

基于深度学习的学生课堂行为识别系统的开发

闫琳英

西安培华学院 陕西西安 710125

摘要: 本研究提出了一种基于深度学习的学生课堂行为识别系统,旨在通过智能化手段实现对学生课堂行为的精准检测与分析。系统采用模块化设计,涵盖数据采集、预处理、模型训练、行为识别及可视化交互五大模块。以 YOLOv8 算法为核心,通过调整迭代次数、学习率等参数优化模型性能,采用 8:2 比例划分 3200 张训练集与 800 张测试集,结合 Mosaic 数据增强、自适应锚框计算及 SPP/PAN 特征融合等技术,提升模型对多尺度目标的检测能力。损失函数包含边界框损失、类别损失与置信度损失,确保检测精度^[1]。

关键词: 深度学习; 学生课堂行为; 识别系统

引言

信息技术跨越式发展趋势下,教育领域逐渐朝向数字化、智能化趋势发展,传统课堂教学模式逐渐无法满足当前教育改革提出的要求,如何在课堂教学中精准掌握学生课堂学习专注度,对学生学习状态进行精准评估,成为广大教师亟待解决的关键问题^[2]。

1. 系统功能模块设计

本次系统设计以模块化设计为基础,系统框架涵盖数据采集模块、数据预处理模块、模型训练模块、行为识别模块、可视化交互模块。

1.1 数据采集模块

数据采集模块从多元信息源收集课程学生表现图片,基于 GitHub 等开源项目抓取开放图像数据的同时,从学校各个监控摄像头中抓取课堂图像数据,保证采集样本数据的丰富性和客观性。本次系统拟采集学生“举手”“低头”“转身”“站立”等课堂行为。

1.2 数据预处理模块

数据预处理模块主要对获取数据进行一系列加工与处理,提高数据时效性。将采集数据进行清理操作,提出清晰度、完整度较差的无用图像后,运用数据增强方法,经过图像亮度变换、模糊变换、图像旋转等方法增加模型的鲁棒性。进而使用 LabelImg 图像标注方法,对图片中的“举手”“低头”“转身”“站立”等课堂行为详细标注,为模型训练提供充足的高质量标注信息。

1.3 模型训练模块

本次模型训练子程序以 YOLOv8 行为检测模型作为模型算法,训练模型中,通过调整模型迭代次数、学习率等模型参数,优化模型及提高模型检测准确率。

1.4 行为识别模块

行为识别模块集成训练好的 YOLOv8 模型,可对输入的图片、视频进行实时行为检测,精准识别学生课堂行为类型,并对每种行为的出现频次进行统计。

1.5 可视化交互模块

PyQt5+Gradio 设计的可视化交互模块可为教师提供检测结果可视化、可交互的人机界面,教师可在界面中选择上传图片或视频、使用摄像头进行检测,系统直接将检测结果以可视化形式显示出来,包括行为类别、置信度、位置信息以及统计信息等,并能够导出检测结果生成报表^[3]。

2. 学生课堂行为识别模型设计

2.1 数据采集与预处理

标注训练集作为模型训练的前期工作,本次研究根据“举手”“低头”“转身”“站立”等课堂行为,具体如下:



图 1 课堂行为图

借助标注工具 LabelImg 对增加后的 4000 张图片进行逐张标注。具体如下：



图 2 标注图

标注过程完全遵照标注规则，避免对目标行为框标注时发生偏离，标注完毕后，同时生成对应的标注 txt 文件，其中包含目标行为名称，标注框中心点坐标，标注框长宽比等数据信息。并生成 class.txt 文件，记录所有行为名，为下一步进行模型训练提供正确的类别对应关系。

2.2 模型设计与训练

本次系统实验环境如操作系统、CPU、编程语言、框架以及 IDE：

表 1 实验环境信息表

项目	配置
操作系统	Windows 10 专业版
CPU	Intel(R)Core(TM)i5.10300H CPU @ 2.50GHZ
编程语言	Python 3.8
框架	Pytorch 1.2.1
IDE	Pycharm 2023.3.4

2.2.1 YOLOv8 网络模型

本系统采用的课堂行为识别方法以改进的 YOLOv8 算法为核心，融合多种先进技术，实现对学生课堂行为的高效、精准识别。

输入数据方面，采用 Mosaic 的数据增强方法，用 4 张随机图片拼接后作为一个新的训练样本，并增加训练样本数量提升难度，使得训练好的模型有较好的稳定性和抗干扰能力。另外加入了适应性的锚框计算和适应性的图片缩放，使模型适应不同的目标行为检测不同尺度的任务，使得检测更加精准。Backbone 部分，由 Focus 结构和 CSP 结构组成。Focus 结构通过间隔采样处理输入图像，在降采样同时保留图像细节信息；CSP 结构将特征图分为两部分，一部分进行卷积处理，另一部分直接传递，随后重新组合处理与未处理

的特征图，降低计算量，提升特征提取效率^[4]。Neck 部分使用 SPP、PAN。SPP 结构对于特征图进行多尺度池化，完成多尺度融合，有利于提高模型对多尺度目标的探测能力；PAN 结构是 FPN 的改进网络，将网络中不同层的特征进行融合以进一步提升模型检测精度。损失函数方面，YOLOv8 的损失函数构成包括三个主要部分：边界框损失、类别损失和置信度损失。

(1) 边界框损失：这部分损失计算的是预测边界框与实际边界框之间的位置偏差。这种损失是为了确保模型能准确地预测目标物体的位置。如公式 1 示：

$$CIoU_{loss} = 1 - CIoU = 1 - \left(IOU - \frac{Distance_2^2}{Distance_C^2} - \frac{v^2}{(1-IOU)+v} \right) \quad (1)$$

在计算机视觉的目标检测任务中，评估预测边界框与实际边界框的匹配程度通常涉及以下几种度量方法：

C：这是指围绕预测边界框和实际边界框形成的最小封闭矩形。Distance_2：这代表预测边界框中心点和实际边界框中心点之间的直线欧几里得距离。Distance_C：此项是指矩形 C 对角线的长度。v：该参数用于评价宽度和高度比例的一致性。IOU：交并比，是指预测边界框与实际边界框的重叠部分面积与两者总面积的交集之比。

(2) 类别损失：表示框中物体类别的准确性。默认使用交叉熵函数来计算。如公式 2 示。

$$p \begin{cases} p, y=1 \\ 1-p, y=0 \end{cases} \quad (2)$$

公式 2 可以简化为公式 3：

$$L = -\log p(3)$$

在上述的 4.2 和 4.3 中，变量 y 代表输入样本的标签，其中正样本标记为 1，负样本标记为 0。而变量 p 则表示模型对输入样本为正样本的预测概率。

(3) 置信度损失：反映预测框中目标存在概率，值在 0 至 1 之间变化，值数较高时，表明在对应的预测框中发现目标的可能性也较高。

通过上述多种技术的协同应用，本系统实现对学生课堂行为的精准识别，为课堂教学管理与教学质量提升提供有力技术支撑^[5]。

2.2.2 模型训练

训练过程中，按照前面整理好的数据集按 8 : 2 分别对训练集和测试集进行划分，共计 3200 张作为训练集数据，800 张作为测试集数据，存储在 VOC 格式的文件夹中，供训练和测试集读取和处理使用。

此外,应在主函数中调整训练所需要的参数,针对需求进行调整迭代次数、图像大小等参数。通过对模型进行多次训练和参数调整,保证在训练集上模型的识别精度始终维持较高水准,关注边界框损失、类别损失和置信度损失等损失函数在训练过程中的变化情况,动态调整模型参数以保证模型的收敛性和推广性,训练出的最终模型在测试集上表现较好的识别效果,为课堂行为准确识别提供依据。

3. 实验分析

本次实验对深度学习学生课堂行为识别系统的整体表现进行评估,详细讨论该系统性能的深度学习学生课堂行为识别系统分析主要包含以下方面的内容:

3.1 识别准确率分析

除测试集 94.9% 的平均准确率外,在真实课堂场景进行现场测试,选择了不同班级不同课的多节课进行实时的行为判断,系统的准确率能达到 93% 以上,对于大多数学生的大部分行为都能够被准确的识别,如 50 人教室的课堂中,有 47 人被系统正确识别,包括举手行为识别 12 次,低头行为识别 18 次,转脸行为识别 7 次,站立行为识别 10 次,只有 3 人由于遮挡等因素行为判断有误。

3.2 识别响应时间分析

针对响应时间,系统平均每张图的检测时间在 0.1s 左右,视频检测帧率为 25fps 以上,可适应实时监测的要求。真实课堂教学中,学生上课时,教师上传照片或打开摄像头后,可在较短的时间获得检测结果,实时了解学生上课情况。

3.3 系统泛化能力分析

为检验系统在各种场景的泛化性,分别对光线不足的普通教室、人数较多的教室以及教室存在背景等场景进行实验测试,结果显示,系统在各场景下均具有较高的识别率,系统具有泛化性较好。如在光线不足环境下,系统对举手行为的识别准确率均能达到 90% 以上,相较于传统方法具有较明显的优越性^[6]。

上文的实验结果表明,本系统利用深度学习进行学生课堂教学行为识别的正确识别率、反应时间、泛化性能均可以满足课堂教学中行为识别的需要。此系统可为课堂管理提供一定的技术支持。

4. 结果与讨论

4.1 实验结果

本实验实现了用深度学习算法对课堂学习行为的学生

进行行为分析,同时通过测试与实验说明其具有较好的可行性与实用性,能够对课堂中常见的学生举手、低头、转头、起立行为进行准确识别,综合准确率达到 94.9%,在实际课堂中的学生课堂行为识别准确率达到 93% 以上,具有较好的实际应用价值。

4.2 实验讨论

本系统开发在教育教学使用方面具有一定的开发意义与创新性,基于深度学习的学生课堂行为识别系统使用实现了学生课堂行为自动化、智能化分析,规避传统教师依靠主观观察课堂学习状态的方式,为教师提供客观、正确学生的课堂学习状态分析数据,帮助教师深入了解学生课堂表现情况,及时优化、调整教学方法,提高教学质量;此外,该系统可实现学生课堂行为的实时统计分析 with 数据统计表显示,通过其显示的表格分析数据,教师可根据学生的行为情况进行有针对性的教学辅导,实现个性化教学,如教师可通过该系统统计数据了解学生课堂发言人数、低头走神人数等相关信息,针对该问题予以关注^[7]。

5. 结论

综上所述,文章以基于深度学习的课堂行为识别系统开展研究,从上述分析可以看出,基于深度学习的学生课堂行为识别系统为教育教学提供新的途径。文章认为,应拓展行为识别范畴,对算法性能进行优化的同时,强化隐私保护,持续完善系统功能,推动深度学习技术在教育领域中纵深发展,为实现个性化教育与教学质量提升奠定良好基础。

参考文献:

- [1] 王禹钧,马致明.基于深度学习的学生课堂行为识别研究[J].软件工程,2023,26(07):40-43+62.
- [2] 闫兴亚,匡娅茜,白光睿,等.基于深度学习的学生课堂行为识别方法[J].计算机工程,2023,49(07):251-258
- [3] 韦宁燕,卢静,韦洪浪,等.基于深度学习的学生课堂专注度识别研究综述[J].信息技术与信息化,2022,(07):192-195.
- [4] 王泽杰,沈超敏,赵春,等.融合人体姿态估计和目标检测的学生课堂行为识别[J].华东师范大学学报(自然科学版),2022,(02):55-66.
- [5] 白佳,谷林.基于改进 YOLOv7-tiny 算法的学生课堂行为检测[J].计算机与现代化,2025,(06):9-15+50.
- [6] 郑一方,陈兴文.基于深度学习的学生课堂行为检测

方法研究 [J]. 山西电子技术 ,2025,(03):92-94.

[7] 史二颖 . 基于深度学习的网络安全异常行为检测技术研究 [J]. 电脑编程技巧与维护 ,2025,(05):167-169.

基金项目: 陕西省教育厅科研专项计划项目“融合

多尺度注意力机制的人体姿态检测模型的研究” (编号: 24JK0580)。

作者简介: 闫琳英 (1990-), 女, 硕士研究生, 讲师, 研究方向为深度学习、图像识别。