

# 机器学习实验--人脸识别

成员：     兰丰    1611640226    

    张浩歌    1611640225    

    孙健    1611640302    

    姜浩东    1611640310    

    解邦旭    1611640118

# 一、简介

## 1.1 人脸识别:

人脸识别, 是基于人的脸部特征信息进行身份识别的一种生物识别技术. 用摄像机或摄像头采集含有人脸的图像或视频流, 并自动在图像中检测和跟踪人脸, 进而对检测到的人脸进行脸部识别的一系列相关技术, 通常也叫做人像识别、面部识别.

## 1.2 FaceNet:

FaceNet 是一个通用的系统, 采用 CNN 神经网络将人脸图像映射到 128 维的欧几里得空间, 我们可以根据两幅人像的余弦相似性去判断两个人像的相似程度. 两个人像之间的余弦相似性越高, 说明它们越相似.

# 二、实验环境与分工

## 2.1 实验环境:

- 机器学习框架: TensorFlow
- 编译环境: Spyder

## 2.2 实验分工:

- 兰丰: 结论分析与论文书写及汇报演讲
- 张浩歌: 环境搭建
- 孙健: 收集资料和协助
- 姜浩东: 代码调试
- 解邦旭: 作图及论文排版

# 三、工作流程

## 3.1 FaceNet 结构:



**batch**：是指输入的人脸图像样本，这里的样本是已经经过人脸检测找到人脸并裁剪到固定尺寸（例如 160x160）的图片样本。

**Deep architecture**：指的是采用一种深入学习架构

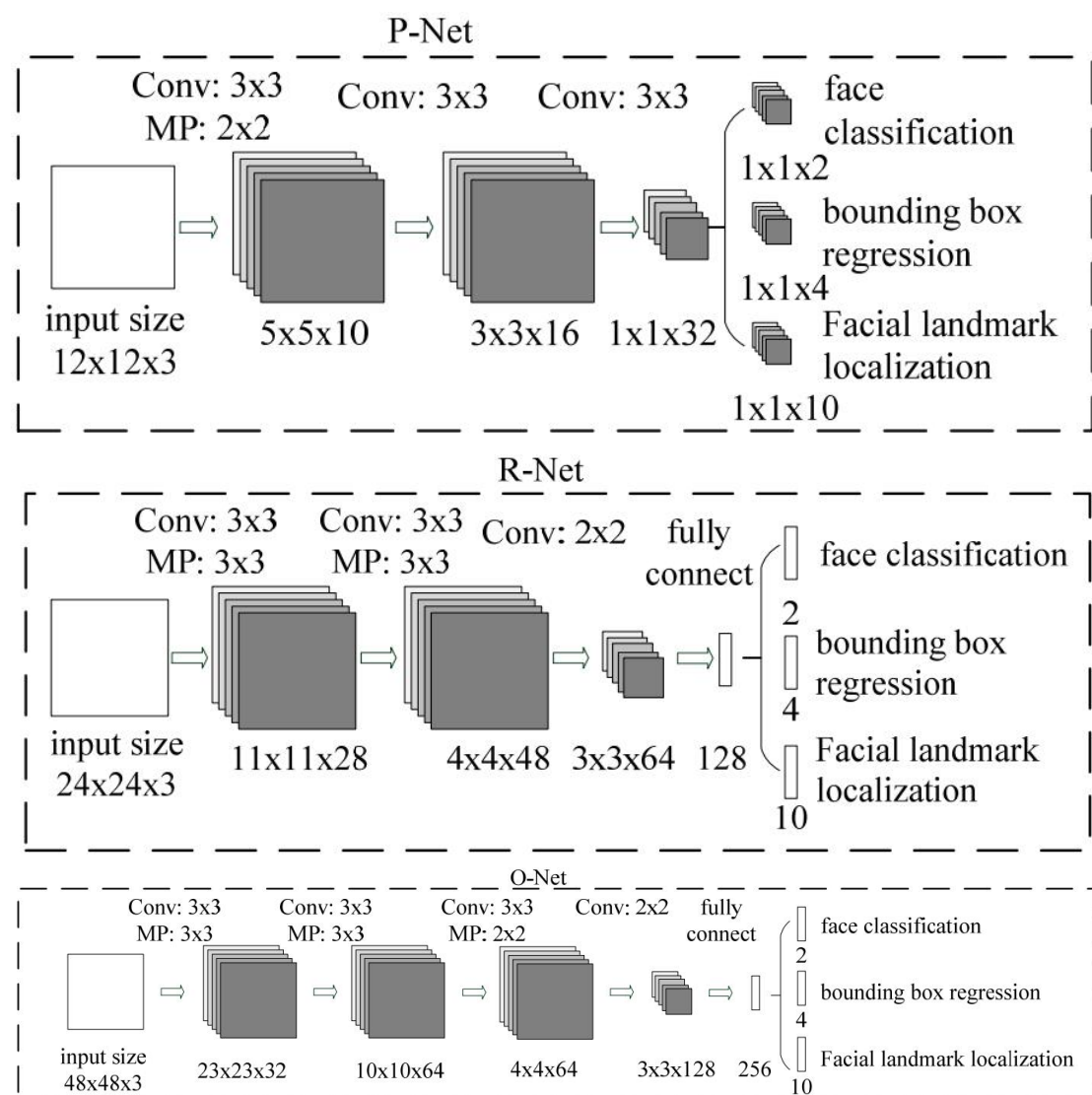
**L2**：是指特征归一化（ $\|f(x)\|_2=1$ ，这里是 2 次方的意思）

**Embeddings**：就是前面经过深度学习网络，L2 归一化后生成的特征向量（这个特征向量就代表了输入的一张样本图片）

**triplet Loss**：就是有三张图片输入的 Loss

### 3.2 图片预处理：

首先检测一张图片中的人脸并且将它进行提取出来。这里采用的是 MTCNN 神经网络进行人脸检测。



### 3.3 卷积网络选择模型 GoogLeNet

先将该图片库的每张图片训练并映射成为 128 维向量矩阵后保存成 npy. 当比对的时候只需要读取这个 npy 得到每张图片的 128 维特征向量再进行计算即可.

### 3.4 余弦相似性:

严格来讲余弦距离不是距离, 而只是相似性. 其他距离直接测量两个高维空间上的点的距离, 如果距离为 0 则两个点“相同”; 余弦的结果为在  $[0, 1]$   $[0, 1]$   $[0, 1]$  之中, 如果为 1, 只能确定两者完全相关、完全相似.

假设两用户同时对两件商品评分, 向量分别为(3,3)和(5,5), 这两位用户对两件商品的喜好其实是一样的, 余弦距离此时为 1, 欧式距离给出的解显然没有余弦值直观.

相对于标准化后的欧式距离, 余弦距离少了将数据投影到一个均值为 0 的区间里这一步骤. 对于点 X 和点 Y, 其余弦距离:

$$d = \frac{\sum_{i=1}^N x_i y_i}{\sqrt{\sum_{i=1}^N x_i^2} \sqrt{\sum_{i=1}^N y_i^2}}$$

利用余弦相似性判断人脸与库中模型训练的人脸相似度.

## 五、总结

这次实验我们实现了利用摄像头进行动态人脸检测和人脸识别系统, 该模型的优点是只需要对图片进行很少量的处理 (只需要裁剪脸部区域, 而不需要额外预处理, 比如 3d 对齐等), 练的损失函数直接针对实际误差. 在实验过程中遇到诸多困难, 如图片不能正常识别, 模型无法加载等. 在我组成员互相研讨及查阅大量资料和反复调试后逐一解决. 本次实验让我们学到很多机器学习在实践中应用的知识, 让我们更进一步的了解机器学习.

## 六、参考文献

[1] Jake VanderPlas. Python 数据科学手册. 人民邮电出版社. 2018

[2]郑泽宇.TensorFlow 实战 Google 深度学习框架. 电子工业出版社. 2017