

基于 MTCNN 和 ResNet 的人脸识别考勤应用

岑香玉 周 丽

四川大学锦城学院 计算机与软件学院 四川 成都 611731

【摘要】随着人工智能的发展,机器学习技术的不断前进,人脸识别技术也在智能识别系统当中广泛应用。在校园中考勤系统仍是传统的人工点名以及签到,这会带来时间成本以及影响上课效率。为解决这一问题,提出通过 MTCNN 算法进行人脸检测,ResNet 算法迁移学习进行人脸识别来进行课堂考勤。本实验对比了 GoogLeNet、AlexNet、VGG 算法,从对比实验中 ResNet 算法在识别人脸的准确率可达,具有较高准确率。

【关键词】MTCNN 算法; ResNet 算法; 人脸检测; 人脸识别; 迁移学习

1 引言

神经网络的应用随着科技的发展,也逐渐落在各个领域,其中在人脸检测以及识别的领域有着领先的优势。目前有各种各样的考勤系统,放眼到大学校园当中,老师往往采用的是人工点名签到,但是存在耗时性、不真实性、低效性等缺点,如果是使用指纹扫描或虹膜识别等设备,往往成本又会很高。人脸作为一种生物特征,它是唯一的、不可复制的^[1],因此人脸识别可以很

好运用到身份鉴定当中,采用神经网络算法来实现人脸识别,可以大大地降低成本以及时间,来提高考勤效率。

2 人脸检测和识别

人脸检测的过程基本分成 6 个步骤:图像采集;通过神经网络算法进行人脸检测;检测到的图片进行预处理;提取图像的人脸特征;将特征与图像采集得到的人脸进行匹配;最后得到最大可能的人脸,完成人脸识别。人脸检测以及人脸识别的流程如下图 1 所示:

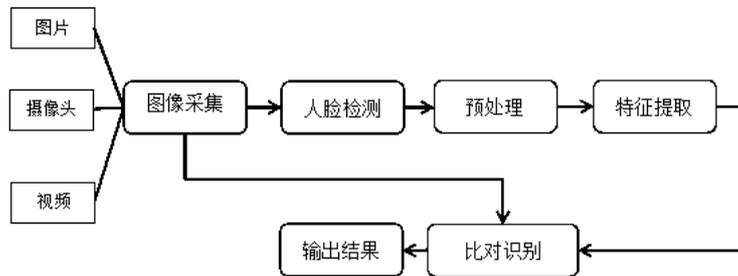


图 1 人脸识别的流程

本实验中采用 MTCNN (Multi-task convolutional neural network) 多任务卷积神经网络模型实现人脸检测,提取到特征之后由 ResNet (Residue Network) 深度残差网络算法实现人脸识别。

2.1 MTCNN 网络

MTCNN 是一种基于级联架构的多任务 CNN 网络 [2],其中 CNN 是卷积神经网络模型,CNN 神经网络类似于人的大脑识别人图像的过程,来源于生物大脑视觉中的神经网络结构,输入一张图片,通过卷积层神经网络提取到图像的特征,使用激活函数对数据做映射得到图像的

特征,并且 CNN 广泛运用在计算机视觉以及智能语音识别当中。MTCNN 是在 CNN 的基础上更深层的网络结构,可以同时做到人脸检测和人脸对齐。

2.1.1 三个级联神经网络结构

MTCNN 级联架构的多任务卷积神经网络输入的是一张图片,通过放缩生成不同大小的图像,构建出图像金字塔 [3],得到的图像传入 P-net (Proposal Network)、R-net (Refined Network)、O-net (Output Network) 得到更加精细的图片,三种级联神经网络模型如图 2 所示。

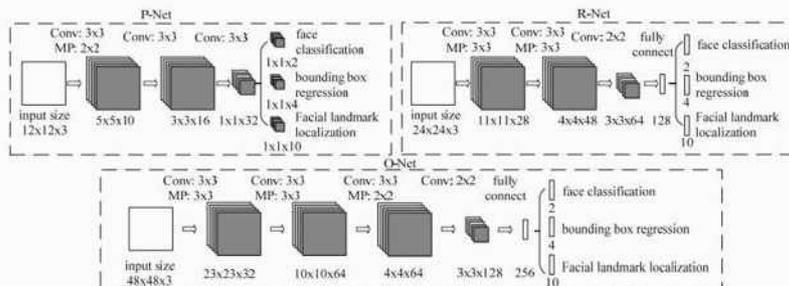


图 2 P-net, R-net, O-net 网络示意图

如图所示由P-net 网络生成人脸候选窗体, 边界回归和人脸候选框 [4] 进行对比校准, 使用非极大值抑制算法 (NMS) 去掉覆盖的边框。上一层的输出是R-net 的输入, 更进一步细化, 跟P-net 处理相同, 最后NMS 合并候选框。在O-net 做更加精细的处理。不同的是最后会输出人脸的5个关键点坐标点, 该网络中卷积通道数加多, 网络深度加深, 对于人脸的检测会更加准确。

2.1.2 MTCNN 算法核心

①人脸特征点定位以及边框回归问题: 两者都是回归问题, 人脸特征点定位问题最后得是计算经过预测的关键点以及实际点的偏差、边框回归问题是计算得到的边框与实际边框之间的偏差, 这个偏差通常用欧式距离作为回归问题的损失函数。

②人脸检测的损失函数: 人脸检测属于分类问题, 交叉熵损失函数 (Cross Entropy), 人脸检测的交叉熵损失函数如下:

③级联架构的损失函数权重问题: 在神经网络当中, P-net、R-net、O-net 损失函数都会存在, 但是并不是所有都会发挥作用, 其中的权重会在P-net、R-net、O-net 中有不同的取值侧重。

2.2 ResNet 网络

随着网络不断地加深堆叠, 意味着可以提取到的特征更加的丰富, 同时可以更好地体现图像、得到更深层次的特征, 实际上, 反而得到的模型的效果并不理想, 为了解决神经网络深度加深, 网络模型退化这一问题 [5], 由Kaiming He 提出了深度残差网络 (ResNet), 通过这个结构, 能够使得最后的准确率提高, 残差网络的基本结构如图3所示。

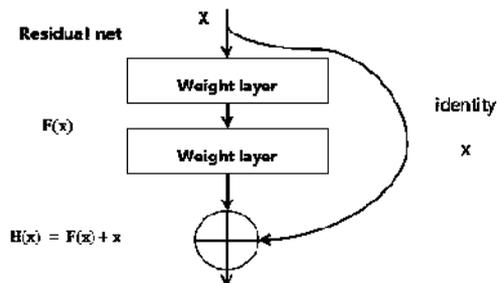


图3 深度残差网络示意图

可以看出 ResNet 并没有将每一个卷积层相互直接相连起来, 而是引入了残差单元 [6]: 残差是指预测值与观测值之间的差, 输出由两部分组成, 一种是卷积运算, 一种是将输入的作为输出, 相加得到, 保留了输入的信息。

在更深层的神经网络模型验证得到, 不是由于过拟合以及梯度下降导致了模型效果变差, 而是在更深层次的网络层中没有学习到恒等映射 [7], 加深网络的效果消失了。但是有残差网络中加入了, 就相当于加入了非恒等映射, 就可以让后面的深层网络开始进行恒等映射, 让深层的网络发挥作用, 最后能够提高模型的准确率。

3. 实验与结果分析

本实验使用 python 编程, 调用 OpenCv 对图像处

理, 以及获取摄像头写入图像, 使用由 Facebook 发布的 PyTorch 框架。图像采集的方式有很多种, 最常见的是图像上传、以及摄像头采集, 为得到更加真实的图片, 本实验通过摄像头采集人脸的图片以及上传图片的方式。

3.1 图像采集

实验使用摄像头采集到正脸, 侧脸, 抬头, 戴眼镜, 戴口罩, 歪头等图片, 以及上传图片得到数据集, 再通过图像转换处理成灰度图, 并且将数据按照 8: 2 划分为训练集以及测试集, 得到数据和标签的文件如图4所示:

| | |
|---------------------|---|
| bqs_gray/bqs_00.jpg | 1 |
| bqs_gray/bqs_01.jpg | 1 |
| bqs_gray/bqs_02.jpg | 1 |
| bqs_gray/bqs_03.jpg | 1 |
| bqs_gray/bqs_05.jpg | 1 |
| bqs_gray/bqs_06.jpg | 1 |
| cxy_gray/cxy_00.jpg | 2 |
| cxy_gray/cxy_01.jpg | 2 |
| cxy_gray/cxy_02.jpg | 2 |
| cxy_gray/cxy_03.jpg | 2 |
| cxy_gray/cxy_04.jpg | 2 |
| cxy_gray/cxy_05.jpg | 2 |
| cxy_gray/cxy_06.jpg | 2 |



图4 数据集中的图片以及对应的标签

3.2 MTCNN 人脸检测步骤

①创建人脸数据集: 通过爬取网络明星的照片集制作人脸数据库, 通过MTCNN 提取人脸 ROI 区域, 包括左眼、右眼、鼻子、左嘴角、右嘴角。

②在课堂人脸考勤应用当中加入了5个同学的人脸图像, 保持与原文件夹相同的组织结构, 组成 data.csv 文件。

③划分数据集: 划分的数据用来做训练集, 用 OpenCv 写入到 train.csv 文件中, 划分的数据做测试集写入到 test.csv 文件中。

④数据预处理: 用 transform 对图像做均值归一化处理得到特征向量, 通过MTCNN 模型检测人脸, 结果如图5。

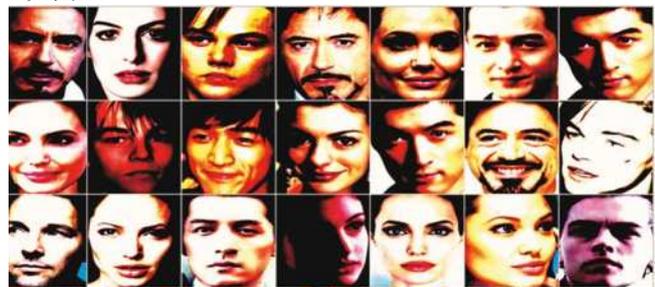


图5 MTCNN 检测效果图片

如图3所示, 在图像中人物有表情以及侧脸部分, MTCNN 都能够很好地检测到人脸的特征。

3.3 ResNet 模型训练

本实验所使用的人脸识别的神经网络是 ResNet-18 进行迁移学习, 损失函数为交叉熵, 优化器为随机梯度下降, 学习率为 0.1, 权重衰减为 0.05, batch_size 为 37, 使用 Softmax 做分类。并且使用 ResNet、GoogLeNet、AlexNet、VGG 做了对比实验。

3.3.1 迁移学习

深度学习中的迁移学习是指已经很好地解决了一个领域的问题，可以将学到的参数迁移到与之相似的问题上，从而可以快速解决问题，而不需要重新训练一个模型，即有举一反三的能力。由于图像结构是相似的，可以使用相同的参数训练，因此常应用到计算机视觉中，本实验在 ResNet-18 模型上对人脸数据集进行了迁移学习，很大程度上节省了训练的时间。

本实验通过先冻结预训练模型的卷积层，修改最后的全连接层的神经元数量为本实验的人脸类别数，之后再使用微调后的 ResNet-18 模型做人脸识别。

| 网络模型 | 准确率/% |
|-----------|-------|
| ResNet | 90.62 |
| GoogLeNet | 84.83 |
| AlexNet | 70.00 |
| VGG | 68.75 |

表 1 四种常见网络测试的结果

从表 1 当中也看出，ResNet 模型在测试集上识别出人脸的准确率高于其他的模型，能够有效地识别到人脸。识别的效果图如图 6。

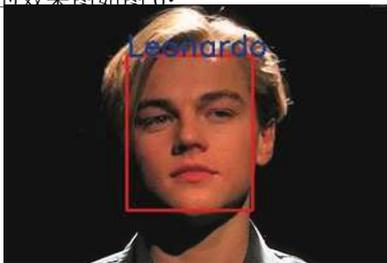


图 6 ResNet 人脸识别出 Leonardo

使用 ResNet 迁移学习模型来识别人脸图片，如图所示可以很好地识别出 Leonardo，人脸的识别精度较高。

4 结束语

本文在校园中考勤系统仍是传统的人工点名以及签到，为减少时间成本以及降低点名带来的影响，基

于 MTCNN 和 ResNet 算法实现了人脸检测和人脸识别。MTCNN 通过级联架构的神经网络提取到人脸 ROI 区域，在对比实验可以得到 ResNet 优于其他的人脸识别算法，可以很好地应用到课堂的考勤系统当中。人脸识别的应用广泛，可扩展到其他的应用场景中，通过建立较大的数据库，优化神经网络，提高人脸识别精度，从而应用到智能识别系统中。

【参考文献】

[1] 贾小硕, 曾上游, 潘兵, 周悦. 基于改进 MTCNN 网络的目标人脸快速检测 [J]. 计算机工程与科学, 2020, 42(07):1262-1266.

[2] J. Xiang and G. Zhu, "Joint Face Detection and Facial Expression Recognition with MTCNN," 2017 4th International Conference on Information Science and Control Engineering (ICISCE), 2017, pp. 424-427, doi: 10.1109/ICISCE.2017.95.

[3] 王灵珍, 赖惠成. 基于多任务级联 CNN 与中心损失的人脸识别 [J]. 计算机仿真, 2020, 37(08):398-403.

[4] 陈凯, 祖莉, 欧屹. 基于 YOLOv3 与 ResNet50 的摄影机器人人脸识别跟踪系统 [J]. 计算机与现代化, 2020(04):30-36+41.

[5] 何伟鑫, 邓建球, 刘爱东, 丛林虎. MTCNN 和 RESNET 的人脸识别弹库门禁系统研究 [J]. 单片机与嵌入式系统应用, 2020, 20(04):51-54.

[6] 李博健. 基于 Resnet 的人脸识别系统的实现 [D]. 哈尔滨理工大学, 2020.

[7] 张翠芳, 金天恺, 苏博艺, 龙芳, 呼亚楠, 谢兰欣. 基于残差神经网络的人脸识别及仿真平台 [J]. 徐州工程学院学报 (自然科学版), 2019, 34(01):33-37.