

基于深度学习的课堂学习状态评价量化研究*

傅翠娇** 王翼飞

北京航空航天大学计算机学院, 北京 100191

摘要 针对课堂教学中学生人数较多教师很难在授课过程中兼顾每个学生学习状态的问题, 基于卷积神经网络的人脸图像表情分类模型以及人脸识别技术, 对课堂上学生的学习状态进行自动检测, 设计了课堂学习状态评价策略, 综合考虑了PAD三维情感理论、抬头率和侧脸比例等指标, 把指标映射为学生学习的理解度、投入度、兴趣度、专注度和活跃度等学习状态, 提出了一种打分与评级模式, 对学习状态进行综合评价, 采用了真实的学生上课视频进行实例分析, 从时间维度和教学效果两个角度综合评价了学生的学习状态, 验证了评价策略的有效性。

关键字 深度学习, 课堂学习, 学习状态, 评价, 量化

Quantitative Research on Classroom Learning Status Evaluation Based on Deep Learning

Fu Cuijiao Wang Yifei

School of Computer Science and Engineering
Beihang University, Beijing 100191, China;
fucuijiao@buaa.edu.cn, 2274534164@qq.com

Abstract—It is difficult for teachers to balance the learning status of each student during the teaching process in response to the problem of a large number of students in classroom teaching. The learning status of students in class is automatically detected based on a convolutional neural network facial image expression classification model and facial recognition technology. A classroom learning status evaluation strategy is designed, which comprehensively considers indicators such as PAD three-dimensional emotion theory, head up rate, and profile ratio. The indicators are mapped to students' learning understanding, engagement, interest, focus, and activity. A scoring and rating model is proposed to comprehensively evaluate the learning status. Real student class videos were used for case analysis, and the learning status of students was comprehensively evaluated from both time and teaching effectiveness perspectives, verifying the effectiveness of the evaluation strategy.

Keywords—Deep Learning, Classroom Learning, Learning Status, Assessment, Quantification

1 引言

课堂教学是教学活动中最为关键的部分, 学生的课堂学习状态一直是家长与教师关注的重点, 在授课过程中, 教师需要时刻关注学生的学习状态, 人数较多时, 教师很难兼顾每位学生的学习状态, 也无法及时准确地了解课堂的整体效果, 从而很难据此及时调整自己的教学方式以改善教学质量。因此, 对学生的课堂学习状态进行量化研究, 显得尤为重要。

表情是人内心情感的重要体现。在自然状态下, 人的面部五官形态变化, 能够反映出人内心的情感状况。因此, 通过获取学生上课时的表情, 可以分析出他们的内心状态。表情为高兴说明学生喜爱本节课

程, 积极融入其中并发现了有趣之处, 学习状态极好; 表情为中立说明学生的内心情感趋于平静, 没有太大的起伏, 可能专注听课, 也可能发呆走神, 整体表现为一般; 表情苦闷则说明学生对该堂课程的教学不感兴趣, 不愿意认真听讲, 学习状态低下。因此, 只要能够实时收集课堂中学生的表情类别情况, 就可以在在一定程度上分析出他们的学习状态, 这样有利于教师调整上课节奏, 优化教学质量。除表情外, 其他维度的特征也能较好的反映学生的学习状态, 比如抬头、侧身、喝水、看书等常见的行为动作, 也可以体现学生的学习专注度、活跃度等。通过采用图像识别技术, 对这些行为动作进行检测统计, 并映射为学生的学习状态, 进而生成对课堂效果的量化评价, 提高课堂评价的丰富度与准确度。

孙绍涵^[1]等人基于YOLOv4算法对学生的行为动作进行识别, 如玩手机、听讲和交头接耳等等, 用这

*基金资助: 北京航空航天大学教改项目(4006138): 基于深度学习的课堂学习行为量化研究。

**通讯作者: 傅翠娇 fucuijiao@buaa.edu.cn。

些事件的持续时长来反映学生的课堂学习状态。王国琿^[2]等人综合考虑了学生的考勤、面部表情和肢体动作，并分析这些数据，供任课教师参考与总结。郭贇^[3]采用卷积神经网络实现了人脸表情的分类算法，并对学生进行头部姿态估计，综合反映学生的学习情况。针对人体面部这一特定维度，研究人员采取了不同的方式来提高表情分类的准确性。现有的较为出色的表情数据集主要是 Fer2013 与 CK+，大致包含了高兴、中立、难过、生气、厌恶、轻蔑、惊讶和恐惧等八类^[4]，不同表情之间的差别往往十分微小，不易区分。苏悦^[5]等人使用改进后的 ini_xception 网络对表情数据集进行训练，准确率较 ResNet-18 网络提高了 5% 左右，能够更好的检测学生的学习状态与表情变化。钟瑞^[6]等人采用 ResNet-50 进行特征提取，运用嵌入注意力机制的残差网络对表情进行分类，在 RAF-DB(Real-world Affective Faces Database)真实世界面部表情数据库和 Fer2013 表情数据集上分别取得了 87.65% 和 73.57% 的准确率。郑伟^[7]提出了一种基于多尺度注意力机制的密集连接网络表情识别方法，在 CK+ 和 Fer2013 数据集上取得了较高的识别率，分别达到 96.2% 和 85.5%，与 DenseNet121 网络相比提高了 8.4% 和 8.6%。但从目前的相关研究来看，基于表情识别的课堂学习状态评价还存在如下问题^[8]：

(1) 现有的表情识别算法在运行速度上过于依赖系统硬件配置，可能在应用推广时受到限制，如面对表情实时检测等场景。

(2) 学生面部表情与学习状态之间的映射并不绝对，没有规范的标准将课堂行为、表情等因素衡量为学生的学习效果。

(3) 人脸表情的种类较多，差距也十分微小，识别起来并不容易，标注正确且适用广泛的表情数据集、设计高效的算法提升表情识别准确率都是十分重要的事情。

(4) 人脸图片及视频如果不够清晰，或者人脸在其中显示得很小，可能导致系统无法检测出来，从而不能进行后续分析，以至于影响到整体效果。

2 基于卷积神经网络的人脸图像表情分类模型

表情分类技术也称为表情识别技术，其主要目的在于获取图像中人脸的表情类别，常见的人脸表情大致分为高兴、惊讶、伤心、生气、恐惧、蔑视、中立和厌恶等八类，因而表情识别通常属于八分类任务。表情识别技术通过分析图片和视频中人物的面部表情状态，如瞪眼、皱眉、张嘴等微动作，来确定被识别对象的心理状态。卷积神经网络(Convolutional Neural Networks, CNN)是一种具有卷积计算的前馈神经网络，它能够进行表征学习，完成分类、预测等多种任务，

目前已经是深度学习的代表算法之一。利用神经网络对人脸表情数据集进行训练，通过不断的计算图像中的像素值与表情类别的映射函数关系，调整神经网络的参数与权值，从而得到识别准确率较高的表情分类模型^[9]。人脸表情分类算法大致包括人脸数据的获取与处理、高级特征提取和表情种类判断这三个过程，具体设计与实现流程如图1所示。

使用卷积神经网络在优化后的Fer2013plus数据集上进行实验，其中训练集的准确率为87%，测试集的准确率为76%，能够对人脸表情进行相对正确的分类。

3 课堂学习状态评价策略

在课堂上，学生的学习状态可以通过其情绪与行为进行体现，只要能够时刻关注学生的情绪变化与行为方式，并且明晰这些表情和动作的真实含义，就能合理的评估出他们此时的学习状态，进而供教师参考，制定教学对策。

本文采用面部表情PAD值、抬头率以及侧脸比例三类指标，作为评价学生学习状态的依据。通过计算各个指标的真实值，并判断所在区间，即可获取对应的程度值，并加以打分，得到最终的总体评价结果。各个指标的数据值依靠人脸识别技术与表情分类技术加以检测完成。

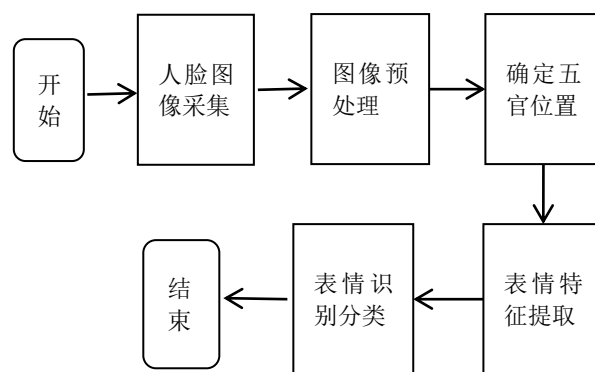


图1 算法设计与实现流程图

3.1 基于表情PAD值的学习状态评价

PAD三维情感模型将情感状态划分为愉悦、激活和优势三个维度，分别对应着P值、A值和D值。PAD值中，P值代表愉悦度(Pleasure)，用于表示情感状态的积极与消极程度；而A值代表激活度

(Arousal)，体现了人体的生理因素强度，如激素分泌等等；D值表示优势度(Dominance)，表示个人对外界事物的影响水平。使用P、A、D这三个维度的值大小，就可以表达出人体特定的情感状态。

PAD值应用于教学心理学中，其含义可以衍生为理

解度、投入度和兴趣度，即学生对课堂上教师讲解的内容是否理解，在上课时学生是否投入其中，以及学生对本堂课程是否感兴趣。

由表1^[8]可知，对于面部表情来说，八类基本表情都有对应的PAD取值，比如快乐，其P值为2.77，A值为

1.21，D值为1.42。通过对表情进行三维量度的转化，就能较好的计算出图像中所有学生的PAD均值。PAD值的大小对应着不同的理解程度、投入程度和兴趣程度，同时根据程度的轻重，本文对其进行了评分，1至5分依次递进，见表2^[8]。理解度、投入度和兴趣度各占5分，能够从多个维度反映出课堂上的学习状态效果。

表 1 八种基本表情的PAD值

序号	表情	P	A	D
1	happy	2.77	1.21	1.42
2	neutral	1.57	-0.79	0.38
3	surprised	1.72	1.71	0.22
4	sad	-0.89	0.17	-0.70
5	fear	-0.93	1.30	-0.64
6	disgust	-1.80	0.40	0.67
7	anger	-1.98	1.10	0.60
8	contempt	-1.85	0.70	0.65

表 2 PAD值和学习状态的对照表

类别	取值范围	程度	评分
P	$(-\infty, -1.35)$	非常不理解	1
	$[-1.35, -0.45)$	比较不理解	2
	$[-0.45, 0.45)$	一般理解	3
	$[0.45, 1.8)$	比较理解	4
	$[1.8, +\infty)$	非常理解	5
A	$(-\infty, 0.3)$	非常不投入	1
	$[0.3, 0.5)$	比较不投入	2
	$[0.5, 0.9)$	一般投入	3
	$[0.9, 1.1)$	比较投入	4
	$[1.1, +\infty)$	非常投入	5
D	$(1.35, +\infty)$	非常不感兴趣	1
	$(0.45, 1.35]$	比较不感兴趣	2
	$(-0.45, 0.45]$	一般感兴趣	3
	$(-1.35, -0.45]$	比较感兴趣	4
	$(-\infty, -1.35]$	非常感兴趣	5

3.2 基于抬头率的学习状态评价

课堂上的抬头次数能够较好的体现学生的专注程度，抬头次数多，则说明学生一直在注视黑板或者课件，跟随着教师的讲解，专心的聆听着，而学生如果一直低头，除去做题或是看书，则很有可能是在走神或者做与本堂课程无关的事情，学习状态不佳。

课堂抬头率，顾名思义是指认真听课的学生数与班级总人数之间的百分比。因此，只要获取课堂上某一时刻的学生抬头人数，并除以班级的总人数，就能计算出这一瞬间的班级抬头率，进而反映学生此刻的专注程度，计算公式见公式(1)，其中H表示抬头率，H_Count表示当前的抬头人数，People表示班级总人数或研究对象总数。本文主要通过采用人脸检测技术，识别图像中的人脸张数，来代表课堂上学生的抬头人数。

$$H = H_Count \div People \quad (1)$$

课堂抬头率与专注度的对照表见表3，抬头率越高，表明学生的专注程度越好，反之则越差，对应的评分为1分至5分，作为衡量总体学习状态的评分之一。

表 3 抬头率与课堂专注度的对照表

抬头率	程度	状态得分
$[0.0, 0.2)$	非常不专注	1
$[0.2, 0.4)$	比较不专注	2
$[0.4, 0.6)$	一般专注	3
$[0.6, 0.8)$	比较专注	4
$[0.8, 1.0]$	非常专注	5

3.3 基于侧脸比例的学习状态评价

课堂上学生出现侧脸，反映的是他们在和同桌互相讨论问题、交头接耳，是课堂活跃度的体现。一般在教师提出某个问题，需要学生进行交流探讨时，侧脸的比例会升高，而当教师授课，学生认真听讲时，侧脸比例则会下降，此时课堂的活跃度也相对降低。侧脸比例与抬头率类似，是利用侧脸总人数除以班级总人数计算得到，见公式(2)，其中F表示侧脸比例，F_Count表示侧脸总人数，People表示班级总人数。本文通过使用OpenCV提供的侧脸级联分类器haarcascade_profileface.xml来完成图像侧脸个数的检测。

$$F = F_Count \div People \quad (2)$$

课堂侧脸比例与活跃度的对照表见表4，侧脸比例越高，表明课堂的活跃度越好，反之则越差，对应的评分为1分至5分，作为衡量总体学习状态的评分之一。

3.4 学习状态综合评价

最终的课堂学习状态评价结果，需要综合表情PAD值、抬头率和侧脸比例三类指标评分得到，具体表现为理解度、投入度、兴趣度、专注度和活跃度，各占5

分, 相加后最低得分为5分, 最高得分为25分。课堂总体得分的具体计算见公式(3), 其中Score表示总得分, P_S表示P值对应的理解度得分, A_S表示A值对应的投入度得分, D_S表示D值对应的兴趣度得分, H_S表示抬头率对应的专注度得分, F_S表示侧脸比例对应的活跃度得分。

$$\text{Score} = P_S + A_S + D_S + H_S + F_S \quad (3)$$

表 4 侧脸比例与课堂活跃度的对照表

侧脸比例	程度	状态得分
[0.00, 0.05)	非常不活跃	1
[0.05, 0.10)	比较不活跃	2
[0.10, 0.20)	一般活跃	3
[0.20, 0.30)	比较活跃	4
[0.30, 1.00]	非常活跃	5

表 5 课堂总体得分评价表

总得分	学习状态等级
5~9	较差
10~14	一般
15~19	良好
20~25	优秀

表5显示了课堂总体得分Score所对应的学习状态等级, 根据分值的高低, 本文将课堂学习状态划分为优秀、良好、一般和较差四个等级, 其中20~25分为优秀, 15~19分为良好, 10~14分为一般, 5~9分为较差。

4 教学实践效果

为了进一步阐释课堂学习状态评价策略的含义, 并验证其有效性, 本文采用了真实的课堂学生上课视频进行实例分析并与测试成绩进行对比, 截取图像示例见图2。视频来自于我校一节大学计算机基础理论课堂, 本文截取了40分钟左右的时长用于具体分析, 研究对象为图中的20名同学。分别从时间维度和教学效果两方面进行分析, 收集的指标结果为人脸表情、PAD值、抬头率和侧脸比例。

4.1 时间维度分析

为了体现学习状态的时间变化特征, 本文将40分钟的视频每隔2分钟截取一次, 共计得到20幅图像。针对这些图片, 本文采用卷积神经网络训练得到的表情分类模型对学生的表情进行提取, 并根据PAD值对照表得到PAD均值, 同时使用OpenCV的正脸检测器和侧脸检测器识别图像, 得到抬头率以及侧脸比例, 最后依据课堂学习状态评价策略获得最终的评价结果。



图 2 学生上课视频截取图像

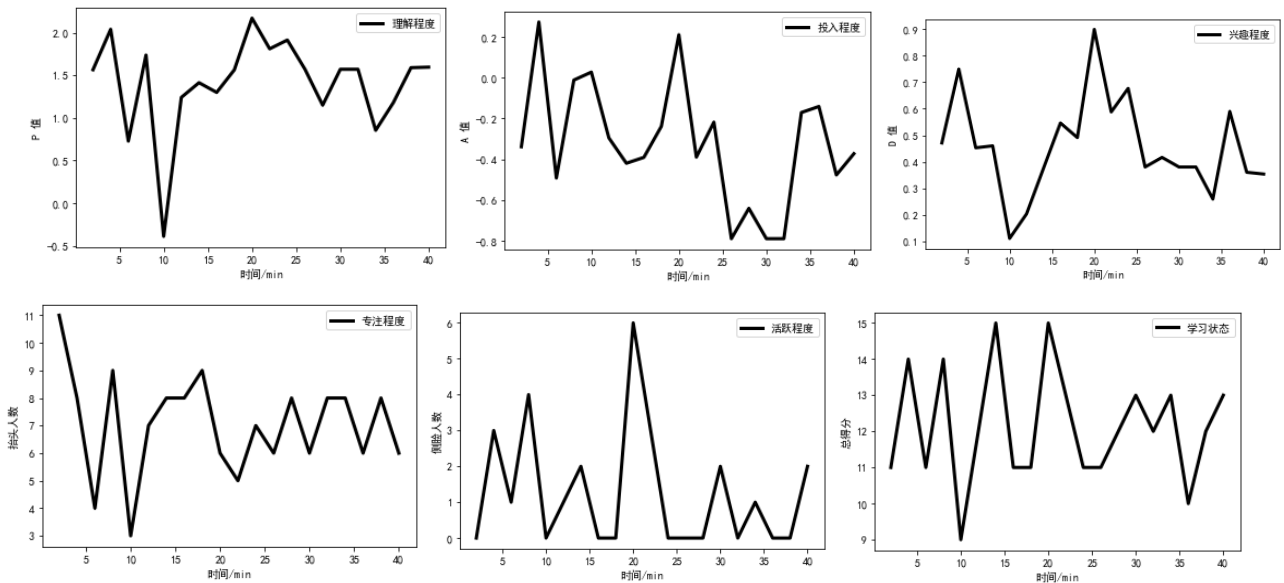


图 3 时间维度各指标统计结果 (a, b...)

图3显示了PAD值、抬头人数、侧脸人数以及总得分的变化趋势, 分别反映出学生的理解程度、投入程

度、兴趣程度、专注程度、活跃程度以及总体学习状态。由图3可以发现, 学生的各项指标往往处于不断变

化之中，而观察整体趋势，却大致相同，比如在第10分钟、第35分钟左右，许多指标都出现了一次低谷，而在20分钟处，却出现了一次峰值。

从原因上分析，学生在最开始上课时，情绪高涨但未完全进入上课状态，因此学习状态不算稳定，往往出现较大波动。在第10分钟左右时，教师正在讲解一道上周的实验课算法题，难度较大，因此可能较多同学没能听懂，故而出现厌倦、痛苦等情绪，导致学习状态得分较低。随后教师开始讲解新内容，各项指标又有了一定的回升。



图 4 作为比照对象的学生

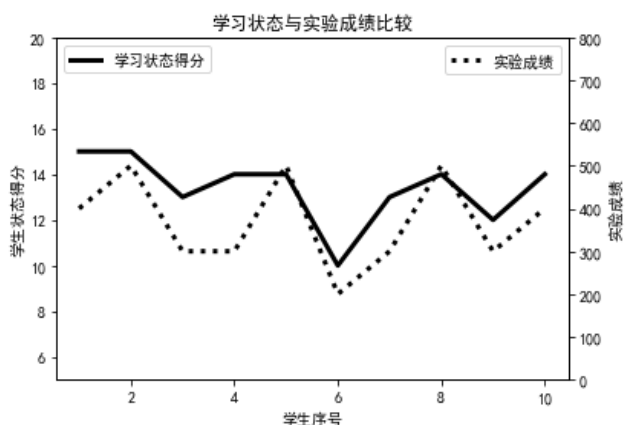


图 5 学习状态与实验成绩比较

在20分钟时，教师给出了一道关于辗转相除法的交流讨论题，班级的活跃度大幅度提升，学习积极性被调动，学习状态良好。

之后，学生可能产生困乏，各状态指标又开始缓慢下滑，而接近下课时，却又开始提升。

通过采用上述时间维度的评价分析方式，教师可以及时了解课堂上学生学习状态变化情况，从而制定对策，以改善教学模式，调整教学节奏，比如可以在学生疲倦时提出问题，增添一些趣味活动，对于较难的知识点，可以多花一些时间讲解或者采取更生动形象的方式展示，这样通过观察图中的时间轴进行点对点分析，能够有效的提高学生的学习状态。

4.2 教学效果分析

(1) 与课后实验成绩对比

为了验证课堂学习状态评价策略的有效性，一共选择了10名具有代表性意义的学生作为实验对象，见图4。从图4中可以了解到，这10位学生的学习状态得分均处于10至15分之间，实验成绩均处于200至500分（满分为500分）之间。对比两条折线图，可以发现，这二者的趋势接近，一方面较好的反映了课堂学习状态与课程成绩的相关性，另一方面也验证了本文所提出的课堂学习状态评价方案的有效性，这对于学生的培养、教学的研究都具有较好的参考意义。

(2) 教学实践前后学生学习成绩的对比

选作实验的教学班级学生共20人，在学生期中考试后开始根据学生的学习状态数据精准地调整教学方式，以实验前学生的期中上机考试成绩和实验后的期末上机考试成绩作为实践新模式下教学效果改善的依据进行对比，虽然学生的成绩受很多因素的影响^[10]，但在其他条件没有改变的情况下，期末考试试题在难度和考试范围都增大的情况下，从图6可以看出教学实践后每个学生的成绩均有提高，也从一个侧面反应出通过精准获取学生的学习状态并据此调整教学方式的方法可以取得更好的教学效果。

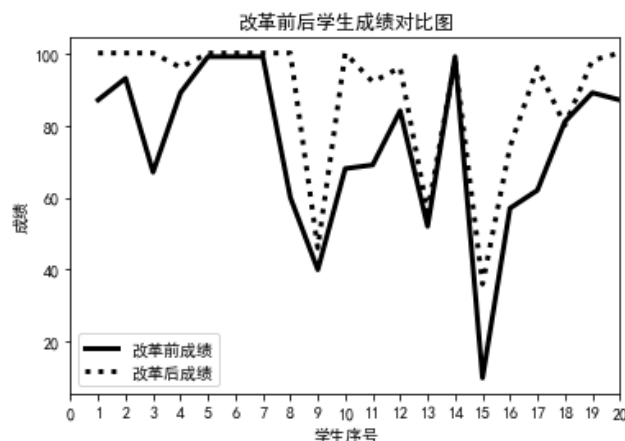


图 6 改革前后学生成绩对比图

5 结束语

本文使用卷积神经网络对表情数据集进行训练，生成了精度为76%的表情分类检测模型，能够识别出人脸的面部表情，并显示为快乐、中立、生气、害怕等8个类别之一。之后设计了课堂学习状态评价策略，综合考虑了PAD三维情感理论、抬头率和侧脸比例，并映射为学生学习的理解度、投入度、兴趣度、专注度和活跃度。然后提出了一种打分与评级模式，在课堂教学的实际应用中，从时间维度和教学效果两个角度综合评价了学生的学习状态和学习效果。后续还可以进

一步提升表情类别、抬头率等指标的检测精度，更好的服务于教学。

参考文献

- [1] 孙绍涵, 张运楚, 王超, 张汉元. 基于深度学习的学生课堂注意力评价[J]. 计算机系统应用, 2022, 31(06): 307-314.
- [2] 王国琿, 张璇, 郑浩. 基于深度学习的学生课堂学习状态分析[J]. 高教学刊, 2022, 8(31): 1-5.
- [3] 郭赞. 基于深度学习的学生课堂注意力评价研究[D]. 山西: 山西师范大学, 2020.
- [4] 林秋晨. 基于面部图像的学生课堂学习状态分析方法研究[D]. 重庆: 西南大学, 2021.
- [5] 苏悦, 杨春金, 王建霞. 基于表情识别技术的学生课堂状态检测[J]. 河北工业科技, 2022, 39(05): 396-403.
- [6] 钟瑞, 蒋斌, 李南星, 崔晓梅. 嵌入注意力机制残差网络的人脸表情识别方法[J]. 计算机工程与应用, 2023, 06(03): 1-13.
- [7] 郑伟. 多尺度注意力机制DenseNet网络的表情识别方法[J]. 软件导刊, 2023, 22(02): 81-86.
- [8] 尹鹏博. 基于面部表情识别的课堂学习状态监测研究[D]. 新疆: 新疆师范大学, 2022.
- [9] 唐强, 张璐平, 夏志远, 彭俊, 符子扬. 表情识别在课堂教学评价中的应用研究[J]. 现代信息科技, 2022, 6(20): 191-195.
- [10] 石程, 赵明华, 黑新宏, 苗启广. 成果导向的教学评价方法探索[J]. 计算机技术与教育学报, 2023, 11(2): 66-69.