

Анализ и сравнение методов распознавания лиц на примере датасета Labeled Faces in the Wild.

Трофимов Иван

December 2022

Постановка задачи.

На практике часто возникает задача о распознавании лиц. В настоящей работе производится сравнение и исследование методов распознавания лиц на примере датасета Labeled Faces in the Wild.

Методы.

Будем сравнивать два метода:

1. SVM – метод опорных векторов. Поскольку изображения задаются пространством большой размерности, то перед этим сделаем метод главных компонент PCA.
2. CNN – сверточная нейронная сеть.

Имплементация датасета.

Исходный датасет есть в `sklearn.datasets`. Загрузим датасет, каждое изображение имеет свою метку. Выберем такую загрузку, чтобы минимальное число изображений на каждого человека было равно 10.

```
1 lfw_people = fetch_lfw_people( resize=0.5, color=True, min_faces_per_person=10)
```

Посмотрим на картинку:



Число разных классов при такой имплементации равно 158.

Метрика.

Выберем соответствующую метрику качества, по которой будет происходить сравнение моделей. В качестве такой метрики выберем *roc-auc* – площадь под кривой аус. Выбор такой метрики обусловлен следующим:

1. *roc – auc* – является адекватной метрикой для оценки модели классификации, Она отображает истинную и постоянную способность модели к прогнозированию. Это связано с тем, что она учитывает еще и вероятности для предсказания всех меток.

2. В реализации `sklearn.metrics` реализация этой метрики есть и в случае многоклассовой классификации.
3. В задаче с дисбалансом классов (коей наша задача является) есть реализация метрики, не чувствительная к дисбалансу классов. Согласно документации это реализация метрики с параметрами *multiclass* = "ovo" и *average* = "macro".

Эксперименты.

PCA + SVM.

Исходные изображения в `sklearn` хранятся в виде numpy-массивов, у которых число столбцов равно 8742. Конечно, это колоссальная размерность. Будем снижать размерность методом PCA до 200. А затем в этом пространстве применим *SVM*. У нас будут два разных ядра: полиномиальное и RBF.

Для поиска оптимальных гиперпараметров будем применять `GridSearchCV` на 5-ти фолдах.

Результат с метрикой качества на тестовом датасете приведен ниже в таблице:

SVM - ядро	roc-auc
RBF	0.9271
POLY	0.9321

Таблица 1: Результаты метрики для лучших моделей с разными ядрами в методе SVM

CNN.

Следующая модель, которую мы будем исследовать – это сверточная нейронная сеть (CNN). В настоящей работе были проделаны разные эксперименты: менялась конфигурация сети, гиперпараметры, число эпох. В зависимости от этого получались разные результаты. Ниже будет приведена статистика с экспериментов.

Общая схема работы такая: сначала делается несколько сверточных слоев, после этого несколько полносвязных слоев так что в результате мы имеем логиты. Для получения из логитов вероятностей нужно использовать softmax.

Более детальное описание архитектур нейросетей представлено в ноутбуке *LFW_analysis.ipynb*. Результаты экспериментов представлены ниже в таблице.

Номер эксперимента	Число эпох	learning rate	roc-auc	roc-auc	Комментарий
			train	test	
1	30	1e-3	1.0	0.9792	3Conv+2Linear
2	100	1e-3	1.0	0.9802	3Conv + 2Linear
3	30	1e-3	1.0	0.9819	3Conv+3Linear
4	50	1e-3	0.9999	0.9753	4Conv+3Linear
5	70	1e-3	1.0	0.9806	4Conv+4Linear
6	200	1e-4	1.0	0.9815	4Conv+4Linear
7	500	1e-4	1.0	0.9790	4Conv+4Linear

Таблица 2: CNN эксперименты

Ниже приведены графики обучения:

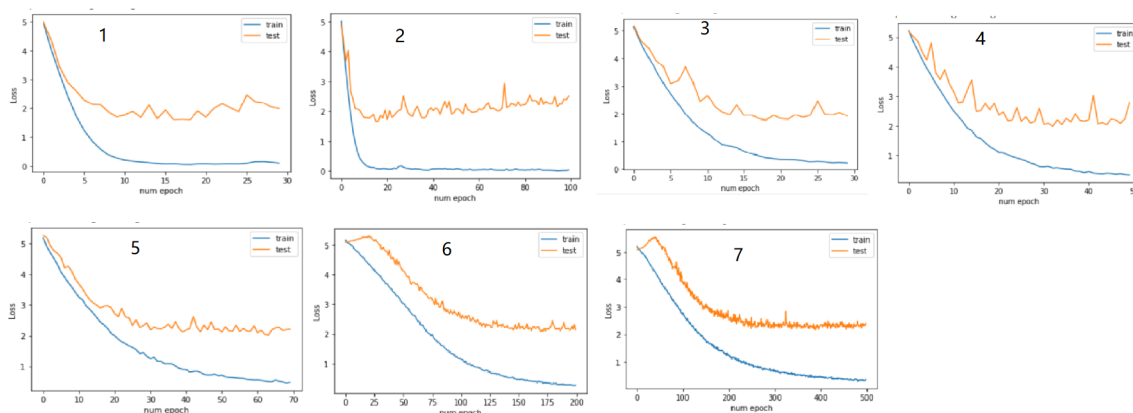


Рис. 1: Графики обучения. По вертикальной оси – loss, по горизонтальной оси – число эпох. Номер на картинке соответствует номеру эксперимента. Желтый цвет – тестовая выборка. Голубой цвет – тренировочная выборка.

Обсуждение результатов.

Исходя из экспериментов, сделанных выше, можно судить о следующем.

Во-первых, результаты, которые были получены с помощью CNN превосходят результаты, полученные с помощью SVM. Это видно из сравнения таблиц 1 и 2. Заметим, что метрика качества на тестовой выборке с хорошей точностью стремилась к единице. При этом, вероятно, в некоторых экспериментах происходило переобучение. Здесь видятся следующие предложения для улучшения: расширить датасет, нужно больше изображений, изменение архитектуры сети, возможно, нужно больше поиграться с dropout.

Во-вторых, как видно из таблицы 2 усложнение архитектуры сети не всегда ведет к улучшению качества этой модели на тестовой выборке.

В-третьих, лучшее качество, которого удалось добиться по этой

метрике есть 0.9819, что было достигнуто в эксперименте 3. Что есть довольно неплохой результат.

Предложения по улучшению.

Как мы видим, предложенное решение задачи распознавания лиц содержит изъяны. В ходе экспериментов не получилось достичь на тестовом датасете метрики, превышающей 0.99. Есть несколько предложений по улучшению моделей:

1. Работа с данными. А именно разделение датасета на тестовую и тренировочную выборки. Думаю, что можно сделать разбиение на тестовую и тренировочную выборки более однородным.
2. Более тщательный анализ архитектуры сети. В частности более тщательный подбор гиперпараметров.

Выводы.

В настоящей работе была исследована задача распознавания лиц на примере датасета Labeled Faces in the Wild. В ходе решения было сравнено две модели машинного обучения: PCA+SVM (classic ML) и CNN (Deep Learning).

Для PCA+SVM были исследованы два ядра: радиальное (RBF) и полиномиальное (POLY). Для сверточной нейронной сети были проведены 8 экспериментов (разные архитектуры и гиперпараметры сети) и была установлена лучшая модель, дающая на тестовой выборке метрику качества: 0.9819.

Анализ результатов привел к следующему: CNN показывает метрику качества лучше, чем PCA+SVM.