

基于2D和3D图像处理技术的在线学习评估

姚远,徐晶晶,朱小倩

(华中师范大学物理科学与技术学院,湖北武汉430079)

摘要:新冠疫情导致全球在线教育异军突起,如何打破时间、空间的限制实时监督学生的在线学习状态随时调整教学策略进而提高学习效率具有重要意义。基于2D与3D图像处理技术,提出了一种在线评估学习状态的方法,将学习者的学习状态分为无人、多人、用户未授权、分心以及疲劳五种。利用AdaBoost算法与ResNet模型实现人脸检测和识别,并加入质心跟踪算法提高人脸识别检测效率;利用RGB-D图像实时获取人脸三维模型,通过EPNP算法获取学习者头部姿态进而评估学习姿态;提取学习者眼睛和嘴巴的实时图像特征,获取学习者眼睛与嘴巴纵横比值,实现学习者学习疲劳状态的实时预警。测试表明学习者的身份识别、姿态评估以及疲劳预警有效可行,为提高在线学习的质量提供了积极思路。

关键词:图像处理;在线学习;身份识别;姿态评估;疲劳预警

中图分类号:TP317.4

文献标识码:A

文章编号:1673-629X(2021)12-0128-07

doi:10.3969/j.issn.1673-629X.2021.12.022

Online Learning Evaluation Based on Processing Technologies in 2D and 3D Images

YAO Yuan, XU Jing-jing, ZHU Xiao-qian

(School of Physical Science and Technology, Central China Normal University, Wuhan 430079, China)

Abstract: In light of the emerging global online education resulted from the outbreak of COVID-19, how to break the time and space limits to supervise students' online learning conditions on timely units and adjust teaching strategies freely to facilitate their learning efficiency is of paramount significance. Based on the processing technologies in 2D and 3D images, we put forward an online evaluation method of learning conditions, which divides the learning condition of learners into five categories: No Person, More than One Person, Failed Authorization from the User, Distraction and Fatigue. AdaBoost algorithm and ResNet model are used to realize face detection and recognition, and centroid tracking algorithm is added to improve the efficiency of face recognition detection. RGB-D image can be adopted to acquire 3D face models at any time while EPNP algorithm can help collect their head gesture and evaluate their learning conditions. Real-time image characteristics about their eyes and mouth is extracted to get an aspect ratio between the two parts, making it possible to achieve infinite warning of their fatigue in learning. As suggested in the test, the practice proves feasible in recognizing the learner's identity, evaluating his gestures and warning of their fatigue, which provides a positive thought on improving the quality of online learning.

Key words: image processing; online learning; identity recognition; posture assessment; fatigue warning

0 引言

2020年初新冠肺炎疫情爆发以来,国内外学校纷纷关闭,打破时空限制满足不同群体的远程在线教学呈井喷式发展,大学MOOC、腾讯课堂、网易云课堂、51talk、猿辅导、新东方在线、学而思等为学习者提供了丰富的教学资源^[1]。与传统的课堂相比,在线学习缺乏一定的监督机制,在线教学系统无法自动评估学习者的学习状态进而影响了教学效果,因此实时分析评

估进而监督在线学习者的学习状态具有重要意义。

Gupta S K等人^[2]通过分析学习者的面部表情,将学习者情绪划分为高积极情绪、低积极情绪、高消极情绪、低消极情绪四种,教师可利用分析结果改进教学策略,提高学习效率。A. Tzanavari等人^[3]基于多媒体技术分析学生在学习过程中的情绪、行为和状态等,建立了一个基于虚拟现实的在线学习系统。D' Mello S^[4]通过捕捉学习者学习过程中椅背上的压力变化、语音

收稿日期:2020-12-13

修回日期:2021-04-14

基金项目:2019年湖北省知识产权运用示范工程(高校院所)后补助项目(35)

作者简介:姚远(1974-),男,副教授,博士,硕士,从事信息处理、嵌入式系统应用、测控网络等相关领域的研究;徐晶晶(1996-),女,硕士,从事图像处理相关领域的研究。

信息以及面部特征,判断学生处于厌烦、高兴、疑问等情感状态中的哪一种。Tseng 等人^[5]提出了一种基于人脸标志点位置变化率的学生注意力检测算法。吴沧海等人^[6]通过对学习者表情、眼部特征、身体姿态的分析,进行情感计算得出学习者的状态。张建利等人^[7]结合情境感知、面部特征提取与图像识别技术,对学习者的学习状态进行研究。熊碧辉等人^[8]基于眼睛闭合、头部偏转以及视线特征,提出了一种融合头姿检测与视线检测的注意力检测方法。但是以上方法存在一定的弊端,一些基于捕捉人体生理信号的方法,侵入性较强,影响学生学习状态,且对采集设备要求较高;基于图像处理技术的方法多是基于 2D 图像处理技术,受环境因素(如光线)影响较大,缺乏对 3D 信息的捕捉和分析。

因此,文中基于 2D 和 3D 图像处理技术,提出了一种在线评估学习状态的方法。该方法利用 AdaBoost 算法与 ResNet 模型检测与识别人脸,利用 EPNP 算法获取学习者头部姿态评估学习者的分心状态,通过眼睛和嘴巴的特征预警学习者的疲劳状态,为在线学习者提供积极的个性化服务。

1 基于 2D 图像处理技术的在线学习者身份识别

优质线上内容的打造需要耗费大量的人力、物力和财力,当前盗版、非法获取在线资源等问题影响和制约了在线教育的发展。文中基于 2D 图像处理技术对学习者的身份鉴别,当识别到观看视频的人员身份异常时,发出预警信息随之停止播放课程,有效避免了在线课程内容的泄露。

1.1 基于 AdaBoost 算法的人脸检测

基于 Haar-like 特征的 AdaBoost 算法可实现人脸的快速检测^[9]。Haar-like 特征模板内有白色和黑色两种矩形,将该模板的特征值定义为白色矩形像素和与黑色矩形像素和的差值。AdaBoost 算法利用 Haar-like 特征将 2D 彩色图像中人脸特征量化,从而区分人脸和非人脸。算法的核心思想为:首先对样本集中的每个训练样本进行权值初始化,随后进行训练并且更新权值,得到不同的弱分类器,然后将弱分类器根据其分类能力从低到高进行排列级联,从而将单个弱分类器组合成更加复杂的强分类器,最终得到检测器。

AdaBoost 算法对不同表情、是否佩戴眼镜以及不同光照等状态的人脸图片检测表现出了良好的性能,但在人脸偏转角度过大时会出现漏检。文中在此基础上增加了肤色特征^[10],基于简单高斯模型建立肤色模型,随后在肤色模型的基础上利用 Otsu 自适应阈值分割算法选取分割阈值,对人体脸部肤色进行分割,对

于漏检图像进行二次检测。根据不同亮度与色度的图像,分割阈值范围在 $[0.45, 0.55]$ 之间。

1.2 基于 ResNet 模型的人脸识别

针对 CNN 网络因模型深度的不断加深而出现的梯度爆炸和梯度消失现象,何凯明等人提出了基于残差网络结构的 ResNet 模型^[11]。ResNet 模型把残差网络引入到深度卷积神经网络中,通过残差块把训练好的浅层网络与自身映射的增加层巧妙地连接在一起,使得当网络深度的增加而导致模型准确率下降时,返回之前的浅层网络,残差网络结构如图 1 所示。

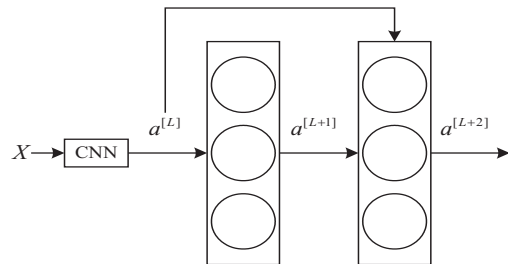


图 1 残差网络结构

图 1 中网络的输入为 x , 输出为第 L 层的激活值 $a^{[L]}$, 权重矩阵为 $W^{[L+1]}$, 偏差因子为 $b^{[L]}$, 根据公式(1)可得线性激活输出为 $Z^{[L+1]}$, 再通过线性整流函数(rectified linear unit, ReLU) g , 由公式(2)可得非线性激活输出 $a^{[L+1]}$ 。

$$Z^{[L+n]} = W^{[L+n]} a^{[L+n-1]} + b^{[L+n-1]}, n = 1, 2 \quad (1)$$

$$a^{[L+1]} = g(Z^{[L+1]}) \quad (2)$$

$$a^{[L+2]} = g(Z^{[L+2]} + a^{[L]}) \quad (3)$$

同理,根据公式(3)得到下一轮的非线性激活值 $a^{[L+2]}$,在进行第二轮非线性激活之前加入了第 L 层的激活值 $a^{[L]}$,即将上一层网络输出信息直接传送到网络深层,实现了非相邻层之间的线性传递,加速网络收敛,从而避免了梯度消失和爆炸问题。

文中利用 Dlib 库的 ResNet 人脸识别模型进行人脸识别,获取待测人脸 128 维特征向量时耗费大量资源,无法实现实时监测,所以在人脸识别模块加入质心跟踪算法^[12],利用欧氏距离最小的原则,将相邻帧获取的人脸框质心联系在一起,获取目标在连续帧中的变化轨迹。对于同一个目标,只对初始帧进行特征描述算子提取,其余后续帧只进行人脸检测工作,大大提高了人脸识别检测效率。

2 基于 3D 图像处理技术的在线学习者姿态评估

学习者注意力是否集中与其头部姿态密切相关。当学习者以注意力集中的状态学习时会不自主地靠近并且正视计算机屏幕,反之,当学习者分心时,会出现左顾右盼、长时间低头、抬头等动作,此时头部不再正

对电脑,因此可通过识别人的头部姿态判断学习者的注意力集中程度。

头部姿态角由俯仰角(pitch)、偏航角(yaw)与翻滚角(roll)3个自由度定义。头部姿态检测算法的步骤分为:2D人脸关键点检测;3D人脸模型匹配;求解3D人脸关键点对应2D人脸关键点转换关系;根据旋转矩阵求解欧拉角。其中3D人脸模型常用的获取方法有:(1)利用不可变形的三维人脸模型;(2)利用模型匹配算法拟合三维人脸模型。方法1无法获取精确世界坐标系人脸三维特征点,误差较大;方法2模型匹配算法的重建精度严重依赖于所使用模型,而模型的训练对数据的采集及处理要求很高。因此文中利用深度摄像头获取RGB-D图像,基于3D图像处理技术构建人脸部三维模型。姿态计算使用EPNP迭代算法将从彩色图像获取的二维人脸特征点映射到构建好的三维人脸模型中,获取头部姿态旋转矩阵 \mathbf{R} 与平移矩阵 \mathbf{T} 。

2.1 基于RGB-D图像的人脸三维特征点提取

人脸特征点定义为人脸面部的关键区域位置,例如眼角、嘴角、鼻尖、唇、脸颊轮廓等。文中首先利用级联回归树算法从RGB图像中提取人脸68个二维特征点位置向量^[13],然后利用RGB-D图像将二维人脸特征点映射到三维空间中,以获取高精度三维人脸特征点坐标。

RGB-D数据是将同一时刻采集的深度数据和彩色数据对齐和融合的结果。假设物体表面上一点在深度相机与彩色相机坐标系下的空间坐标分别为 \mathbf{P}_{de} 、 \mathbf{P}_{rgb} ,该点在深度图与彩色图像上的像素坐标分别为 \mathbf{p}_{de} 、 \mathbf{p}_{rgb} ,深度相机与彩色相机内参分别为 \mathbf{H}_{de} 、 \mathbf{H}_{rgb} ,则变换关系如式(4)所示:

$$\begin{cases} \mathbf{p}_{de} = \mathbf{H}_{de} \mathbf{P}_{de} \\ \mathbf{p}_{rgb} = \mathbf{H}_{rgb} \mathbf{P}_{rgb} \end{cases} \quad (4)$$

\mathbf{P}_{rgb} 与 \mathbf{P}_{de} 之间的关系可通过式(5)表示:

$$\mathbf{P}_{rgb} = \mathbf{R} \mathbf{P}_{de} + \mathbf{T} \quad (5)$$

式中, \mathbf{R} 与 \mathbf{T} 为深度相机坐标系变换到彩色相机坐标系的旋转矩阵与平移矩阵。在同一场景下,通过标定获取深度相机与彩色相机的内外参矩阵,由公式(4)、公式(5)即可通过RGB-D图像中某一点的像素值获取该点的深度信息,进而获取该点的三维坐标。

2.2 EPNP算法

EPNP算法是一种非迭代的位姿估计算法。算法的核心是将世界坐标系的3D坐标表示为4个不共面虚拟控制点的加权和^[14],求解4个虚拟控制点在相机坐标系下的坐标,将求解2D-3D的PnP问题转化为求解3D-3D的刚体变化问题。

世界坐标系的 n 个3D参考点坐标为 $\mathbf{p}_i^w (i=1,2,$

$\dots, n)$,对应相机坐标系坐标为 $\mathbf{p}_i^c (i=1,2,\dots,n)$ 。4个虚拟控制点在世界坐标系的坐标为 $\mathbf{C}_j^w (j=1,2,3,4)$,对应相机坐标系的坐标为 $\mathbf{C}_j^c (j=1,2,3,4)$,则有如下关系成立:

$$\begin{cases} \mathbf{p}_i^w = \sum_{j=1}^4 a_{ij} \mathbf{C}_j^w, \sum_{j=1}^4 a_{ij} = 1 (i=1,2,\dots,n) \\ \mathbf{p}_i^c = \sum_{j=1}^4 a_{ij} \mathbf{C}_j^c, \sum_{j=1}^4 a_{ij} = 1 (i=1,2,\dots,n) \end{cases} \quad (6)$$

设 \mathbf{p}_i^w 对应的像素坐标系坐标为 (u_i, v_i) ,相机坐标系下的第 $j (j=1,2,3,4)$ 个虚拟控制点坐标为 (x_j^c, y_j^c, z_j^c) ,结合相机坐标与像素坐标的转化关系^[15],当有 n 对空间参考点与像素坐标点一一对应时,得到 $2n$ 个方程:

$$\mathbf{M} \mathbf{x} = 0 \quad (7)$$

其中, \mathbf{M} 是 $2n \times 12$ 的矩阵; $\mathbf{x} = [\mathbf{C}_1^{c^T}, \mathbf{C}_2^{c^T}, \mathbf{C}_3^{c^T}, \mathbf{C}_4^{c^T}]^T$ 为 12×1 维向量,其中包含4个虚拟控制点在相机坐标系下的非齐次坐标。

方程组式(7)的解 $\mathbf{x} \in \ker(\mathbf{M}^T \mathbf{M})$,其中 $\mathbf{M}^T \mathbf{M}$ 的零特征值维数与空间参考点的特征以及相机焦距有关,求解方程通常选择4组以内的零特征值对应的特征向量,进而将求解2D-3D的PnP问题转化为绝对定位问题,获得摄像头的位姿参数,即文中所求的头部姿态参数。

3 基于特征分析的在线学习者疲劳预警

在线学习者疲劳预警主要是通过分析学习者的眼睛、嘴部的疲劳特征判断人是否处于疲劳状态,当学习者处于疲劳状态时,进行自动预警。

3.1 眼部特征判定疲劳状态

人眼在正常状态下每分钟眨眼15~20次,每次眨眼大概需要0.4s,眨眼间隔约为3s^[16]。当人处于疲劳状态时,人眼持续闭合时间增长,眨眼频率会明显增高或减少。文中采用眼睛纵横比(eye aspect ratio, EAR)作为评价指标,判断眼睛处于睁开或闭合中的哪种状态。当人眨眼时,眼睛纵横比会先减小后增大,因此可通过眼睛纵横比在一段时间内的变化规律判断是否为眨眼动作。眼睛纵横比值的计算步骤为:

(1)通过彩色图像帧获取人眼部特征点(如图2所示)。

(2)计算出左右眼睛纵横方向相应位置的欧几里得距离,获取左右眼睛的纵横比。计算左右眼睛平均纵横比值,并以此代表眼睛纵横比值,计算公式如下:

$$\text{EAR}_{\text{left}} = \frac{\|\mathbf{P}_{37} - \mathbf{P}_{41}\| + \|\mathbf{P}_{38} - \mathbf{P}_{40}\|}{2 \|\mathbf{P}_{36} - \mathbf{P}_{39}\|} \quad (8)$$

$$\text{EAR}_{\text{right}} = \frac{\|\mathbf{P}_{43} - \mathbf{P}_{47}\| + \|\mathbf{P}_{44} - \mathbf{P}_{46}\|}{2 \|\mathbf{P}_{42} - \mathbf{P}_{45}\|} \quad (9)$$

$$EAR_{avg} = \frac{EAR_{left} + EAR_{right}}{2} \quad (10)$$



图 2 眼睛特征点分布(左眼特征点模型、睁开、闭合)

3.2 嘴部特征判定疲劳状态

嘴部状态主要有闭合、说话以及打哈欠三种。说话过程中人的嘴巴上下张开幅度较小,且变动频繁,嘴型持续时间较短,与打哈欠时的嘴巴状态正好相反。研究资料表明,当人在打哈欠时其嘴部将处于张大嘴的状态 3 s 以上^[17],当嘴部状态改变时,嘴巴纵横比值(mouth aspect ratio, MAR)随之改变。文中嘴巴纵横比值的计算步骤为:



图 3 嘴部特征点分布(嘴部特征点模型、闭合、半张开、完全张开)

4 仿真实验分析

学习者的学习视频采集摄像头:英特尔公司 Intel® RealSense™ D400, D415, 其采样频率为 30 帧/s, 图像分辨率为 640 × 480; 图像仿真分析计算机配置: Intel (R) Core (TM) i5, 1.8 GHz, 8G 内存; 仿真软件为 PyCharm; 编程语言为 Python。

4.1 学习者身份识别实验

文中利用 D415 中的 2D 彩色摄像头进行学习者身份识别实验。实验利用 Dlib 库的人脸框位置检测器 `dlib.get_frontal_face_detector()` 检测摄像头前的人脸数 n , 利用 Dlib 库的人脸识别模型 `dlib_face_recognition_resnet_model_v1.dat` 获取待测人脸 128 维特征向量, 使用欧氏距离与已录入数据库人脸特征数据进行匹配, 使用最近邻分类器返回样本的标签。

通过测试发现实时提取当前帧人脸特征向量平均耗时 0.2 s, 导致图像帧率低于 4 帧/s, 影响后续模块检测结果, 因此文中添加质心目标跟踪。当摄像头前人脸数 $n \geq 1$ 时, 提取初始帧图像进行人脸识别, 利用质心追踪算法判断视频流目标变化情况, 利用欧氏距离最小的原则, 将不同帧获取的人脸框的质心联系在一起, 在此过程中有人加入或退出, 则提取变化的初始帧进行人脸识别, 当摄像头前人脸数不变时, 只进行人脸检测。通过实验发现, 加入质心目标跟踪算法后, 界面输出的帧率从 2 帧/s ~ 3 帧/s 提高到 17 帧/s ~ 20

(1) 通过彩色图像帧获取人嘴部特征点(如图 3 所示)。

(2) 计算出嘴巴对应点位置的欧几里得距离, 获取嘴巴的嘴巴纵横比值。嘴巴纵横比值的计算公式为:

$$MAR = \frac{\|P_{61} - P_{67}\| + \|P_{62} - P_{66}\| + \|P_{63} - P_{65}\|}{3 \|P_{60} - P_{64}\|} \quad (11)$$

帧/s, 识别效率明显提高。

为验证 ResNet 模型算法的有效性, 将 ResNet 模型算法在人脸数据集 LFW 上进行测试, 识别率达到 99.38%, 满足了身份鉴权防止学习内容泄露的需求。

4.2 学习者异常姿态检测实验

文中利用 Dlib 库官方提供的分类器 `shape_predictor_68_face_landmarks.dat` 提取人脸二维特征点, 获得彩色图像像素坐标系下 68 个二维特征点的位置信息。通过 D415 的 3D 红外摄像头获取深度图, 将彩色图与深度图对齐融合获得 RGB-D 图像, 利用 Realsense D415 SDK 中的 `rs.rs2_deproject_pixel_to_point(intrinsics, (x, y), d)` 获取人脸 68 个三维特征点坐标值。

为减少计算量, 在上述 68 个特征点中选取 14 个 2D-3D 特征点对作为 EPNP 算法的输入。实验结果如图 4 所示, 图中实心矩形点表示实时采集的 14 点人脸二维特征点, 空心原点表示 EPNP 算法得出的位姿参数对人脸三维特征点反投影得到的二维点。图 4(a) 表示利用文献[18]中的方法, 将来自科英布拉大学系统与机器人研究所的 3D 人脸模型特征点代替世界坐标系三维特征点位置获取的仿真结果。图 4(b) 表示利用 D415 获取的人脸三维特征点作为真实世界坐标系三维坐标位置的仿真结果。

由仿真结果可知, 利用文中方法 EPNP 算法获取头姿一致性更好。

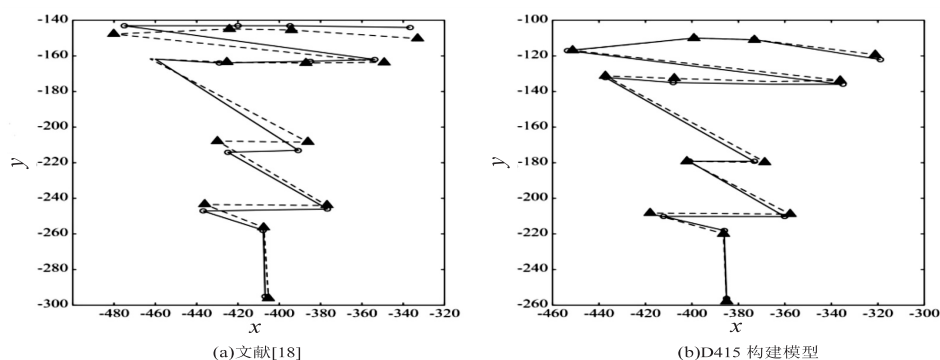


图 4 三维人脸模型仿真

学习者姿态预警信息的获取与设定的头部偏转欧拉角阈值紧密相关。文中进行了头部俯仰偏转与左右偏转动作测试以获取阈值。测试结果显示,头部俯仰偏转主要改变俯仰角(pitch)的大小,头部左右偏转主要改变偏航角(yaw)的大小。根据测试数据设定头部俯仰偏转阈值为 $\text{pitch} = \pm 15^\circ$, 头部左右偏转阈值为 $\text{yaw} = \pm 30^\circ$ 。当 $\text{pitch} > 15^\circ$ 或 $\text{pitch} < -15^\circ$ 时,判断学习者处于头部俯仰偏转状态;当 $\text{yaw} > 30^\circ$ 或 $\text{yaw} < -30^\circ$ 时,判断学习者处于头部左右偏转状态。在真实的学习场景中,学习者以正常状态学习时,也有可

能出现头部姿态异常的情况,例如学生暂时低头看书、轻微点头或摇头等。为防止误测,文中设定异常头部姿态时间阈值为 4 s,当学生头部处于异常姿态的时间超过阈值时,则判定为分心状态。

为验证文中所用姿态检测算法的有效性,以 1 分钟为单位采集 5 组视频图像,图像连续 4 s 以上处于姿态异常状态,则判定分心,测试结果如表 1 所示。根据测试结果可知,文中算法头部姿态检测可以实现异常姿态导致的分心状态检测。

表 1 分心状态测试

状态	第 1 组	第 2 组	第 3 组	第 4 组	第 5 组
是否超出 4 s 低头抬头	是	否	是	否	否
是否超出 4 s 左右转头	是	是	否	否	是
是否分心	是	是	是	否	是

4.3 学习者疲劳状态预警实验

人的眨眼频率与哈欠等疲劳特征的获取与人眼睛纵横比、嘴巴纵横比的阈值设定密切相关,对两名实验者分别进行了眨眼以及哈欠动作检测以获取所需阈值。分析检测结果可知,当 $\text{EAR} < 0.23$ 时人的眼睛处于闭合状态,所以设定 EAR 的阈值为 0.23。当学习者眼睛在 0.4 s 时间内 EAR 先减小到小于阈值再增大时,则判断眨眼一次,每分钟眨眼频率 $f_{\text{blk}} < 10$ 或 $f_{\text{blk}} > 20$ 时,则判断学生已处于疲劳状态。当 $\text{MAR} > 0.6$ 时人处于张大嘴状态,所以将 0.6 设定为张大嘴的阈值,

当嘴部超过 3 s 时间处于张大嘴状态时,则判断打哈欠一次,当每分钟内哈欠次数达到 2 次及以上,则判定学生出现疲劳。

为验证该算法的有效性,以 1 分钟为单位采集 5 组视频图像进行检测,获取学习者的每分钟眨眼频率与哈欠次数,判定学习者是否处于疲劳状态,测试结果如表 2 所示。根据测试结果可知,综合考虑学习者的眨眼频率与哈欠次数,可比较有效地实现学习者疲劳状态的判定。

表 2 疲劳状态测试

状态	第 1 组	第 2 组	第 3 组	第 4 组	第 5 组
眨眼频率	5	15	22	15	25
哈欠次数	2	2	1	3	2
是否疲劳	是	否	是	是	是

4.4 系统实验与结果分析

文中由上述身份识别、姿态检测、疲劳检测三个模块构成在线学习检测系统,并且将系统检测结果数据上传至 Web 端平台,对学生的上课异常状态进行实时记录,教师可通过平台实时查看学生学习异常状态。

系统部分实验结果如图 5 所示,Web 端平台数据显示如图 6 所示。图 5(a) 为学生处于正常上课状态;图 5(b) 为摄像头前出现未授权用户时,异常状态报警框显示“未授权用户”;图 5(c) 中,摄像头前人脸数 $n = 0$,异常状态报警框显示“脱离学习座位”;图 5(d) 摄

像头前出现人脸数 $n > 1$, 异常状态报警框显示“多人”; 图 5(e) 为学生长时间东张西望时, 异常状态报警框显示“东张西望, 分心”; 图 5(f) 为学习者眨眼频率

过快或过慢, 异常状态报警框显示“眨眼频繁, 疲劳预警”。



图 5 在线学习检测系统部分实验结果

在线学习监测系统	
查询	位置: 首页 > 预警查询 > 查看学生异常状态
教师信息	
预警信息查询	
学生学习异常状态信息	
学生姓名	异常学习状态
徐晶晶	东张西望, 分心!!!
徐晶晶	哈欠频繁, 疲劳预警!!!
徐晶晶	眨眼过快或过慢, 疲劳预警!!!
徐晶晶	眨眼过快或过慢, 疲劳预警!!!
徐晶晶	眨眼过快或过慢, 疲劳预警!!!
徐晶晶	眨眼过快或过慢, 疲劳预警!!!
徐晶晶	眨眼过快或过慢, 疲劳预警!!!
徐晶晶	哈欠频繁, 疲劳预警!!!

图 6 Web 端数据显示

综合分析实验结果, 将本系统与文献[7]与文献[19]中的方法进行对比, 结果见表 3。文中方法相对于文献[7]方法功能更加完善, 相对于文献[19]方法在检测准确度方面更有优势。本系统基本可以实现学生学习状态的评估, 并且在学生出现异常状态时进行实时预警提醒, 但是该方法假设学习者的学习环境背景简单, 且不需要长时间低头做笔记、长时间大声诵读等, 所以在环境背景复杂的情境下一定程度上影响了

系统检测结果, 这也是接下来要研究的重点内容。

表 3 实验对比结果 %

状态	文献[7]	文献[19]	文中
无人	100	80	100
多人	100	80	100
用户未授权	无	无	100
分心	87.5	60	84
疲劳	75	60	82

5 结束语

文中利用深度摄像机 D415 实时获取学习视频,通过人脸识别、头部姿态识别、疲劳检测三个模块,基于 2D 和 3D 图像处理技术开展了无人、多人、用户未授权、分心以及疲劳等学习状态的评估,仿真实验测试了该方法的可行性,为在线教育随时调整教学策略提高在线学习的质量提供了积极思路。

参考文献:

- [1] 朱悦,黄荣伟,黄盛月. 在线学习环境下大学生持续线上学习意愿的影响因素[J]. 浙江师范大学学报:自然科学版,2020,43(4):388-395.
- [2] GUPTA S K, ASHWIN T S, GUDDATI R M R. Students' affective content analysis in smart classroom environment using deep learning techniques[J]. Multimedia Tools and Applications, 2019, 78(18): 25321-25348.
- [3] VESIN B, IVANOVIC M, MILICEVIC A K, et al. Rule-based reasoning for building learner model in programming tutoring system[C]//Advances in web-based learning - ICWL 2011-10th international conference. Hong Kong, China: Springer-Verlag, 2011.
- [4] GRAESSER A, D' MELLO S. Automatic detection of learner's affect from gross body language[J]. Applied Artificial Intelligence, 2009, 23(2): 123-150.
- [5] CHEN Yung-Hui, TSENG Chun-Hsiung. A camera-based attention level assessment tool designed for classroom usage[J]. Journal of Supercomputing, 2018, 74(11): 5889-5902.
- [6] 吴沧海,熊焕亮,王映龙. 远程学习中学习状态判断的情感计算研究[J]. 软件导刊:教育技术,2013,12(7):24-27.
- [7] 张建利,段富. 基于面部特征的学习状态的研究[J]. 计算机工程与设计,2013,34(7):2460-2464.
- [8] 熊碧辉,周后盘,黄经州,等. 一种融合视线检测的注意力检测方法[J]. 软件导刊,2018,17(7):31-36.
- [9] 李佩,汪红娟,李业丽,等. 基于 AdaBoost 人脸检测算法的分析研究[J]. 北京印刷学院学报,2020,28(1):128-132.
- [10] 唐云龙,何麒,陈平. 基于肤色特征的人脸检测方法研究[J]. 电脑知识与技术,2019,15(19):198-200.
- [11] 吴宇豪,陈晓辉. 基于改进的 ResNet 的人脸表情识别系统[J]. 信息通信,2020(7):37-39.
- [12] 梁习卉子,陈兵旗,李民赞,等. 质心跟踪视频棉花行数动态计数方法[J]. 农业工程学报,2019,35(2):175-182.
- [13] 傅杰,黄树成. 多特征融合的级联回归人脸对齐方法研究[J]. 江苏科技大学学报:自然科学版,2020,34(3):54-59.
- [14] 张慧娟,熊芝,劳达宝,等. 基于 EPNP 算法的单目视觉测量系统研究[J]. 红外与激光工程,2019,48(5):190-195.
- [15] 汪佳宝,张世荣,周清雅. 基于视觉 EPNP 加权迭代算法的三维位移实时测量[J]. 仪器仪表学报,2020,41(2):166-175.
- [16] 龚淑娟. 基于面部多指标的货运司机疲劳驾驶检测[J]. 物流技术,2020,39(9):74-81.
- [17] 谢国波,陈云华,张灵,等. 基于嘴巴特征点曲线拟合的哈欠检测[J]. 计算机工程与科学,2014,36(4):731-736.
- [18] 武君,白刚,张彩霞. EPNP 和 POSIT 算法在头部姿态估计上的实验比较与分析[J]. 北方工业大学学报,2017,29(2):19-27.
- [19] 卫晓娜. E-Learning 系统中基于 Adaboost 算法的注意力识别[J]. 电脑知识与技术,2010,6(10):2453-2456.

CNCC2021 坚定走来! 让我们深圳一“会”!

因疫情防控需要,CNCC 首次出现延期。经过多方努力,CNCC2021 定于 2021 年 12 月 16-18 日在深圳举行。本年度大会以“计算赋能加速数字化转型”为主题,设有 18 个特邀报告、3 场大会主题论坛,113 场前沿技术论坛和丰富的活动及展览展示。



(中国计算机学会 <https://www.ccf.org.cn/Focus/>)