

Кластеризация

Беляков Дмитрий

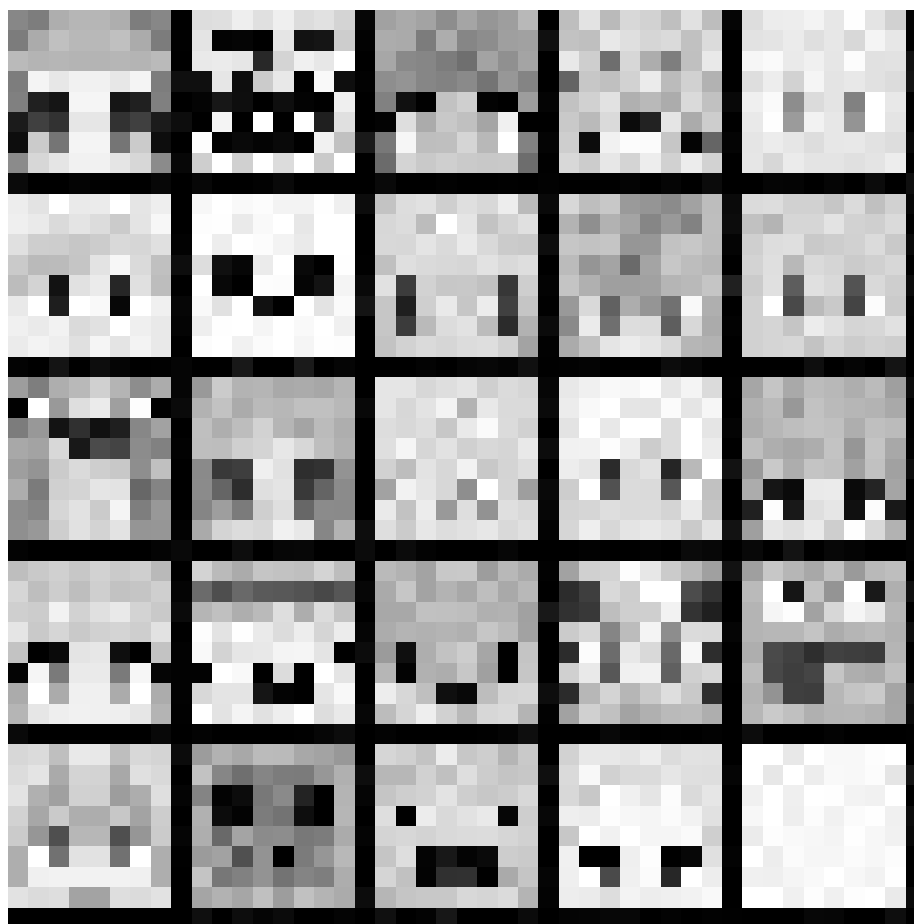
April 2, 2022

1 Описание задачи

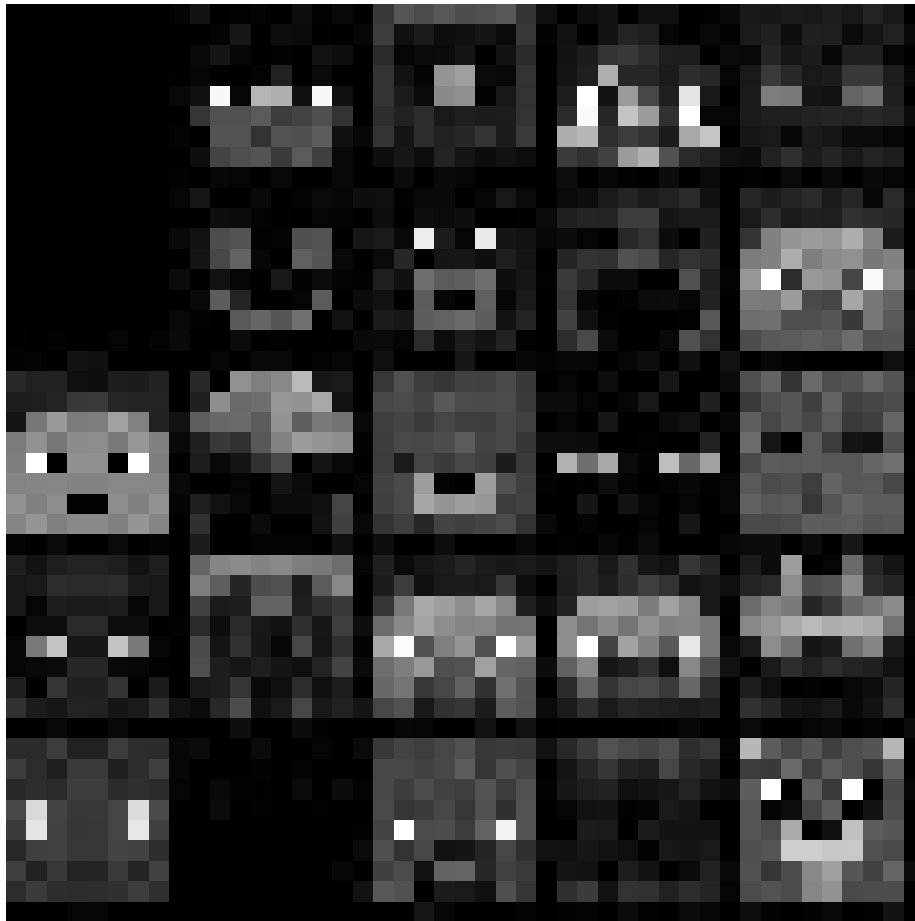
Попробуем провести кластерный анализ текстур лиц игроков из игры minecraft. Используем датасет Minecraft skins (<https://www.kaggle.com/datasets/alxmamaev/minecraft-skins>). Датасет предоставляет 7930 изображений с текстурами моделей игроков minecraft. Вырежем из каждого файла текстуру лица - квадрат 8 на 8. Конвертируем его в черно-белое изображение и запишем как массив из 64 элементов. Затем с помощью иерархической кластеризации из sklearn разобьем изображения на кластеры. Рассмотрим кластеризацию с одиночной усредненной и ward типами связей.

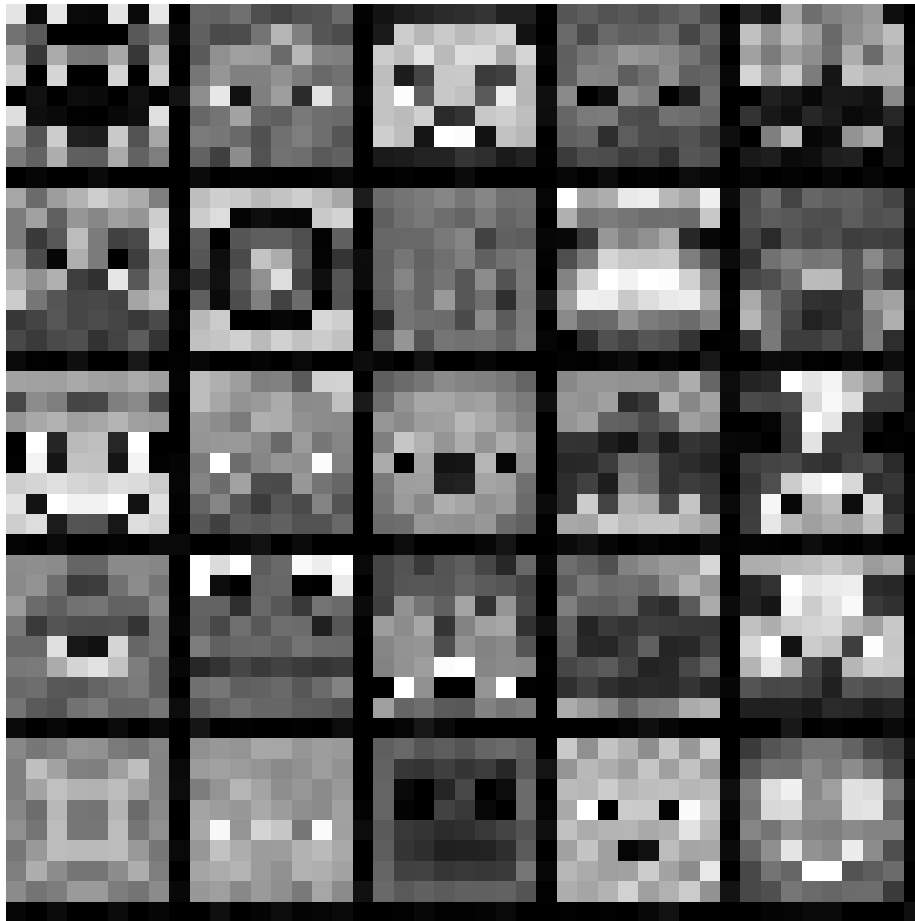
2 Описание решения

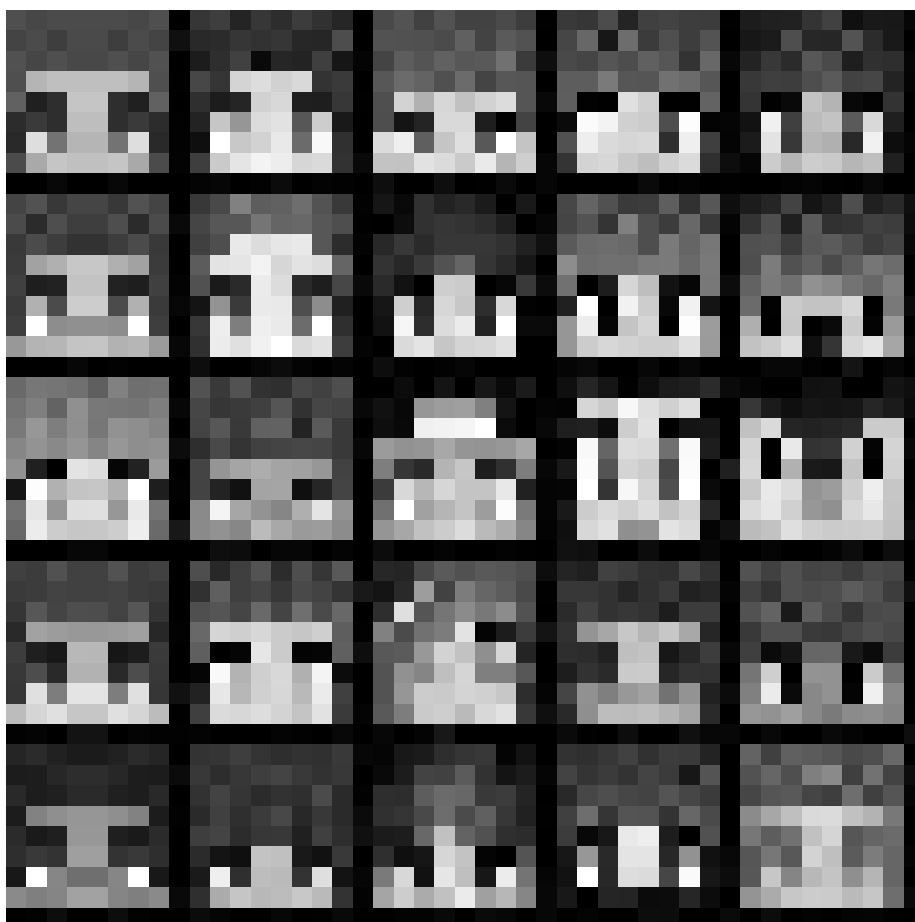
Кластеризация с одиночной и усредненной типами связи оказались малоэффективными и определили большинство изображений к одному кластеру. Однако кластеризация с ward-связью разбила изображения на пять равных кластеров. Рассмотрим по 25 случайных изображений из каждого кластера











3 Вывод

Как видно из полученных выборок, в каждом кластере изображения обладают весьма значительными сходствами. В первом кластере - светлые лица. В третьем темные. Во втором преимущественно "мужские" (волосы занимают середину лица, глаза близко к центру). В пятом преимущественно "женские" (Волосы занимают больше половины лица, глаза внизу). В четвертом лица которые не попали в другие кластеры (преимущественно одноцветные без отличительных характеристик)