

Отчет

Face landmarks detection

Repository: <https://github.com/akanametov/face-landmarks-dlib>

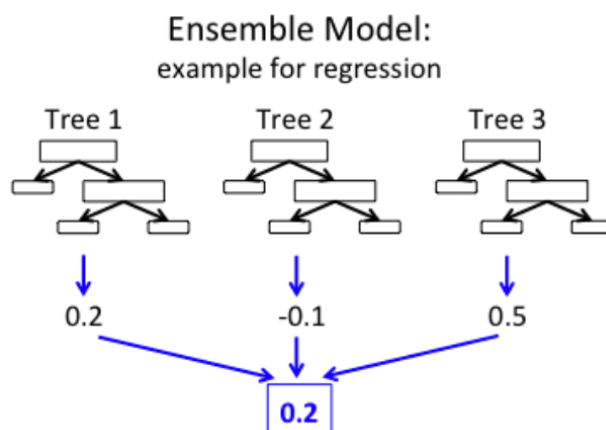
Постановка задачи

Задача заключается в реализации алгоритма обнаружения 68 особых точек на лице человека (face alignment), и тестировании данного алгоритма на общедоступных датасетах и сравнении с аналогами.

Обучающие данные представляют собой разметку 68 точек лица, в формате [IBUG](#). Данные для обучения и тестирования можно найти по [ссылке](#). Предоставленные данные состоят из датасетов 300W и Menpo, на тестовой части которых (папки test) нужно измерить точность работы обученной модели. Обучающие данные можно использовать любые, например можно взять train часть 300W и Menpo датасетов (из Menpo нужно выкинуть профильные изображения, на которых размечено только 39 точек).

Эксперименты

В качестве детектора точек используется Ensemble Regression Tree (ERT) [1] реализованный в библиотеке dlib:

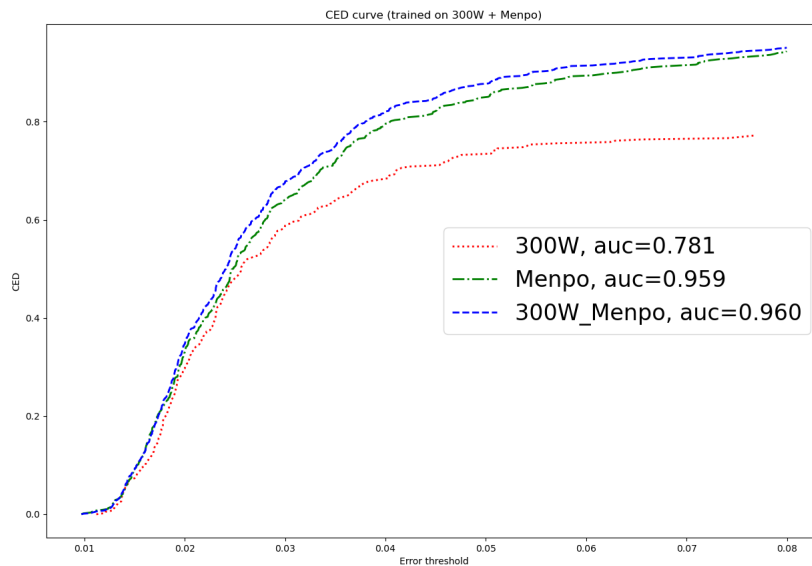


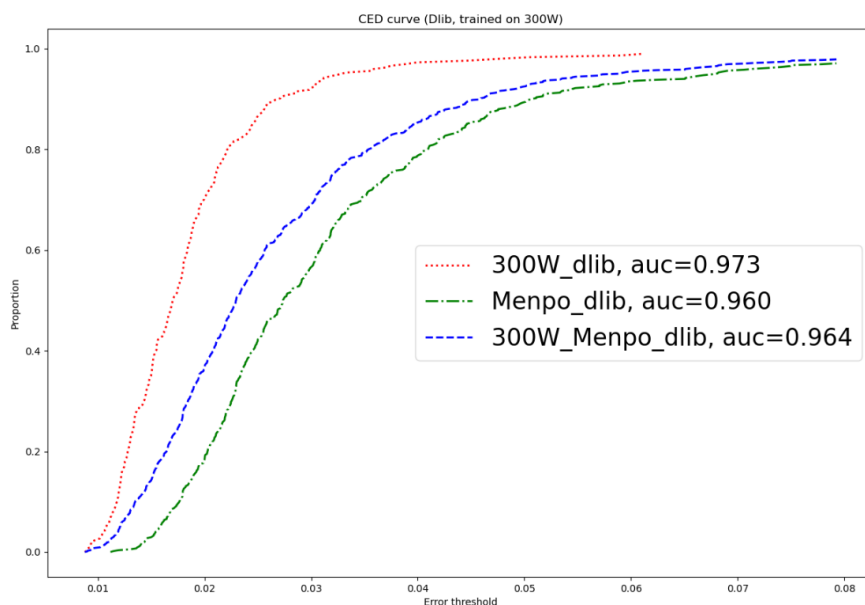
Модель обучается на train части датасетов: 300W, Menpo и 300W + Menpo (для этого сначала нужно подготовить аннотации в формате .xml). Ниже приведены гиперпараметры обученных моделей:

Данные обучение	300W	Menpo	300W + Menpo
Гиперпараметры			
<i>mu</i>	0.05	0.05	0.05

<i>tree_depth</i>	4	4	5
<i>cascade_depth</i>	15	15	20
<i>feature_pool_size</i>	512	512	512
<i>num_test_splits</i>	50	50	50
<i>oversampling_amount</i>	5	5	5
<i>oversampling_translation_jitter</i>	0.2	0.2	0.2
<i>num_threads</i>	4	4	4

В качестве метрики используется CED (cumulative error distribution) и график CED-AUC на тестовой выборке. Ниже приведены графики тестирования модели обученной на train части датасетов 300W и Menpo, а также график для модели из dlib:





Пример работы детектора точек:



Выводы

В качестве алгоритма детектора был выбран ERT из dlib и обучен отдельно на данных 300W и Менпо, а также на 300W + Менпо. Лучший результат был получен на 300W + Менпо с глубиной дерева $tree_depth=5$ и $cascade_depth=20$. Качество модели может быть улучшено за счет доразметки данных (координаты точек присутствуют не для всех лиц). Кроме того, обучение детектора bounding box тоже может привести к уточнению координат лиц.

Обученные модели, чекпоинты, скрипты для обучения, валидации а также примеры можно найти в репозитории [4].

Reference

1. Kazemi, Bahid, and Josephine Sullivan. "[One millisecond face alignment with an ensemble of regression trees.](#)" *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014.
2. [300 Faces in-the-wild challenge \(300W\)](#). 2013.
3. Deng et al. "[The Menpo Benchmark for Multi-pose 2D and 3D Facial Landmark Localisation and Tracking.](#)" *International Journal of Computer Vision*. 2018.
4. Face-landmarks detector trained on 300W and Menpo datasets
<https://github.com/akanametov/face-landmarks-dlib>