

DOI: 10.13382/j.jemi.B2307023

基于深度学习的学生课堂学习状态监测系统设计*

张立军¹ 曹江涛¹ 姬晓飞² 王天昊¹

(1. 辽宁石油化工大学信息与控制工程学院 抚顺 113001; 2. 沈阳航空航天大学自动化学院 沈阳 110136)

摘要:现阶段,学生课堂学习状态的研究大多集中在单人的在线监测,对于多人且环境复杂的线下课堂的监测还处于探索阶段。该研究针对线下教育设计了学生课堂学习状态监测系统,对学生课堂出勤情况及学生面部出现的疲劳状态进行实时监测。首先,使用DSFD人脸检测算法结合ResNet深度残差网络对学生进行人脸识别,记录学生出勤情况;然后,使用ERT回归树集合算法结合头部姿态估计对打哈欠和低头瞌睡的疲劳行为进行检测;再使用加入CBAM模块改进的YOLOv5目标检测算法对学生闭眼行为进行检测;最后,形成一套完整的集合出勤、疲劳检测的学生课堂学习状态监测系统。该系统在实际课堂的测试环境下,可以准确的对学生的出勤进行统计,并且可以实时的监测学生面部出现的打哈欠、低头瞌睡、闭眼的疲劳状态,检测的准确率均超过90%,检测速度约为14.1 fps,证明该系统具有重要的使用价值。

关键词:人脸识别;计算机视觉;疲劳检测;出勤检测;课堂监测

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Design of classroom learning state monitoring system for students based on deep learning

Zhang Lijun¹ Cao Jiangtao¹ Ji Xiaofei² Wang Tianhao¹

(1. School of Information and Control Engineering, Liaoning Petrochemical University, Fushun 113001, China;

2. School of Automation, Shenyang Aerospace University, Shenyang, Liaoning 110136, China)

Abstract: At present, most of the research on students' classroom learning status focus on single-person online monitoring, and the monitoring of offline classroom with multiple students and complex environment is still in the exploratory stage. A monitoring system for students' classroom learning status was designed for offline education to monitor students' classroom attendance and fatigue state of students' faces in real time. First, DSFD face detection algorithm combined with ResNet deep residual network was used to recognize students' faces and record students' attendance. Then, ERT regression tree set algorithm combined with head pose estimation was used to detect the fatigue behavior of yawning and drowsiness. Then, the improved YOLOv5 object detection algorithm added CBAM module was used to detect students' closed eyes behavior. Finally, a complete set of integrated attendance, fatigue detection of student classroom learning state monitoring system is formed. In the actual classroom test environment, the system can accurately calculate the students' attendance, and can real-time monitor the fatigue state of yawning, lower head and closed eyes on the face of the students. The detection accuracy rate is more than 90%, and the detection speed is about 14.1 fps, which proves that the system has important use value.

Keywords: face recognition; computer vision; fatigue detection; attendance detection; classroom monitoring

0 引言

课堂作为知识传播的“载体”具有重要的作用,教师

通过课堂讲授、讨论和互动,将学科知识传授给学生。随着人工智能算法的发展,可以使用计算机视觉技术对学生课堂学习情况进行监测,这不仅可以帮助教师了解学生的学习状态,还有利于教师调整与改进授课方式,提高

授课质量。

学生课堂学习状态监测的研究主要包含两个方面:一方面是基于人脸识别^[1]算法的学生课堂出勤情况监测,另一方面是基于学生面部表情识别和面部关键点检测的疲劳行为监测。其中,人脸识别是指将检测到的人脸与已知的人脸数据库中的信息进行匹配,以确定个体身份。例如 Zhang 等^[2]使用 Dlib 库中的基于方向梯度直方图(histogram of oriented gradients, HOG)特征的人脸检测算法与 ResNet(residual network)深度残差网络结合进行人脸识别。屈东东等^[3]设计了一种轻量化人脸识别算法 L-MobileNet V2,该算法将传统分类任务中的 Softmax 损失改为 Am-Softmax 作为辅助损失函数,优化了原有网络结构。此类方法网络简单,但是将其应用在多人课堂场景下时,由于监控摄像头距离学生远近不同,存在人脸图像像素占比不一,导致人脸检测和人脸识别失败。针对以上问题, Schroff 等^[4]通过嵌入学习引入三元组损失函数(triplet loss, TL)增大类间距与缩小类内距,并将人脸图像映射到一个紧凑的特征空间,以提高系统对多人人脸识别的性能。Deng 等^[5]提出可加性角边缘损失(additive angular margin loss, ArcFace)来稳定人脸识别算法的训练过程,进一步提高人脸识别模型对复杂人脸的判别能力。此类算法鲁棒性强,但模型复杂、计算量大,存在实时性差的问题。目前,基于面部表情识别的监测主要是对课堂中学生面部出现的困惑、惊讶、难过、开心等情绪进行识别。例如 Mohamad 等^[6]为了克服数据稀缺的挑战,提出了一种深度学习模型与两步策略训练法,对在线课堂中学生的表情进行识别。于婉莹等^[7]构建了多路深度注意力网络,通过自注意力机制分配不同权重,并使用损失函数限制权重的分配,以解决复杂课堂场景存在遮挡情况下的学生表情识别问题。但此种方法因学生面部表情种类繁多,且往往是关注整张脸的变化,缺乏关注面部关键部位信息变化,导致识别效果不佳。相较之下,基于学生面部关键点的疲劳检测算法更加直观,只需关注眼部与嘴部位置变化,以此判断是否有闭眼、打哈欠等行为,并可以结合头部姿态估计判断是否存在低头瞌睡行为。例如 Zhuang 等^[8]提出一种基于瞳孔和虹膜分割的眼睛状态疲劳检测方法。Yi 等^[9]提出同时检测眼睛宽高比、眼睑闭合超过瞳孔的百分比、眨眼频率和瞳孔闭塞率的多特征融合算法判断眼睛疲劳程度。Yang 等^[10]提出采用三维卷积和双向长短期记忆网络进行时空特征提取的深度学习网络检测打哈欠行为。Mou 等^[11]开发了一种各向同性自监督学习方法判断眼睛与嘴巴的状态。基于面部关键点的疲劳检测算法在疲劳驾驶和单人听课状态检测方面已经展现出一定的鲁棒性。但在线下课堂的应用背景下,由于学生人数多,且像眼部这种关键部位的数据占比偏小,导致闭眼检测效果不佳。

针对以上问题,该研究设计了学生课堂学习状态监测系统,此系统可以实现实时监测线下课堂学生的出勤情况和疲劳行为。该系统使用具有高检测精度与实时性优势的双重人脸检测器(dual shot face detector, DSFD)算法结合 ResNet 网络实现复杂环境下的精确人脸识别。此外该系统使用加入轻量级卷积注意力模块(convolutional block attention module, CBAM)的改进 YOLOv5 目标检测算法提升对人眼小目标的检测精度。

1 系统的总体框架

本文的总体设计框架如图 1 所示,分为学生课堂出勤情况监测和学生课堂疲劳状态监测两部分。具体实现如下:

1) 学生课堂出勤情况监测。首先使用 DSFD 人脸检测算法定位所有学生人脸,对定位到的人脸采用 ResNet 网络提取 128 维特征向量,再将现有的 128 维特征向量与人脸库中的信息进行对比,判断学生是否出勤并对出勤情况进行记录。

2) 学生课堂疲劳状态监测。在 DSFD 人脸检测算法定位到学生人脸后,使用 ERT 算法提取面部特征点并获取嘴部标志索引,以此判断嘴部疲劳特征,同时结合头部姿态估计判断头部疲劳特征。使用改进的 YOLOv5 目标检测算法判断眼部疲劳特征。最后依据不同的疲劳特征给出不同的疲劳警示。

2 基于学生课堂出勤监测算法

2.1 人脸检测算法介绍

人脸检测是实现学生课堂学习状态监测的“基石”。为准确实时地定位到学生人脸,本文应用 Li 等^[12]提出的 DSFD 人脸检测算法,此算法以 VGG16 为主干网路,采用特征增强模块(feature enhance module, FEM)增强原始特征映射,使网络具有更强的识别性和鲁棒性。同时将单镜头检测器扩展到双镜头检测器,使用两组不同的锚计算渐进损失,使其能够更好地学习特征。与其他算法不同的是此算法利用感受野扩大和新的锚设计策略后,不再需要步长、锚和感受野 3 种尺寸满足等比例区间原则。因此,DSFD 人脸检测算法相比于其他人脸检测算法更加灵活,对低分辨率人脸或遮挡严重的人脸有更高的准确率与实时性,适用于该研究中的学生人数多,面部所占像素比少且人脸图像像素占比不一的课堂状态监测场景。

2.2 人脸识别算法介绍

人脸识别是一种生物特征识别技术,旨在识别和验证个体身份。本文使用的是 Dlib 的基于卷积神经网络

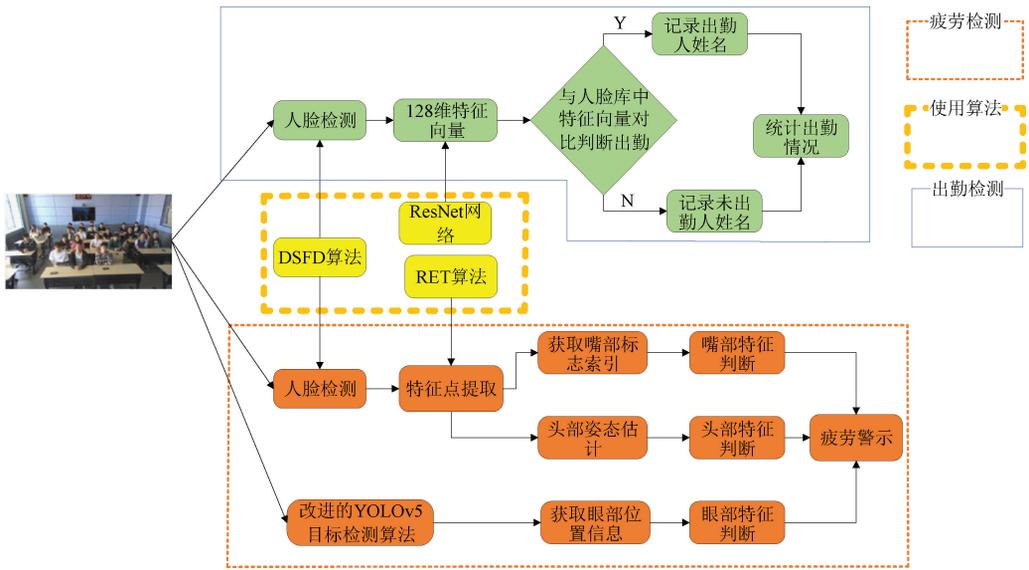


图 1 算法框架

Fig. 1 Algorithmic framework

(convolutional neural network, CNN) 的人脸识别算法。该算法对 DSFD 人脸检测算法检测到的人脸图像使用 ResNet 深度残差网络提取 128 维特征向量,再通过设置阈值,比较两张人脸特征向量的欧氏距离进行人脸识别。该算法具有较高的人脸识别精度与速度,满足课堂出勤检测的需求。人脸识别流程如图 2 所示。

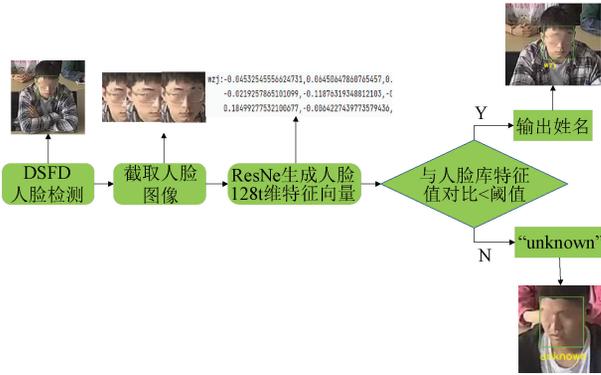


图 2 人脸识别流程图

Fig. 2 Face recognition flow chart

3 基于学生课堂疲劳状态监测算法

3.1 基于 ERT 的人脸关键点定位算法

人脸疲劳状态可以根据面部区域特征分析进行判定,面部区域是基于人脸关键点而定^[13]。本文采用回归树集合算法(Ensemble of Regression Trees, ERT)建立预训练关键点检测模型进行人脸关键点定位^[14]。基于 ERT 算法建立的预训练模型原理如式(1)所示。

$$\hat{S}^{(t+1)} = \hat{S}^{(t)} + r_t(I, \hat{S}^{(t)}) \quad (1)$$

式(1)为级联回归器, $\hat{S}(t+1)$ 为当前人脸估计值, $r_t(I, \hat{S}^{(t)})$ 为回归器。若有训练数据集 $(I_1, S_1), (I_2, S_2), \dots, (I_n, S_n)$ 其中 I_n 为人脸图像, S_n 为图像的形状向量。设训练数据为 $\{(I_{\pi_i}, \hat{S}_i^{(t)}, \Delta S_i^{(t)})\}_{i=1}^N$, 其中 $\pi_i \in \{1, \dots, n\}$, 学习率为 $0 < \gamma < 1$, R^2P 表示图像中所有人脸的 x, y 坐标,对回归器进行初始化为:

$$f_0(I, \hat{S}^{(t)}) = \operatorname{argmin}_{\gamma \in R^{2P}} \sum_{i=0}^n \|\Delta S_i^{(t)} - \gamma\|^2 \quad (2)$$

$$\Delta S_i^{(t)} = S_i - \hat{S}_i \quad (3)$$

对回归器中 γ_t 根据式(4)进行迭代并更新 $f_k(I_{\pi_i}, \hat{S}^{(t)})$, $k = 1, \dots, K$:

$$r_{ik} = \Delta S_i^{(t)} - f_{k-1}(I_{\pi_i}, \hat{S}_i^{(t)}) \quad (4)$$

$$f_k(I_i, \hat{S}_i^{(t)}) = f_{k-1}(I_i, \hat{S}_i^{(t)}) + \gamma g_k(I, \hat{S}^{(t)}) \quad (5)$$

式中: $g_k(I, \hat{S}^{(t)})$ 为弱回归函数, k 为迭代次数,最终训练完成的树的回归器为:

$$r_t(I, \hat{S}^{(t)}) = f_k(I_i, \hat{S}_i^{(t)}) \quad (6)$$

该算法的思想是训练多个回归树,然后将它们的预测结果组合起来以获得更准确的回归模型。通过预训练模型获得的人脸图像关键点原理如图 3 所示。

3.2 打哈欠状态检测

嘴部的状态主要有 3 种:说话、打哈欠与闭合。闭合的状态容易区分,说话与打哈欠两种状态都使嘴部纵横比(mouth aspect ratio, MAR)发生变化,说话时嘴巴张开

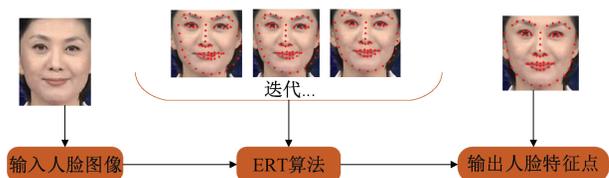


图 3 人脸关键点原理

Fig. 3 Face key point schematic

程度较小且维持的时间较短,打哈欠时嘴巴张开程度较大且维持时间较长。根据这一特点,本文对嘴部 3 种状态的 MAR 值进行记录,记录结果如图 4 所示。由图可知嘴部闭合时 $MAR = 0$;说话时 $0 < MAR < 0.8$;而打哈欠时 $MAR \geq 0.8$ 。因此,本文结合张口度与张口时间设计了一种双阈值检测法对打哈欠进行检测。

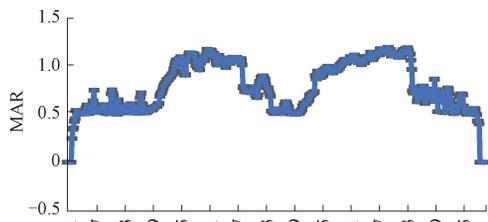


图 4 嘴部状态 MAR 值变化

Fig. 4 The MAR value of the mouth state changes

根据双阈值检测法,结合持续时间判定学生是否出现打哈欠的疲劳迹象,判定公式如式(7)所示:

$$Yawn = \frac{M_i}{N} \times \% \quad (7)$$

$M_i (i = 0, 1, 2, \dots)$ 为符合打哈欠的帧数, N 为 1 min 内总帧数。其中,在帧率为 22 fps 的设备上对实验室的 12 位同学进行打哈欠时长测试,发现平均每位同学打一

个哈欠的帧率大约在 121~154 fps,因此按照式(7)可计算得出 $9.1\% \leq Yawn \leq 11.7\%$ 之间为打哈欠状态,为能快速关注到困意较深的同学,将其阈值设置阈值为 10%,当 $Yawn \geq 10\%$ 时认为打了一个哈欠。

3.3 基于改进的 YOLOv5 闭眼检测算法介绍

在 3.2 节中根据人脸特征点定位到嘴巴区域对打哈欠行为进行检测,此外,该研究也考虑了利用特征点定位眼睛区域来检测闭眼状态。然而,在实际的课堂场景中,发现由于眼部特征点占面部特征点的比例偏小且聚集,导致无法依据眼脸纵横比来确定眼睛的睁开与闭合状态。因此,本文使用改进的 YOLOv5 目标检测算法对课堂中学生的闭眼情况进行检测。

YOLOv5 是一种轻量级、高效率的目标检测算法,它在较小的模型尺寸和更快的推理速度下实现了高精度的目标检测^[15]。YOLOv5 的模型结构分为主干网络 (Backbone)、特征融合模块 (Neck)、检测头 (Detection Head) 及输出层。在 Backbone 部分使用 CSPDarknet53 (cross stage partial Darknet 53) 网络进行特征提取,Neck 部分的设计类似于 FPN (feature pyramid network),融合来自不同层级的特征信息,用于在多尺度目标检测任务中取得更出色的检测效果。Head 部分是由一系列卷积层组成,负责输出目标检测的结果。在卷积神经网络^[16]中,底层神经网络包含更多的语义信息,深层神经网络包含更多的位置信息,由于语义信息对小目标的检测尤为重要,因此,在主干网络与特征融合模块之间引入轻量级的卷积注意力模块 (convolutional block attention module, CBAM) 提高网络对小目标的检测精度,整体的网络结构如图 5 所示。同时根据数据集的尺寸大小对锚框的宽高比进行了调整,以保证在检测眼睛这类小目标时具备更高的检测率。

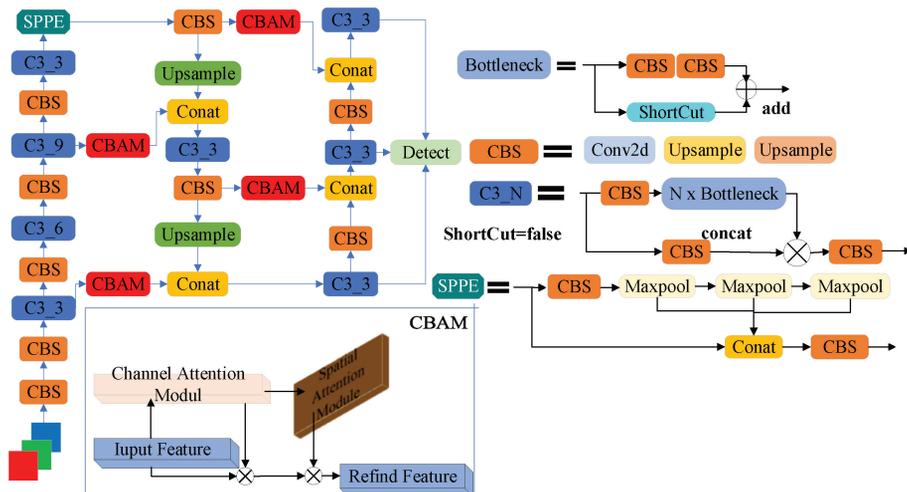


图 5 改进 YOLOv5 整体结构

Fig. 5 Improve the overall structure of YOLOv5

3.4 头部姿态估计判断

当学生听课处于疲劳状态时会不自觉的向前后倾斜或者左右倾斜,不能全神贯注的听讲,影响学习效率。头部姿态有3种表现形式,如图6所示,即分别沿X轴、Y轴、Z轴旋转从而形成俯仰角(Pitch)、偏航角(Yaw)、滚转角(Roll)。

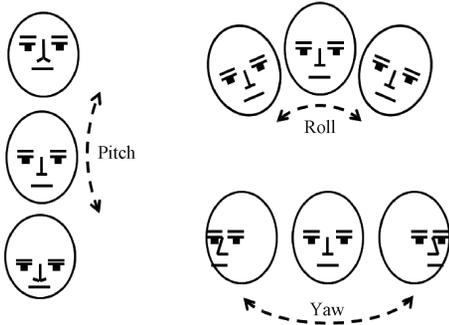


图6 头部姿态估计
Fig. 6 Head pose estimation

偏航角和俯仰角的变化会影响左右眼角与左右嘴角到鼻尖的距离。因此,本研究采用眼角和嘴角作为4个特征点来计算它们与鼻尖之间的距离。计算公式如式(8)所示:

$$d_x = |P_x - P_{NoseTip}| \quad (8)$$

其中, d_x 为特征点到鼻尖的距离, P_x 为特征点, $P_{NoseTip}$ 为鼻尖点。 L_1 、 L_2 分别为左右眼与鼻尖的距离,如图7所示。而滚转角的变化是通过左右嘴角水平方向与嘴角到鼻尖方向的夹角来识别,计算公式如式(9)所示:

$$\vec{u} = P_x - P_{NoseTip} \quad (9)$$

式中: \vec{u} 为鼻子点的特征向量。

$$A_x = \cos^{-1} \frac{\vec{u}}{|\vec{u}|} \times \frac{\vec{v}}{|\vec{v}|} \quad (10)$$

式中: \vec{v} 为水平向量, A_x 为鼻尖与特征点之间的水平夹角,图7虚线为水平矢量, A_1 和 A_2 为鼻尖与左右嘴角两侧的夹角。根据得出的特征向量求出欧拉角,再设置阈值,就可以判断学生是否出现低头瞌睡行为。

4 系统总体功能实现及测试

4.1 系统总体功能实现

依据改进的人脸识别算法识别学生身份,依据人脸关键点和头部姿态估计检测嘴部、头部状态,使用YOLO检测眼部状态。这些特征检测与识别之间独立工作互不影响、相互补充,使得模型在某一特征检测受阻时,也能较好的对学生课堂学习情况进行监测。对于课堂出勤

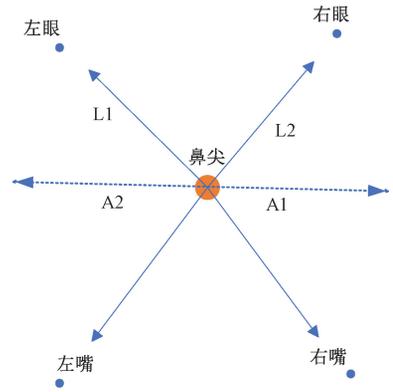


图7 人脸方位角估计特征示意图
Fig. 7 Face azimuth estimation feature diagram

及疲劳状态的判定如下:

1) 学生课堂出勤情况检测。系统在上课开始前的2 min内,自动抓取前50帧图像进行人脸识别,一旦识别完成,系统会在人脸框左下方显示检测到的学生姓名,并将出勤名单以Excel表格形式输出,如果检测到的人脸未在人脸库则会显示“unknown”。教师只需对未识别的人脸进行二次检测,以防出现错检、漏检或替课情况发生。

2) 学生眼部疲劳预警。使用YOLOv5对眼睛睁、闭合进行检测,眼部疲劳阈值设置为40 s,即学生闭眼时间超过40 s则将眼部检测框转变成白色,以此进行疲劳预警。

3) 学生嘴部疲劳预警。根据双阈值检测法, $MAR \geq 0.8$, $Yawn \geq 10\%$ 时认为打了一个哈欠。设置阈值为4,即检测到4次打哈欠,检测框转变成蓝色给出疲劳预警。

4) 学生头部姿态疲劳预警。根据求出的实时欧拉角,判定学生是否进行低头瞌睡行为,即pitch和roll变化大于一定值时检测框变成红色进行疲劳预警。

4.2 设备平台设置

本次实验的配置如表1所示,硬件设备为GeForce RTX 3060(GPU) 12 G显存和Windows11系统,使用Python3.8编程语言和Pytorch1.13.1框架。该研究在这个平台上做人脸检测、人脸识别、疲劳检测相关测试,验证系统性能。

表1 设备平台配置

Table 1 Equipment platform configuration

实验过程	测试阶段
测试平台	GeForce RTX 3060(GPU)
编程语言	Python3.8
系统	Windows11
学习框架	Pytorch1.13.1

4.3 人脸检测算法测试

该研究分别对 DSFD、RetinaFace^[17]、YOLOv5 算法在数据集 WIDER FACE 进行测试。3 种算法对人脸的检测准确率和检测帧率如表 2 所示。

表 2 人脸检测算法性能对比

Table 2 Performance comparison of face detection algorithms

算法模型	准确率/%	帧率/fps
DSFD	90.1	22.54
RetinaFace	91.4	7.96
YOLOv5	82.9	30.82

由表 2 可知 RetinaFace 人脸检测算法的准确率最高达到了 91.4%，YOLOv5 算法对人脸检测的准确率最低为 82.9%，而 DSFD 人脸检测算法准确率处于两者之间偏上为 90.1%。从测试的帧率数据中可知，YOLOv5 的帧率最高，达到了 30.82 fps，其他两个人脸检测算法 RetinaFace 与 DSFD 的帧率分别为 7.96、22.54 fps。对于课堂状态监测而言，既需要保证准确性又需要保证实时性，RetinaFace 算法在准确性方面表现出色，但其检测速度未能达到令人满意的水平。YOLOv5 在检测速度上表现出色，但准确率较低。因此，经综合考虑，选择准确率高且检测速度较快的 DSFD 算法实现对学生人脸的检测。

4.4 人脸识别算法测试

课堂的出勤率既能反映学生对知识的渴望程度又能反应学校的教学质量。本文应用改进的 dlib 人脸识别算法记录学生出勤情况，并在相同条件下，对比改进前与改进后的人脸识别算法的准确率与实时性，对比结果如表 3 所示。

表 3 改进的人脸识别算法对比

Table 3 Comparison of improved face recognition algorithms

算法模型	阈值	准确率/%	帧率/fps
HOG + ResNet	0.4	75.0	8.4
	0.6	80.2	8.2
DSFD + ResNet	0.4	90.5	14.4
	0.6	96.1	14.1

由表 3 可知使用 DSFD 算法替换原 Dlib 中使用 HOG 算法的方法，在阈值为 0.4 和 0.6 的准确率分别提高 15.5% 和 15.9%，帧率由 8.4 和 8.2 fps 增加到 14.4 和 14.1 fps。人脸识别对比结果如图 8、9 所示。

由图 8 中可知，使用 HOG 人脸检测算法与 ResNet 网络结合进行人脸识别的方法，对存在遮挡与人脸图像



图 8 HOG + ResNet 进行人脸识别

Fig. 8 HOG + ResNet for face recognition



图 9 DSFD + ResNet 进行人脸识别

Fig. 9 DSFD + ResNet for face recognition

像素占比较小的人脸的检测效果不佳，而改进的使用 DSFD 算法检测人脸的方法(图 9)对人脸的识别效果明显提升。因此，充分证明了，改进的人脸识别算法的可行性。

4.5 眼睛检测算法的训练和结果测试

1) 数据集制作

本文在南京航空航天大学提供 CEW (Closed Eyes in the Wild)^[18] 的眼睛数据集的基础上，加入了真实课堂采集的学生眼睛图片数据，对数据集进行扩充。同时为了提高网络训练效果，增加模型的鲁棒性，对数据集进行水平翻转、随机亮度调整、高斯模糊等操作来增强数据集。最后得到 7 200 张图片，其中睁眼图片 3 500 张，闭眼图片 3 700 张，按照 YOLO 格式进行标注，并按照 6 : 2 : 2 划分训练集、测试集和验证集。

2) 实验设置

本次实验的硬件设备为 4.1 节介绍所示，实验代码是基于 Ultralytics 公司 YOLOv5 第 6 代版本中的 YOLOv5s 进行改进。训练迭代次数 (epoch) 均为 200；批次大小为 16；采用随机梯度下降方法。

3) 实验结果

将改进算法与其他检测算法进行对比，对比结果如表 4 所示。由表 4 可知，在相同数据集的条件下，本文改进的网络与 YOLOv7m 的准确率相同，都为 97.5%，比 YOLOv8s 准确率略低，但本文改进的算法在参数量、训练时间和检测时间方面均优于 YOLOv7m，YOLOv8s。因此，综合考虑，本文算法符合实际应用的要求。改进前较改进后的整体效果对比如图 10、11 所示。

表 4 算法模型改进对比

Table 4 Comparison of algorithm model improvement

算法模型	准确率/%	Params×10 ⁶	训练时间/h	检测时间/ms
YOLOv5s	96.6	7.0	2.917	14
本文	97.5	7.0	2.924	15
YOLOv5m	96.6	20.8	3.698	27
YOLOv7s ^[19]	96.6	10.2	3.093	22
YOLOv7m	97.5	20.9	3.615	25
YOLOv8s ^[20]	95.1	14.1	3.287	20
YOLOv8m	97.0	25.8	11.777	36

由图 10 中,仅使用 YOLOv5 网络对眼睛进行检测,能清楚地看到最后一排的靠窗和前排几名学生的眼睛未能检测到,而加入 CBAM 模块后(图 11)检测效果明显变好,证明了改进方法的可行性。



图 10 YOLOv5 眼睛检测
Fig. 10 YOLOv5 eye detection



图 11 YOLOv5+CBAM 眼睛检测
Fig. 11 YOLOv5+CBAM eye test

4.6 疲劳检测算法测试

为了验证系统的可行性,将本系统与高校智慧教室的教育跟踪摄像系统相结合,以检测学生上课时的疲劳情况。教育跟踪摄像系统内置高速嵌入式处理器,能够集成人体检测和行为识别的图像分析算法,实现对学生的跟踪与定位功能。在这个基础上,将疲劳检测算法进行封装,增加对学生疲劳状态检测的功能。这种整合使得系统能够在已有的学生跟踪和定位功能基础上,同时实现对学生疲劳状态的监测。通过这一方式,能够更全面地了解学生的上课状态,为教师提供更有价值的信息,以优化教学效果,系统检测结果如图 12、13、14 所示。

由 12 可知检测到中间第 2 排右数第 2 个女生存在长时间低头行为,检测框变成红色;图 13 中检测到中间第 3 排右数第 3 个男生存在打哈欠行为,检测框变成蓝色;图 14 检测到中间排右数第 2 个男生存在闭眼行为,眼睛检测框变成白色。



图 12 低头行为检测
Fig. 12 Phubbing behavior detection



图 13 打哈欠行为检测
Fig. 13 Yawning behavior detection

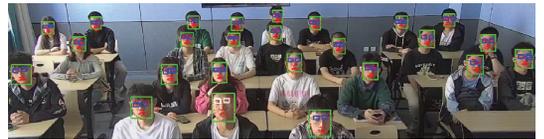


图 14 闭眼行为检测
Fig. 14 Eye closing behavior detection

5 结 论

本文针对线下课堂学生的出勤情况及疲劳行为设计了学生课堂学习状态监测系统。该系统采用 DSFD 人脸检测算法结合 ResNet 深度残差网络对学生人脸进行识别;使用 ERT 算法以及头部姿态估计技术检测学生打哈欠和低头行为;使用改进的 YOLOv5 目标检测算法识别学生的闭眼行为。经实际环境测试,该系统不仅能够准确、实时地检测出学生的出勤情况,还能捕捉到他们的疲劳行为,证明了该系统在教育领域具有重要的应用潜力。

参考文献

[1] YU Z, ZHAO C, WANG Z, et al. Searching central difference convolutional networks for face anti-spoofing[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 5295-5305.

[2] ZHANG D, LI J, SHAN Z. Implementation of Dlib deep learning face recognition technology [C]. 2020 International Conference on Robots & Intelligent System (ICRIS). IEEE, 2020: 88-91.

[3] 屈东东, 贺利乐, 何林. 改进的轻量化人脸识别算法[J]. 智能系统学报, 2023, 18(3): 544-551.

QV D D, HE L L, HE L. Improved lightweight face recognition algorithm[J]. Journal of Intelligent Systems,

- 2023, 18(3): 544-551.
- [4] SCHROFF F, KALENICHENKO D, PHILBIN J. Facenet: A unified embedding for face recognition and clustering[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 815-823.
- [5] DENG J, GUO J, XUE N, et al. Arcface: Additive angular margin loss for deep face recognition [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 4690-4699.
- [6] MOHAMAD NEZAMI O, DRAS M, HAMEY L, et al. Automatic recognition of student engagement using deep learning and facial expression [C]. Joint european Conference on Machine Learning and Knowledge Discovery in Databases. Springer, Cham, 2020: 273-289.
- [7] 于婉莹, 梁美玉, 王笑笑, 等. 基于深度注意力网络的课堂教学视频中中学生表情识别与智能教学评估[J]. 计算机应用, 2022, 42(3): 743-749.
YU W Y, LING M Y, WANG X X, et al. Expression recognition and intelligent teaching evaluation of middle school students in classroom teaching videos based on deep attention network [J]. Computer Application, 2022, 42(3): 743-749.
- [8] ZHUANG Q, KEHUA Z, WANG J, et al. Driver fatigue detection method based on eye states with pupil and iris segmentation [J]. IEEE Access, 2020, 8: 173440-173449.
- [9] YI Y, ZHOU Z, ZHANG W, et al. Fatigue detection algorithm based on eye multifeature fusion [J]. IEEE Sensors Journal, 2023, 23(7): 7949-7955.
- [10] YANG H, LIU L, MIN W, et al. Driver yawning detection based on subtle facial action recognition [J]. IEEE Transactions on Multimedia, 2020, 23: 572-583.
- [11] MOU L, ZHOU C, XIE P, et al. Isotropic self-supervised learning for driver drowsiness detection with attention-based multimodal fusion [J]. IEEE Transactions on Multimedia, 2021, 25: 529-542.
- [12] LI J, WANG Y, WANG C, et al. DSFD: Dual shot face detector [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 5060-5069.
- [13] 陆荣秀, 张笔豪, 莫振龙. 基于脸部特征和头部姿态的疲劳检测方法 [J]. 系统仿真学报, 2022, 34(10): 2279-2292.
LU R X, ZHANG B H, MO ZH L. Fatigue detection method based on facial features and head posture [J]. Journal of System Simulation, 2022, 34 (10): 2279-2292.
- [14] KAZEMI V, SULLIVAN J. One millisecond face alignment with an ensemble of regression trees [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014: 1867-1874.
- [15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779-788.
- [16] DU J, LU H, ZHANG L, et al. A spatial-temporal feature-based detection framework for infrared dim small target [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-12.
- [17] DENG J, GUO J, ZHOU Y, et al. Retinaface: Single-stage dense face localisation in the wild [J]. arXiv preprint arXiv:1905.00641, 2019.
- [18] SONG F, TAN X, LIU X, et al. Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients [J]. Pattern Recognition, 2014, 47(9): 2825-2838.
- [19] 吴兴辉, 何赞泽, 周辉, 等. 改进 YOLO v7 算法下的监控水域环境人员识别研究 [J]. 电子测量与仪器学报, 2023, 37(5): 20-27.
WU X H, HE Y Z, ZHOU H, et al. Research on environmental personnel identification of monitored waters based on improved YOLO v7 algorithm [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(5): 20-27.
- [20] WANG G, CHEN Y, AN P, et al. UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios [J]. Sensors, 2023, 23(16): 7190.

作者简介



张立军, 2021 年于营口理工学院获得学士学位, 现为辽宁石油化工大学硕士研究生, 主要研究方向为计算机视觉。

E-mail: 1822896247@qq.com

Zhang Lijun received her B. Sc. degree from Ying kou Institute of Technology in 2021. Now she is a M. Sc. candidate in Liaoning Petrochemical University. Her main research interest includes computer vision.



曹江涛, 2009 年于英国普茨茅斯大学获得博士学位, 现为辽宁石油化工大学教授、硕士生导师, 主要研究方向为智能方法及其应用、视频分析与处理等。

E-mail: cigroup@126.com

Cao Jiangtao received his Ph. D from the

University of Portsmouth in 2009. He is currently a professor and master supervisor at Liaoning Petrochemical University. His research interests include intelligent methods and applications, video analysis and processing, etc.



姬晓飞 (通信作者), 博士, 副教授, 主要研究方向为视频分析与处理、模式识别理论等。

E-mail: jixiaofei7804@126.com

Ji Xiaofei (Corresponding author) Ph. D., associate professor and M.Sc. supervisor

at Shenyang Aerospace University. Her main research interests include video analysis and pattern recognition theory, etc.



王天昊, 2020 年于辽宁石油化工大学获得学士学位, 现为辽宁石油化工大学硕士研究生, 主要研究方向为计算机视觉。

E-mail: 1936883214@qq.com

Wang Tianhao received his B. Sc. degree from Liaoning Petrochemical University in 2020. Now he is a M. Sc. candidate in Liaoning Petrochemical University. His main research interest includes computer vision.