

一、项目介绍

MTCNN 提出了一种 Multi-task 的人脸检测框架，将人脸检测和人脸特征点检测同时进行。提出一个新的基于 CNN 的级联型框架，用于联合 (joint) 人脸检测和对齐；还设计轻量级的 CNN 架构使得速度上可以达到实时；提出一个有效的 online hard sample mining 方法来提高表现能力；在人脸检测和人脸对齐上提高了不少精度。论文原文采用 caffe 实现，本项目用 keras/tensorflow+python 实现。

二、论文与数据集

论文地址：https://kpzhang93.github.io/MTCNN_face_detection_alignment/

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

CelebA 数据集：<http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html>

三、论文理解要点

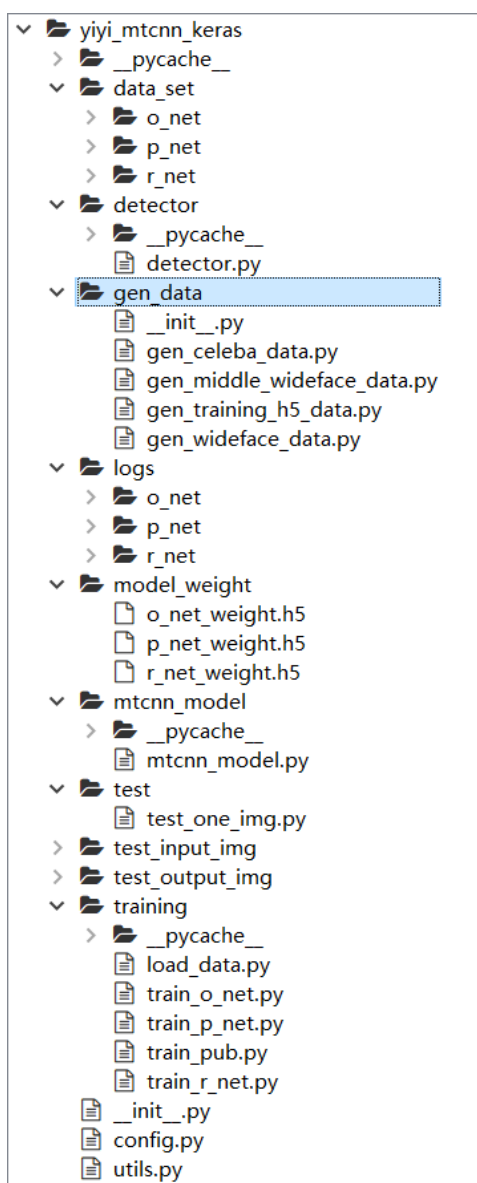
- 1、不是端到端的，模型分 p_net/r_net/o_net 三个网络，分三步；
- 2、训练数据样本分为 正样本、负样本、部分样本、关键点样本，它们在训练时的比例控制为 1: 3: 1: 2；
- 3、数据样本来自 Wider_face 和 CelebA 数据集，Wider_face 数据集生成正样本、负样本、部分样本三类样本(通过 iou 来区别)，CelebA 数据集生成关键点样本；
- 4、每个网络的损失函数均由三部分组成，这三部分分别对应网络的三种任务：分类任务、边框回归任务、关键点检测任务；三部分损失在 p_net/r_net 的比重是: 1: 0.5: 0.5，在 o_net 中是 1: 0.5: 1，比重在损失函数中体现为 a 参数上；
- 5、正负样本用于分类任务，正样本和部分样本用于边框回归任务，关键点样本用于人脸标注点定位，这些在损失函数中体现为 b 参数上；
- 6、人脸分类任务损失函数中需要实现在线困难样本选择
- 7、r_net/o_net 的训练数据正样本、负样本、部分样本，是通过前面已经训练好网络进行预测生成，即 r_net 的正样本、负样本、部分样本 由 p_net 生成，o_net 的正样本、负样本、部分样本 由 p_net、r_net 生成；关键点样本由 CelebA 数据集生成；
- 8、p_net/r_net/o_net 预测的数据，均要实现 nms，对输入图片要实现图片金字塔；(整体作为一个 detector)

四、环境

为方便初学者使用，环境在 window 下，tensorflow 用 CPU 版

- 1、OS Windows 10 x64
- 2、Python 3.6.2
- 3、Tensorflow 1.8.0 CPU 版
- 4、Keras 2.2.0
- 5、Opencv 3.4.2

五、项目目录文件结构



注：

- 1、data_set：三个网络的训练数据；
- 2、detector：各网络的 detector，用于 p_net/r_net/o_net 预测的数据，均要实现 nms，对输入图片要实现图片金字塔；（整体作为一个 detector）
- 3、gen_data：1) gen_wideface_data：Wider_face 数据集生成正样本、负样本、部分样本三类样本；2) gen_celeba_data：CelebA 数据集生成关键点样本；（p_net/r_net/o_net 用）；3) gen_training_h5_data：将生成的四类样本数据分别生成 h5 文件；4) gen_middle_wideface_data：r_net/o_net 的训练数据正样本、负样本、部分样本，通过前面已经训练好网络进行预测生成；
- 4、log：存放检测点；
- 5、model_weight：模型训练结果
- 6、mtcnn_model：模型网络实现
- 7、test：模型测试

- 8、test_input_img: 模型测试输入图片
- 9、test_output_img: 模型测试输出图片
- 10、training: 1) load_data: 数据 generator , 模型训练时使用 fit_generator; 2) train_pub: 模型训练公共文件, 主要实现 loss, 各模型 compile; 3) train*_net: 各网络模型训练入口
- 11、config: 网络超参数, 项目参数, 各目录等
- 12、utils: iou、nms 等函数实现

六、注意事项

- 1、Loss 函数复杂, 需要仔细理解, 三部分损失对不同样本不同网络均不一 ;
- 2、r_net/o_net 的训练数据正样本、负样本、部分样本, 是通过前面已经训练好网络进行预测生成, 不是像 p_net 那样, 数据直接来自数据集;
- 3、训练数据样本分为 正样本、负样本、部分样本、关键点样本 , 它们在训练时的比例控制为 1: 3: 1: 2;

七、参考

- 1、https://kpzhang93.github.io/MTCNN_face_detection_alignment/
- 2、https://blog.csdn.net/qq_14845119/article/details/52680940
- 3、<https://blog.csdn.net/u011808673/article/details/78954654>
- 4、https://blog.csdn.net/App_12062011/article/details/78836320