

Отчет по проекту на тему:

«Детектирование лиц на фотографиях на основе сверточных нейронных сетей»

Подготовили:

Кетков С.

Кондратьев Н,

гр. 16 МАГ ИАД

Введение

В проекте мы рассматриваем задачу детектирования лиц на основе сверточных нейронных сетей. Для тестирования мы используем две обученные сверточные сети:

- MTCNN (<https://medium.com/wassa/modern-face-detection-based-on-deep-learning-using-python-and-mxnet-5e6377f22674>)
- Tiny face detector (<https://www.cs.cmu.edu/~peiyunh/tiny/>)

Для тестирования используется часть набора данных Wider face, доступного по следующей ссылке:

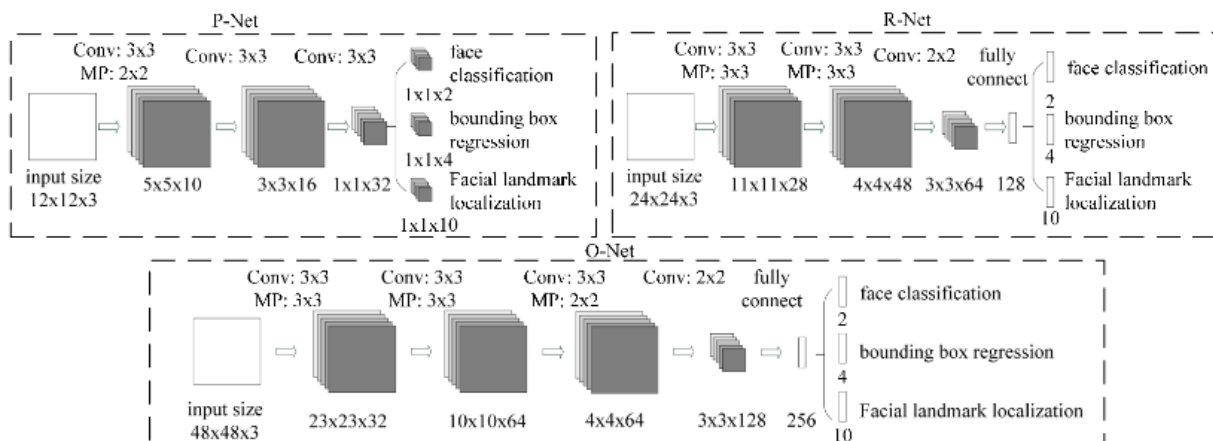
<http://mmlab.ie.cuhk.edu.hk/projects/WIDERFace/>.

Для сравнения двух обученных сетей мы используем ошибки на тестовой выборке и вычислительную эффективность каждой из сетей.

Архитектура

Для каждой из рассматриваемых сетей мы приводим основные характеристики и блоки, их составляющие, а также используемые функции потерь.

1) Сеть MTCNN



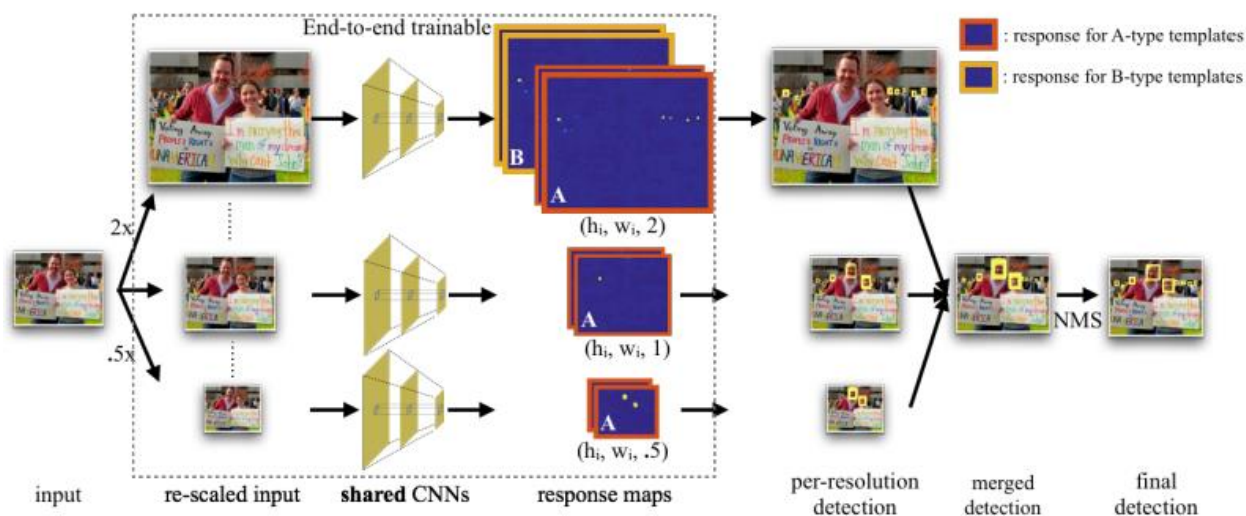
В сети используется три основных блока:

- Proposal Network
- Refine Network
- O – Net

Первый блок необходим для выделения прямоугольных областей, которые наиболее вероятно содержат изображение лица. Для выделения границ областей, наложенных друг на друга, используется non-maximum suppression. Для задачи восстановления регрессии в данном случае используется евклидова (L_2) норма. Второй блок предназначен для классификации полученных изображений и разделения их на те, которые содержат лицо человека и те, которые лица не содержат. Для этих целей формулируется задача двухклассовой классификации и в качестве функции потерь используется кросс-энтропия. Третий блок предназначен для выделения специальных точек на лицах (facial landmarks). В качестве функции потерь используется L_2 норма.

В данной модели также используется online hard sample mining. Тем самым изображения тренировочной выборки ранжируются согласно полученным значениям функции потерь и изображения с наибольшими потерями (первые 70% списка) и только по ним считается градиент в процессе обратного распространения ошибки.

2) Tiny face detector



Данная сеть предназначена специально для распознавания “маленьких” изображений лиц. Для этого используется дообучение (fine tuning) одной из моделей ImageNet. Для того, чтобы сеть могла лучше обучиться распознавать небольшие изображения лиц исходное изображение сэмплируется в разных масштабах, подается на вход сети, а затем результат сливается в один и к нему применяется non-maximum suppression.

Также к тренировочной выборке применяется следующая операция предобработки. Для каждого изображения выделяются области, которые имеют, по меньшей мере, 70% пересечения с настоящими изображениями лиц. Остальные части изображения игнорируются, т.е. градиент от них приравнивается к нулю.

Тестирование

Для тестирования сетей мы исследуем ошибку на тестовой выборке и вычислительную эффективность. Для подсчета ошибки для одного изображения используется следующий алгоритм:

1. Для каждой пары найденное – настоящее изображение подсчитывается мера схожести Жаккарда, которая равна площади пересечения отнесенной к площади объединения.
2. Все пары ранжируются по величине индекса и выбирается первые N пар, где N - число настоящих лиц на изображении.
3. Для всех N пар, если индекс Жаккарда > 0.5 , то лицо считается детектированным верно (+1), иначе – неверно (+0). Таким образом, картинка характеризуется отношением правильно детектированных лиц к общему числу лиц на фотографии.

Таким образом, мы вычисляем среднюю ошибку для всех изображений тестовой выборки.

В качестве тестовой выборки используется 3500 фотографий из набора данных WiderFace. В качестве оценки вычислительной эффективности мы приводим время работы каждой из сетей на приведенном наборе данных.

Результаты экспериментов

Результаты экспериментов представлены в следующей таблице:

Model	Dataset	Test error	Testing time
MTCNN	Wider face (1000 photos)	0.323	3506 s
TFD		0.128	40435 s

В результате можно сделать вывод, что Tiny Face Detector имеет меньшую ошибку на тестовой выборке, однако обладает меньшей вычислительной эффективностью, чем MTCNN. Таким образом, в зависимости от цели детектирования, можно использовать модель с необходимыми свойствами.

Список литературы:

1. “Joint face detection and alignment using multi-task cascaded convolutional networks”, Zhang K., Zhang Z., Li Z., 2015
2. “Finding tiny faces”, Hu P., Ramanan D., 2017