę jakości rozwiązania. Nie mniej z reguły użycie większej liczby komponentów pozwala lepiej zachować wariancję zbioru początkowego, a metoda KPCA dzięki swym właściwościom pozwala lepiej aproksymować zbiory o nieliniowych zależnościach. Najprawdopodobniej z tego powodu obie metody dały podobne rezultaty.