

基于人脸检测和模糊综合评判的在线教育专注度研究

钟马驰 张俊朗 蓝扬波 何月铎

华南师范大学计算机学院 广州 510631

摘要 针对在线教育中对学生的监督手段较为匮乏的问题,基于人脸检测的模糊综合评判算法通过人脸图像检测头部的左(右)转头转角、头部抬(低)头转角、眼睛闭合度、嘴巴闭合度和面部表情,使用头部的左(右)转头转角和头部抬(低)头转角对头部姿态进行评分,在眼睛闭合度和嘴巴闭合度的检测结果的基础上对疲劳度进行评分,结合面部表情的检测结果对情绪进行评分,再根据三个评分使用模糊综合评判方法量化评判其学习专注度。该算法应用于在线教育平台的学生课堂专注度评估,有助于授课老师及时获取网络课堂中学生的课堂专注度,为改进教学计划和督促学生学习提供助力。基于此算法设计的在线教育平台的学习专注度检测系统经模拟场景使用测试,能有效地根据人脸检测结果评估学生的课堂专注度,提升了课堂质量和学生的学习效果。

关键词: 在线教育;课堂监督;人脸检测;模糊综合评判;专注度

中图法分类号 TP183

Study on Online Education Focus Degree Based on Face Detection and Fuzzy Comprehensive Evaluation

ZHONG Ma-chi, ZHANG Jun-lang, LAN Yang-bo and HE Yue-hua

College of Computer Science, South China Normal University, Guangzhou 510631, China

Abstract Aiming at the problem of less supervised means for students in online education, the fuzzy comprehensive evaluation algorithm based on face detection detects the head's left (right) turning angle, head lifting (low) angle, eye closure, mouth closure and facial expression through face image detection, uses the left (right) turning angle of the head and the head lifting (low) angle to score the head posture. Based on the results of eye closure and mouth closure, the fatigue score is evaluated. Combined with facial expression detection results, the emotion score is evaluated. Then the fuzzy comprehensive evaluation method is used to quantitatively judge the concentration degree of learning according to the scores of head posture, fatigue and emotion. The algorithm is applied to the evaluation of students' classroom concentration on the online education platform, helping the instructors to timely acquire the classroom concentration of the students in the online classroom, and provide assistance for improving the teaching plan and urging students to learn. The learning concentration detection system of the online education platform designed based on this algorithm is used to simulate the scene use test, which can effectively evaluate the classroom concentration of the students according to the face detection result, and improve the classroom quality and the learning effect of the students.

Keywords Online education, Classroom supervision, Face detection, Fuzzy comprehensive evaluation, Concentration

1 引言

在线教育作为一种全新的教育模式,由于其开放性、多元化等特点,逐渐受到了学生和家长的欢迎,在线见教育市场也迎来了蓬勃发展,但在实际使用过程中还有许多不足之处。最大的不足在于课堂中教师对学生的监督效果差。在传统的教学过程中,师生都处于一个有限的空间中,讲台的位置非常有利于快速观察整个教室的情况,在上课过程中,教师只需目光一扫,便可掌控全场,有效监督学生。但在在线教育中,空间的限制被打破,教师大部分注意力都用在讲课上,剩余的注意力压根无法有效掌控全部学生的状态和保证学生的上课质量质量,更无法发挥监督作用;对于学生而言,在线教育虽然选择多、时间灵活,但是对学生的自律程度要求非常高,而部分在线教育平台的统计数据的结果显示,在线课堂存在明显的学习者自主性不佳的现象,导致各类课程普遍是注册

率高但课程通过率低^[1-2]。2020年新冠肺炎疫情期间,不少家长反映,由于网课的任课老师无法有效监督学生,孩子在家上网课的学习效果非常低。这不仅使得许多学生浪费了宝贵的学习时间,同时也不利于老师对教学进度和质量的把控。

这次疫情是在线教育第一次短时间内取代传统教育的大规模应用,虽然目前它还无法完全替代传统教育,但是随着5G技术的产生和逐步投入使用,我们看到了在线教育成为传统教育的重要补充的发展趋势。为此,针对在线教育中老师无法有效监督学生的缺点,探索监督系统成为在线教育发展过程中绕不开的问题。

2 相关工作

在对在线教育的监督系统的探索中,Wang^[3]和Yuan^[4]分别使用情感计算和表情分类来检测被检测者的情绪和表情,但是他们并没有对检测解雇,Li^[5]首先采用Adaboost算

法检测人脸,定位人脸的眼睛和嘴巴,并分析眼睛闭合状态和嘴巴闭合状态,再根据眼睛闭合状态和嘴巴闭合状态的特征数据,采用模糊推理的方法把学习者的学习状态归为专注、正常、疲劳3种专注度类型。Pan^[6]使用 Adaboost 算法检测人脸,若算法失效,则判断无人,否则定位眼睛和嘴角位置并提取特征数据,进而计算眼睛间距和嘴角弧度,根据计算结果判断学习者的表情和是否瞌睡,并把检测数据和处理结果展示给学习者。Xiong^[7]使用 Adaboost 算法检测人脸,若算法失效则判断无人,否则截取眼部区域,进行眼睛开闭检测,根据 PERCLOS 原理评估学习者的学习状态,提取眼部轮廓信息和定位虹膜中心位置计算视线,判断分心状态,另外还使用 OpenCV 中 3D 姿态评估算法 POSIT 来计算学习者的面部偏转角,判断分心情况,最后综合以上 3 种分析来判断学习者的专注度。

通过以上的相关研究可以发现,在线教育现阶段的检测因子可归结为头部姿态、面部表情、眼睛状态、嘴巴状态和视线,虽然以上研究中每一个系统出于各自的考量都不能同时兼顾这 5 个因子,但检测因子都是这些因子的组合,这在一定

程度上说明这 5 个因子被广泛认定为是影响在线教育学习者专注度的重要因素。在以上的研究中(如表 1 所示),Li 同时对面部表情、眼睛状态和嘴巴状态 3 个因子进行检测,评估因子多,有利于提高评估结果的真实性,但是其最终评判结果仅将专注度划分为专注、正常、疲劳 3 个等级,分级太少,反馈信息准确度低,不利于学习者和老师更加准确地把握其学习专注度,实际应用效果会打折扣。Pan 对学习者的面部表情、眼睛状态和嘴巴状态都进行了检测,但他也仅仅提取了学习者的表情特征和判断是否瞌睡,把结果呈现给学习者。这些结果过于生硬,普通用户很难根据这些结果一目了然地获知学习者当前的学习状态,因此其实用性还有所不足。Xiong 对学习者的脸部偏转角、眼睛状态和视线进行了检测,做了无人检测、闭眼检测、面部偏转角检测和视线偏转检测,检测因子多而全面,评估结果真实性较高,但是其最终评判结果仅为学习者是否分心,过于单一,纵使是不分心,仍然存在不同学习者的学习状态有很大差距的情况。该系统的评测结果不利于教师准确掌握学习者的学习情况和实现监督的目的。

表 1 应用系统列表

Table 1 Application system list

对象	检测因子					评判步骤	评判结果
	头部姿态	面部表情	眼睛	嘴巴	视线		
王济军		✓				情感计算	判断和理解情绪
袁源		✓				表情分类	表情
李文倩		✓	✓	✓		表情模糊推理	3 个专注度类型
潘翔		✓	✓	✓		提取表情特征、判断瞌睡	表情特征和瞌睡判断结果
熊碧辉	✓		✓		✓	无人评判、闭眼检测、面部偏转角、视线偏转	是否分心
本文	✓	✓	✓	✓		模糊综合评判	量化评分

针对以上系统的不足,本文提出了一种基于头部姿态、面部表情、眼睛状态和嘴巴状态 4 个因子检测,使用模糊综合评判,把学习者的学习专注度量表示为 0~100 之间的分数的算法。算法包括图像预处理模块、人脸特征定位与提取模块、头部姿态评估模块、面部表情识别模块、眼睛状态评估模块、嘴巴状态评估模块和模糊综合评判模块。在人脸特征定位与提取中,包括检测、截取图片中的最大人脸区域的图像和采集人脸 68 个特征点。在头部姿态评估中,根据提取的人脸特征点,采用 OpenCV 的 solvePnP 函数来评估头部姿态。在面部表情识别中,首先使用 CNN 算法基于 fer2013 数据集训练表情识别模型,再使用模型检测截取的人脸图像的表情。在眼睛状态评估中,根据眼部特征点计算眼睛的闭合度。在嘴巴状态评估中,根据嘴部特征点计算嘴巴的闭合度。在模糊综合评判中,设定检测周期,根据周期内的每一次检测结果计算头部偏转角均值,评估头部姿态评分,并使用模糊综合评判计算情绪综合评分,使用 PERCLOS 疲劳度检测算法计算疲劳度,最后采用模糊综合评判,把头部姿态评分、情绪评分和疲劳度组成模糊集,使用层次分析法来确定各个因素权重,量化评判具有模糊性的评判目标——学习专注度。

本文之所以没有检测视线,一方面是受光线、设备摄像头分辨率低、人脸距离摄像头远等因素的影响,图像中虹膜区域图像的面积较小且模糊,检测难度较大;另一方面,通过对检测对象的头部姿态评估,能大致判断其注意力区域,从而替代视线检测。

3 算法的提出与设计

3.1 头部姿态评估

根据心理学的研究,注意是一种内部心理状态,但是它可以通过人的外部行为表现出来。当一个人注意一个对象时,其感觉器官往往会朝向该对象,以便获得更清晰、更全面的信息^[8]。

在学习的过程中,学生面部朝向可以在很大程度上反映学生注意力的关注点,故检测脸部朝向就能大致判断出视线方向和关注点。因此本文采用检测面部朝向的方法来判断学习者把注意力放在电脑屏幕上的程度,进而判断其课堂专注度。

本文采用 OpenCV 和 Dlib 库实现头部姿态评估。首先使用 Dlib 库检测出人脸图像中左眼的左角、右眼的右角、笔尖、下巴和左(右)嘴角 6 个特征点(如图 1 所示);再使用 OpenCV 的 solvePnP 函数计算出头部的回转向量和翻转向量;在这两个向量的基础上,使用 OpenCV 的 norm 函数计算绕旋转轴的旋转角度和在 X、Y、Z 方向上的分向量 $\cos(\beta_x)$, $\cos(\beta_y)$, $\cos(\beta_z)$ 。构造四元素表示如下:

$$w = \cos\left(\frac{\alpha}{2}\right) \quad (1)$$

$$x = \sin\left(\frac{\alpha}{2}\right) \cos(\beta_x) \quad (2)$$

$$y = \sin\left(\frac{\alpha}{2}\right) \cos(\beta_y) \quad (3)$$

$$z = \sin\left(\frac{\alpha}{2}\right) \cos(\beta) \quad (4)$$

转换为欧拉角:

$$\begin{bmatrix} \phi \\ \theta \\ \psi \end{bmatrix} = \begin{bmatrix} \text{atan2}(2(wx + yz), 1 - 2(x^2 + y^2)) \\ \arcsin(2(wy - zx)) \\ \text{atan2}(2(wz + xy), 1 - 2(y^2 + z^2)) \end{bmatrix} \quad (5)$$

其中, ϕ, θ, ψ 分别为头部分别围绕相机设定的 X 轴、 Y 轴和 Z 轴的旋转夹角 $pitch, yaw$ 和 $roll$ (如图 2 所示), 其中 $pitch$ 角可判断头部的抬(低)头, yaw 角可判断头部的左(右)转头。通过这两个角度即可评估头部姿态。

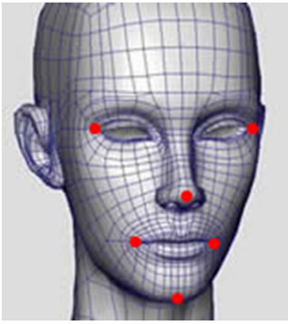


图 1 面部的 6 个特征点 (图源: Free3D)

Fig. 1 6 feature points of face (Source: Free3D)

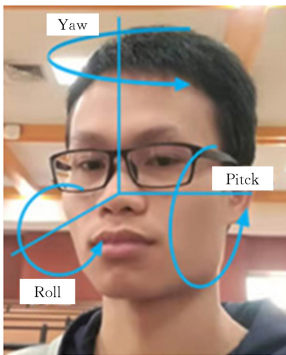


图 2 头部的三维旋转坐标

Fig. 2 3 dimensional rotation coordinates of the head

3.2 疲劳度检测

Wierwille 等^[9]通过实验发现, 眼睛的闭合时间在一定程度上反映了疲劳度, 提出了 PERCLOS 疲劳度检测算法, 根据单位时间内闭眼的时间与单位时间之比判定疲劳度, 这个比值被称为 PERCLOS 值, 其越接近 1, 表示疲劳程度越高。本文的疲劳度检测以 0.22 作为闭眼的判断阈值, 用 0.17 作为哈欠的判断阈值, 统计一个专注度评估轮回中的 PERCLOS 值、平均闭眼时长、打哈欠频率, 并分别赋予 1, 0, 0.5 的权重计算疲劳度评分。

3.2.1 闭眼检测

闭眼检测采用计算眼睛闭合度判定。开源库 libfacedetection 可检测人脸图像中的 68 个特征点, 其中特征点 36~41 为右眼的 6 个特征点 (如图 3 所示), 特征点 42~47 为左眼的 6 个特征点。每组特征点中包括 2 个眼角特征点、2 个上眼皮特征点和 2 个下眼皮特征点。计算上下眼皮的距离与左右眼角距离之比, 结果即为 2 只眼睛的闭合度 R 值:

$$t_1 = \sqrt{(x_{37} - x_{41})^2 + (y_{37} - y_{41})^2} \quad (6)$$

$$t_2 = \sqrt{(x_{38} - x_{40})^2 + (y_{38} - y_{40})^2} \quad (7)$$

$$t_3 = \sqrt{(x_{36} - x_{39})^2 + (y_{36} - y_{39})^2} \quad (8)$$

$$R_1 = \frac{1}{2} \frac{t_1 + t_2}{t_3} \quad (9)$$

$$t_4 = \sqrt{(x_{43} - x_{47})^2 + (y_{43} - y_{47})^2} \quad (10)$$

$$t_5 = \sqrt{(x_{44} - x_{46})^2 + (y_{44} - y_{46})^2} \quad (11)$$

$$t_6 = \sqrt{(x_{42} - x_{45})^2 + (y_{42} - y_{45})^2} \quad (12)$$

$$R_2 = \frac{1}{2} \frac{t_4 + t_5}{t_6} \quad (13)$$

取均值:

$$R = \frac{1}{2} (R_1 + R_2) \quad (14)$$

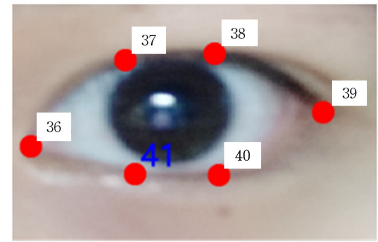


图 3 眼睛特征点

Fig. 3 Eye feature points

本文通过团队 4 名成员的实验测试, 统计总结出眼睛状态和闭合度 R 值的关系, 如表 2 所列。

表 2 眼睛状态和闭合度 R 值的关系

Table 2 Relationship between eye state and R value of closure

序号	眼睛状态	R 值
1	正常状态	≥ 0.28
2	微闭状态	$0.22 \sim 0.28$
3	闭眼	≤ 0.22

统计结果为, 当眼睛闭合度 R 小于等于 0.22 时, 判定为闭眼。

3.2.2 哈欠检测

打哈欠检测采用计算嘴巴闭合度判定。在开源数据库 libfacedetection 检测的 68 个特征点中, 48~68 为嘴巴的特征点 (如图 4 所示), 考虑到嘴型、嘴唇厚度等影响因素, 本文采用嘴巴的上下内嘴皮的距离与嘴角距离之比来计算嘴巴闭合度 R 值:

$$t_1 = \sqrt{(x_{50} - x_{58})^2 + (y_{50} - y_{58})^2} \quad (15)$$

$$t_2 = \sqrt{(x_{52} - x_{56})^2 + (y_{52} - y_{56})^2} \quad (16)$$

$$t_3 = \sqrt{(x_{48} - x_{54})^2 + (y_{48} - y_{54})^2} \quad (17)$$

$$R = \frac{1}{2} \frac{t_1 + t_2}{t_3} \quad (18)$$

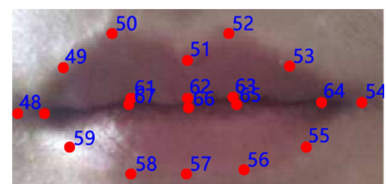


图 4 嘴巴特征点

Fig. 4 Characteristic points of mouth

本文通过团队 4 名成员的实验测试, 统计得到嘴巴状态

和闭合度 R 值的关系,如表 3 所列。

表 3 嘴巴状态和闭合度 R 值的关系

Table 3 Relationship between mouth state and R value of closeness

序号	嘴巴状态	R 值
1	正常状态	$0 \sim 0.05$
2	微开状态	$0.06 \sim 0.16$
3	打哈欠	≥ 0.17

当嘴巴闭合时,闭合度为 0,当闭合度 R 大于等于 0.17,判定为打哈欠。

3.3 表情检测

面部表情是指人体通过面部肌肉的变化来表现各种情绪状态。面部表情是人体在不刻意掩饰时的内在情绪的外在表现形式,旁观者通过面部表情可以直观地了解一个人的情绪^[10]。随着心理学的发展,早在 20 世纪 50 年代就已经开始研究情绪的认知理论。阿诺德和拉扎勒斯在 20 世纪 50 年代提出了著名的情绪认知-评价理论,该理论认为情绪刺激必须通过认知评价才能引起一定的情绪^[10],而情绪可以反映出一个人对事物的主观喜好。Feng 等^[11]通过实验发现,在积极情绪下,大学生的认知抑制功能下降,认知水平提升,而认知水平的提升对于提升学习效率有正向促进作用。综合以上研

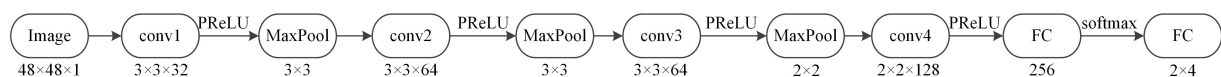


图 5 CNN 卷积层示意图

Fig. 5 CNN convolution diagram

训练阶段,以 6000 个数据作为一轮训练,设定为 100 轮迭代训练。训练得到模型,模型经验证集检测,准确率达到 72%。

表情检测时,首先从摄像头获取图像,规格化图像为 48x48,使用 OpenCV 截取图片中包含人脸的最大尺寸的图像区域,对截取的图像灰化处理,降低运算复杂度,并把处理后的图像代入模型检测,从计算结果中取拟合度最高的标签对应的表情,认定为图像中人脸的表情。

3.4 模糊综合评判

在实际问题中,目标决策往往是多个因素的综合决策,有的属性带有模糊性,不能量化分析,也无法简单用“好”与“不好”来评判。模糊综合评判决策是一种基于模糊数学的综合评判方法,其基本思想是利用模糊线性变换原理和最大隶属度原则,使用隶属函数将带有模糊性的属性量化,再使用传统的量化评判方法评判,这样既能统一处理问题中定量和定性的因素,又能兼顾模糊性属性之间的差异,因此实用性和可靠性非常高^[12]。

本文的评判目标——学习专注度,目前在学界并不能被量化衡量,属于模糊性目标,并且面部朝向、疲劳度、情绪这 3 个因素在评判时也无法直接确定各个因素对评判目标的影响权重的大小,同样具有模糊性,故本文的评判方法采用模糊综合评判。在本文中,单位时间内水平面部偏转角均值、垂直面部偏转角均值、闭眼次数、哈欠次数、7 种面部表情出现的次数是本文评判的第一层因素,第一层因素对应评判的面部朝向、疲劳度、情绪是第二层因素。故本文采用面向多层因素评判的模糊综合评判,即多级模糊综合评判(如表 4 所列)。

究,我们认为面部表情在很大程度上能够反映学习者的学习专注度,通过研究学习者的面部表情可以衡量其学习专注度。

面部表情识别分为高兴、厌恶、愤怒、平静、伤心、恐惧、快乐 7 种常见表情。识别方法采用当下热门并且成熟的卷积神经网络(CNN),使用了谷歌开源数据集 fer2013 数据集作为训练集,使用 keras 框架搭建卷积神经网络的卷积层。本文将前 30000 个数据划分为训练集,将位置处于 30000~35000 的 5000 个数据划分为验证集,其余的数据并入训练集。在模型训练开始前,通过数据增强将训练集规模扩充到 60000 多个,提高数据集规模有助于提高模型的准确性。在数据预处理阶段,先把数据规格化为 48x48 的图像,并进行灰化处理,从而降低计算复杂度。

对输入的 48x48x1 的数据,第一层卷积层采用 32x32 提取信息,损失函数采用 PreLUde,池化层为 3x3,添加损失函数为 0.25,防止过拟合。第二层为 64x64 的卷积层,损失函数为 PreLU,池化层为 3x3,损失函数为 0.25。第三层为 64x64 的卷积层,损失函数为 PreLU,池化层为 3x3,损失函数为 0.25。第四层为 128x128 的卷积层,损失函数为 PreLU。最后把数据全连接,设定损失函数为 softmax(如图 5 所示)。

表 4 综合评判的因素组成

Table 4 Factor composition of comprehensive evaluation

第二层因素		第一层因素	
序号	因素名称	序号	因素名称
1	头部姿态	1	左(右)转头转角均值
		2	抬(低)头转角均值
		3	PERCLOS 值
2	疲劳度	4	平均闭眼时长
		5	哈欠频率
		6	高兴表情次数
3	情绪	7	厌恶表情次数
		8	愤怒表情次数
		9	平静表情次数
		10	伤心表情次数
		11	恐惧表情次数
		12	快乐表情次数

3.4.1 第一层因素评判

头部姿态评分分为左(右)转头评分和抬(低)头评分,综合评分取二者的最小值。首先计算周期内的左(右)转头转角均值 θ_1 和抬(低)头转角均值 θ_2 。

根据人体颈椎关节的左右旋转范围为 $\pm(60^\circ \sim 80^\circ)$,本文取 70° 为最大左(右)转头转角度,转角 θ_1 越小,评分越高,大于 70° 皆为 0 分。

$$k_1 = \begin{cases} 1 - \frac{|\theta_1|}{70^\circ}, & 0^\circ \leq |\theta_1| \leq 70^\circ \\ 0, & |\theta_1| > 70^\circ \end{cases} \quad (19)$$

人体颈椎关节的抬(低)头角度为 $-60^\circ \sim +60^\circ$,大于 0° 为抬头,小于 0° 为低头。转角 θ_2 越接近 0,头部与地面越接近水平,表示学习者正平视前方,所以转角 θ_2 越接近 0,评分越高。

$$k_2 = \begin{cases} 1 - \frac{|\theta_2|}{60^\circ}, & 0^\circ \leq |\theta_2| \leq 60^\circ \\ 0, & |\theta_2| > 60^\circ \end{cases} \quad (20)$$

综合评分取两者的最小值:

$$p_1 = \min(k_1, k_2) \quad (21)$$

3.4.2 疲劳度评分

疲劳度评判。本文采用 PERCLOS 值、平均闭眼时长、打哈欠频率的综合评分^[13], 其中每轮专注度综合评估中包含 τ 次检测闭眼、哈欠、头部转角和面部表情, τ 次检测总共耗费的时间为 T , 根据闭眼次数 t_1 和检测次数 τ 计算 PERCLOS 值 m_1 :

$$m_1 = \frac{t_1}{\tau} \quad (22)$$

根据 τ 次检测耗费的总时间 T 计算平均闭眼时长 m_2 :

$$m_2 = \frac{T}{t_1} \quad (23)$$

根据哈欠次数 t_2 计算打哈欠频率 m_3 :

$$m_3 = \frac{t_2}{\tau} \quad (24)$$

取三者的权重分别为 1, 0.8, 0.5, 计算疲劳度综合评分 p_2 :

$$p_2 = m_1 + 0.8 m_2 + 0.5 m_3 \quad (25)$$

3.4.3 情绪评判

层次分析法是美国工程院院士 Thomas L. Saaty 在 1971 年提出的, 主要应用于具有多个不确定因素的决策问题^[14]。模糊综合评判在确定各个因素的权重时引入了该方法。在确定各个因素的权重时, 该方法对所有的 n 个因素两两比较优先级并量化表示一个因素相对另一个因素的优先级, 经过 $n!$ 次比较之后, 得到每一个因素相对于其余因素的优先级, 最后把所有因素的优先级组合构成 $n \times n$ 的判断矩阵, 通过判断矩阵计算每个因素的权重^[15], 如表 5 所列。

表 5 判断矩阵标度及其含义

Table 5 Judgment matrix scale and its meaning

标度	含义
1	表示因素 u_i 与 u_j 比较, 具有同等重要性
3	表示因素 u_i 与 u_j 比较, u_i 比 u_j 稍微重要
5	表示因素 u_i 与 u_j 比较, u_i 比 u_j 明显重要
7	表示因素 u_i 与 u_j 比较, u_i 比 u_j 强烈重要
9	表示因素 u_i 与 u_j 比较, u_i 比 u_j 极端重要
2, 4, 6, 8	2, 4, 6, 8 分别表示相邻判断, 1-3, 3-5, 5-7, 7-9 的中值
倒数	表示因素 u_i 与 u_j 比较得判断句, 则 u_i 与 u_j 比较得判断 $u_{ji} = 1/u_{ij}$

表情评分使用模糊综合评判模型的一级综合评判模型, 把周期内生气、厌恶、恐惧、快乐、伤心、惊讶、平静 7 种表情出现的次数组成评价因素 $U = [u_1, u_2, \dots, u_7]$ 。并用层次分析法确定各个评价因素的权重。

构造 $A-U$ 判断矩阵

$$H = \begin{bmatrix} 1 & 3 & 3 & \frac{1}{6} & 1 & \frac{1}{5} \\ \frac{1}{3} & 1 & \frac{1}{2} & \frac{1}{9} & \frac{1}{3} & \frac{1}{8} \\ \frac{1}{3} & 2 & 1 & \frac{1}{6} & 1 & \frac{1}{5} \\ 6 & 9 & 7 & 1 & 6 & 2 \\ 1 & 3 & 2 & \frac{1}{6} & 1 & \frac{1}{5} \\ 5 & 8 & 6 & \frac{1}{2} & 5 & 1 \\ 5 & 5 & 5 & \frac{1}{3} & 5 & \frac{1}{2} \end{bmatrix}$$

方根法计算特征向量和最大特征根。计算判断矩阵每一行元素的乘积 M_i , 即:

$$M_i = \prod_{j=1}^7 u_{ij}, i, j = 1, 2, \dots, 7 \quad (26)$$

计算 M_i 的 7 次方根 \overline{W}_i , 即:

$$\overline{W}_i = \sqrt[7]{M_i} \quad (27)$$

对向量 $\overline{W}_i = [\overline{W}_1, \overline{W}_2, \dots, \overline{W}_7]^T$ 做归一化处理, 即:

$$W_i = \frac{\overline{W}_i}{\sum_{i=1}^7 \overline{W}_i} \quad (28)$$

则 $W = [W_1, W_2, \dots, W_7]^T$ 即为所求特征向量:

$$W = [0.06312521, 0.02693035, 0.03981097, 0.36753161, 0.05957265, 0.25979793, 0.18323128]^T$$

计算判断矩阵的最大特征根 λ_{\max} , 即:

$$\lambda_{\max} = \frac{1}{7} \sum_{i=1}^7 \frac{(HW)_i}{W_i} \quad (29)$$

其中, $(HW)_i$ 表示向量 HW 的第 i 个元素。

$$HW = \begin{bmatrix} (HW)_1 \\ \vdots \\ (HW)_7 \end{bmatrix} = \begin{bmatrix} h_{11} & \cdots & h_{17} \\ \vdots & \ddots & \vdots \\ h_{71} & \cdots & h_{77} \end{bmatrix} \begin{bmatrix} W_1 \\ \vdots \\ W_7 \end{bmatrix} \quad (30)$$

一致性指标检验 (RI 值如表 6 所列)

$$CI = \frac{1}{6} (\lambda_{\max} - 6) \quad (31)$$

$$CR = \frac{CI}{RI} \quad (32)$$

表 6 1-9 阶 RI 值

Table 6 1-9 order RI value

n	1	2	3	4	5	6	7	8	9
RI	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45

计算得到 $CR \approx 0.04373 < 0.10$, 判断矩阵具有满意一致性, 说明权数分配是合理的。情绪综合评分为各个表情符出现次数和权重积之和, 即

$$p_3 = U \cdot W \quad (33)$$

3.4.4 第二层因素评判

计算专注度综合评分使用模糊综合评判模型的多级综合评判模型, 第一层因素的评分已经计算完毕。把面部朝向评分、疲劳度评分、情绪评分组成评价因素 $P = [p_1, p_2, p_3]$, 使用层次分析法计算各个因素的权重。

构造 $A-U$ 判断矩阵

$$H = \begin{bmatrix} 1 & 3 & 7 \\ 1/3 & 1 & 5 \\ 1/7 & 1/5 & 1 \end{bmatrix}$$

方根法计算特征向量 $W = [0.649118, 0.27895457, 0.07192743]^T$ 和最大特征根 $\lambda_{\max} \approx 3.0649$, 一致性指标检验 $CR \approx 0.05594 < 0.10$, 权数分配合理 (如表 7 所列)。

表 7 评判因素与权重

Table 7 Evaluation factors and weights

第二层因素			第一层因素			
序号	权重	因素名称	序号	权重	因素名称	
1	0.649118	面部朝向	1	1	左(右)转头转角均值	
			2		抬(低)头转角均值	
			3		PERCLOS 值	
2	0.27895457	疲劳度	4	0.8	闭眼次数	
			5		0.5	打哈欠次数
			6		0.059785	生气表情次数
			7		0.027026	厌恶表情次数
			8		0.039953	恐惧表情次数
3	0.07192743	情绪	9	0.368842	快乐表情次数	
			10		0.059785	伤心表情次数
			11		0.260724	惊讶表情次数
			12		0.183885	中性表情次数

专注度综合评分为3个因素评分与权重积之和,即:

$$p = P \cdot W \tag{34}$$

4 系统开发与测试

4.1 系统开发

本系统的电脑开发系统为 win10,开发编程语言选取 Python,开发工具使用 Pycharm,编程环境选择 Python3.7。本系统设定每隔 1 秒进行一次检测,每 5 次为一个检测周期。

4.1.1 获取图像与图像预处理

本系统使用 OpenCV 自带的功能打开电脑摄像头,获取图像,若获取失败,则再次获取。对获取成功的图像判断分辨率,若其长(宽)大于 700,则把图像缩小为原图像的 1/3。对分辨率处理过的图像灰化处理,降低后续计算的复杂度。使用预先加载的 Dlib 库检测图像的 68 个人脸特征点,若检测成功,则返回 68 个特征点和灰化处理的图像,否则重新获取图像,再执行一遍以上步骤(如图 6 所示)。

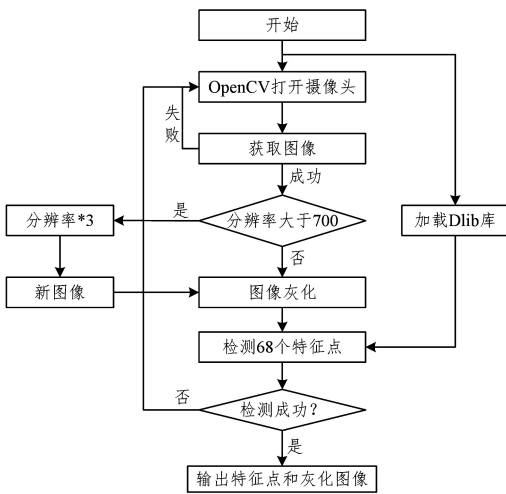


图 6 获取图像与图像预处理流程图

Fig. 6 Flow chart of image acquisition and preprocessing

4.1.2 头部姿态评估

此模块首先获取图像获取与预处理模块输出的 68 个人脸特征点,在特征点中选择左眼左角点、右眼右角点、鼻尖、左(右)嘴角点和下巴 6 个特征点,调用 OpenCV 的 solvePnP 方法,计算头部的旋转向量和翻转向量,根据两个向量分别可计算出头部围绕 Pitch、Yaw 和 Roll 的旋转角度,其中 Pitch 大于 0 表示抬头角度,Pitch 小于 0 表示低头角度,Yaw 大于 0 表示左转头角度,Yaw 小于 0 表示右转头角度(如图 7 所示)。

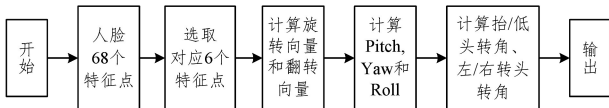


图 7 头部姿态评估流程图

Fig. 7 Flow chart of head posture assessment

4.1.3 闭眼检测和哈欠检测

此模块首先获取图像获取和预处理模块输出的 68 个人脸特征点,在特征点中选择左(右)眼的 12 个特征点和嘴部的 6 个特征点,计算左(右)眼的闭合度,取其均值,判断闭合度是否小于等于阈值 0.22,若满足条件,则闭眼次数加 1;根据嘴部 6 个特征点计算嘴巴闭合度,判断闭合度是否大于等于阈值 0.17,若满足条件,则哈欠次数加 1(如图 8 所示)。

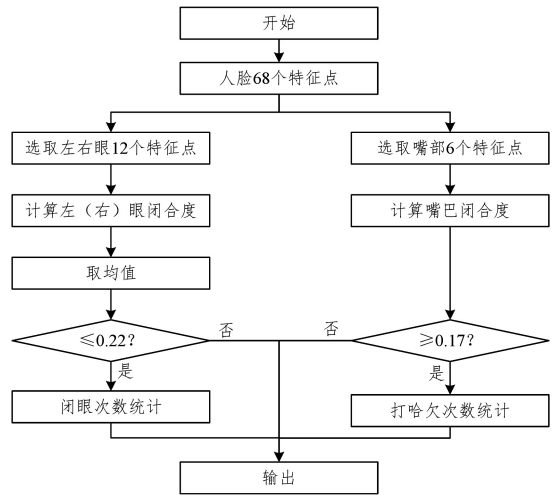


图 8 闭眼和哈欠检测流程图

Fig. 8 Flow chart of closed eyes and yawn detection

4.1.4 表情识别

此模块首先获取图像获取与预处理模块中输出的图像,使用预先加载的表情识别模型识别图像,根据识别结果增加对应表情出现的次数(如图 9 所示)。

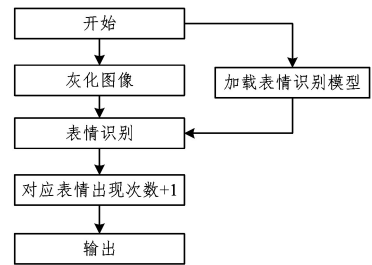


图 9 表情识别流程图

Fig. 9 Expression recognition flowchart

4.1.5 模糊综合评判

模糊综合评判模块首先判断检测周期是否结束,若未结束,则继续监测,否则根据检测的头部抬(低)头转角、头部左(右)转头转角、PERCLOS 值、平均闭眼时长、哈欠频率、周期内 7 种表情出现的次数分别计算出头部姿态评分、疲劳度评分、情绪评分,再基于模糊综合评分计算学习专注度评分(如图 10 所示)。

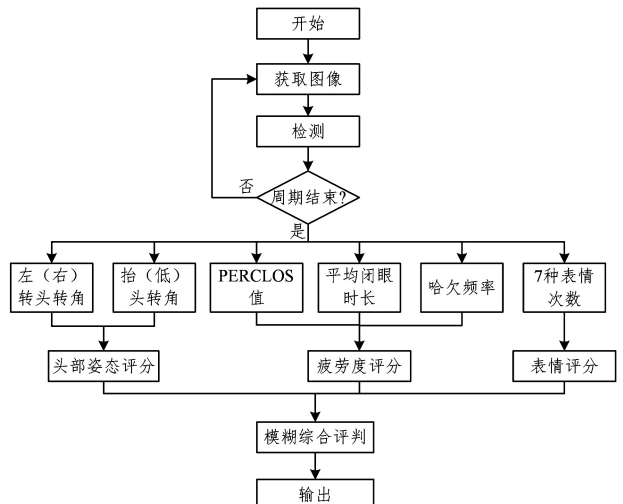


图 10 模糊综合评判流程图

Fig. 10 Flow chart of fuzzy comprehensive evaluation

4.2 系统测试

在以上的研究中,本文针对现阶段的在线教育中,有效的学生监督手段未能得到大范围应用的问题,结合现阶段人工智能的方法,提出了一个基于人脸检测和模糊综合评判的在线教育专注度检测方案。该方案的有效性还需要在实际应用中进行检测。为此,本文设计了一个检测方案:被测人员正对电脑观看 10 min 教学视频,检测系统同时开始工作。在这个过程中,被测人员随机模拟正常状态、打瞌睡状态、哈欠状态、头部轻度偏转状态和头部严重偏转状态,计算在不同状态下系统给学习者的学习专注度评分的均值(如图 11 所示)。



图 11 学习者的 5 种学习状态

Fig. 11 5 learning states of learners

由表 8 可以看出,对于不同的实验者模拟的相同的学习状态,系统给学习者的评分之间的差异并不大。相对于正常状态,其他状态下的学习专注度分数都有不同程度的降低,特别是头部严重偏转状态和瞌睡状态,其降幅更大,与实际情况相结合,头部严重偏转和瞌睡都会对学习者的学习专注度产生非常大的负面影响。轻度偏头时,说明学习者的注意力有一定的转移,但是程度不大,所以系统评分相对正常状态有所下降,但是降幅较小。打哈欠说明学习者的精神状态出现了疲劳,这会对其学习专注度产生较大的负面影响,所以系统评分相对正常状态有较大降幅。

表 8 测试者不同状态的系统评分

Table 8 System scores of different states of testers

实验者 序号	正常 状态	轻度偏头 状态	严重偏头 状态	哈欠 状态	瞌睡 状态
1	87.2	80.4	54.2	66.9	58.7
2	86.6	81.4	53.6	65.4	49.4
3	91.9	77.1	53.1	56.7	61.4
4	88.8	78.4	52.7	59.7	63.7

通过对本文设计的系统的模拟实验,说明该系统在一定程度上能够实现对学习者的学习专注度的检测。通过该系统,授课教师可在远程端实时获取学生在线上上课的学习专注度,掌握课堂状态,及时调整教学计划,提高课堂质量,同时对学生有很大的监督,从而帮助自制力差的学生更好地学习。

结束语 为了填补在线教育中对学生的监督手段的空白,本文提出了基于人脸检测的模糊综合评判算法,用于评估网络课堂的学生的学习专注度。为了提高评估结果的客观真实性,本文基于人脸检测在设定时间周期内采集用户的面部的水平和垂直方向偏转角度、闭眼和打哈欠检测、表情 3 个因子,分别转换为对应的面部朝向评分、疲劳度评分和情绪评分。把 3 个评分合并为模糊集,使用二元对比排序法确定隶

属函数,把模糊集代入隶属函数计算得到专注度综合评分。本文采用多因子检测分析,3 个因子与学习专注度都有紧密联系,在客观程度上能较好地反映学习专注度,减少了检测结果的偶然性,增加了可靠性。其中,对于 3 个因子的综合评分阶段不同因子的权重难题,本文采用了成熟、实践性强的二元对比排序法,通过问卷收集数据,确定 3 个因子的相对优先度函数,并计算得到其隶属函数。最后把模糊集代入隶属函数,计算得到综合专注度评分。实地测试结果表明:在网络课堂中,基于人脸检测的模糊综合评判算法对学习专注度的评估效果良好,可帮助授课老师及时获取课堂中学生的课堂专注度,掌控课堂状态,以便及时调整教学计划和督促学生学习,有效提高课堂质量和学生的学习效果。

受限于目前人工智能技术的不足,检测系统的使用条件较高,当图像特征不明显时,检测系统的检测能力会大幅下降,无法有效检测。另外面部表情识别、疲劳度检测和头部姿态检测这 3 种技术还不十分成熟,其检测精度还不是非常高,这在很大程度上限制了检测系统的准确率,并影响了最终的检测结果。

在未来的研究中,本系统将致力于改进面部表情识别、疲劳度检测和头部姿态检测的深度学习算法,以提高面部表情识别模型、疲劳度检测模型和头部姿态检测模型的准确率。此外,还需提高整个系统在图像特征不明显时的检测和处理水平,提高系统的健壮性,增强系统在实际使用过程中应对复杂情况的能力,以保证整个系统的实用效果。

致谢 在本次论文的撰写中,蒋运承老师和詹捷宇老师为选题、研究内容、研究方法等方面给予了非常多宝贵的建议,在此,我们对蒋运承老师和詹捷宇老师表示诚挚的感谢以及真心的祝福。

参考文献

- [1] WU L Q. A comparative study on the effect of online classroom teaching and traditional classroom teaching[J]. Education Modernization, 2019, 6(9): 276-277.
- [2] TU Y, ZHANG Y J. An empirical study on the supervision model of knowledge construction effect in online classroom[J]. Education Modernization, 2018, 41(3): 64-73.
- [3] WANG J J. Research on the application of emotion computing based on expression recognition technology in Modern Distance Education[D]. Tianjin: Tianjin Normal University, 2005.
- [4] YUAN Y. Research on the recognition of bad expression in Distance Education [D]. Chengdu: Sichuan Normal University, 2014.
- [5] LI W Q. Research on learning focus recognition method for Distance Education[D]. Hangzhou: Hangzhou Dianzi University, 2018.
- [6] PAN X. Research on face detection and expression feature extraction of learners in online teaching context[D]. Beijing: Capital Normal University, 2009.
- [7] XIONG B H. Research and application of attention detection method integrating sight analysis in online learning [D]. Hangzhou: Hangzhou University of Electronic Science and technology, 2018.
- [8] PENG D L. General psychology [M]. Beijing: Beijing Normal University Press, 2001: 169-193.

- [9] WIERWILLE W W, ELLSWORT L A, WREGGIT S S, et al. Research on vehicle-based driver status/performance monitoring; development, validation, and refinement of algorithms for detection of driver drowsiness, final report[R]. Office of Crash Avoidance Research National Highway Traffic Safety Administration, 1994, 12: 247.
- [10] LIN C D, YANG Z L, HUANG X T. Dictionary of Psychology (Volume I) [M]. Shanghai: Shanghai Education Press, 2003: 947-948.
- [11] FENG M N, LIU X M. The relationship between cognitive inhibition and emotion and action control style of college students [J]. Chinese Journal of mental health, 2019(11): 862-867.
- [12] SHEN J C, DU S X, LUO Y, et al. Fuzzy comprehensive evaluation method and application based on cloud model [J]. Fuzzy System and Mathematics, 2012, 26(6): 115-123.
- [13] LU Y, WANG Z. Detecting Driver Yawning in Successive Images[C] // The International Conference on Bioinformatics and Biomedical Engineering. IEEE Xplore, 2007, 07: 581-583.
- [14] SAATY, THOMAS L, PENIWATI, et al. Group Decision Making: Drawing out and Reconciling Differences[M]. Pittsburgh, Pennsylvania: RWS Publications, 2013: 1-385.
- [15] LI X C. Fuzzy mathematics method and application[M]. Beijing: Chemical Industry Press, 2017: 83-113.



ZHONG Ma-chi, born in 1997, undergraduate. His research direction is computer vision.

(上接第 174 页)

些低资源语种等,选择音节为识别单元往往能取得更好的效果。综合目前的研究进展,不难发现,如何将深度学习更好地应用到音节识别的研究中是一个值得重点关注的问题。

参 考 文 献

- [1] LI J, XU M X, ZHANG J Y, et al. Comparison of Acoustic Model Elements in Chinese Continuous Speech Recognition: Syllable, Phoneme and Vowel[C] // Proceedings of the 6th National Conference on Man-machine Voice Communication. 2001: 391-395.
- [2] MA Z X, WANG H, LI X. Overview of Speech Recognition Technology[J]. Journal of Changji University, 2006(3): 93-97.
- [3] ZHAN X Y, WU J P, ZHANG Y W. Optimum Vector Quantization Codebook Design for Speaker Recognition [J]. International Conference on Signal Processing Proceedings, 2004, 7: 14-16.
- [4] HOU L M, ZHANG J Q. Design of a Speaker Recognition System Based on Chinese Monosyllables [J]. Journal of Lanzhou University, 1996(4): 81-85.
- [5] JIA Z J. Research and Implementation of Chinese Speech Recognition Technology[D]. Tianjin: Tianjin University of Technology, 2013.
- [6] CHU S C, TSAI P W, PAN J S. Cat swarm optimization[C] // 9th Pacific International Conference on Artificial Intelligence. Berlin: Springer, 2006: 854-858.
- [7] LU Y H. Research on Key Issues of Speech Recognition Technology[D]. Shaanxi: Shaanxi Normal University, 2014.
- [8] CAO G B. Research on Continuous Speech Recognition Technology Based on HMM[D]. Nanjing: Nanjing University of Science and Technology, 2018.
- [9] DEIVAPALAN P G. A segmented syllable-based isolated word recognizer for indian languages [D]. Indian Institute of Technology Madras, 2008.
- [10] THANGARAJAN R, NATARAJAN A M, SELVAM M. Word and triphone based approaches in continuous speech recognition for Tamil language[J]. WSEAS Transactions on Signal Processing, 2008, 4(3): 76-85.
- [11] WANG Z Y, XIAO X. HMM speech recognition model based on segment length distribution [J]. Acta Electronica Sinica, 2004 (1): 46-49.
- [12] CHAO H, YANG Z L, LIU W J. Improved algorithm of syllable-based acoustic model in Chinese speech recognition [J]. Journal of Computer Applications, 2013, 33(6): 1742-1745.
- [13] CAO H. A New Chinese Monosyllable Recognition Method [J]. Journal Publishing Center of Tsinghua University Press, 1990 (4): 87-92.
- [14] WU Y. Design and Implementation of Chinese Speech Syllable Recognition Algorithm Based on SVM and HMM [D]. University of Electronic Science and Technology of China, 2010.
- [15] ZHOU N, ZHAO Y, LI Y Q, et al. Study on Continuous Speech Recognition Based on Bottleneck Features for Lhasa-Tibetan Dialect [J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2018, 54(2): 249-254.
- [16] WU J F. Research and Application of RNN-DNN Speech Recognition System [D]. South China University of Technology, 2018.
- [17] ZHAO Y B, SUN S H. A TTRNN-based Chinese syllable full syllable recognition method [J]. Journal of Harbin Institute of Technology, 2001(2): 213-216.
- [18] PAUL J W. Backpropagation through time what it does and how to do it [J]. Proceedings of the IEEE, 1990, 78(10): 575-580.
- [19] ZHU S, CHEN D W, HUANG T Y. Feature parameter curve methods for high-performance NN-based speech recognition [C] // IEEE International Conference on Acoustics. IEEE, 2002.



ZHANG Jing, born in 1997, postgraduate. Her research interests include speech recognition and deep neural network.



YANG Jian, born in 1976, Ph.D, associate professor, is a member of China Computer Federation. His research interests include speech recognition and deep neural network.