

基于深度学习的人脸识别课堂考勤应用

陈泓池 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】日常课堂的考勤是学校课堂管理中不可或缺的组成部分之一，在日常生活中常用的课堂考勤方式存在着大大小小的许多问题，如考勤数据的统计速度缓慢、代签等。文章基于深度神经网络 MTCNN 和 ResNet 实现了人脸识别，并将其应用于课堂考勤，抛弃了过去使用人力进行统计管理的后进形式，制胜了考勤中的不规范行为，解决学校以往考勤管理工作中存在的大量问题，为学校有效施行考勤制度提供了科学的依据，极大地节省了课堂考勤所占用的时间。

【关键词】人脸识别；深度学习；课堂考勤

1 引言

1.1 研究背景与意义

目前，高校学生的教育管理依旧是通过课堂考勤的形式来实现的，而在很多大学，学生上课不出勤的现象十分严重。教师利用上课时间进行课堂点名或小程序签到为现有大学考勤管理的主体形式，如此一来消耗人力、时间，代签等情况就会无法杜绝，无法利用考勤管理来对学生起到有效监督作用。而在当前社会，人脸识别技术发展得十分迅速，基于人脸识别的考勤系统具有效率高、精度高等特点^[1]，可快速反映出学生该门课程是否出勤，更加便捷、透明地进行考勤管理，而且人脸识别技术的实现是可以直接通过普通摄像头进行，无需其他高成本设备，这种低成本的方式能够被广大用户接受，迅速普及到各大高校，帮助学校加强考勤管理。

1.2 人脸识别研究现状

人脸识别^[2]是生物特征识别技术的一种，它使用相关人类面部的特征信息来执行身份识别，它是数字图像处理、计算机和模式识别等学科交叉的产物。

刚开始面部识别主要是凭借提取面部特征点和它们之间的拓扑关系来进行识别，此用法容易施行并且一目了然。1991年提出“特征脸”^[3]，首次把人脸识别与统计特征技术、主成分分析结合在一起，致使人脸识别效果取得了特大的进步。2000年后随着机器学习算

法的大范围推广，研究人员开始将遗传算法、boosting 和 SVM 等方法用于人脸识别。而后由于计算机硬件性能在科技的发展中得到了极大提高，使得深度学习技术^[4]发展十分快速，在大规模数据训练的神经网络基础上，图像分类、手写识别、语音识别等应用获得了极为显著的效果。

2 人脸检测及识别技术的研究

2.1 MTCNN 人脸检测算法

MTCNN^[5]是一种多任务级联形式的卷积神经网络，此网络采用了级联的方式，与 HAAR 级联检测^[6]在某些地方有一定的类似之处，都于初期便开始筛选并拒绝一大部分的图像区域，使得进入后期 CNN 网络阶段时，其计算量与时间得到了较大程度的降低。它包含了 P-Net、R-Net、O-Net 三个级联的多任务卷积神经网络，而其中任意一个卷积神经网络均有人脸分类、边框回归和关键点定位的学习任务^[7]。

但是三个阶段的侧重点各有不同：

第一阶段的网络模型称为推荐网络 P-Net，该网络将有机会出现人类面部的窗口与边界 Box 回归大略检测出，再依据回归结果对得取的窗口实施校正，然后将其中出现较高程度重合的窗口使用非最大压制 (NMS) 兼并掉，其网络结构如图 1 所示。

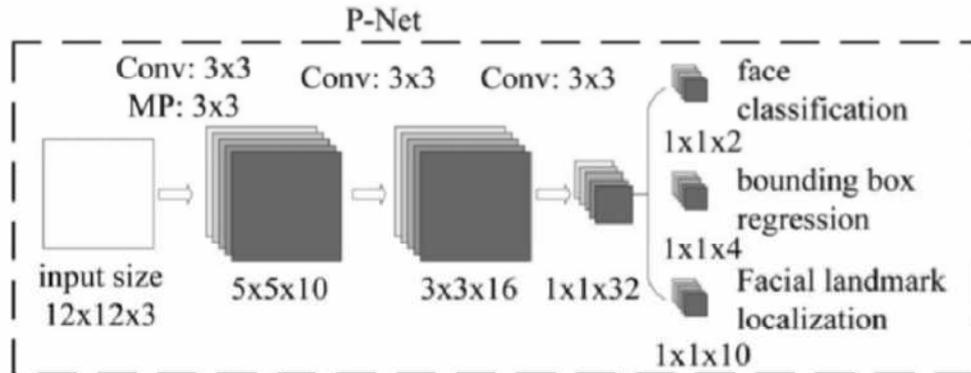


图 1 P-Net 结构图

第二阶段的网络模型称为优化网络 R-Net，利用一个更有能力的 CNN 网络使得大部分不是人脸的待选框被

消除掉，而后将边框回归的结果继续进行调整，应用 NMS 让重叠的窗口被合并掉，其网络结构如图 2 所示。

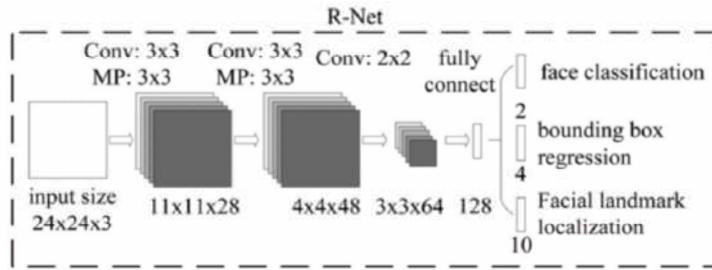


图 2 R-Net 结构图

第三阶段的网络模型称为输出网络 O-Net, 将 R-Net 的输出数据作为输入因而从事愈加深入的提取, 通过一

个效能更好的监督来识别人类相貌的区域, 而且会对相貌区域的特征点位置进行回归预测, 最后输出人脸上面的五个标记点, 其网络结构如图 3 所示。

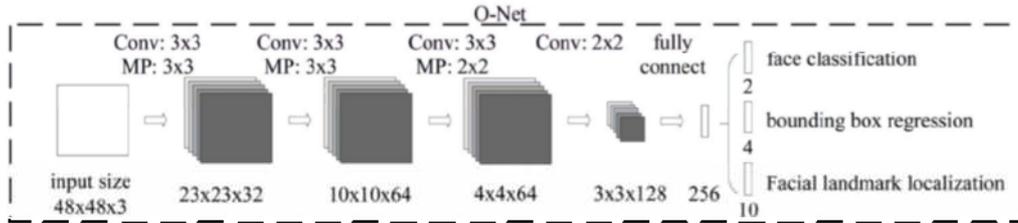


图 3 O-Net 结构图

以上就是 MTCNN 算法的基本结构。在整个的人脸检测过程中, 由 P-Net 至 R-Net, 再至最后的 O-Net, 鉴于逐渐增添的输入图像, 逐渐添多的卷积层的通道数, 逐渐变深的神经网络, 人脸识别的精度也逐渐变高, 但是网络参数也因此变得越来越多。O-Net 网络运行的效果最好那么为何不从一开始就直接使用该网络而是使用这三个网络呢。这是因为要是一开始就直接使用 O-Net 网络对图像进行识别, 其运行速度会非常慢, 得到结果的时间也会变得十分漫长。而先使用 P-Net 过滤掉一部分非人脸区域, 再传给 R-Net 进行二次过滤, 最后由 O-Net 识别过滤后的结果。通过多次过滤的方式, 一步步减少了最后需要辨别的数量, 使计算时间得到有效的降低。

问题在网络层中得到有效抑制, 得到了对于图像更抽象化的特征表达。

2.2 深度残差网络人脸识别算法

深度残差网络 (ResNet)^[8] 是于 2015 年被提出的一种深度卷积网络, 出现之时便夺得了图像分类、检测、定位三项的魁首^[9]。它拥有深度卷积神经网络本身的自适配特征提取特点, 同时解决了由于网络深度加添而出现的网络性能退化的矛盾。

ResNet 网络采取了如图 4 所示的一种短路的形式, 通过输入跳跃连接到输出, 实现了对输入内容的重用, 如果网络的输入值设为 x 、某一网络层设为 H , 则该层的输出为 $H(x)$, $H(x)+x$ 则作为下一层的输入。此过程能够让因为网络层数增加而引起的梯度扩散与梯度爆炸

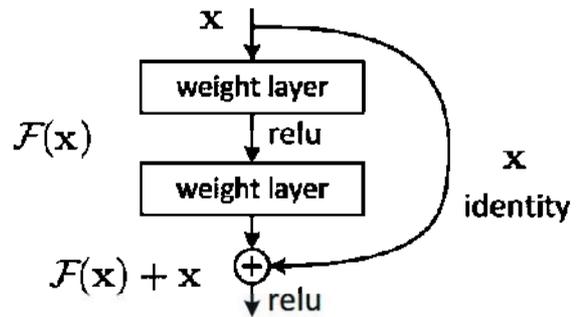


图 4 ResNet 网络形式

3 系统实现

3.1 人脸检测

人脸检测是人脸识别与处理的基础, 其目的是在视频流中检测和定位出人脸图像, 并返回人脸边界框和人脸特征点的坐标 [10]。文章将目前被广泛使用的 MTCNN 网络作为检测模型, 用以检测和对齐人脸关键点。

MTCNN 人脸检测流程:

- 迅速产生待选框。
- (2) 过滤待选框, 重叠窗体合并。
- (3) 达成窗体去留, 面部关键点输出。

具体流程图如图 5 所示。

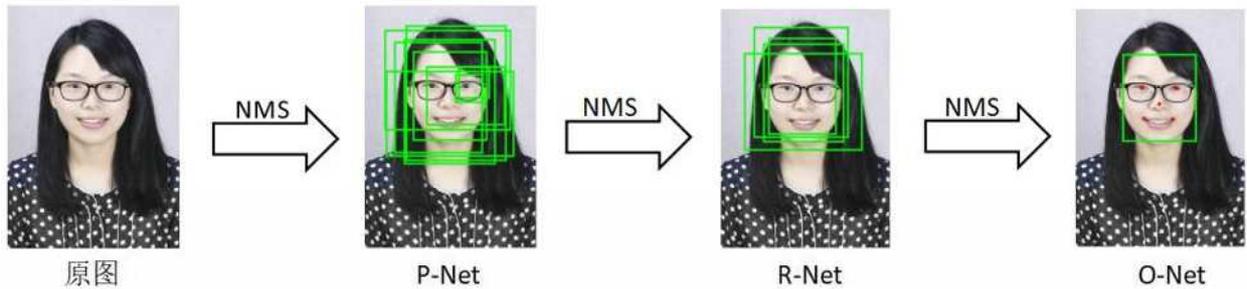


图5 MTCNN神经网络训练流程图

3.2 人脸识别

在从视频流中检测到人脸后，需要为其加上标签，即依照人脸识别出该名学生的姓名以及其他相关信息。

人脸识别具体流程：

加载人脸识别模型

计算特征向量

计算欧氏距离，它是指两个向量于空间中的直线距离。两向量之间的欧氏距离越小则其越相近，不同处越少。若两个向量分别为 $x(x_1, x_2, \dots, x_m)$ 、 $y(y_1, y_2, \dots, y_m)$ ，则它们之间欧氏距离的计算公式如公式 1。

(公式 1)

识别人脸，两个向量的欧氏距离越小，则其越相近的结论放在人脸识别中，可得出不同人脸特征向量的欧氏距离越小，二者越相似的结论。若距离小于某个值，则可判定为同一个人，以此达成人脸识别的判别。

```

1.66237190e-01, 2.42377426e-02, -2.61276402e-02, 2.16263369e 01,
1.79047301e-01, 1.06509954e-01, -6.94652051e 02, -2.40043849e-02,
1.92848854e-02, -7.76724517e-03, -2.77894381e-02, 7.16376344e-02,
7.04587847e 02, 3.37971359e 01, 1.14788155e-01, 3.23747359e 02,
-2.01553810e-02, 3.57827032e-03, -6.88013507e-02, 5.15632257e-02,
-2.89753605e-01, -9.39028114e-02, 2.12965626e-02, 1.23104863e-01,
-4.96431701e-02, -1.29149228e-01, 1.23441298e-01, -6.47879079e-02,
-2.59986947e-01, -1.22779766e-02, 1.14170514e-01, 2.25154549e-01,
2.27058575e-01, -8.69142413e-02, -4.28985804e-03, -9.53668132e-02,
1.58290684e-01, -1.53770819e-01, -9.52070765e-03, 7.95241445e 02,
5.13708703e-02, -9.97092500e-02, 8.01974386e 02, -6.80193231e-02,
1.17156416e-01, 1.98554953e-01, -1.05778746e-01, -1.21705560e-02,
1.21431649e-01, -9.37339276e-02, 3.03095744e-02, -6.42748326e-02,
2.10248798e-01, 6.84872866e-02, -8.35196674e-02, -2.06961483e-01,
    
```

图6 人脸特征向量

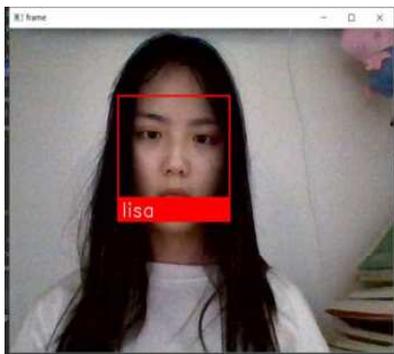


图7 人脸识别结果

3.3 考勤数据库

在完成了核心功能人脸检测和人脸识别之后，剩下的工作主要为如何实现自动化考勤。

考勤数据库具体实现形式：

创建一张 excel 表格，在其中录入班级所有学生的姓名。

利用前面研究的人脸检测、识别算法将视频流中的人脸检测识别出来并与其姓名相对应。

检测到学生后在运行结果栏打印出检测到的学生姓名、人数以及当时的时间，结果栏显示结果如图 8 所示。

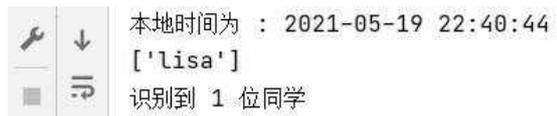


图8 结果栏显示结果

在 excel 表格中录入检测到该名学生的时间以确定其当日是否出勤，考勤数据表结果如图 9 所示。最终根据考勤表中的数据即可直接反映当堂课学生的出勤状况。

| | A | B | C |
|---|--------|---|--------------------|
| 1 | rose | 0 | |
| 2 | lisa | 1 | 2021-5-19 22:40:44 |
| 3 | jisoo | 2 | |
| 4 | jennie | 3 | |

图9 考勤数据表

4. 总结与展望

4.1 本文总结

目前在课堂考勤方面，人工点名仍是主流，既占用课堂时间又不能解决代签的问题，也有一些课堂考勤采用了物联网技术的签到系统，但是依旧不能很好地解决考勤问题。文章在绪论部分将本研究的背景意义、国内外自动化考勤现状和人脸识别技术的研究现状做了系统介绍。第二章是人脸检测与识别技术有关算法的简单介绍，人脸检测部分，解释了深度学习的 MTCNN 算法的大致结构。人脸识别方面，讲述了深度残差网络的大致形式。第三章介绍了人脸检测、人脸识别以及考勤打卡三种功能的实现步骤和实验结果图片。本文的人脸检测与识别模型利用了 MTCNN 和 ResNet 算法，对于摄像头中的视频流数据则用 OpenCV 来处理，实现了对学生信息的管理，促进人脸检测和识别功能的实现，提高课堂上考勤效率。

4.2 研究展望

本文采用 dilb 库中现有的 ResNet 模型接口实现

了一个基于监控视频流的课堂自动化考勤系统,而人脸识别考勤的应用场景丰富多样,本文考虑的只是各方面条件都较适合进行人脸识别的应用场景。对人脸检测与识别功能产生影响的元素有很多,核心问题是当光照条件不好时,人脸检测的效果会大打折扣。除此之外,课堂考勤时可能会存在众多的小区域人脸并通常伴随遮挡的情况。人脸检测在这种应用场景中还需要继续研究。人面部的表情变化是影响识别准确率的一个重要因素。有些时候由于面部表情的剧烈波动,面部特征的位置发生变化,再加上环境因素的干扰会对识别系统带来极大的干扰,因此多种人脸特征融合可以有效的整合人脸的各种信息,有助于开发一个更加稳定的识别系统。

【参考文献】

- [1] 赵志强,袁恩昌,张智雄. 基于dlib库的人脸识别考勤系统设计与实现[J]. 无线互联科技,2020,17(20):72-74.
- [2] 简讯. 人脸识别技术综述[J]. 电脑知识与技术,2019,15(02):171-172.
- [3] Sirovich L, Kirby M. Low-Dimensional Procedure for the Characterization of Human Faces[J]. Journal of the Optical Society of America. A, Optics and image science, 1987, 4(3):519-524
- [4] 郑远攀,李广阳,李晔. 深度学习在图像识别中的应用研究综述[J]. 计算机工程与应用,2019,55(12):20-36.
- [5] Zhang K, Zhang Z, Li Z, et al. JointFaceDetection and Alignment Using Multitask Cascaded Convolutional Networks[J]. IEEE Signal Processing Letters, 2016, 23(10):1499-1503.
- [6] Lienhart R, Maydt J. An Extended Set of Haar-like Features for Rapid Object Detection[C]. Image Processing International Conference on IEEE, 2002
- [6] 王慧. 基于人脸识别的智能教育管理平台的设计与实现[D]. 北京交通大学,2019.
- [8] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. 2015
- [9] 邹鑫. 基于卷积神经网络的图像分类算法研究[D]. 西北师范大学,2018.
- [10] 徐建峰,孙浩,陆萍,董虎胜. 基于深度学习的人脸识别会议考勤系统开发[J]. 现代计算机,2020(13):86-89.