

基于人脸检测的大学课堂关注度研究*

唐康, 先强, 李明勇

(重庆师范大学 计算机与信息科学学院, 重庆 401331)

摘要:【目的】利用深度学习的人脸检测技术为课堂教学评价提供新的方案。【方法】构建适用于课堂教学的人脸检测级联卷积神经网络模型,并进行相应优化,提出了基于统计人脸检测的抬头率来量化课堂关注度的方法。【结果】通过检测和统计课堂视频中的人脸,计算出学生的抬头率,统计出学生的课堂专注度以及时间分布,帮助教师及时准确地了解课堂教学情况。【结论】通过大量的测试和优化,该系统在人脸检测中具有较好的有效性和可靠性,可以为学生提供个性化教学,同时为教师提升课堂教学质量和教学效率提供参考。

关键词:深度学习;人脸检测;抬头率;课堂专注度;教学评价

中图分类号:TP391.4

文献标志码:A

文章编号:1672-6693(2019)05-0123-07

课堂教学是衡量教师教学质量的主要内容。学生作为课堂学习的主体,他们的课堂专注度可以评价学生是否有效地参与课堂学习,这是实现教师有效教学的一个重要的因素。教师希望在课堂教学过程中能够实时、全面、准确地掌握学生的学习情况,以便调整教学进度和内容,达到更加有效的教学效果。然而,目前教师普遍通过课堂观察以及提问等方式来实时了解学生的课堂学习情况,这会造成课堂信息传递与反馈的滞后性和片面性。并且,随着智能手机等电子设备的普及,目前的课堂教学过程中,出现了大批“低头族”。因此,通过统计“抬头率”可以在一定程度上判断学生的课堂专注度,从而有助于有效地提升课堂教学效率^[1]。

人脸检测是人脸识别的前提和组成部分,也是计算机视觉技术的重要分支。传统的人脸检测技术受限于算法和设备条件,实时性和准确率都无法满足现实场景的需求,且在光线暗、有遮挡等不利条件下表现不佳。近些年来,随着机器学习、深度学习在图像和计算机视觉领域的进一步发展和应用,人脸检测技术取得了显著的进步。基于深度学习的人脸检测技术具有精度高、鲁棒性强等特点,因此在教室等较大场景中也能表现出优秀的时效性和准确性。

本文基于高校的教室场景,利用深度学习建立人脸检测系统,通过实时检测和统计课堂上学生的人脸数,计算出“抬头率”等课堂特征,并对学生的课堂专注度进行分析,以期评估学生听课效率和改善课堂教学质量提供科学的解决方案。

1 相关研究

现代人脸检测技术主要分为两类:基于手工特征的人脸检测和基于深度学习的人脸检测^[2-9]。基于手工特征的人脸检测技术的主要代表是 Viola 和 Jones 在 2001 年提出的“V-J 算法”^[2],该算法利用 Haar-like 特征和 Adaboost 算法,将“级联”的概念引入到人脸检测方法中,保证了人脸检测的实时性和较好的准确性,在当时取得了巨大的突破。2012 年,以 AlexNet^[3]为代表的卷积神经网络(CNN)方法被广泛应用在目标检测领域,精度取得了显著提升。2015 年,FacenessNet^[4]结合人脸特征及目标检测方法,将深度学习应用在人脸检测中,在人脸数据集 Fddb 中创下新的精度记录。2016 年,Zhang 等人^[5]将多任务学习、级联与 CNN 相结合,将人脸检测和人脸对齐技术融合,提出了 MTCNN,使用了三级网络结构,在保证检测速度的同时,将精度再次提高。2018 年初,Zhang 等人^[6]提出 FaceBoxes,结合了目标检测中 Faster-RCNN^[7]和 SSD^[8]的方法,提出了基于两个不同量

* 收稿日期:2018-12-20 修回日期:2019-08-30 网络出版时间:2019-09-26 11:24

资助项目:国家自然科学基金(No. 11471063);重庆市教育委员会科技项目(No. KJ1600322;No. KJQN201900520);重庆市自然科学基金(No. cstc2014jcyjA40011)

第一作者简介:唐康,男,研究方向为深度学习与人脸检测,Email:632649396@qq.com;通信作者:先强,男,讲师,Email:113618998@qq.com

网络出版地址: <http://kns.cnki.net/kcms/detail/50.1165.N.20190926.1123.016.html>

级的 CNN 的人脸检测算法,在保证精度的前提下,实现了在 CPU 环境的模型训练和检测过程。

计算机、互联网技术的进步以及图像处理、模式识别等理论的发展,推动了智能视频监控系统的应⤿用^[10]。目前的智能视频监控系统中应用的人脸检测技术大多使用传统检测技术,广泛应用在银行、交通运输等行业。而教育行业的人脸检测和人脸识别等技术主要应用于远程教育、在线课堂等领域^[11]。例如:冯满堂等人^[12]在 2011 年的工作中将表情识别技术带入远程课堂中,首次实现了网络环境下的表情识别和感情判断系统,提高了远程教学的效率和用户满意度。程萌萌等人^[13]构建了包括教师、学生、课程、情感等 4 部分的智能教学系统结构模型,创新性地利用表情识别与视线跟踪技术实现情感模块的反馈,为网络学习中情感反馈机制的发展提供技术支持。孙波等人^[14]使用图像特征提取技术将普遍的人脸表情特征和个人面部特征相分离,排除了个人因素对普通表情识别效果的干扰,提高了表情识别的准确性,并将此应用在三维虚拟学习平台 Magic Learning 的师生情感交互子系统上,成功地实现了基面部表情的学习者情感识别及情感干预功能。詹泽慧等人^[15]结合 Facial recognition 和眼球追踪技术,开发了远程学习者的情感与认知识别模型,该模型很大地提高了网络环境下学习者状态判断的准确率,并且添加了模型对学习者的情感和认知支持。侯立刚等人^[16]利用传统的 V-J 人脸检测方法建立了基于学生抬头率的课堂评估系统。韩丽等人^[17]通过分析学生的课堂状态,总结出了倾听、理解、疑惑、抗拒和不屑这 5 种学生听课状态,并对每种状态对应的面部表情进行具体分析,初步建立了基于面部表情的课堂评估标准。

Kevin 和 Joyc^[18]的研究中对课堂专注做了如下定义:上课注意听课,当个案正在被记录时要眼睛看老师或相关的学习材料(如课本、练习题),参与到课堂的学习中,具体表现为:1) 一直坐在椅子上;2) 保持安静;3) 注意看老师或学习资料。本文在已有的研究基础上,针对教室场景和教学课堂环境对 MTCNN 和 FaceBoxes 进行了相应的优化和融合,提出了一种基于深度学习的人脸检测方法。通过对人脸检测结果的处理和分析得到班级和学生个人的抬头听课情况,从而提出了一种分析学生专注程度、评价课堂教学质量的方法。

2 系统设计

基于深度学习的人脸检测方法,通过视频监控课堂教学过程中学生的面部情况,获得学生的“抬头率”等数据。然后融合 MTCNN 和 FaceBoxes 来构建人脸检测的 CNN 模型,使用网络人脸数据集对面模型进行训练和调试参数;将本地的课堂教学视频输入模型,对模型进行测试和优化,使模型对本地数据满足精度和实时性要求。最后,通过统计学生个人和课堂整体的学生“抬头率”数据,设计出合理的课堂教学评价方案,帮助教师全面准确的掌握课堂教学情况。

本系统的人脸检测算法部分使用 Python 作为主要的编程语言,使用了 OpenCV-Python、PIL、Matplotlib 等作为视频和图片的读取、处理和输出工具,使用 Numpy 等作为系统中的数据处理工具,采用 Facebook 团队开发的深度学习框架 Pytorch 0.4.0 搭建和训练本系统中的人脸检测 CNN 模型。系统的实现流程如图 1 所示。

通过使用 OpenCV 自带的 Haar-like 人脸检测模块以及网络上开源的人脸检测 CNN 模型,对教室多人场景的图片进行测试,发现两者分别存在检测精度过低和速度过于缓慢的问题。因此,本文借鉴了 MTCNN 和 FaceBoxes 来进行人脸检测模型搭建,并针对教室课堂场景进行了优化和改进。

2.1 CNN 人脸检测模型

为了权衡人脸检测的精度和实时性,本文采用了将不同尺度和规模的 3 个 CNN(CNN1, CNN2, CNN3)级联的方式组成最终的人脸检测模型。CNN1 至 CNN3 的网络尺度和规模逐步加大, CNN1 迅速检测出大量候选的人脸框和置信度,通过 CNN2 使用非极大抑制方法(Non-maximum suppression, NMS)进行初步筛选^[19],筛选后的结果再通过 CNN3 进行最终的筛选和边框回归(Bounding box regression)^[20],从而标定出最终的人脸框。本文的级联 CNN 人脸检测模型和人脸检测过程如图 2、图 3 所示。

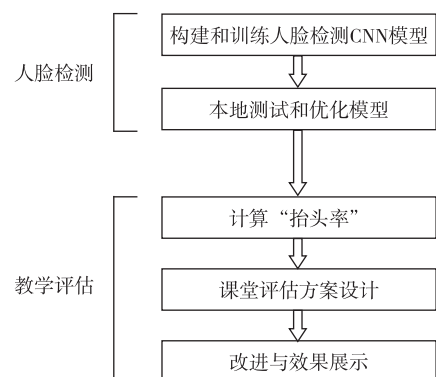


图 1 系统实现流程

Fig. 1 System implementation process

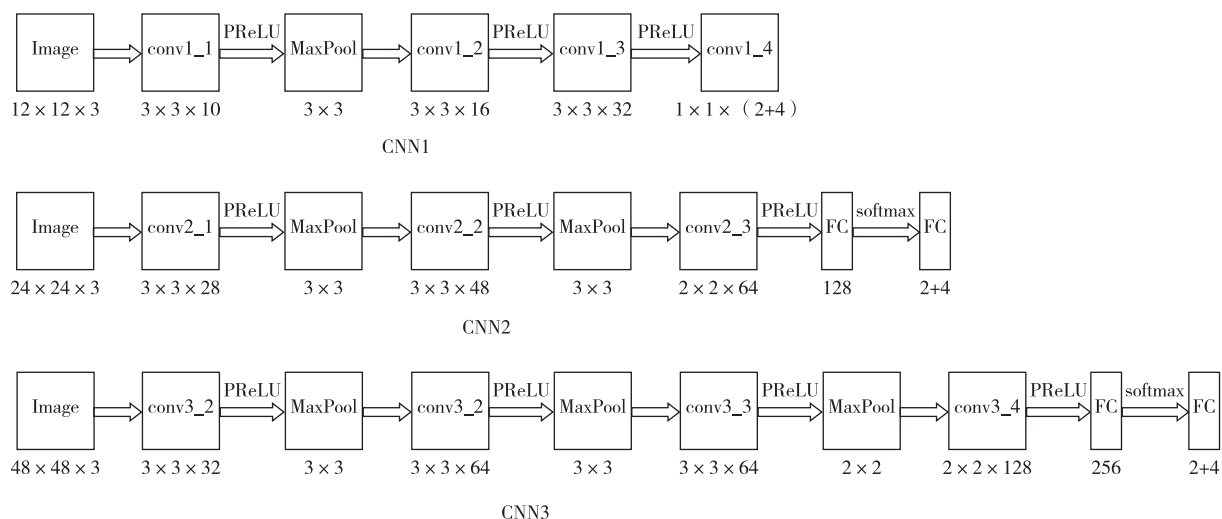


图 2 人脸检测网络模型

Fig. 2 Face detection network model

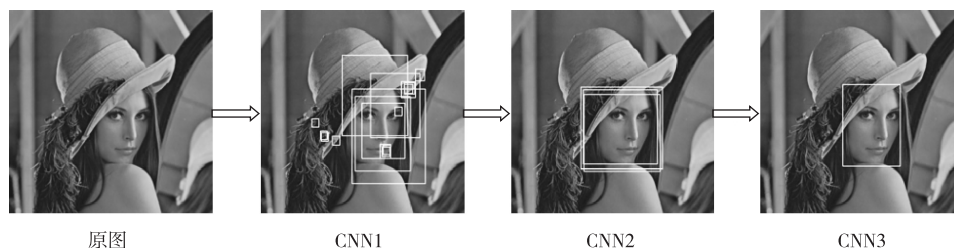


图 3 人脸检测过程

Fig. 3 Face detection process diagram

本文使用网络下载的开源有标记的人脸数据集 FDDB^[21]对 3 个 CNN 网络进行训练。FDDB 训练集为 12×12 的 RGB 图片,输入 CNN1 进行训练,输出尺度为 1×1×2 和 1×1×4 的多维数组,分别表示人脸分类信息(置信度)和定位信息;将 CNN1 得到的人脸定位框输出为图片,尺度变换后输入 CNN2 进行处理,输出经过筛选的人脸分类信息和定位信息;CNN3 的处理过程和 CNN2 类似,提高了分类和定位的精度,并加入了边框回归,输出最终的检测结果。其中,为了减少 CNN1 的网络参数,使用了带有(2+4)个 1×1 卷积核的卷积层代替了全连接层作为输出层,并对分类信息使用 softmax 作为激活函数。

对于人脸分类信息,使用交叉熵方程作为损失函数:

$$L^{class} = \sum_i (-y_i^{class} \ln(\hat{y}_i^{class}) + (1 - y_i^{class})(1 - \ln(\hat{y}_i^{class})))$$

对于人脸的定位信息,使用欧式距离方程作为损失函数 $L^{locate} = \sum_i \|\hat{y}_i^{box} - y_i^{box}\|_2^2$,其中: \hat{y}_i^{class} 和 \hat{y}_i^{box} 分别表示模型输出的分类信息和定位信息数据, y_i^{class} 和 y_i^{box} 是对应的标记数据(Ground truth)。对两个损失函数使用随机梯度下降(Stochastic gradient descent)求取最优解,得到 3 个 CNN 的网络参数。

2.2 模型优化

使用本地拍摄的课堂图片,对训练好的人脸检测模型进行测试。在使用较大场景、人数较多的图片进行测试时,人脸检测模型主要存在漏检率高、速度慢等问题。针对这些问题,做出了以下调整和优化:

1) 图片预处理:根据具体场景,生成相应的图像金字塔(Image pyramid),大量减少了输入 CNN 模型的数据量。一般地,基于 CNN 的人脸检测方法需要对待检测侧图片进行大量的尺度变换,构成图像金字塔,以检测出不同大小的人脸。在课堂场景中,学生位置相对固定,人脸尺度差距不大,根据实验设备拍摄的视频中学生的人脸尺度,生成了较小规模的图像金字塔,加快了模型的运行速度。图像金字塔的生成公式为 $I = p_{size} \times \frac{d_{size}}{f_{size_{min}}} \times \eta^k$,其中: I 表示 Image pyramid 的尺度; p_{size} 为图片尺度; d_{size} 表示检测尺度,本模型取 12; $f_{size_{max}}$ 和 $f_{size_{min}}$ 分别表示

最大人脸和最小人脸,通常为 20; η 为缩放率,本模型取 $\frac{\sqrt{2}}{2}$; $k=0,1,2,\dots$ 表示缩放次数。缩放次数由最大人脸决定,当 $I \leq \frac{p_{\text{size}}}{f_{\text{size}_{\text{max}}}}$ 时不再缩放,图像金字塔生成完成。

2) 超参数调试:经过大量的实验,对迭代次数、学习率、筛选阈值等超参数进行多次调整和尝试,重新训练模型并在验证集上测试。最终,将通过筛选的人脸偏移角度范围大致确定在正负 30° 之间,以获得“有效听课”人脸。

3) 在本地的课堂照片中截取人脸图像进行标记,作为训练集对 CNN3 进行迁移训练^[22]。本研究对本地课堂照片进行了随机分割、镜像、改变色调等预处理,使用工具对有效人脸进行标记。通过迁移学习,在 CNN3 现有参数的基础上进行再次优化训练,增强了人脸检测模型的适应性。

通过优化,人脸检测模型的漏检率显著降低,准确率和检测速度提高,在测试图片上的优化效果见图 4。

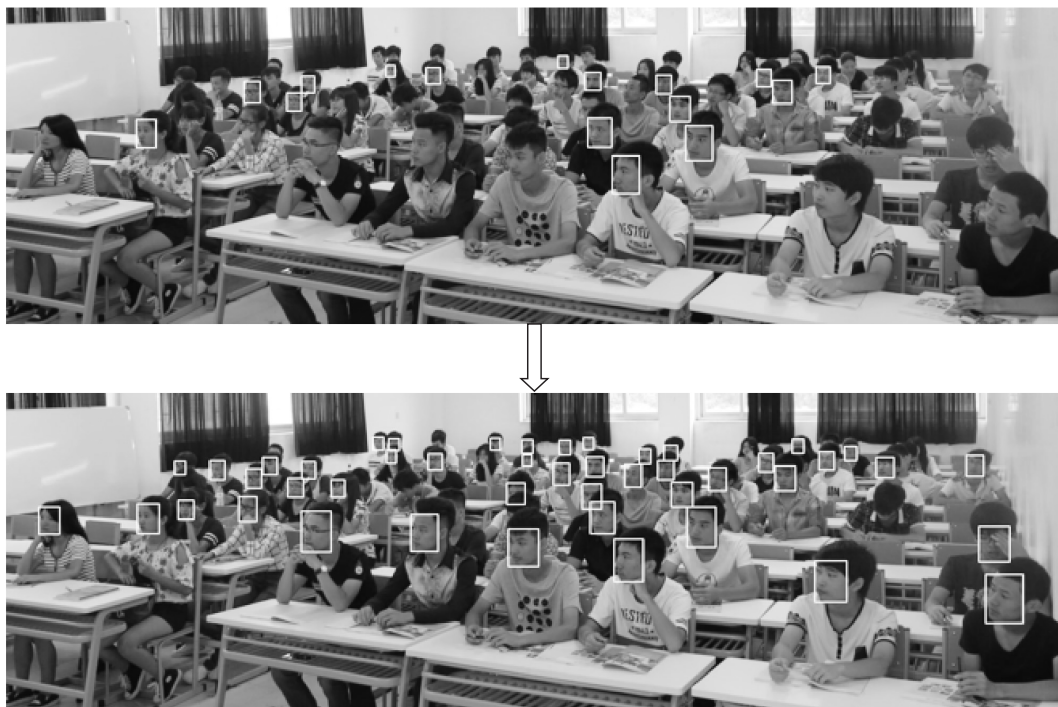


图 4 模型优化前后的人脸检测效果对比

Fig. 4 Comparison of face detection results before and after model optimization

本研究使用了 100 张本地的课堂场景图片对优化后的人脸检测模型进行了测试,并且通过人为对比和统计,得到在人数大于 40 人的教室场景中,模型在每一张图片中的误检数(False positives)为 $\sigma \leq 2$ 。在这 100 张图片中的平均查全率(Average recall ration)为 $\bar{x}_{\text{recall}} = 94.6\%$ 。因此,经优化后的 CNN 模型对于大型教室场景中的人脸检测具有相当的准确性和可靠性,能够基本满足实验要求。

2.3 课堂专注度分析

在课堂上,每隔一定时间间隔对学生进行一次人脸检测,记录每一次检测到的人脸数量和位置信息,通过统计方法计算出课堂上学生的抬头率,并根据学生个人和班级整体的抬头率数据分析学生的课堂专注度情况。

2.3.1 个人抬头率 可以通过学生个人的抬头率了解课堂上每个学生的听课情况。每隔一定时间间隔对课堂上的所有学生进行一次人脸检测,对检测到的人脸进行编号,代表每位同学,但由于每一次检测到的人脸数可能不同,所以同一编号可能不代表同一个人。因此,在基于人脸检测的个人抬头率计算中,对同一个人的定位是一个关键问题。在课堂环境中,由于学生个体以及拍摄设备在教室里的位置都相对固定,可以认为在同一个位置,不同时间检测到的人脸为同一名同学。因此,本文采用计算交并比(Intersection over union)的方法实现同人定位,

公式为 $I_{AB} = \frac{A \cap B}{A \cup B}$ 。

遍历每个时间间隔检测到的所有人脸框,分别与目标人脸范围计算 I 值,当 I 值大于阈值时,该人脸框与目

标人脸表示同一人。计算这个人脸出现的次数在总检测次数的占比,即这名学生在该课堂上的个人抬头率 $\alpha_i^{ind} = \frac{T_i}{T} \times 100\%$ 。其中: T_i 表示学生 i 人脸被检测到的次数, T 表示该课堂上总的人脸检测次数。

2.3.2 整体抬头率 整体抬头率是指每个时间间隔内,全班抬头的人数对总人数的占比,可以反映课堂上学生整体听课的即时情况。本文计算每一次检测到的人脸数对学生总人数的占比,再计算平均值,可得到课堂的平均整体抬头率 $\alpha^{ent} = \frac{1}{T} \sum_{t=1}^T \frac{n_t}{N} \times 100\%$ 。其中: n_t 表示第 t 次检测到的人脸数, N 表示学生总人数。

2.3.3 基于抬头率的课堂专注度分析 通过计算所得的学生个人和班级整体的抬头率,可以从一定程度上对课堂中每个学生以及班级整体的专注度情况做出大致的判断。经过多次实验,并结合对课堂视频的人为分析结

果,可以得出以下规律: $\alpha^{ent} \in \begin{cases} (0, 0.5) \\ [0.5, 0.7] \\ (0.7, 1) \end{cases}$ 。

对于一般的课堂,当班级整体抬头率低于 0.5 时,学生的课堂专注度较差;当抬头率在 0.5 到 0.7 之间,学生的课堂专注度良好;而当抬头率高于 0.7 时,学生整体通常表现得十分专注。

而对于学生个人来说,除了统计他在整个课堂中的抬头率,在各个时间点低头、抬头行为与班级整体行为的契合程度也是判断学生个人专注度的一个要点。在某一时间段内,如果某个学生一直处于低头状态,而班级整体的抬头率大于 0.7,可判断该学生在这段时间内不够专注;而如果在这段时间内班级整体的抬头率低于 0.5,此时可能为低头笔记等行为,则可认为该学生的专注程度较高。

3 基于人脸识别的课堂专注度应用

在实验中,通过长达约 40 min 课堂视频进行分析,视频格式为 .mp4, 帧宽 1 280, 帧高 720, 帧率为每秒 30 帧。将视频输入系统中,每隔 2 s 进行 1 次检测和计算,总共进行了约 1 200 次检测。选取了第 1 位同学(记为“student 0”)作为学生个人抬头率的计算样例。经计算可得这位同学在课堂中的抬头率为 $\alpha_0^{ind} = 56.04\%$, 班级整体的抬头率为 $\alpha^{ent} = 58.24\%$, 整体专注度良好。该名同学和班级整体的抬头率具体检测情况如图 5 所示。

同时,将编号为 student 0 的学生低头行为与班级整体行为的契合度进行计算。在整堂课程的 1 200 多次检测中该生低头总次数为 534 次,同时统计该学生低头时刻班级整体的抬头率情况。如果整体抬头率低于 0.5,该学生的状态

与整体契合,视为专注;如果这些时刻的班级整体抬头率高于 0.5,视为不专注。统计结果如表 1 所示。

由表 1 可知,该学生在 1 200 多次检测中,有 534 次未检测到人脸,其中有 221 次班级整体的抬头率低于 0.5,即超过半数学生没有抬头,此时该学生的低头行为与整体契合,视为专注;97 次班级的整体抬头率高于 0.7,即有超过 70% 的学生抬头听讲,而该学生此时可视为不专注。

对各个时段班级的整体抬头率进行分析,可以得到整堂中的不同时段学生整体的专注情况,见图 6。根据图 6 的数据,对学生整体抬头率每隔 10 min 计算一次平均数,得到学生的抬头率、专注度的分时段表,见表 2。

从表 2 中可以看到,学生整体的课堂专注度是随课堂的时间进程变化的。课堂的前 20 min 整体表现良好,在 20~30 min 时间段学生的专注程度达到顶峰,随后开始下滑,在课堂最后的 10 min 达到最低点。实验结果与相关研究所述的大多数大学生在上课时注意力集中时间保持在 60%~80% 的规律相吻合^[23]。教师也可根据此规律,调整授课方案,尽量不要把重要的教学内容安排到课堂最后的 10 min。

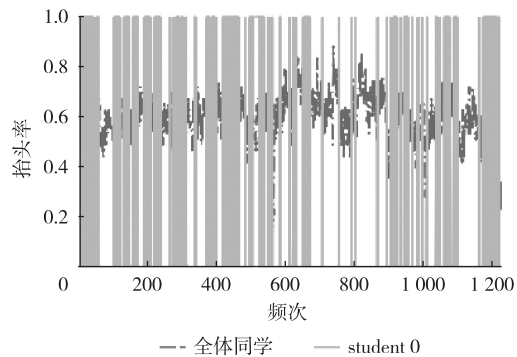


图 5 学生个人和整体的抬头率检测情况

Fig. 5 Rate detection of student individual and overall head-lift

表 1 student 0 与班级整体行为契合度分析

Tab. 1 Fit analysis of student 0 and class behavior

| | 整体抬头率低于 0.5 | 整体抬头率高于 0.7 |
|--------|-------------|-------------|
| 该生低头次数 | 221 | 97 |
| 是否专注 | 专注 | 不专注 |

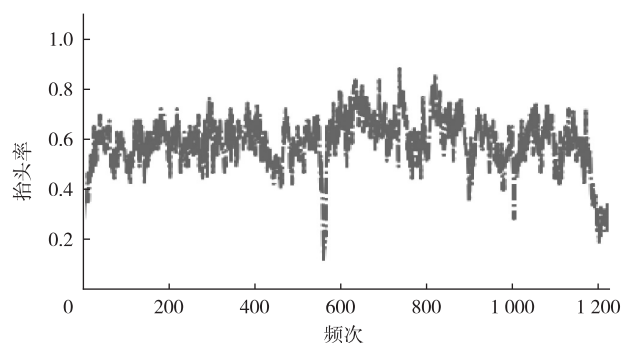


图 6 学生整体的抬头率检测情况

Fig. 6 Detection of overall head-lift rate

表 2 抬头率、专注度分时段表

Tab. 2 Head-lift rate and concentration timetable

| 时段/min | 整体抬头率 | 课堂专注度 |
|--------|---------|-------|
| 0~10 | 0.560 7 | 良好 |
| 10~20 | 0.608 9 | 良好 |
| 20~30 | 0.743 3 | 优秀 |
| 30~40 | 0.374 2 | 差 |

4 总结

在系统实现和实验的过程中发现,对于类似本文的大场景多目标检测,MTCNN等多网络级联算法速度较慢,而端到端(End to end)的算法在训练和运行速度上均有很大的优势,精度也能符合大部分的需求,如YOLO算法。针对课堂教学的应用需求,结合当前的深度学习和人脸检测技术,本文提出了在课堂环境中具有高精度和及时性的人脸检测算法和“抬头率”的计算方法,以及基于此的学生专注度分析方法。该方法具有有效性和可靠性,为评估和改善学生听课效率和课堂教学质量提供了科学的解决方案,尤其是对课堂“低头族”有较大的监督和警示作用,对于教师的教学方案也具有建议和指导意义。今后考虑借鉴YOLOv3等方法搭建训练新的人脸检测系统,提升系统运行速度,并加入人脸识别、表情识别等内容,作为新的教学评价指标,增加系统的准确性和鲁棒性。

参考文献:

- [1] 郭秀兰,赵佳敏.本科课堂教学“出勤率、抬头率、满意率”的调查报告[J].改革与开放,2016(19):108-110.
GUO X L,ZHAO J M. Investigation report on “attendance, head-up rate, satisfaction rate” for undergraduate classroom teaching[J]. Reform & Opening, 2016(19): 108-110.
- [2] VIOLA P,JONES M. Rapid object detection using a boosted cascade of simple features[EB/OL]. [2018-12-10]. https://cse.buffalo.edu/courses/cse725/peter/Viola_2001.pdf.
- [3] KRIZHEVSKY A, SUTSKEVER I, HINTON G. Image net classification with deep convolutional neural networks [J]. Advances in neural information processing systems, 2012,25(2):1097-1105.
- [4] YANG S, LUO P, LOY C C, et al. From facial parts responses to face detection: a deep learning approach[EB/OL]. (2015-09-22)[2019-03-25]. <https://arxiv.org/pdf/1509.06451v1.pdf>.
- [5] ZHANG K P,ZHANG Z P,LI Z F, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. IEEE Signal Processing Letters,2016,23(10):1499-1503.
- [6] ZHANG S F,ZHU X Y,LEI Z, et al. Face boxes:a CPU real-time face detector with high accuracy[EB/OL]. (2018-12-25)[2019-03-25]. <https://arxiv.org/pdf/1708.05234.pdf>.
- [7] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [C]//International Conference on Neural Information Processing Systems. Cambridge, US: MIT Press, 2015:91-99.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]. //Computer Vision-ECCV 2016. Amsterdam, Netherland, 2016.
- [9] 王丹,赵宏伟,戴毅,等.基于回归的人脸检测加速算法[J].重庆邮电大学学报(自然科学版),2019,31(4):550-555.
WANG D,ZHAO H Y,DAI Y, et al. A face detection acceleration algorithm based on regression[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2019,31(4):550-555.
- [10] 范亚男,葛卫丽.智能视频监控系统发展及应用[J].价值工程,2010,29(17):97-98.
FAN Y N,GE W L. Development and application of intelligent video surveillance system[J]. Value Engineering, 2010,29(17):97-98.
- [11] 杨金朋,薛耀锋,李佳璇,等.基于人脸表情识别的在线学习情感计算研究[J].中国教育技术装备,2017(18):35-36.
YANG J P,XUE Y F,LI J X, et al. Research on online learning emotion calculation based on facial expression

- recognition[J]. *China Educational Technology & Equipment*, 2017(18):35-36.
- [12] 冯满堂, 马青玉, 王瑞杰. 基于人脸表情识别的智能网络教学系统研究[J]. *计算机技术与发展*, 2011, 21(6):193-196.
- FENG M T, MA Q Y, WANG R J. Research on intelligent network teaching system based on facial expression recognition[J]. *Computer Technology and Development Comput Technol Dev*, 2011, 21(6):193-196.
- [13] 程萌萌, 林茂松, 王中飞. 应用表情识别与视线跟踪的智能教学系统研究[J]. *中国远程教育*, 2013(3):59-64.
- CHENG M M, LIN M S, WANG Z F. Research on intelligent teaching system based on expression recognition and sight tracking[J]. *Distance Education in China*, 2013(3):59-64.
- [14] 孙波, 刘永娜, 陈玖冰, 等. 智慧学习环境中基于面部表情的情感分析[J]. *现代远程教育研究*, 2015(2):96-103.
- SUN B, LIU Y N, CHEN J B, et al. Emotional analysis based on facial expressions in smart learning environment [J]. *Modern Distance Education Research*, 2015(2):96-103.
- [15] 荆洲, 权伟, 唐杰, 等. 基于人脸识别的智能课堂点名系统[J]. *软件工程*, 2017, 20(5):43-46.
- JING Z, QUAN W, TANG J, et al. Intelligent classroom name system based on face recognition[J]. *Software Engineering*, 2017, 20(5):43-46.
- [16] 侯立刚, 王远阳, 张仕尔, 等. 人脸检测技术在教学评估中的应用[J]. *电子世界*, 2016(24):37-38.
- HOU L G, WANG Y Y, ZHANG S E, et al. The application of face detection technology in teaching evaluation [J]. *Electronics World*, 2016(24):37-38.
- [17] 韩丽, 李洋, 周子佳, 等. 课堂环境中基于面部表情的教学效果分析[J]. *现代远程教育研究*, 2017(4):97-103.
- HAN L, LI Y, ZHOU Z J, et al. Analysis of teaching effect based on facial expressions in classroom environment[J]. *Modern Distance Education Research*, 2017(4):97-103.
- [18] CALLAHAN K, RADEMACHER J A. Using self-management strategies to increase the on-task behavior of a student with autism[J]. *Journal of Positive Behavior Interventions*, 1999, 1(2):117-122.
- [19] SUN C M, VALLOTTON P. Fast linear feature detection using multiple directional non-maximum suppression[J]. *Journal of Microscopy*, 2009, 234(2):147-157.
- [20] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[EB/OL]. (2014-10-22) [2018-12-10]. <http://export.arxiv.org/pdf/1311.2524>.
- [21] JAIN V, LEARNED-MILLER E G. Fddb: a benchmark for face detection in unconstrained settings [EB/OL]. [2018-12-01]. <https://people.cs.umass.edu/~elm/papers/fddb.pdf>.
- [22] 米硕, 孙瑞彬, 李欣, 等. 迁移学习在卷积神经网络中的应用[J]. *中国新通信*, 2018(7):121.
- MI S, SUN R B, LI X, et al. Application of transfer learning in convolutional neural networks[J]. *China New Telecommunications*, 2018(7):121.
- [23] 杨伟文, 蒋月婷, 肖洪广, 等. 影响大学生上课注意力集中原因的调查和对策[J]. *中国高等医学教育*, 2012(12):44-45.
- YANG W W, JANG Y T, XIAO H G, et al. Investigation and countermeasure of affecting the causes of college students' attention[J]. *China Higher Medical Education*, 2012(12):44-45.

The Research of College Classroom Focus Based on face Detection

TANG Kang, XIAN Qiang, LI Mingyong

(College of Computer and Information Science, Chongqing Normal University, Chongqing 401331, China)

Abstract: [Purposes] Based on the technology of deep learning face detection, a new project is provided for classroom teaching evaluation. [Methods] Combined with face detection, it constructs classroom appropriate face detection optimized cascade convolutional neural mode, and proposes a new method to analyze class focus based on head-lift rate. [Findings] Through the detection and numeration of the students' faces in the video, the head-lift rate is calculated to analyze the class focus and time allocation, in order to help the teachers to be acquainted with the class teaching situation. [Conclusions] With a multitude of tests and optimization, this system possesses validity and reliability in face detection, which also offers an individualized learning for students and enhances the classroom teaching quality and efficiency for teachers.

Keywords: deep learning; face detection; head-lift rate; class focus; instructional evaluation

(责任编辑 黄 颖)