



计算机科学

COMPUTER SCIENCE

基于视频的在线学习情感识别研究

魏艳涛, 罗洁琳, 胡美佳, 李文昊, 姚璜

引用本文

魏艳涛, 罗洁琳, 胡美佳, 李文昊, 姚璜 [基于视频的在线学习情感识别研究](#) [J]. 计算机科学, 2022, 49(11A): 211000049-6.

WEI Yan-tao, LUO Jie-lin, HU Mei-jia, LI Wen-hao, YAO Huang. [Online Learning Emotion Recognition Based on Videos](#) [J]. Computer Science, 2022, 49(11A): 211000049-6.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于面部视频的非接触式心率检测方法研究](#)

Non-contact Heart Rate Detection Based on Facial Video

计算机科学, 2022, 49(11A): 211000182-6. <https://doi.org/10.11896/jsjcx.211000182>

[融合频率和通道卷积注意的脑电\(EEG\)情感识别](#)

EEG Emotion Recognition Based on Frequency and Channel Convolutional Attention

计算机科学, 2021, 48(12): 312-318. <https://doi.org/10.11896/jsjcx.201000141>

[一种基于Q-学习算法的增量分类模型](#)

Incremental Classification Model Based on Q-learning Algorithm

计算机科学, 2020, 47(8): 171-177. <https://doi.org/10.11896/jsjcx.190600150>

[语音任务下声学特征提取综述](#)

Survey of Acoustic Feature Extraction in Speech Tasks

计算机科学, 2020, 47(5): 110-119. <https://doi.org/10.11896/jsjcx.190400122>

[在线影响力最大化研究综述](#)

Survey on Online Influence Maximization

计算机科学, 2020, 47(5): 7-13. <https://doi.org/10.11896/jsjcx.200200071>

基于视频的在线学习情感识别研究

魏艳涛^{1,2} 罗洁琳^{1,2} 胡美佳^{1,2} 李文昊¹ 姚 璜¹

1 华中师范大学人工智能教育学部 武汉 430079

2 华中师范大学湖北省教育信息化研究中心 武汉 430079

(yantaowei@ccnu.edu.cn)

摘要 随着疫情防控的常态化开展,在线学习已成为日常教学活动的主要形式之一。然而,随着在线学习活动的大规模开展,“情感缺失”问题日益凸显,已成为阻碍在线学习成效的主要原因之一。针对上述问题,探讨利用视频数据的非侵入式在线学习情感状态识别方法。首先采集了22名学生在线学习的面部视频和心率数据,构建了双模态在线学习情感数据库。其次,采用帧注意网络从学习视频中提取表情特征,识别在线学习情感状态,识别精度达到了87.8%。最后,探讨了视频心率识别方法在在线学习情感分析中的应用,研究结果表明,困惑状态下的心率水平具有显著性。本文从学习视频数据挖掘出发,重点探讨了基于表情和视频心率的在线学习情感识别,为提高在线学习情感状态感知提供了新思路。

关键词: 学习情感;在线学习;面部表情;心率;情感识别

中图分类号 G40-057

Online Learning Emotion Recognition Based on Videos

WEI Yan-tao^{1,2}, LUO Jie-lin^{1,2}, HU Mei-jia^{1,2}, LI Wen-hao¹ and YAO Huang¹

1 Faculty of Artificial Intelligence in Education, Central China Normal University, Wuhan 430079, China

2 Hubei Research Center for Educational Informationization, Central China Normal University, Wuhan 430079, China

Abstract With the normalization of epidemic prevention and control, online learning has become one of the main forms of daily teaching activities. However, with the large-scale development of online learning activities, the problem of “emotional loss” is increasingly prominent, which has become the main reason for the low completion rate of online learning. Aiming to deal with the above problems, the non-invasive online learning emotion state recognition method using video data is discussed. Firstly, the facial videos and heart rate data of 22 students learning online are collected to construct a bimodal online learning emotion database. Secondly, the frame attention network is used to extract facial expression features from the learning video and recognize the emotional state of online learning, and its recognition accuracy reaches 87.8%. Finally, the application of the video heart rate recognition method in online learning emotion analysis is discussed. Research results show that the heart rate level in the confused state is significant. Starting from learning video data mining, focusing on learning emotion recognition based on facial expressions and video heart rate, which provides a new idea for improving the perception of emotional state in online learning.

Keywords Learning emotion, Online Learning, Facial expression, Heart rate, Emotion recognition

随着互联网、人工智能等新一代信息技术的快速发展,在线学习已成为一种重要的学习方式。在全球疫情蔓延的情形下,在线学习的优势和必要性进一步凸显。然而,在线学习在为学习者提供跨时空支持和资源共享保障的同时,其“情感缺失”的问题却日益突出,导致部分在线学习者产生孤独感,继而进一步引发学习倦怠感^[1-2]。教育并非一个单方向的机械式知识传输过程,学生自身的情感和动机会在很大程度上影响教育的成果。而大部分的智能教育技术缺乏对用户的情感理解,导致在教育应用中的效果受到限制^[3]。因此,如何精准感知学习者的情感状态并实施有效干预是目前教育领域关注的焦点^[4-6]。

近年来,智能技术赋能教育为有效监控在线学习提供了有力的抓手,物联网技术、传感器技术和大数据计算能力的不断

提升为学习行为分析提供了技术支撑,基于智能技术的学习情感检测识别也引起了学者们的广泛关注^[7]。学生在在线学习的过程中,产生了丰富的学习数据,除了论坛交互、学习日志等阶段性数据,还包括视频、点击流等实时数据。随着深度学习算法的快速发展,基于计算机视觉的学习情感感知识别方法,因具有非接触、非侵扰、实时性好以及易规模化实施等显著优势,成为教育与信息科学交叉领域的研究热点^[8-9]。该方法通过学习终端的摄像头采集学习者视频,并依据视频中的多种视觉线索提取学习者特征,实现相关学习状态的感知。值得注意的是,在线学习中终端设备采集的学习者的学习视频中蕴含了丰富的学习情感信息。这种方法具有非接触、非侵扰以及易规模化等优势,可感知自然学习情景下的真实情感状态,且视频中包含了丰富的动态情感特征。因此,

基金项目:教育部人文社会科学研究项目:基于人工智能的在线学习参与度识别研究(20YJC880100)

This work was supported by the Humanities and Social Sciences of China MOE(20YJC880100).

通信作者:胡美佳(meijiahu@mails.ccnu.edu.cn)

学习者视频为获取其学习情感状态、缓解在线学习“情感缺失”问题提供了一条有效途径,进而有助于促进在线学习高质量发展。

本文在学习分析理论的指引下,立足于在线学习环境,紧紧围绕视频数据源,展开在线学习情感识别研究,以期为学生的情绪感知提供辅助,为在线学习情感支持和学习预警提供技术支撑。首先,考虑到深度学习算法进行训练时依赖于大数据样本,本文借鉴了心理学研究的情绪诱导法,构建了一个结合生理数据和面部表情数据的视频情感数据库。其次,围绕学生在线学习情感识别的目的,本文采用了帧注意力网络进行学生视频情感的识别,最后,本文考虑将基于视频的非接触式远程生理信号检测方法引入在线学习情感识别研究中,这是一种新的尝试,为在线学习分析提供了新思路。

1 在线学习情感识别研究现状

目前在线学习情感研究的关键点在于对学生在线学习情感的精准感知,即学习情感有哪些表征形式,这些情感表征或情感提示中哪些部分是可量化的,具体的量化方法有哪些。Pekrun 等提出要对认知活动中的情感进行测量,并编制了一个认知—情感量表。该量表将认知情感分为积极的情感、消极的情感和中立的情感^[10]。积极的情感包括愉悦和好奇,消极的情感包括沮丧和厌倦,惊讶被当成是一种中立状态的情感。但是基于量表对情感进行测量的方法是回顾性的,这种方式无法及时地感知学生的情感状态,且不太适用于在线学习环境的测量。

日常生活中,人类可以通过语言、姿态和面部表情感知他人的情感。在线学习一定程度上是一个人机交互的过程,因此,让机器能够识别学生的情感是一种高效可行的方法。近年来,教育研究者基于数据挖掘、计算机视觉等方法对学生的学习状态进行测量。其中,基于计算机视觉的情感识别方法的有效性已得到大量验证,Ashwin 搭建了一个卷积神经网络对学生的无聊、投入和中性 3 种情感状态进行了识别,对于扮演的的情感识别精度达到了 86%,对于自发情感的识别精度为 70%^[11]。Xu 为识别智慧学习环境下学习者的情感,设计了 7 层卷积神经网络结构,在自建学习者情感数据库上训练,结果表明,该模型能较好地识别学习者情感^[12]。

尽管已有许多学者依据面部表情实现了学习者的投入度检测或情感识别,但由于在学习过程中无法保证人脸可以全程被摄像头捕获,而单一模态的数据缺失会影响情感识别或投入度检测的结果。因此,有研究者提出了融合面部表情和其他模态的数据进行多模态识别的方法。例如,国内研究者 Ma 结合面部表情、肢体动作和在线学习者评论的文字和表情设计和实现了多模态的在线学习情感分析模^[13]。Jia 等^[14]采用人工智能方法,基于课堂视频,对学生的位置分布、面部表情和姿态特征进行统计分析,分析课堂的活跃度。Chen 等基于视频中的面部表情和心率检测写作过程中的学习者的投入程度,该研究首次将远程心率识别用于投入度的自动检测,但是由于当前的视频心率检测技术尚不成熟,其鲁棒性和精准性还有待提升^[15]。

学习情感识别已经形成了一定规模,但仔细梳理可以发现,我国目前有关中小学生学习情感的研究集中在传统课堂,关于在线学习情感识别的研究并不多,主要原因在于:疫情

之前,我国在线教育并未得到普及,业界对于此类研究的需求较少;其次,基于计算机视觉的方法通常建立在大样本数据库上,而当前公开的真实在线学习场景下的情感数据库较为缺乏,限制了此类研究的推进。

机器学习人类情感的灵敏度和准确度取决于情感数据集的质量,由于情感的表征有多种形式,情感数据库所涵盖的数据各不相同。Lucey 等收集了来自 123 个对象的五百多个视频帧序列,制作了 CK+ 数据集^[16]。该数据集将视频转化成帧序列进行存储,每个帧序列都被标记了 7 种基本情感中的一种。需要注意的是,该数据集是实验室中收集的扮演情感,并非真实情感的流露。AFEW 数据集是从各国影片中剪辑出的带有情感语义的视频片段,该数据集标签为 6 种基本情感和中立。由于视频中带有自然的表情、各种头部姿态、遮挡和照明,AFEW 极具挑战。OMG 是 Barrors 等在 YouTube 上征集的数据库,该数据库收集人们在自然无约束环境下的视频,注释者通过语音提示等信息将视频剪辑成片段,并基于二维情感模型为片段进行情感标注^[17]。每个片段都有两个数值标签,分别代表情感激活程度和正负性。所有的数据中 2441 个作为训练集、617 个作为测试集、2229 个作为验证集。Aff-Wild 同样是在 YouTube 上收集的,与 OMG 不同的是,Aff-Wild 的注释是帧级的,视频的每一帧都有两个标签^[18]。其中训练集 1 008 650 帧,测试集 215 450 帧。基于面部表情进行情感识别是常见的方法,但是,情感的表征方式具有多样性,不同情感伴随不同的生理反应,不少研究者提出结合生理数据、表情、姿态等多模态数据的情感计算方法。与此同时,多模态情感数据库的出现正迎合了现实的需要。MAHNOB-HCI 是一个记录情感刺激反应的多模态数据库,旨在进行情感识别和内隐标记研究^[19]。它基于多个传感器同步记录音频信号、眼睛凝视数据、面部视频和周围/中枢神经系统生理信号,通过情感视频诱导的方式诱发受试者的情感体验,数据标签由受试者使用唤醒、效价、可预测性和情感关键词进行自我标注和自我二次标注。录制的视频、生理数据和注释被分割并存储在数据库中。国内公开通用的大规模数据集较少,研究者们通常依据研究的需要自建数据集,数据集的情感类别定义上具有较大差异。例如,在面向学生情感识别时,Xu 开发了在线学习环境中的学生表情数据库,该数据库包含 70 名研究生的 17456 幅图像,图像标签包括常态、专注、走神、悲伤、惊恐、高兴、愤怒^[20]。He 等构建了自发学习表情数据库,该数据库共包含 22 名学习者的 979 张面部表情图像,所有面部表情图像被标记为惊讶、疲倦、困惑、快乐和中性 5 个表情标签^[21]。目前基于图像的情感计算研究早期多集中在单帧人脸图像的静态识别,因此相关研究最初以人脸图片制作数据库,随着研究者将视角转向基于帧序列的动态情感识别,研究者开始构建视频情感数据库以满足研究需求。

综上所述,随着计算机视觉与教育研究的交叉融合,基于计算机视觉的学生情感自动识别方法受到了教育界的认可。然而,目前基于视频的在线学习情感识别研究还处于起步探索阶段,在在线学生情感表征、特征提取等方面还存在众多深层次问题有待进一步研究,例如如何利用多视觉线索对学生情感进行表征、如何挖掘多视觉线索中内隐的学习情感特征,以及如何实现基于多维度多视觉线索融合的多粒度学习情感识别等,这些均是亟需解决的关键科学问题。在此背景下,

本研究创建了一个学生在线学习情感数据库,为学生情感识别提供了有力的数据支撑。随后,本文考虑将基于视频的非接触式远程生理信号检测方法引入在线学习情感识别研究中,为在线学习分析提供了新思路。最后,本文紧紧围绕视频这一数据源,面向学生的在线学习过程开展情感识别研究,为在线教育提供了感知学习状态的工具,有利于在线教育开展学习预警和个性化教学服务。

2 数据库构建

2.1 数据采集

为了采集在线学习环境下真实学习情感,本研究创设了4种不同的在线学习情境:观看视频、游戏型学习、答题和阅读文本。结合真实课堂调研和专家访谈,并考虑到不同情感的发生机制不同、个体的情感阈值差异等,本文针对不同的被试对象和学习情境分别设计诱导材料,主要对愉悦、专注、困惑和厌倦4种情绪进行诱导,这4种情绪状态来源于 Baker 对于认知情感状态的定性研究^[22]。尽管学习过程中可能还存在其他情绪状态,但本文选择了4种典型的认知情感状态进行检测。本文选取湖北省武汉市 H 小学四、五年级学生和湖南省湘潭市 M 教育培训机构的五、六年级和高一年级的学生为研究对象,共计采集了22名学生在4种在线学习情境下的学习过程数据。通过情绪诱导实验,本研究共采集有效视频数据104份,并同时记录了学生的心率数据。我们将两种数据基于时间戳进行了对齐,视频记录为MP4格式,心率值记录为CSV格式。

2.2 数据处理

本研究对采集到的数据进行了处理,包括剪切、对齐和标注。首先将视频剪切为6s左右的视频片段,并记录视频对应时间的心率值。随后为学习视频打上情感标签。情感标签和

心率值均在视频片段的文件名中标出,并将其记录在 Excel 表中。例如,对编号为8的学生进行数据采集,该数据是第三次实验采集的,数据采集使用了编号为2的摄像头,包含了看视频(编号为01)、游戏型学习(编号为02)、答题(编号为03)、阅读文本时(编号为04)的学习过程数据,因此该学生在看视频时的数据在 Excel 表中存储名为03-02-08-01.csv。

2.3 数据标注

为了保证数据标注的可信度,本研究采用了自我报告和专家标注并行的方式对视频进行标注。视频标签包括愉悦、专注、困惑和厌倦4种学习情感状态。

首先,我们对标注者进行培训,培训内容围绕视频的编码方式、数据标记规则、学习情感的定义和情感的表征等方面进行。然后,完成培训的标注者参与数据标记,赋予每个短视频以情感标签(图1为情感标注实例)。最后,根据情感诱导实验后的自我报告进行数据库的二次标记。当两次标记出现冲突时表示样本无效,从数据库中删除样本。经过数据筛选,最终得到1211个时长为6s左右的情感标注视频,我们对这些视频的情感标签、心率值都进行了详细记录,数据记录示例如表1所列。



图1 情感标注实例

Fig. 1 Examples of emotion annotation

表1 数据标记记录

Table 1 Records of emotion annotation

| 视频编码 | 保存的文件名 | 开始时间 | 结束时间 | 心率值 | 情感标签 |
|------|-------------|------------------|----------|----------|---------|
| 2 | 02-01-01-03 | 02-01-01-03-01-0 | 00:25.28 | 00:30.44 | 109 |
| 2 | 02-01-01-03 | 02-01-01-03-02-2 | 01:17.48 | 01:24.00 | 114 |
| 2 | 02-01-01-03 | 02-01-01-03-03-2 | 02:02.36 | 02:08.84 | 114,115 |
| 2 | 02-01-01-03 | 02-01-01-03-04-2 | 02:18.00 | 02:26.08 | 112 |
| 3 | 02-01-01-03 | 02-01-01-03-05-3 | 02:26.08 | 02:34.00 | 112,113 |
| 3 | 02-01-01-03 | 02-01-01-03-06-3 | 02:45.32 | 02:51.36 | 114 |
| 3 | 02-01-01-03 | 02-01-01-03-07-3 | 03:05.40 | 03:11.84 | 118 |
| 3 | 02-01-01-03 | 02-01-01-03-08-3 | 03:13.00 | 03:21.12 | 118 |
| 2 | 02-01-01-03 | 02-01-01-03-09-2 | 03:24.64 | 03:29.20 | 113 |
| 1 | 02-01-01-03 | 02-01-01-03-10-1 | 03:34.76 | 03:41.28 | 113,115 |
| 0 | 02-01-01-03 | 02-01-01-03-11-0 | 00:28.68 | 00:34.68 | 109 |
| 0 | 02-01-01-03 | 02-01-01-03-12-0 | 00:37.60 | 00:43.60 | 109 |
| 0 | 02-01-01-03 | 02-01-01-03-13-0 | 01:09.64 | 01:15.64 | 114 |
| 3 | 02-01-01-04 | 02-01-01-04-01-3 | 00:11.52 | 00:17.52 | 106 |
| 4 | 02-01-01-04 | 02-01-01-04-02-4 | 00:34.72 | 00:40.72 | 105 |
| 4 | 02-01-01-04 | 02-01-01-04-03-4 | 00:50.00 | 00:58.32 | 104 |

3 基于视频的在线学习情感识别

3.1 基于表情的在线学习情感识别

3.1.1 迁移学习

当前,深度网络结构的学习能力越来越强大,但是当训练数据过少时,易导致过拟合。尽管教育大数据时代已经到来,但是标注诸如学习情感之类的视频大数据还十分困难。为了

解决学习情感数据标注不足的问题,本文采用了迁移学习方式对模型进行训练,这种方式能够减少模型训练的时间,加快模型收敛的速度,减少过拟合的风险。本研究使用了在 MS-Celeb-1M 人脸识别数据集^[23]和 FER+表情数据集^[24]上预先训练的 ResNet18 作为预训练模型。

3.1.2 改进的帧注意网络

尽管视频蕴含了丰富的动态情感特征,但是在在线学习

中,表情情感的面部表情可能稍纵即逝,视频中的所有帧并非都有助于情感识别。因此,本文使用帧注意力网络进行在线学习情感识别^[25],该网络包含特征嵌入和特征聚合两个模块。在特征嵌入模块,将生成的帧特征表示传入特征聚合模块;在

特征聚合模块,有两个注意模块,通过这两个注意模块聚合视频的所有帧特征,得到一个紧致的视频特征表示,用于4种学习情感识别。本研究按照 Meng 等^[26]的方法修改了 ResNet18 的输出层,添加了一个特征聚合模块,修改后的网络结构如图2所示。

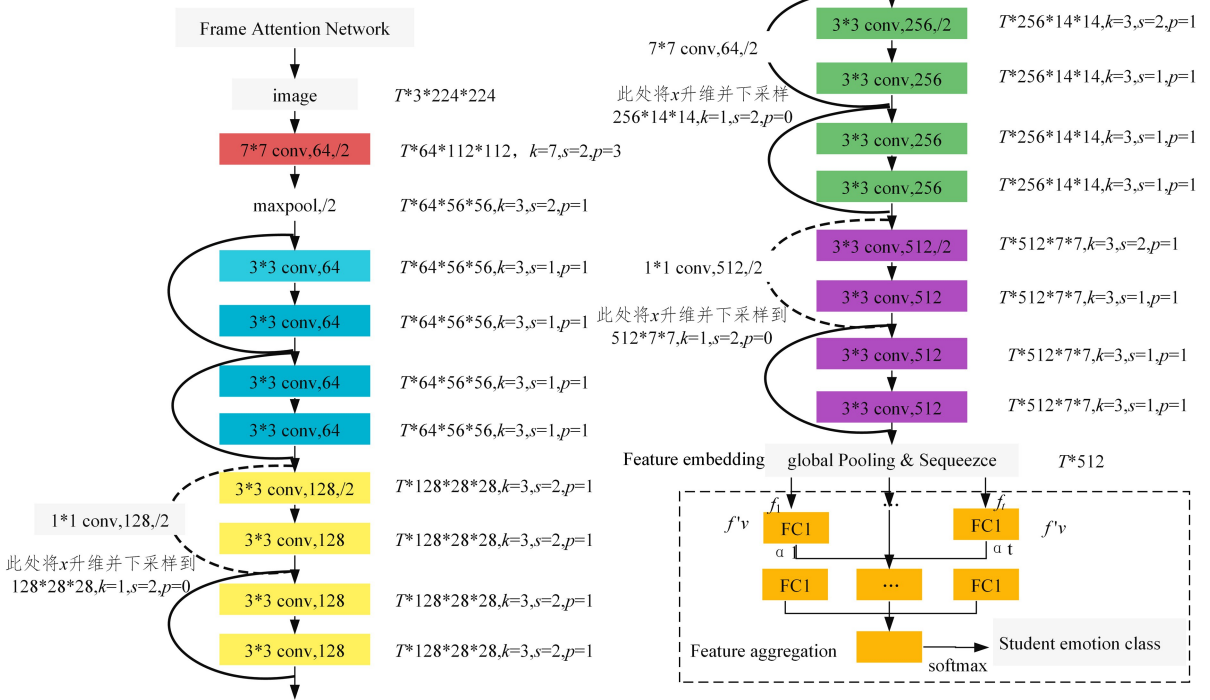


图2 改进的帧注意力网络结构

Fig. 2 Improved frame attention network structure

3.1.3 基于表情的在线学习情感识别

从情感表征的角度而言,人脸表达的情感信息最丰富;从视觉感知的角度出发,相较于静态的图像帧,视频蕴含动态的情感变化。因此,本研究紧紧围绕视频数据,运用深度学习方法挖掘面部特征进行学习情感识别(见图3),该方法结合了静态与动态特征,能更好地识别学习情感状态。本文首先对数据集进行预处理,对数据集进行文件夹分类,从视频中提取相应的关键帧,并进行人脸检测与对齐。其次,将数据集分为训练集、验证集与测试集来设计视频帧采样方式。随后基于 ResNet18 预训练模型进行迁移学习,使用帧注意力网络学习情感特征,训练在线学习情感识别模型。最后,实现基于帧注意力网络的在线学习情感识别。

进行测量,但同样需要紧密接触皮肤。接触式的测量方式虽然较为准确,但必然会带来很多不适与不便,尤其在学习场景下,无接触测量更加重要。近年来,基于远程光电容积脉搏波描记法(Remote Photoplethysmography, r-PPG)的心动周期信号测量方法受到广泛关注^[27]。该方法通过实时测量面部因心脏搏动而产生的对不同波长光照吸收/反射量的周期性变化,以此来推断心动周期情况以及心率、呼吸频率和心跳变异性等生理指标^[28]。本文采用了 Wang 等^[29]的 POS 方法,实现了基于视频的学生心率检测,并根据心率变化识别学习情感状态。

4 在线学习情感识别结果分析

4.1 实验环境简介

本文实验所用的数据来源于前期通过诱导实验采集构建的学习情感数据库。硬件环境方面,本实验的主要设备配置包括 Intel Xeon E5-2560 处理器, NVIDIA GeForce GTX 1080Ti, Windows10 64 bit;软件环境方面,本实验所用编程语言为 Python,使用 PyTorch 框架。本文将数据集按照 8:1:1 的比例划分为训练集、测试集和验证集,由于本文构建的学生情感视频数据集的视频样本为 1 211 个,因此训练集、测试集和验证集的数据样本为 976, 122, 123 个。

4.2 基于表情的在线学习情感识别结果分析

4.2.1 模型参数选择

本文采用交叉熵损失作为模型的损失函数,并采用随机梯度下降算法进行优化。动量设置为 0.9,权重衰减为 1×10^{-4} 。为了得到最佳的模型,本文进行了调参实验,对每种

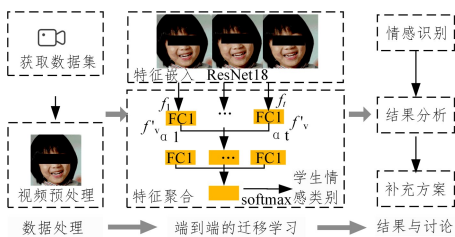


图3 学习情感识别流程图

Fig. 3 Flow chart of learning emotion recognition

3.2 基于视频心率的在线学习情感识别

心率估计和监测对于确定学习者的学习情感状态非常重要,传统的测量心率的方法都是接触式的,如电极式心电图,通过电极片感应人体的心动电流来测量心率,在可穿戴设备如智能手表上,最常用的是利用光电容积脉搏波(PPG)来

参数设置都训练了 100 个 epochs。为了能够保存训练中的最佳模型,每训练一个 epoch,使用验证集验证一次,当验证集精度升高时保存模型,选择训练得到的最佳模型在测试集上评估。学习率,本文先后设置了 1×10^{-4} , 4×10^{-4} 和 1×10^{-5} 。设置 batch size 为 8, 32 和 64, 调参结果如图 4 所示。其中横坐标代表不同的学习率,纵坐标代表不同的批处理样本大小。可以看出,当学习率设置为 10×10^{-5} , batch size 为 32 时,训练得到的模型精度最高。

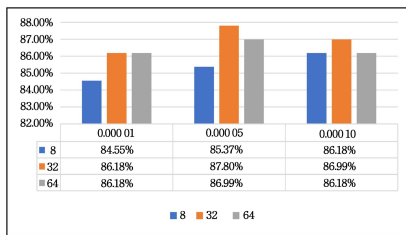


图 4 调整超参数实验结果

Fig. 4 Experimental results of adjusting super parameters

4.2.2 结果分析

经上述分析,本文得到了性能最佳的学习情感识别模型,该模型对 4 种学习情感状态的识别精度达到了 87.804%,从图 5 中可以看出,愉悦和厌倦的精确度最高,达到了 96.30% 和 97.02%,与常用的二维情感模型对应,这两种情绪分别位于唤醒度的两级,对于人类的视觉感知而言,也是易于识别的情感状态。但是专注和困惑两种状态的识别精度较低,主要是因为处于这两种状态的面部表情较为相似,区分度不高。因此,如何设计相应的算法提高这两种在线学习情感状态具有重要意义。

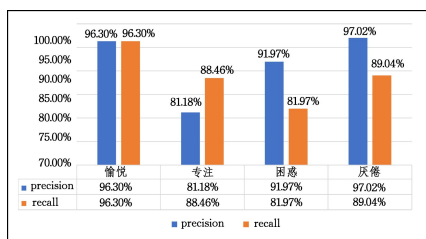


图 5 不同情感状态识别准确度

Fig. 5 Recognition accuracy of different emotional states

4.3 基于视频的在线学习情感识别分析

上一节实验结果表明,本研究提出的基于表情的方法对于困惑和专注两种情绪状态的识别精度比愉悦和厌倦低,且这两种情绪易于混淆。因此,本文尝试使用一种新颖的方式——r-PPG 法来分析不同情感状态的特征,图 6 给出了其中一名学生的视频心率检测结果。

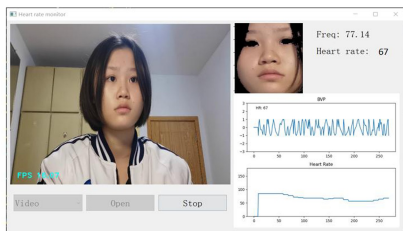


图 6 学生视频心率检测示例

Fig. 6 Example of student video heart rate detection

10 个,共计 200 个视频,并从每个视频中采样两个心率值,代表该状态的心率值。本文依据不同情感状态下的视频心率值来分析学生不同情感状态下的心率特征。

对于不同情感状态下的心率特征,本文进行了方差分析,分析结果如表 2 所列。对于每种情感状态,本文均采样了 100 个心率值数据。4 种不同情感状态下的心率平均值显示,困惑状态的心率值最高,厌倦状态的心率值最低。通过方差分析,F 分布为 10.416,其对应的 P 值为 0.0001。

表 2 不同情感状态的心率水平差异分析

Table 2 Analysis of heart rate differences in different emotional

| 情感状态 | N | 心率值 | |
|----------|-----|-------------|--------|
| | | 平均值 | 标准差 |
| 愉悦 | 100 | 85.46 | 12.567 |
| 专注 | 100 | 84.07 | 13.661 |
| 困惑 | 100 | 92.83 | 17.415 |
| 厌倦 | 100 | 81.9 | 14.687 |
| F 值 | | 10.416 | |
| LSD 事后检验 | | 3 > 1, 2, 4 | |

经过 LSD 事后分析,发现困惑与其他 3 种情感状态之间的显著性均小于 0.001,因此认为困惑与其他 3 种情感状态的心率值具有显著性差异。进一步观察表明,愉悦与专注的显著性为 0.504,专注和厌倦的显著性为 0.297,愉悦与厌倦的显著性为 0.087,继而得出结论:愉悦与专注、专注与厌倦之间不存在显著性差异,而愉悦和厌倦之间存在显著性差异。

在基于面部视频的情感识别中,困惑和专注易混淆,而视频心率与情感状态的方差分析显示,困惑的心率水平极其显著。因此,可以将视频心率检测作为在线学习困惑检测的补充方案,在多模态数据驱动下,基于视频数据、融合心率和面部表情特征,对在线学习情感状态做进一步的判别。

结束语 本文在学习分析理论的指引下,紧紧围绕视频数据开展在线学习情感识别研究。本文借鉴心理学研究的情绪诱导法,构建了包含心率数据和面部表情数据的学习情感视频数据库。围绕在线学习情感识别,本文采用改进的帧注意力网络进行表情识别,该模型对于愉悦、专注、困惑、厌倦这 4 种学习情感状态的识别精度分别为 96.296%, 81.18%, 81.97%, 97.01%。同时,本文验证了视频心率对于识别困惑情感的可行性。本文的研究可为无接触式多模态学习情感分析研究奠定基础^[12-13],后续我们将结合面部表情和视频心率特征开展多模态的学习情感识别的研究。

参考文献

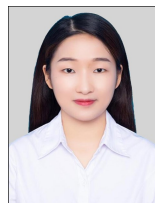
- [1] HUANG C Q, TU Y X, YU J H, et al. Research and implementation of Data-driven Online Learning Burnout Early Warning Model[J]. E-education Research, 2021, 42(2): 47-54.
- [2] GUO C L. Solutions to the Lack of Teaching Emotion between Teachers and Students in Online Teaching under the Epidemic Situation[J]. The Chinese Journal of ICT in Education, 2021 (7): 65-69.
- [3] HUANG C. Research on Learner Emotion Modeling and Application for E-learning[D]. Wuhan: Central China Normal University, 2014.
- [4] WEI R J, DING Y Z, ZHANG L, et al. Design and Implementa-

本文从每位同学的 4 种情感视频中各取出视频序列共

- tion of Emotion Recognition Module in Online Learning System [J]. *Modern Educational Technology*, 2014(3):115-122.
- [5] CAO X M, ZHANG Y H, PAN M, et al. Research on Recognition Method of Learning Participation from the Perspective of Artificial Intelligence-Analysis of Deep Learning Experiment based on Multimodal Data Fusion[J]. *Journal of Distance Education*, 2019(1):32-44.
- [6] TAO X M, CHEN X Y. Research on emotion recognition based on eye movement features in online learning environment[J]. *Computer Technology and Development*, 2021, 31(3):186-190.
- [7] YU Z T, LI X B, ZHAO G Y. Emotion Recognition and Education[J]. *artificial intelligence*, 2019(3):29-36.
- [8] PISE A, VADAPALL H, SANDERS I. Facial emotion recognition using temporal relational network; an application to E-learning[J]. *Multimedia Tools and Applications*, 2020(2):1-12.
- [9] DINDAR M, JARVELA S, AHOLA S, et al. Leaders and followers identified by emotional mimicry during collaborative learning; A facial expression recognition study on emotional valence[J]. *IEEE Transactions on Affective Computing*, 2020, 13(3):1390-1400.
- [10] PEKRUN R, VOGL E, MUIS K R, et al. Measuring emotions during epistemic activities: the Epistemically-Related Emotion Scales[J]. *Cognition & emotion*, 2017, 31(6):1268-1276.
- [11] ASHWIN T S, RAM M R G. Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks[J]. *Education and information technologies*, 2020, 25(2):1387-1415.
- [12] XU Z G. Emotion Recognition of Learning Picture in Intelligent Learning Environment and Its Application[D]. Jinan: Shandong Normal University, 2019.
- [13] MA Y T. Design and Implementation of Emotion Analysis Model for Online Learning based on Multimodality [D]. Nanjing: Nanjing Normal University, 2019.
- [14] JIA L Y, ZHANG Z H, ZHAO X Y, et al. Classroom Student State Analysis Based on Artificial Intelligence Video Processing [J]. *Modern Educational Technology*, 2019, 29(12):82-88.
- [15] CHEN X, CHENG J, SONG R, et al. Video-Based Heart Rate Measurement; Recent Advances and Future Prospects[J]. *IEEE Transactions on Instrumentation and Measurement*, 2018, 68(10):3600-3615.
- [16] LUCEY P, COHN J F, KANADE T, et al. The Extended Cohn Kanade Dataset(CK+): A complete dataset for action unit and emotion-specified expression[C]// *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. 2010:94-101.
- [17] BARROS P, CHURAMAN N, LAKOMKIN E, et al. The omg-emotionbehavior dataset[C]// *International Joint Conference on Neural Networks (IJCNN)*. 2018:1-17.
- [18] LI J, CHEN Y, XIAO S, et al. Estimation of affective level in the wild with multiple memory networks[C]// *IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2017:1-8.
- [19] WIEM M B H, LACHIRI Z. Emotion classification in arousal valence model using MAHNOB-HCI database[J]. *International Journal of Advanced Computer Science and Applications*, 2017, 8(3):318-323.
- [20] XY Z G, ZHANG G W, MENG X Z, et al. Learners' Emotion Recognition and Application Based on Deep Learning[J]. *e-EDUCATION RESEARCH*, 2019, 40(2):87-94.
- [21] HE X K, GAO Q, LI Y Y, et al. Research on Spontaneous Learning Expression Recognition Method Based on Deep Learning Model [J]. *Computer Applications and Software*, 2019, 36(3):180-186.
- [22] BAKER R S, D'MELLO S K, RODRIGO M, et al. Better to be frustrated than bored: The incidence and prevalence of affect during interactions with three computer-based learning environments [J]. *International Journal of Human-Computer Studies*, 2010, 68(4):223-241.
- [23] GUO Y, ZHANG L, HU Y, et al. Ms-celeb-1m: A dataset and benchmark for large-scale face recognition[C]// *European Conference on Computer Vision*. Cham: Springer, 2016:87-102.
- [24] BARSOUM E, ZHANG C, FERRER C C, et al. Training deep networks for facial expression recognition with crowd-sourced label distribution [C]// *Proceedings of the 18th ACM International Conference on Multimodal Interaction*. 2016:279-283.
- [25] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016:770-778.
- [26] MENG D, PENG X, WANG K, et al. Frame attention networks for facial expression recognition in videos[C]// *2019 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2019:3866-3870.
- [27] CHEN X, CHENG J, SONG R, et al. Video-based heart rate measurement: Recent advances and future prospects[J]. *IEEE Transactions on Instrumentation and Measurement*, 2018, 68(10):3600-3615.
- [28] NIU X S, HANG H, SHAN S G. Review of Physiological Index Measurement Methods Based on R-ppg[J]. *Journal of Image and Graphics*, 2020, 25(11):2321-2336.
- [29] WANG W, DENG B A C, STUIJK S, et al. Algorithmic principles of remote PPG[J]. *IEEE Transactions on Biomedical Engineering*, 2016, 64(7):1479-1491.



WEI Yan-tao, born in 1983, Ph.D, associate professor, is a member of China Computer Federation. His main research interests include educational artificial intelligence, computer vision and machine learning.



HU Mei-jia, born in 1997, postgraduate. Her main research interests include educational artificial intelligence, computer vision and machine learning.