

基于面部表情的学习困惑自动识别法

江波 李万健 李芷璇 叶韵

(浙江工业大学教育科学与技术学院, 杭州 310023)

[摘要] 学习情绪是学习者模型的重要内容,如何识别学习者的学习情绪是下一代智能教辅系统的关键技术。困惑是最常见的学习情绪之一,及时识别并解决困惑有助于提高学习效果,然而困惑情绪内隐性较强,识别难度较大。本研究设计了一组基于在线测评的困惑诱导实验,提出了一种基于面部表情的学习困惑自动识别算法。研究人员通过设定不同难度的测试题诱导被试产生困惑情绪,同时利用摄像设备实时捕捉学习者的面部表情,提取面部重要特征点,进而利用机器学习算法进行困惑识别。在实验中,本研究使用逻辑回归、支持向量机、K近邻、决策树、随机森林和深度学习等机器学习算法建立学习困惑自动检测模型,并与被试自我报告确定的困惑标签进行对比。实验结果表明,多数分类算法能有效检测学生的学习困惑,随机森林模型的预测性能最佳,平均准确率为71.18%。本研究所提出的方法可为下一代智能教辅系统的学习者情绪建模提供技术支撑。

[关键词] 学习困惑;面部表情检测;情感计算;机器学习;人工智能

[中图分类号] G436 **[文献标识码]** A **[文章编号]** 1007-2179(2018)04-0101-08

一、引言

近年来,大量实证研究表明情绪与学习之间存在关联(Valiente et al., 2012)。生物学研究也发现情绪影响学习的分子机制(Hu et al., 2007)。《地平线报告》(2016 高等教育版)指出,情感计算可能是未来四五年教育领域的重要技术,这一技术旨在开发像人一样思考的计算机,如利用摄像头捕捉面部表情,并利用算法检测和解释学习交互(Johnson et al., 2016)。学习者在学习过程中出现最频繁的情绪有困惑、投入、沮丧、厌倦等(D' Mello & Graesser, 2012)。为了区别人类的六种基本情绪(Ekman et al., 2013),雷曼等(Lehman et al., 2010; Lehman et al., 2008)分析学习者与智能教辅系统交互50小时所经历的情绪,发现困惑是最常见的。蒂梅洛(D' Mello,

2013)对21项研究分析后发现,困惑在最高频次的15种情绪中排名第二。这些研究共包含1430名学习者在多种不同学习情境进行1058小时互动学习产生的学习情绪。

学习情绪是由学习行为的后果引起的,不同时期情绪会发生转化(D' Mello & Graesser, 2012)。学习困惑通常会对学习产生消极影响。例如,在MOOC学习过程中,持续的困惑会增加学生的消极情绪,影响学习成绩,甚至最终导致学生辍学(Yang et al., 2015)。反之,学习困惑及时得到解决,会将情绪过渡到积极状态,学生对问题的回答会趋于正确并得到正面反馈,进而提升教学效果(D' Mello, Person & Lehman, 2009; D' Mello et al., 2014)。因此,教育工作者对学习困惑进行有效检测,并提供有效干预,将有助于改进学生的学习表现。

[收稿日期]2018-05-07 [修回日期]2018-06-22 [DOI编码]10.13966/j.cnki.kfjyjj.2018.04.011

[基金项目]国家自然科学基金项目“基于多目标稀疏优化的多视图聚类方法”(61503340)。

[作者简介]江波,博士,副教授,硕士研究生导师,浙江工业大学教育科学与技术学院,研究方向:学习分析、教育数据挖掘、情感计算(bjiang@zjut.edu.cn);李万健,硕士研究生,浙江工业大学教育科学与技术学院(wanjian_li@163.com);李芷璇,硕士研究生,浙江工业大学教育科学与技术学院;叶韵,硕士研究生,浙江工业大学教育科学与技术学院。

在情感计算中,面部表情是识别情绪的重要途径之一。学生有困惑情绪时,常常伴随有特定的面部表情(Grafsgaard et al., 2011)。早期研究发现,困惑是一种较普遍的不对称表情,特别是皱眉肌的不对称性是识别学习过程中困惑情绪是否出现的有效信号。困惑情绪的面部活动主要在眼部周围,如眼睛缩小,眉毛下降汇聚等(Durso, Geldbach, & Corballis, 2012; 马惠霞等, 2015)。

西尔维娅(Silvia, 2010)将困惑归类为“知识情感”,认为这是一种涉及理解和思考的情感,是一种元认知信号:反映人们对于正在发生事物的不理解。当学生在学习过程中遇到异常、矛盾或僵局,且不确定如何继续下去时,困惑会随之产生。困惑对学习的主要影响体现在两个方面:一方面,学习者长期处于困惑状态,会造成学习积极性减退,导致学习成绩下降以及课程的保留率降低(Lee et al., 2011)。另一方面,当学生能有效调节自身困惑,或者学习环境提供了足够的脚手架帮助学习者解决困惑时,困惑将产生积极影响。对学生情感的准确理解可能有助于改进情感检测,以及设计增加学习参与度的干预措施(Yang et al., 2015)。

学习困惑的检测方法可分四类:基于文本信息、学习行为、生理信号和面部表情。杨等(2015)使用MOOC论坛文本信息和点击数据自动识别表示困惑的帖子后发现,学生表现或暴露于论坛中的困惑越多,他们继续学习MOOC的可能性越低。斯坦福大学的研究人员利用斯坦福Mocpost语料库,训练了一组异构的分类器对论坛发帖内容进行多维度分类,并使用信息检索技术为困惑的帖子推荐对应知识的视频片段(Agrawal et al., 2015)。贝克等(Baker et al., 2012)基于ASSISTments平台的学习行为数据,应用八种数据挖掘算法建立分类模型,自动识别困惑等四种常见学习情绪(Wang, Heffernan & Heffernan, 2015)。博特略等(Botelho et al., 2017)使用循环神经网络(RNN)、长短时记忆网络(LSTM)和门口循环单元(GRU)等深度学习算法,实现基于网络学习行为的学习困惑识别。王等(Wang et al., 2013)发现利用EEG信号区分学生观看MOOC视频片段是否感到困惑的能力很弱,分类器预测学生困惑的准确率较低(平均为51%至56%),与人类观察的效果相差不大。杜尔索等

(Durso et al., 2012)用肌电图监测受试者的面部肌肉(皱眉肌和降口角肌),同时要求被试举手表示困惑,结果发现肌电记录的面部肌肉运动规律与被试自我报告的困惑规律之间存在显著相关;后测中被试做出相应的困惑表情,该困惑表情与肌电图检测的结果一致。

然而,上述四种方法存在诸多局限。基于文本信息的检测方法仅针对留下评论记录的学习者,无法对不参与评论的学生有效检测;基于学习行为的识别方法需要针对不同学习系统开发专用的细粒度学习行为采集系统;利用生理信号检测学习困惑需要学习者佩戴采集脑电信号的专业仪器设备,这些设备通常较为昂贵且需要在特定实验环境下才能实施。由于电脑和移动设备的摄像头非常普及,且硬件设备不贵,近年来基于面部表情的学习情绪识别广受关注(Durso et al., 2012; Whitehill et al., 2014; Grafsgaard et al., 2011)。

目前,基于面部表情的学习情绪识别研究主要集中在关注投入(Engagement),涉及困惑的研究相对较少。同时,这类研究主要集中在国外,国内相关研究刚刚起步,且不同人种的面部特征存在差异,国际同行提出的模型不一定适用于我国的教育情境。为解决以上问题,本研究设计了一组基于在线测评的困惑诱导实验,提出了一种基于面部表情的学习困惑自动识别算法,并通过设定不同难度的测试题诱导被试产生困惑情绪,同时利用摄像头实时捕捉学生的面部表情,提取面部重要特征点,进而利用机器学习算法进行表情识别。本研究提出基于面部表情的学习困惑识别方法,应用逻辑回归、支持向量机、K近邻、决策树、随机森林和深度前馈神经网络等常见的机器学习算法建立分类模型,自动检测学习困惑。

二、研究方法

本研究尝试设计实验诱导学生产生困惑情绪,并基于面部表情实现学习困惑的自动识别。整体研究技术路线如下:在实验室环境下,研究人员通过采集学生在英语测试中的面部表情,并经过一系列数据预处理操作提升数据质量,使用六种机器学习分类算法建立学习困惑检测模型,最后对模型进行评估(见图1)。

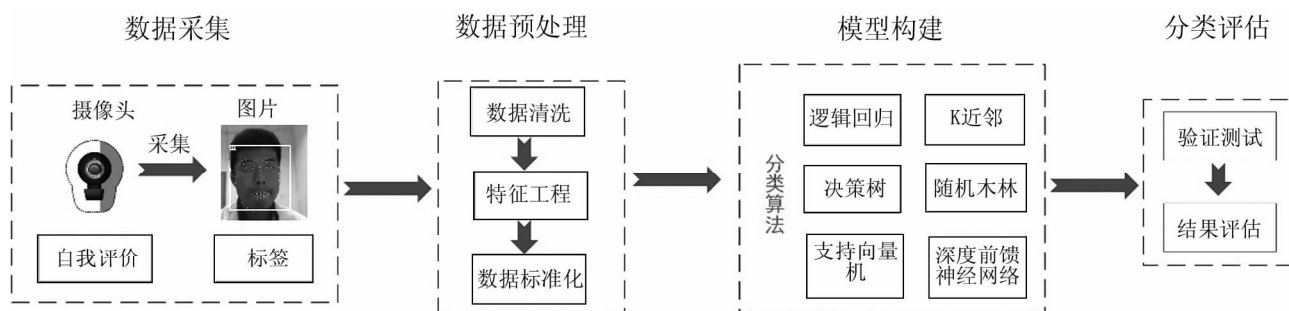


图1 学习困惑检测研究路线

(一) 数据收集

1. 实验流程

本实验选取了某大学 20 名在校研究生进行英语测试,男女比例为 1:1;每位被试需完成 30 道英语测试题,试题来源于知名在线题库“组卷网”(https://www.zujuan.com/)。研究者从题库选择难、中等和容易三种难度的选择题,比例为1:1:1。被试人员英语水平分布为:通过大学生英语六级考试的 40%,已通过四级考试的 50%,未通过四级考试的 10%。本次实验在实验室环境下进行(见图 2),被试在电脑上进行试题作答,研究人员用另外一台电脑通过摄像头(型号: Intel No. VF8010)实时采集被试面部表情。研究者采集学生解答每道题时的 5 张面部表情图片,共采集每个学生 150 张表情图片,整个实验共计采集 3000 张表情图片。



图2 实验室环境下收集学习困惑表情图片

2. 定义数据标签

在基于面部表情识别情绪的相关研究中,国外同行通常根据研究对象的自我报告定义数据标签(Durso et al., 2012; Whitehill et al., 2014)。因此,在测试完成后,被试需要再次浏览所有试题,确定解

答每道试题时的学习困惑状态,并通过自我评价定义困惑标签。为了使被试对困惑标签的定义更客观,本研究仅使用两类学习困惑标签:困惑和不困惑。若被试对英语试题所表达的意思、选项答案、考察知识点等一或多个因素不了解,不知道如何选择则被视为困惑;反之,被试对该道英语试题意思及考察知识点完全清楚,几乎不需太多思考就能选定答案则定义为不困惑。

(二) 数据预处理

1. 数据清洗

实验中,被试思考和解答试题时常会头部晃动,造成面部偏离图像的中央位置,摄像头不能正确捕捉面部特征点,因此,需要对采集的面部特征图像数据进行筛选。摄像头采集的 3000 张图片中,有 492 张未能记录面部特征点或正脸偏离图像中央,删除上述不符合实验要求的数据,共得到 2508 条有效记录。

2. 特征提取

本研究所采集的原始图片规格为 640×480 的彩色照片,共计 307200 个特征,包含大量无关特征和背景噪声,因此需要进行特征提取。本研究利用 ASMLibrary 库直接提取 78 个面部特征点的二维坐标(见图 3)。这些点包括眼睛、眉毛、嘴角等与情绪检测关联较强的面部特征,特征维度为 156。利用特征点构成新的数据集进行困惑检测,能减小运算的复杂度,提高算法效率。

3. 数据标准化

在开始学习困惑自动识别建模任务前,本研究对原始数据进行了标准化处理,利用规范化后的数据构建模型。数据标准化是将数据按比例缩放,使之落入一个小的特定区间。其中最典型的是对数据

做归一化处理,即将数据统一映射到 $[0, 1]$ 区间内。归一化有很多优点,可以提升模型收敛的速度又能提高模型的精度。本文使用的是 z-score 标准化,也叫标准差标准化,即对原始数据的均值和标准差进行标准化。经过处理的数据符合标准正态分布,即均值为 0,标准差为 1,其转化函数为: $x^* = (x - \mu) / \sigma$ 。其中, x^* 为转化后的值, μ 、 σ 分别为所有样本数据的均值和标准差。

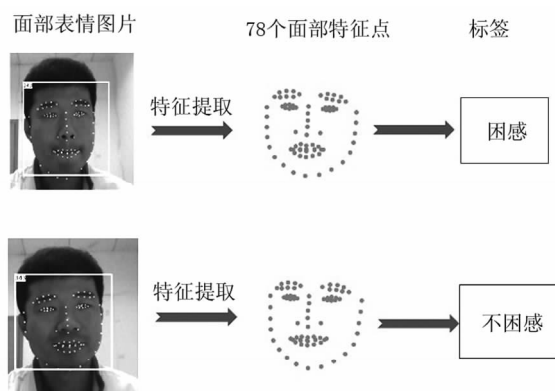


图3 面部特征提取过程

(三) 自动检测模型架构

本研究使用逻辑回归、支持向量机、K 近邻、决策树、随机森林和深度前馈神经网络建立学习困惑自动检测分类模型。逻辑回归 (Logistic Regression, 简称 LR) 是最常见的分类方法之一,常用作预测建模等教育数据挖掘任务(宗阳等,2016)。二元逻辑回归是最简单的形式,即根据一或多个连续或离散自变量预测两个离散因变量的广义线性模型,本研究预测变量只有两个类别,因此使用二元逻辑回归。

支持向量机 (Support Vector Machine, 简称 SVM) 是一种具有坚实的统计理论基础的分类技术,可将数据映射到高维空间进行非线性分类。SVM 应用于高维数据,可有效避免维度灾难问题,本研究采用的支持向量机核函数为高斯核。K 近邻 (K-Nearest Neighbor, 简称 KNN) 是一种基于度量实例间的相似性或距离的“惰性学习”算法,通过距离度量找出训练集中与其最靠近的 K 个训练样本,并使用“投票法”,选择 K 个样本中出现最多的类别作为预测结果。本研究 K 值设置为 3,即通过最邻近的三个样本确定分类标签。

决策树 (Decision Tree, 简称 DT) 是一种解释性很强的分类算法,模型可以通过由节点和分支组成

的树形结构展现出来。常见的决策树算法有 ID3、C4.5、C5.0 和 CART 等,本研究使用的是 CART 算法。随机森林 (Random Forest, 简称 RF) 是一种以决策树为基础分类器的集成方法,其优势是为决策树的训练加入属性的随机选择。具体过程分三步:首先,利用 bootstrap 抽样从原始训练集中抽取 N 个样本,且每个样本的容量与原始训练集相同;其次,对 N 个样本分别建立 N 个决策树模型,得到 N 种分类结果;最后,根据 N 种分类结果对每个记录进行投票表决,决定最终分类(周志华,2016)。

深度前馈神经网络也叫多层感知机 (Multi-Layer Perceptron, 简称 MLP), 是典型的深度学习模型。本研究的深度前馈神经网络由五部分组成(见图 4):一个输入层,五个隐藏层和一个输出层。输入数据的维数为 156,五个隐藏层的神经元数分别为 624、624、312、312 和 156,输出为 0 或 1(0:不困惑;1:困惑)。每个神经元以不同权重与前一层所有神经元相连,数据由输入层经过多个隐藏层传递到输出层。此外,本研究使用反向传播算法最小化当前输出和实际期望输出之间的误差,此过程可降低神经网络在学习迭代过程中的全局误差。为防止反向传播过程中的梯度消失,除输出层的激活函数使用 sigmoid 函数外,其余各层使用 ReLU 函数 (Rumelhart et al., 1986)。

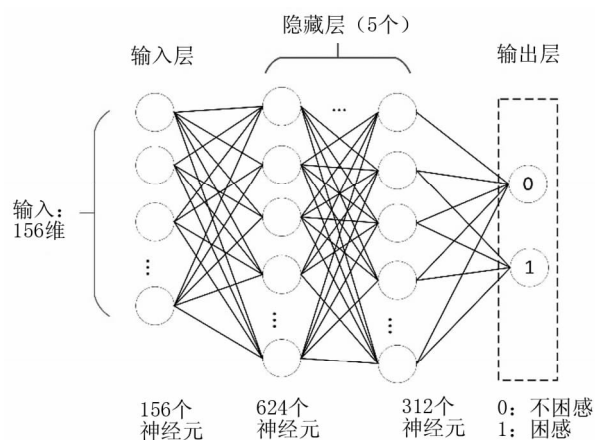


图4 深度前馈神经网络模型结构

三、结果及讨论

(一) 分类性能评估

本研究使用准确率、精度、召回率和 F 值作为评价指标比较各分类器的性能。四种评价指标的计

算公式见表一,其中“真正”(True Positive,简称 TP)对应分类模型正确预测的样本数,“假负”(False Negative,简称 FN)对应分类模型错误预测为负类的正样本数,“假正”(False Positive,简称 FP)对应分类模型错误预测的正类的样本数,“真负”(True Negative,简称 TN)对应分类模型正确预测的负样本数(Tan et al., 2005)。

表一 分类性能度量指标

分类性能度量指标	计算公式
精度(Precision)	$= TP / (TP + FP)$
召回率(Recall)	$= TP / (TP + FN)$
F 值(F-score)	$= 2 * TP / (2 * TP + FP + FN)$
准确率(Accuracy)	$= (TP + TN) / (TP + FN + FP + TN)$

(二) 自动识别学习困惑结果比较

本研究通过 Python 编程语言的 scikit-learn 机器学习库建立学习困惑自动检测模型。由于面部特征点输入属性有 156 维,维度较高,为防止过拟合现象,本研究采用 10 折交叉验证方法进行学习困惑模型训练,通过交叉验证提升模型的泛化能力(Kohavi, 1995)。

1. 六种分类模型学习困惑检测结果

利用逻辑回归、K 近邻、支持向量机、决策树、随机森林和深度前馈神经网络建立的分类模型预测结果见表二,六种分类算法可较好地检测学习困惑,准确率均高于 58.83%。其中,随机森林预测的准确率最高,为 71.8%,K 近邻也取得与随机森林相当的预测准确率,超过 70%;支持向量机、决策树和深度前馈神经网络三种分类模型均可得到约 65% 的预测准确率;逻辑回归模型预测的准确率最低,仅为 58.83%,表现最差。

从表二可以看出,对困惑数据检测的表现整体低于不困惑数据。对于不困惑样本数据,K 近邻取得的精度最高,为 0.74,支持向量机的召回率和 F 值最高,分别为 0.95 和 0.77。对于困惑样本的数据,支持向量机的精度虽最高,召回率却最低。综合表二困惑和不困惑两种样本的平均精度、召回率和 F 值,可得随机森林和支持向量机的平均精度最高,为 0.71;K 近邻和随机森林的平均召回率最高,为 0.71;随机森林的 F 值也是六种分类算法中最高的,为 0.71;逻辑回归的平均精度、召回率和 F 值均最低,分别是 0.58、0.59 和 0.59。

2. ROC 曲线

学习困惑的检测 ROC 曲线见图 5。该曲线是显示分类器真正率(TPR)和假正率(FPR)之间折中的图形化方法(Tan et al., 2005)。真正率沿 y 轴绘制,FPR 显示在 x 轴,在 ROC 曲线中,好的模型靠近左上角部分。图 5 中,随机森林模型在最外侧,最靠近图形的左上角,为最优学习困惑检测模型;相反,表现最差的逻辑回归模型,离左上角距离最远。ROC

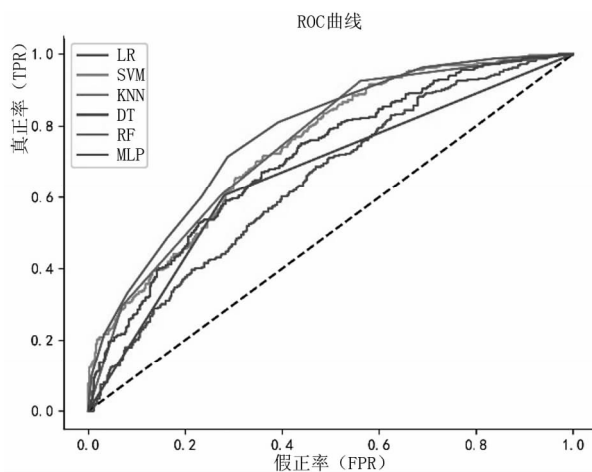


图 5 六种分类算法的 ROC 曲线

表二 六种分类算法的学习困惑检测结果

分类器	准确率	不困惑			困惑		
		精度	召回率	F 值	精度	召回率	F 值
逻辑回归(LR)	58.83%	0.64	0.68	0.66	0.50	0.46	0.48
支持向量机(SVM)	67.33%	0.65	0.95	0.77	0.80	0.28	0.41
K 近邻(KNN)	70.51%	0.74	0.77	0.75	0.65	0.61	0.63
决策树(DT)	66.27%	0.72	0.69	0.71	0.59	0.62	0.60
随机森林(RF)	71.18%	0.72	0.78	0.76	0.66	0.61	0.64
深度前馈神经网络(MLP)	64.44%	0.67	0.78	0.72	0.59	0.45	0.51

曲线下方的面积是另一个评估分类器的指标,面积越大,预测性能越优。总体来看,模型的预测准确率越高,其面积越大。当 FPR 小于 0.2 时,决策树的模型预测能力最弱;当 FPR 大于 0.2 小于 0.5 时,学习困惑分类结果最好的三种算法分别为:随机森林、K 近邻和支持向量机。从总面积看,随机森林面积最大,逻辑回归最小,分别对应最佳和最差的分类效果。因此,从 ROC 的曲线面积和位置可有效比较学习困惑检测模型的性能好坏。

3. 各分类器表现差异的原因

随机森林取得了最高的预测精度,原因在于这是一种集成的方法,将若干决策树模型组合起来,故比单独的决策树模型取得更好的预测表现(周志华,2016)。通过对 K 近邻算法中参数的调整,在 K 为 3 时,取得了超过 70% 的预测准确率。支持向量机对核函数的参数较为敏感,准确率为 67.33%。由于本研究数据的维度较高,决策树的深度较深,可能会造成过拟合现象,最终影响预测的准确率,仅为 66.27%。本研究的深度前馈神经网络有 5 个隐藏层,取得了 65% 左右的准确率。有研究表明,随着深度神经网络隐藏层数目的增加,准确率会有所提高,但模型训练的参数和难度也会随之增加(Hornik et al., 1989)。

(三) 学习困惑检测意义及应用

计算机通过对学习者的情感进行获取、分类和识别,能减轻人们使用计算机的挫败感,帮助理解学习者的情感世界(夏洪文,2008)。使用机器学习分类算法建立学习困惑自动检测模型,及时获取学习者的情感状态,并采取相应的干预手段,可以及时纠正学生的学习行为,对提高教学质量和学习成绩均有积极影响(丁梦美等,2017; Pool & Qualter, 2012)。教学干预在解决学习困惑的同时,有助于学生的知识内化与建构,实现智慧学习。

学习情绪的识别是下一代智能教辅系统的重要组成部分,应用学习困惑检测技术可促进个性化学习(Vail et al., 2016)。学习者在使用智能教学系统时,如果对某些知识点感到困惑,系统将为其推荐相关的学习资源,如与知识点相关的网页、文本和视频片段等,帮助学生摆脱困惑情绪,提升学习效果(陈凯泉等,2017; Agrawal et al., 2015)。此外,将困惑检测技术应用于在线学习平台,有助于降低课程的

辍学率(Yang et al., 2015)。

四、结论与展望

学习困惑检测是设计有效的教学策略和干预措施的重要一步。本研究尝试基于面部表情的困惑识别方法:在英语测试时诱导困惑情绪的产生,并利用摄像头实时采集学生的面部表情,后测中被试通过自我评价确定每道题的困惑标签,对原始数据进行预处理后,应用六种分类算法建立学习困惑自动检测模型。研究表明,六种机器学习分类算法能较好地检测学习困惑,其中,随机森林对学习困惑的检测综合性能最好,是最佳分类模型,这类模型对于未来智能教学系统中学习者困惑的检测有一定参考价值,可为处于困惑状态的学生提供有效的教学干预,对提高学习成绩有一定促进作用。

本研究应用基于面部表情的检测方法,取得了较高的精度,可为学习困惑的识别提供参考,但仍存在较多不足。首先,在图像采集过程中,由于被试答题有时头部会晃动,造成摄像头对特征点的提取不准确,影响数据质量。其次,在后测环节,每位被试需完成 30 道测试题所需时间较长,通过自我评价确定困惑标签准确性会受到时间的影响,可能会造成实验结果的偏差。再次,本研究仅考虑自动检测学习情绪中两种水平的困惑,未来研究可考虑多种学习情绪的多水平识别方法。例如,学习困惑存在多个不同水平(Wang et al., 2013),如参与(Whitehill et al., 2014)、无聊(Kim et al., 2014; Botelho et al., 2017)、沮丧(Cleveland-Innes & Campbell, 2012)等学习情绪也会对学习产生显著影响。最后,本研究所使用的六种常见数据挖掘算法在某些方面仍存在局限,后续研究可考虑使用其他更高效的算法,如卷积神经网络,其优势在于特征的自动提取,对复杂图像的检测精度更高。

[参考文献]

- [1] Agrawal, A., Venkatraman, J., Leonard, S., & Paepcke, A. (2015). Youedu: Addressing confusion in mooc discussion forums by recommending instructional video clips [A]. International Educational Data Mining Society [C].
- [2] Baker, R. S. J. D., Gowda, S. M., Wixon, M., Kalka, J., Wagner, A. Z., & Salvi, A., et al. (2012). Towards sensor-free affect detection in cognitive tutor algebra [A]. International Educational

Data Mining Society[C]. 8.

[3] Botelho, A. F., Baker, R. S., & Heffernan, N. T. (2017). Improving Sensor-Free Affect Detection Using Deep Learning [A]. International Conference on Artificial Intelligence in Education[C] Springer, Cham;40-51.

[4] 陈凯泉,沙俊宏,何瑶,王晓芳(2017). 人工智能2.0重塑学习的技术路径与实践探索——兼论智能教学系统的功能升级[J]. 远程教育杂志, 35(5):40-53.

[5] Clevelandinnes, M., & Campbell, P. (2012). Emotional presence, learning, and the online learning environment[A]. International Review of Research in Open & Distance Learning[C], 13(4): 269-292.

[6] 丁梦美,吴敏华,尤佳鑫,孙众(2017). 基于学习成绩预测的教学干预研究[J]. 中国远程教育:综合版(4): 50-56.

[7] D' Mello, S. (2013). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology [J]. Journal of Educational Psychology, 105(4): 1082.

[8] D' Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning[J]. Learning & Instruction, 22(2): 145-157.

[9] D' Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning [J]. Learning & Instruction, 29(X): 153-170.

[10] DMello, S., Person, N., & Lehman, B. (2009). Antecedent-consequent relationships and cyclical patterns between affective states and problem solving outcomes [A]. Conference on Artificial Intelligence in Education[C]. IOS Press: 57-64.

[11] Durso, F. T., Geldbach, K. M., & Corballis, P. (2012). Detecting confusion using facial electromyography [J]. Human Factors, 54(1): 60-69.

[12] Ekman, P., Friesen, W. V., & Ellsworth, P. (2013). Emotion in the human face: Guidelines for research and an integration of findings[M]. Pergamon Press:55-57.

[13] Grafsgaard, J. F., Boyer, K. E., & Lester, J. C. (2011). Predicting Facial Indicators of Confusion with Hidden Markov Models [A]. Affective Computing and Intelligent Interaction[M]. Springer Berlin Heidelberg:97-106.

[14] Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators[J]. Neural Networks, 2(5): 359-366.

[15] Hu, H., Real, E., Takamiya, K., Kang, M. G., Ledoux, J., & Haganir, R. L., et al. (2007). Emotion enhances learning via norepinephrine regulation of ampa-receptor trafficking [J]. Cell, 131(1): 160-173.

[16] Johnson, L., Adams, S., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). NMC horizon report: 2016 higher education edition [EB/OL]. The New Media Consortium, Austin: 44-45.

[17] Kim, C. M., Park, S. W., & Cozart, J. (2014). Affective

and motivational factors of learning in online mathematics courses[J]. British Journal of Educational Technology, 45(1): 171-185.

[18] Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection[A]. International Joint Conference on Artificial Intelligence [C]. Morgan Kaufmann Publishers Inc:1137-1143

[19] Lee, D. M. C., Rodrigo, M. M. T., Baker, R. S. J. D., Sugay, J. O., & Coronel, A. (2011). Exploring the relationship between novice programmer confusion and achievement[A]. International Conference on Affective Computing and Intelligent Interaction [C] Springer-Verlag:175-184.

[20] Lehman, B., D' Mello, S., & Person, N. (2010). The Intricate Dance between Cognition and Emotion during Expert Tutoring [A]. In International Conference on Intelligent Tutoring Systems[C]. Springer Berlin Heidelberg:1-10.

[21] Lehman, B., Matthews, M., D' Mello, S., & Person, N. (2008). What Are You Feeling Investigating Student Affective States During Expert Human Tutoring Sessions[A]. International Conference on Intelligent Tutoring Systems[C] Springer-Verlag:50-59.

[22] 马惠霞,杨莹莹(2015). 学习困惑:概念、研究范式及其与学习的关系[J]. 心理与行为研究, 13(6):839-845.

[23] Pool, L. D., & Qualter, P. (2012). Improving emotional intelligence and emotional self-efficacy through a teaching intervention for university students [J]. Learning & Individual Differences, 22(3):306-312.

[24] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors [J]. Nature, 323(6088): 533-536.

[25] Silvia, P. J. (2010). Confusion and interest: the role of knowledge emotions in aesthetic experience [J]. Psychology of Aesthetics Creativity & the Arts, 4(2): 75-80.

[26] Tan, P. N., Steinbach, M., & Kumar, V. (2005). Introduction to Data Mining, (First Edition) [M]. Addison-Wesley Longman Publishing Co. Inc.

[27] Vail, A. K., Grafsgaard, J. F., Boyer, K. E., Wiebe, E. N., & Lester, J. C. (2016). Predicting Learning from Student Affective Response to Tutor Questions[A]. International Conference on Intelligent Tutoring Systems[C] Springer International Publishing:154-164.

[28] Valiente, C., Swanson, J., & Eisenberg, N. (2012). Linking students' emotions and academic achievement: when and why emotions matter [J]. Child Development Perspectives, 6(2): 129-135.

[29] Wang, H., Li, Y., Hu, X., Yang, Y., Meng, Z., & Chang, K. M. (2013). Using EEG to Improve Massive Open Online Courses Feedback Interaction[A]. In AIED Workshops [C].

[30] Whitehill, J., Serpell, Z., Lin, Y. C., Foster, A., & Movellan, J. R. (2014). The faces of engagement: automatic recognition of student engagement from facial expressions [J]. IEEE Transactions on Affective Computing, 5(1): 86-98.

[31] 夏洪文(2008). 情感计算及其在人机交互设计中的应用

[J]. 中国电化教育,(1):106-110.

[32] Yang, D., Wen, M., Howley, I., Kraut, R., & Rose, C. (2015). Exploring the Effect of Confusion in Discussion Forums of Massive Open Online Courses [A]. ACM Conference on Learning [C] ACM:121-130.

[33] 周志华 (2016). 机器学习[M]. 北京:清华大学出版社:

179-180.

[34] 宗阳, 孙洪涛, 张亨国, 郑勤华, 陈丽. (2016). Moocs 学习行为与学习效果的逻辑回归分析 [J]. 中国远程教育, 36(5): 14-22.

(编辑:魏志慧)

Automatic Learning Confusion Recognition Using Facial Expressions

JIANG Bo, LI Wanjian, LI Zhixuan & YE Yun

(College of Education Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: *The role of emotion in learning remains largely unexamined and certainly undervalued in the education literature, despite recent molecular research found the evidence of the impact of emotion to learning. Learning emotion recognition is regarded as one of crucial technology in next-generation intelligent tutoring system. Among learning emotion, confusion is one of the most common one, which has a great influence on the learning outcomes. However, automatic recognition of learning confusion is a challenge because of its strong implicitness, compared to basic emotion such as happy and angry. We designed a confusion-inducing experiments within an online test environment and applied machine learning algorithms to detect the confusion emotion reflect from the face of subjects. The subjects were asked to solve a series of test questions with different difficulties within limited time. A high-resolution professional camera was set in front of subjects to capture face image in high frequency. Then, the 78 key feature points of human face were extracted as feature vector, and fed in to a machine learning model to predict confusion. To get a robust prediction model, six popular classification algorithms, i. e., logistic regression, support vector machine, K-nearest neighbor, decision tree, random forest, deep learning, were tested. A 10-fold cross validation was used to compare the six algorithms on five performance metrics that calculated using subjects' self-reported confusion labels. The experimental results show that most of the classification algorithms can detect students' learning confusion effectively and the average accuracy is over 58%, and the random forest is the best prediction model with an average accuracy of 71.18%. The proposed method in this study can provide technical support for the learner affect modeling in the next generation of intelligent tutoring system. The current results are promising, but the evaluation revealed some limitations that could further improve the research. Firstly, the self-reported emotion label is subjective and could be combined with field observation to improve the precision of labeling. In addition, the representation ability of feature point is still very limited. Therefore, in the future, we will attempt to use directly the image to detect confusion with deep neural network technology. Lastly, we also plan to investigate the possibility to simultaneously detect multiple learning emotions, such as confusion, boring and engagement.*

Key words: *learning confusion; detection of facial expression; affective computing; machine learning; artificial intelligence*