

# 基于深度学习的在线课堂学生专注度研究与实践

刘迪昱<sup>1,2</sup>

(1.四川开放大学 四川 成都 610073; 2.四川远程电子出版社 四川 成都 610073)

**摘 要:**专注度是影响学生学习成效的重要因素。专注度与学习的关系非常密切,学生课堂上的专注度不足就会影响学习效果。教师获得学生专注度的准确信息是改善和提高学生课堂学习行为的重要依据。近年来,随着深度学习等人工智能技术的发展,为在线课程学生专注度检测研究提供了新的视角和思路。通过在线课堂获取学生课堂学习行为的影像,并利用深度学习技术对影像进行检测、分类等,得出学生在线课堂学习的专注度,使教师更为快捷、准确地了解到学生在在线课堂中的专注情况,有利于教师及时优化教学方法,提高学生的学习成效。

**关键词:**深度学习;在线课堂;专注度;检测;学习成效

中图分类号:G420;TP391 文献标识码:A 文章编号:1008-3006(2021)03-0041-05

随着智能移动终端的普及,远程教育的学生在线学习过程中使用手机、平板电脑、笔记本电脑进行资料查阅和记笔记的情况已是常态。使用智能终端作为学习辅助工具,有利于学生更快更好地理解学习内容。然而,对于自控力较弱的学生而言,智能终端往往会产生更多的负面影响,使其更容易分心于课程内容之外的其他事物,导致学习专注度下降。在远程教育在线课堂中,学生的专注度是影响其学习效果的一个重要因素。Metcalf 认为,专注度是将注意力持续性地集中在某一特定事物上<sup>[1]</sup>;Mayer 等把专注度解释为将行为及过程的焦点放置于单一或多个刺激上<sup>[2]</sup>;Jong 则从科学的角度指出,专注度是学习行为中的关键角色,会影响信息处理的深度及学习的成效<sup>[3]</sup>。本研究利用深度学习技术对在线课堂学生的面部表情和动作姿态影像进行检测,进而对其数据进行加工处理,分析学生学习的专注度状况,以帮助教师实时调整在线课堂的教学行为,进而提高教学效果。

## 一、深度学习理论概述

近年来,有学者探讨将深度学习理论应用到影

像辨识中,并且构建了许多不同的模型,例如,Convolution Neural Networks 等<sup>[4]</sup>,这些模型也使得深度学习被应用到多个领域<sup>[5]</sup>。深度学习是由一个输入层、一个输出层和多个隐层组成的神经网络系统,是更广泛机器学习方法的一个子域,其目标是使机器能够进行智能化分析和学习,并识别图像、声音和其他数据。深度学习的概念起源于神经网络的研究,它是可以使多层神经网络训练和运行的一系列新的学习模型<sup>[6]</sup>。深度学习使机器模仿人类活动,例如,对视听和思维的模仿就解决了许多复杂的模式识别问题。深度学习已在人工智能相关技术方面取得了巨大进展,特别是在语音和图像识别领域中取得了新的突破<sup>[7]</sup>。深度学习算法的核心是通过多层卷积神经网络建立更加有效的特征分类网络,该算法对目标的识别率以及识别泛化能力远远超过传统的算法。

以卷积神经网络为代表的深度学习技术被广泛运用于目标追踪、动作检测、表情识别等应用场景,并被证实能够大幅提升检测结果的准确性和高效性。事实上,课堂上学生的动作行为、面部表情都可被用于判断其专注程度,这和通过观察法进

本文系四川开放大学 2021—2022 年度教学改革重点项目“媒体融合环境下传统文化融入专业课程教学的策略研究”(课题批准号:XMZXJYXXXTYGL2021002Z)成果。

收稿日期:2021-04-19

作者简介:刘迪昱(1977—),女,四川开放大学工程技术学院,高级工程师;四川远程电子出版社编辑部,副编审。

行专注度检测的原理是一致的。而深度学习技术则为精准、迅速地捕捉和分析学生的动作及表情提供了一种更为有效的算法,因此,可将其运用于课堂专注度的自动化检测,为教学方法的改进和教学质量的提升提供更为准确的依据和参考。

传统的专注度测量方法存在着应用环境的限制、被检测者体验感欠佳、高成本等问题,难以推广到课堂中对学生进行专注度检测。但是,深度学习技术的兴起,特别是该技术在目标追踪、语音识别、运动姿势、人脸识别以及医学影像等研究领域得到了广泛的应用,为探究课堂专注度检测的方法提供了良好的实

践基础,并启发我们将其拓展到教育教学领域。

## 二、专注度检测架构

深度学习相较于传统机器学习方法的优势在于,自动学习给定数据中适用于指定任务的特征,数据集对于基于深度学习的计算机视觉方法非常重要<sup>[8]</sup>。本研究将深度学习技术用于专注度检测,通过影像分析来确定学生的面部表情和动作姿态,并利用深度学习技术对学生在线课堂中的专注度进行检测。整个检测过程的架构如图1所示,主要包括影像采集、面部检测、姿态检测和检测分析四个步骤。

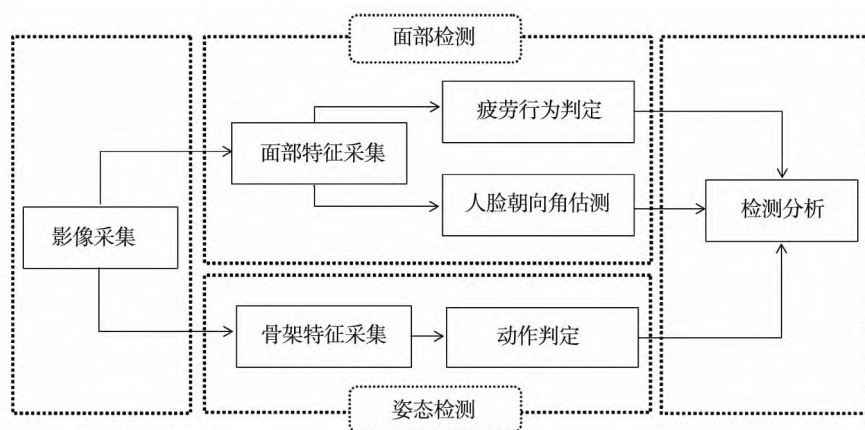


图1 检测过程架构图

### (一)影像采集

在一定间隔时间内,通过摄像机提取一张影像,其间隔时间可依据实验需求调整,时间越短,越有利于观测学生在在线课堂中学习的专注情况,但是,需考虑后续步骤中硬件设备对图像的处理能力。为避免面部和姿态出现在光源不足的情况下而导致辨识效果不佳,故需要在光线充足、明亮的教室进行影像采集,以提高学生面部和姿态辨识的正确率。

此环节所采集的学生在线课堂学习影像分别存放在以学生姓名命名的文件夹中,图像文件实时上传至后端数据库中。

### (二)面部检测

面部表情检测常用的方法是基于局部信息的提取方法,在选取面部关键点或关键区域后提取局部特征并进行分类<sup>[9]</sup>。面部表情特征检测的侧重点是表情特征提取<sup>[10]</sup>;在表情特征提取过程中,图像识别是关键的一环<sup>[11]</sup>。在摄像机采集学生影像

后,经过面部表情的特征点选取、色彩空间转换后,找出人脸在影像中的范围,若无法检测到学生面部表情信息,即表示学生已离开在线课堂;若检测到学生面部表情信息,即表示学生在在线课堂。利用采集到的不同特征判定学生的在线状态,例如,是否存在闭眼、眨眼和打哈欠等疲劳行为,在估测人脸朝向后,判定学生是否专注。

### (三)姿态检测

为了对人体姿态特征进行采样,需要对人体姿态图像进行采集<sup>[12]</sup>。使用影像姿态估测的开源程序OpenPose骨架采取系统取得学生的骨架信息,将骨架信息进行正规化和特征采集,利用采集到的特征作为人工神经网络分类器的输入信息进行动作识别。本研究的动作姿态检测主要针对学生在在线课堂中写笔记、使用笔记本、看手机、举手、歪头、趴睡、单手撑头、接电话等行为动作。

### (四)检测分析

一是眨眼的频率分析。所采集的影像在设定

时间内连续两张有眨眼动作出现,即判定为不专心。二是眼珠移动的方向分析。若所采集到的影像中眼珠方向过偏,即判定为不专心。三是嘴巴张合的频率分析。参与讨论时,学生的嘴巴开合次数越多,表示参与度越高。四是头部摇动的频率分析。利用前后两张影像检测眼睛移动的位置,若移动位置过大,表示学生头部在摇动,头部摇动频率过大,表示精神不佳。五是离开座位的分析。若检测不到人脸,表示学生离开座位或坐姿不良,这些均为学习不专心的表现。六是双手使用物品的分析。将记笔记、使用笔记本电脑和用手机等行为归类在双手使用物品的类别,学生若长时间使用此类物品,即为分心于非课程的内容上。七是单手使用物品的分析。将手撑头和接电话等行为归类在单手使用物品的类别,学生若长时间处于此类状态,即为分心于课程内容之外。

### 三、实践案例分析

将前述设计架构及关键技术应用于检测在线课堂学生学习的专注度,在技术实现上相对较为容易;在其软硬件设备的配置上,也无需大量的资金投入,并且可获得较为理想的检测效果。

#### (一)专注度分析

按上述方法,在取得学生的面部、眼睛、嘴部特征及动作姿态影像基础上,再对学生注意力不足和疲劳状态的行为表现进行分析,从而判定其专注度。本实验将专注度分为三种程度,即,专注、不专注、极不专注,并根据面部特征信息和动作姿态辨识的结果判定专注程度(见表1)。

在专注程度判定中,若面部特征和动作姿态影像符合不专注的情形,则该学生的专注程度被判定为不专注;若符合极不专注的情形,则该学生的专注程度被判定为极不专注;不符合不专注和极不专注的情形时,判定该学生的专注程度为专注。

表1 学生专注程度对应表

专注程度	面部特征	动作姿态
专注	在自然眨眼中,眼睛瞳孔移动轨迹在规定的移动距离范围内,且无左顾右盼和发呆行为。	教师提问时,学生做出响应动作。
不专注	左顾右盼,眼睛瞳孔移动轨迹不在规定的移动距离范围内;眼睛视线方向不在讲台上或长时间集中在智能终端设备上;发呆行为,超过规定的时间范围无任何眼动行为;疲劳行为,眼睛眨眼动作频繁发生;交头接耳,面部移动轨迹偏移规定距离。	长时间使用与学习无关的设备,处于歪头、撑头状态,时间越长表明专注度越低;玩手机、歪头、趴睡、撑头、接电话均为不专注。
极不专注	离开座位,学生面部轨迹未在规定的移动距离内;离开课堂,在规定时间内检测不到学生面部轨迹信息。	离开座位,学生骨架轨迹未在规定的移动距离内;离开课堂,在规定时间内检测不到骨架轨迹信息。

面部特征检测类目与专注度的关系可归纳为五个方面。一是面部信息检测。学生座位区域内若检测不到学生面部信息,说明学生离开、走动或姿势不良,判定学生为极不专注。二是眨眼频率检测。正常人每分钟眨眼约20次,当工作专心在屏幕上时,眨眼率也会随之降低,眨眼检测的影像中连续有眨眼频率超过每分钟40次,即判定为不专注。三是眼珠移动检测。课堂上教师讲授课程时,学生眼珠移动的距离应在相应的范围内,若眼珠移动检测中眼珠移动距离过偏,即判定为不专注。四是嘴巴张合检测。学生多次嘴巴开合打哈欠,说明学生处于疲劳状态,检测到学生嘴巴开合次数越多,表示参与度越低,即为不专心,但在讨论时检测到嘴巴开合次数越多,表示参与度越高,即判定为专注状态。五是头部摇动检测。学生瞌睡、转头时都会出现头部摇动幅度过大现象,利用前后两张影像检测学生眼睛移动的位置,若移动位置过大现象,则表示头

部在摇动,即判定为不专注。

动作姿态检测类目与专注度的关系可归纳为八个方面。一是写笔记检测。在教师讲授课程时,检测到学生长时间连续影像为写笔记骨架,说明学生注意力未在教师讲授的内容上,即判定为不专注,但教师布置练习题后,检测到学生长时间连续影像为写笔记骨架,则判定为专注。二是用笔记本电脑的检测。检测到学生双手在桌面上,且桌面上有笔记本电脑,表示学生可能在使用笔记本电脑,若教师讲授课程时,检测到学生连续敲击笔记本电脑键盘或点击鼠标的时间超过5分钟,即判定为不专注。三是用手机检测。检测到学生用手机,即判定为不专注。四是互动检测。教师提问时,检测到学生做出响应动作,即判定为专注。五是歪头检测。检测到学生连续歪头时间超过2分钟,即判定为不专注。六是趴睡检测。检测到学生连续趴睡时间超过2分钟,即判定为不专注。七是单手撑头检测。检测到学



生连续单手撑头时间超过 2 分钟,表示学生疲倦,即判定为不专注。八是接电话检测。检测到学生抬起单手,且检测到抬起的手臂区域有手机出现,则表示学生在打电话,即判定为不专注。

(二)影像检测

本实验在前述设计架构和关键技术基础上,结合开源软件工具包 Dlib 函数库和 Microsoft Kinect 硬件设备进行学生专注度检测。之所以选用 Dlib,其主要原因在于,Dlib 是一个包含深度学习算法的开源工具包;Dlib 有高效的深度学习算法,能轻松实现面部检测和对像姿态估算,具有高质量的面部识别功能;Dlib 还提供了简捷的图形用户接口。本实验选用的深度影像姿态估测仪是最具代表性的 Kinect for Xbox 360。Kinect for Xbox 360 的周边设备有三个镜头,中间的镜头为 RGB 彩色摄像机,左右两边的镜头分别为红外线发射器和红外线感应器,使其能有效地捕捉学生的面部特征和动作姿态。Kinect 的主要优点表现在捕获的深度影像简化了背景相减的计算量,并能快速地将人的轮廓从背景中分离出来。

在对面部、眼睛和嘴巴的辨识中,本研究以深度学习技术为基础,并结合人脑对面部辨识的模式进行检测。在利用深度学习技术辨识面部过程中,利用与人脑相同的辨识逻辑。采用前述面部特征采集技术对其影像进行搜寻与演算,可确定影像中是

否有人脸存在,进而可标定出其位置、大小与数量等信息,并能够在该面部影像范围内取得眼睛、嘴巴等特征位置进行分析。

本实验选取 20 位受试者开展测试,重点捕获课堂情境下学生的常见面部特征和动作姿态。面部特征检测主要辨识是否存在眨眼、眼珠移动、嘴巴张合、头部摇动等面部特征;动作姿态检测主要辨识写笔记、用笔记本电脑、看手机、举手、歪头、趴睡、单手撑头和接电话等动作姿态特征。在检测中,面部特征和动作姿态按上述角度分别拍摄 6 次,每一阶段拍摄 1 560 帧画面,通过多阶循环录制视频,最后对实验检测数据与人工判断分析数据进行对比,获得检测实验混淆矩阵(见表 2)。

表 2 的每一列代表了预测类目,每一行代表了数据的真实归属类目。例如,最后一行的“97.75%”表示在接电话动作检测中有 97.75%的正确识别率;“2.25%”表示在接电话动作检测中有 2.25%被错误识别为单手撑头。在本实验中,加入了物品辨识的方式来加强辨识效果,分别选择了书本、笔记本电脑和手机作为需要加入的对象辨识来改善辨识效果。加入物品辨识是帮助分辨骨架类似的动作,加入物品辨识后,将骨架类似的动作先分在同一类,再利用物品辨识分辨其中的差异,其辨识率与以往相比有了较大提升。表 2 中从左上角到右下角对角线上的数据为各类目检测的正确率,其整体的正确辨识率达到了良好的效果。

表 2 学生专注程度对应表

		测试类别							
		做笔记	用笔记本	看手机	举手	歪头	趴睡	单手撑头	接电话
分类实测	做笔记	90.11%	1.97%	0	0	0	7.92%	0	0
	用笔记本	0	100%	0	0	0	0	0	0
	看手机	12.15%	0	80.25%	0	7.6%	0	0	0
	举手	0	0	0	97.42%	0	2.58%	0	0
	歪头	2.91%	0	0	0	76.93%	20.16%	0	0
	趴睡	1%	0	0	0	0	99%	0	0
	单手撑头	0	0	0	0	0	0	86.72%	13.28
	接电话	0	0	0	0	0	0	2.25%	97.75%

四、结语

本研究利用深度学习技术与影像处理,对在线课堂中学生面部特征和动作姿态进行检测,进而分析学生课堂上学习的专注度。该检测方法准确度高,且具有不干扰学生学习的特点,还可避免传统方法中主观、耗时、准确度欠佳等问题。同时,在本研究的检测实验中增加了物品辨识内容,使学生在

写笔记、使用笔记本、单手撑头等方面的辨识率比以往有较大提高。但实验中歪头动作的辨识率还处于偏低的水平,究其原因,或因每个人做出歪头的动作有些不同而造成误判。歪头动作共同的特征为双眼形成的直线并不是和桌面呈水平状态,今后的实验可以考虑加入此特征,以提高歪头动作的辨识率。后续研究还可以考虑用同样的方法检测授课教

师的动作行为,并将教师和学生的互动信息结合起来,更好地分析学生的专注度,让教师能更准确地判断是否需要调整教学进度、教学内容和教学方法等,从而不断提高课堂上学生的专注度和学习效率。

参考文献:

- [1]Metcalfe A. Cognitive Psychology: A Student's Handbook(6<sup>th</sup> Ed.) by M.W.Eysenck and M. T.Keane[J].British Journal of Psychology, 2011,(1):138-141.
- [2]Mayer R, Moreno R. Aids to Computer-Based Multimedia Learning[J].Learning and Instruction,2002,(1):107-119.
- [3]Jong T D. Applying Cognitive Science to Education: Thinking and Learning in Scientific and Other Complex Domains[J]. International Journal of Science Education, 2010,(4):561-563.
- [4]Lecun Y, Bottou L, Bengio Y. Gradient-Based Learning Applied to Document Recognition[J].Proceedings of the IEEE,1998, (11):2278-2324.
- [5]Deng L, Yu D. Deep Learning: Methods and Applications[J]. Foundations and Trends in Signal Processing, 2014, (3): 197-387.
- [6]郭丽丽,丁世飞.深度学习研究进展[J].计算机科学,2015, (5):28-33.
- [7]张晶晶,李秋艳,刘硕,等.基于深度学习的人脸识别在身份认证领域应用综述[J].数据通信, 2021,(4):1-6.
- [8]张瑶,卢焕章,张路平,等.基于深度学习的视觉多目标跟踪算法综述[J].计算机工程与应用,2021,(13):55-66.
- [9]Pantic M, Rothkrantz L. Facial Action Recognition for Facial Expression Analysis From Static Face Images[J]. IEEE Transactions on Systems, Man and Cybernetics, Part B, 2004,(3): 1449-1461.
- [10]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.
- [11]王攀.基于 CNN 与 RoELM 的图像分类算法研究[J].计算机与数字工程, 2019,(3):666-671.
- [12]于海鹏,王闻达.基于融合型深度学习的人体动态特征提取[J].河南工程学院学报(自然科学版), 2019,(1):71-76.

(责任编辑:雨楠)

## Research and Practice of Students' Concentration in Online Classroom Based on Deep Learning

Liu Diyu

- (1. Sichuan Open University, Chengdu Sichuan 610073;
2. Sichuan Remote Electronic Press, Chengdu Sichuan 610073)

**Abstract:** Concentration is the most important factor affecting students' learning effectiveness. There is a close relationship between concentration and learning. Students' lack of concentration in class will affect the learning effect. Teachers' obtaining accurate information about students' concentration is an important basis for improving students' classroom learning behavior. In recent years, with the development of artificial intelligence technology such as deep learning, it provides a new perspective and idea for the research of students' concentration detection in online classroom. Images of students' learning behavior in online classroom can be recorded, and then detected and classified with deep learning technology in order to make analysis of students' concentration in online classroom to allow teachers to understand the students' concentration in online classroom more quickly and accurately, which is conducive for teachers to optimize teaching methods and improve students' learning effectiveness in time.

**Key words:** deep learning; online classroom; concentration; detection; learning effectiveness