

分类号：
学号：20222108055

密级：公开
单位代码：10759

石河子大学

硕士学位论文



基于深度学习的师生课堂行为识别系统 研究与实现

学位申请人	马龙
指导教师	周涛 副教授 李志刚 教授
申请学位类别	专业硕士
专业名称	电子信息
研究领域	计算机技术
所在学院	信息科学与技术学院

中国·新疆·石河子
2025年5月

分类号：
学号：20222108055

密级：公开
单位代码：10759

石河子大学

硕士学位论文



基于深度学习的师生课堂行为识别系统 研究与实现

学位申请人	马龙
指导教师	周涛 副教授 李志刚 教授
申请学位类别	专业硕士
专业名称	电子信息
研究领域	计算机技术
所在学院	信息科学与技术学院

中国·新疆·石河子

2025年5月

**Research and implementation of classroom behavior recognition
system for teachers and students based on deep learning**

A Dissertation Submitted to

Shihezi University

In Partial Fulfillment of the Requirements

for the Degree of

Master of Engineering

By

Ma Long

(Electronic Information)

Dissertation Supervisor: Assoc.Prof. Zhou Tao and Prof.Li Zhi-gang

May,2025

石河子大学学位论文独创性声明及使用授权声明

学位论文独创性声明

本人所提交的学位论文是在我导师的指导下进行的研究工作及取得的研究成果。据我所知，除文中已经注明引用的内容外，本论文不包含其他个人已经发表或撰写过的研究成果。对本文的研究做出重要贡献的个人和集体，均已在文中作了明确的说明并表示谢意。

研究生签名：马茂

时间：2025年05月26日

使用授权声明

本人完全了解石河子大学有关保留、使用学位论文的规定，学校有权保留学位论文并向国家主管部门或指定机构送交论文的电子版和纸质版。有权将学位论文在学校图书馆保存并允许被查阅。有权自行或许可他人将学位论文编入有关数据库提供检索服务。有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

研究生签名：马茂

时间：2025年05月26日

导师签名：周涛

时间：2025年05月26日

摘要

人工智能的快速发展推动了深度学习在教育领域的广泛应用。教师和学生是教学过程中不可或缺的主体，其行为状态的识别与分析对教学评估具有重要研究意义。现有的课堂行为研究主要聚焦于学生行为的识别与分析，并且存在课堂行为类别单一、数据挖掘不足、缺乏完整的可视化分析等问题。本文采用深度学习技术，深入研究基于目标检测算法的师生课堂行为识别方法，分别对教师和学生的课堂行为进行有效识别，从而综合性地评估课堂效率。全面的可视化分析直观呈现课堂行为数据，便于教师及时掌握课堂反馈信息，并通过调整和优化教学策略，能够有效提升课堂教学效果。本文的研究工作如下：

(1) 师生课堂行为数据集构建。针对当前缺少教师和学生课堂行为相关的公开数据集问题，本文从新疆某地区的多个学校采集了大量课堂视频数据。通过抽帧、裁剪、打标签等数据处理，分别构建了真实课堂场景下的教师课堂行为数据集和学生课堂行为数据集，支持课堂场景下教师和学生行为识别研究。经过多次的实验和分析，具有较高的实际应用价值。

(2) 面向大目标的教师课堂行为识别方法研究。本文在 YOLOv7 模型中增加了大目标检测层，利用感受野更大的低分辨率特征层提高模型对大目标对象的处理。同时，在原主干网络末端引入 EMA 注意力机制，通过局部特征信息提高模型对相似行为的识别能力。最后，结合 Inner_IoU 思想和 MPDIoU 损失原理，解决了原始边界框损失函数 CIoU 的局限性。改进后的模型在教师行为数据集上的 mAP@.5 和召回率 R 达到了 96.2%和 92.9%，比原模型提高了 1.1%和 2.3%。实验结果表明，该方法不仅提升了识别精度，还加速了模型的收敛过程。

(3) 基于改进 YOLOv8 的学生课堂行为识别方法研究。由于学生目标较多、背景图像复杂，本文使用 EfficientFormerV2 作为特征提取网络，使用 SCConv 卷积优化 YOLOv8 颈部网络中的 C2f 模块。高效的主干网络增强了特征提取能力，同时，通过自校准卷积灵活生成空间位置附近的远程空间及通道间的关联性依赖，进一步挖掘课堂图像中的潜在细粒度特征信息。改进后的模型在学生行为数据集上的识别精度和召回率达到了 94.3%和 88.5%，性能上明显优于其他方法，充分验证了其有效性。

(4) 师生课堂行为识别系统研究与实现。基于上述的改进模型，本文使用 Java 和 Python 语言实现了师生课堂行为识别系统，旨在可视化分析课堂行为数据。通过教师大目标行为识别模型获取教学行为数据，有助于教师快速了解自身的教学方式，为教学反思和改进教学策略提供重要参考，辅助教师优化教学方案。通过学生行为识别模型及时掌握学生听课情况，重点关注听课效率低、注意力不集中的学生，从而提高学生听课质量。

关键词：课堂行为识别；深度学习；注意力机制；MPDIoU；SCConv

Abstract

The rapid development of artificial intelligence has promoted the wide application of deep learning in the field of education. As the core participants of classroom teaching activities, the identification and analysis of teachers' and students' behavior states have important research significance for teaching evaluation. The existing classroom behavior research mainly focuses on the identification and analysis of student behavior, and there are some problems such as single class behavior category, insufficient data mining, and lack of complete visual analysis. In this thesis, deep learning technology is used to deeply study the classroom behavior identification method of teachers and students based on goal detection algorithm, and effectively identify the classroom behavior of teachers and students respectively, so as to comprehensively evaluate the classroom efficiency. Through comprehensive visual analysis, classroom behavior data are presented intuitively, which is convenient for teaching staff to grasp classroom feedback information in time and make targeted adjustments and improvement measures, so as to improve the quality of classroom teaching. The research work of this thesis is as follows:

(1) Construction of classroom behavior dataset for teachers and students. In view of the lack of open data sets related to teachers' and students' classroom behaviors, this thesis collected a large amount of classroom video data from multiple schools in a region of Xinjiang. Through data processing such as frame extraction, clipping and labeling, the teacher's classroom behavior data set and student's classroom behavior data set under real classroom scenes are constructed respectively to support the research on teacher's and student's behavior recognition under classroom scenes. After many experiments and analysis, it has high practical application value.

(2) Research on the recognition method of teacher's classroom behavior for big goal. In this thesis, a large target detection layer is added to the model of YOLOv7, and the low-resolution feature layer with larger receptive field is used to improve the model's processing of large target objects. At the same time, EMA attention mechanism is introduced at the end of the original backbone network to improve the recognition ability of the model through local feature information. Finally, the limitation of the original bounding box loss function CIoU is solved by combining Inner_IoU idea and MPDIoU loss principle. The improved model achieved 96.2% mAP@.5 and 92.9% recall rates on the teacher behavior dataset, 1.1% and 2.3% higher than the original model. The experimental results show that this method can accelerate the model convergence and improve the recognition accuracy.

(3) Research on students' classroom behavior recognition method based on improved YOLOv8. Due to the large number of student targets and complex background images, EfficientFormerV2 was used as the feature extraction network in this thesis, and SConv convolution was used to optimize the C2f module in

the YOLOv8 neck network. The efficient backbone network enhances the feature extraction capability. At the same time, the dependency relationship between the remote space and the channel around the spatial location is constructed by self-calibrating convolutional adaptive, and the potential fine-grained feature information in the classroom image is further mined. The accuracy and recall rates of the improved model on the student behavior data set reached 94.3% and 88.5%, which were significantly superior to other mainstream methods, proving the effectiveness of the proposed method.

(4) Research and implementation of classroom behavior recognition system for teachers and students. Based on the above improved model, this thesis uses Java and Python languages to implement classroom behavior recognition system for teachers and students, aiming at visual analysis of classroom behavior data. The acquisition of teaching behavior data through teachers' big goal behavior recognition model can help teachers quickly understand their own teaching methods, provide important references for teaching reflection and improvement of teaching strategies, and assist teachers to optimize teaching plans. Through the student behavior recognition model, we can grasp the students' listening situation in time, and focus on the students with low listening efficiency and inattention, so as to improve the quality of students' listening.

Key words: classroom behavior recognition; deep learning; attention mechanisms; MPDIoU; SCConv

目录

摘要.....	I
Abstract.....	II
第 1 章 绪论.....	1
1.1 研究的背景及意义.....	1
1.2 国内外研究现状.....	2
1.2.1 行为识别研究.....	2
1.2.2 学生课堂行为识别研究.....	3
1.2.3 教师课堂行为识别研究.....	4
1.3 研究内容及技术路线.....	6
1.3.1 研究内容.....	6
1.3.2 技术路线图.....	7
1.4 论文组织结构.....	8
第 2 章 相关理论与数据集构建.....	9
2.1 深度学习理论基础.....	9
2.1.1 多层感知器.....	9
2.1.2 卷积层.....	10
2.1.3 池化层.....	10
2.1.4 全连接层.....	11
2.2 目标检测算法.....	11
2.2.1 Faster-RCNN 算法.....	12
2.2.2 SSD 算法.....	13
2.2.3 YOLO 系列算法.....	13
2.3 评价指标.....	15
2.4 师生课堂行为数据集构建.....	16
2.4.1 数据采集.....	16
2.4.2 师生课堂行为类别定义.....	16
2.4.3 数据预处理.....	17
2.4.4 数据集构建.....	19
2.5 本章小结.....	20
第 3 章 面向大目标的教师课堂行为识别方法研究.....	21
3.1 改进 YOLOv7 的教师大目标教学行为识别方法.....	21
3.1.1 大目标检测层.....	21
3.1.2 EMA 注意力机制.....	23
3.1.3 Inner_MPDIoU 损失函数.....	24
3.2 实验分析.....	27
3.2.1 教师课堂行为识别分析.....	27
3.2.2 消融研究.....	29
3.2.3 对比实验.....	32
3.3 教学行为量化.....	33
3.3.1 教学互动率量化.....	33

3.3.2 教学运动轨迹量化	34
3.4 教学行为可视化分析	35
3.4.1 教学互动率可视化	35
3.4.2 运动轨迹可视化	36
3.5 本章小结	36
第 4 章 基于改进 YOLOv8 的学生课堂行为识别方法研究	37
4.1 改进 YOLOv8 的学生课堂行为识别方法	37
4.1.1 EfficientFormerV2 网络	37
4.1.2 自校准卷积	39
4.2 实验分析	41
4.2.1 学生课堂行为识别分析	41
4.2.2 消融实验	43
4.2.3 对比实验	44
4.3 学生课堂专注度研究	45
4.4 学生课堂行为量化	46
4.4.1 学生课堂专注度量化的	46
4.4.2 学生课堂讨论状态量化	47
4.4.3 学生课堂自习量化	47
4.5 学生课堂行为可视化	48
4.5.1 课堂专注度可视化分析	48
4.5.2 课堂讨论与自习可视化分析	49
4.6 本章小结	50
第 5 章 师生课堂行为识别系统研究与实现	51
5.1 需求分析	51
5.1.1 功能性需求	51
5.1.2 非功能性需求	52
5.2 系统设计	52
5.2.1 系统开发环境	52
5.2.2 系统总体设计	53
5.2.3 数据库设计	54
5.3 系统功能展示	55
5.3.1 系统管理	56
5.3.2 数据管理	56
5.3.3 行为识别分析	57
5.4 系统测试	60
5.5 本章小结	61
第 6 章 结论与展望	62
6.1 结论	62
6.2 展望	62
参考文献	64
致谢	68
作者简介	69

第1章 绪论

1.1 研究的背景及意义

随着智慧教育时代的到来，信息技术正逐步被应用于课堂教学环境中，通过人工智能技术获取课堂教学数据，能够有效推动智慧课堂的发展。

近年来，世界各国开始重视信息技术与教育的全面融合与创新。2021年德国开启国家教育数字化平台建设，联邦教研部通过推动数字化教学与学习场景的构建和拓展，增强教师在数字化领域的知识储备和技能^[1]。韩国在《2022年教育信息化实施计划》中明确提出，将前沿智能技术运用到教育信息化进程中，构建以人工智能+ICBM为核心的教育信息化体系，促进教育数字信息的共享^[2]。2022年9月，联合国教育变革峰会强调，教育数字化转型是促进教育创新并实现可持续发展目标的关键路径^[3]。

面对各国在教育领域的投入与改革，我国采取了相应的教育实施计划。2018年，我国教育信息化2.0通过“互联网+教育”综合平台，积极探索信息时代背景下教育治理的创新模式^[4]。2023年6月，河南省鹤壁市教育局以“信息化2.0”示范创建为引领，统筹实施“平台+专网+数字校园+大数据”支撑下的课堂创新“四合一”工程，构建“互联网+”条件下教学应用^[5]。教育部部长在2024世界数字教育大会强调，智能化作为教育变革的核心驱动力，能够有效促进科学教育与文化教育的深度融合，实施人工智能赋能行动，积极推动以智助学、以智助教、以智助研，促进智能技术与教育教学的深度融合^[6]。利用信息技术能够解决传统教学面临的问题，能够有效提升课堂教学效果。在教育逐渐进入智能发展的趋势下，若能结合人工智能和计算机技术，智能化分析教师的授课方式和学生的听课情况，为教师优化教学方案提供参考，将有助于提升教学质量。

鉴于我国东西部地区教育发展存在不均衡现象，特别是西北地区受限于经济发展水平，长期面临教育资源匮乏、教育理念更新滞后等现实问题，其表现为优质师资力量短缺、学生自主学习能力薄弱以及课堂教学效能亟待提升等方面。针对上述问题，本研究将深度学习技术引入课堂教学场景，致力于构建智能化课堂行为分析系统。该系统能够自动识别并统计教师教学行为和学生行为状态，实现对课堂互动质量与学生注意力的多维度分析，并通过可视化技术为教师提供实时、直观的课堂数据反馈。

从教师发展维度来看，系统提供的教学行为分析模块能够帮助教师客观评估教学策略的有效性，为教学反思和方案优化提供数据支撑，从而促进教师专业能力的提升。就学生而言，通过精准的学生行为分析，使教师能够全面掌握个体学习状态，针对学习效率低下的学生实施个性干预，有效提升课堂学习效能。相较于传统依赖人工观察、问卷

调查等耗时耗力的教学评估方式,本研究提供的数据驱动型课堂分析机制不仅克服了传统方法存在的滞后性问题,还大幅度降低了教学评估的人力成本,为西北地区教育质量的提升提供了可行的技术路径。

1.2 国内外研究现状

本节主要将国内外学者对行为识别领域的研究现状进行整理与分析,分别从行为识别、学生课堂行为识别和教师课堂行为识别逐层展开描述。根据研究方法的不同,对相关研究进展进行分类和讨论,总结相关领域的研究现状并揭示现阶段存在的问题,基于此提出本文的研究内容。

1.2.1 行为识别研究

人体行为识别技术旨在利用摄像头收集有关外部运动的相关信息,检测并跟踪人体、提取人体行为特征,以及了解和分析人的行为^[7]。国内外针对人体行为识别的研究机构有很多,2021年6月,美国麻省理工学院计算机科学与人工智能实验室提出一种新型触觉感应地毯,在不使用相机情况下通过触觉信息来判断人体姿势。卡内基梅隆大学的CMU Perceptual Computing Lab团队提出了Openpose人体姿态检测模型,可以实现实时的多人姿势估计和人体行为检测。国内有国防科学技术大学、哈尔滨工业大学以及中科院等高校研究团队也在积极研究中,这些高校在人体行为分析与识别中取得了相对不错的成绩。

无论国内还是国外,人体行为识别都以惊人的速度在发展,目前行为识别研究方法主要有基于传统的手工特征方法和基于深度学习的特征学习方法。

(1) 基于手工特征的方法

传统的行为识别流程是首先对所收集的视频进行采样,然后对样本进行特征提取,接着使用编码器对特征编码,对编码得到的特征向量规范化,最后由分类器分类。针对现有方法存在时空依赖性特征提取不足问题,卿入心^[8]设计了一种借助多粒度时空编码器来实现骨架行为识别的方案,通过多粒度空域编码器和多粒度时域编码器提取关节粒度中行为的时间信息和空间信息。为捕捉到更具辨别力的人体运动特征,徐伟尧^[9]提出了一种多级帧选择采样算法,通过提取关键帧得到多个时序层次的样本,再计算多个视图的运动历史和静态历史图。曲梦杰^[10]使用手工特征的传统机器学习方法构建人体行为识别模型,设计了基于基尼不纯度的决策树节点划分标准,能够在特征选择过程中有效平衡行为识别与用户泛化的需求。

(2) 基于深度学习的方法

基于深度学习的方法将特征提取、特征编码、特征分类三个阶段使用一个网络模型

代替,通过反向传播算法自动学习行为特征。Mo 等人^[11]通过多任务热图网络(MTHN)生成共享热图,改进后 MTHN 网络的 AP 达到 73.7%。Zhang 等人^[12]设计了改进时空金字塔卷积网络,在处理特定动作识别任务时通过引入边缘权重分数和多层次特征表示来提高模型的性能,能够有效识别日常健身活动过程中的多种行为姿态。有学者采用多尺度扩张辅助剩余注意网络实现人体异常行为分类,通过对异常行为的自动识别,预防部分犯罪事件,甚至在犯罪发生前将其消除^[13]。

基于手工提取特征的方法工作量大、相对麻烦,且表达能力受手动提取特征的影响。基于深度学习的行为识别方法通过网络通过学习自动提取特征,与传统手工提取特征方法相比,性能上有很大的提升,表达能力更好^[14]。

1.2.2 学生课堂行为识别研究

在课堂教学中,学生的听课行为在一定程度上反映该学生的学习状态^[15],如何捕捉学生的课堂行为并进行分析具有重要的研究意义。目前已有国内外学者对课堂中学生的学习行为作出研究。

(1) 结合注意力机制的学生行为识别方法

计算机视觉中注意力机制的使用,使得模型过滤掉不重要的背景信息,着重于图像中的前景对象。Zhang 等人^[16]在 YOLOv3 中引入注意力机制(CBAM)模块,有效提升了模型精度。成文昊^[17]在 YOLOv5 中引入 CA 注意力机制,获得更多感兴趣的目标区域,降低了教室背景因素的干扰,同时结合其他不同的改进策略增强模型的性能,能够有效识别多种学生课堂行为类别。Wang^[18]和 Tang^[19]在 YOLOv5 中分别使用 SE 模块和 CBAM 模块,降低识别过程中背景信息权重,增加目标对象的识别,提高了模型性能。王禹钧等人^[20]为了提高学生课堂行为识别精度,在现有的模型基础上融合 Ghost 模块和 CA 注意力机制,实现模型轻量化的同时提升检测精度。

(2) 基于骨骼关键点的学生行为识别方法

人体骨骼数据具有更好的抗光照、抗遮挡和抗干扰能力,有学者通过对人体骨骼关键点的研究,实现学生行为姿态识别。胡妍^[21]采用域适应学习框架,提出基于时空 TransFormer 的骨骼域适应行为识别方法,将骨骼序列数据进行序列划分与编码,减小模型复杂度的同时关注子动作变化。朱艳^[22]和徐苏杰^[23]等人使用 Kinect 传感器在课堂环境中提取学生骨骼的 3D 坐标数据,实现基于骨架特征提取和 SVM 分类器的课堂行为识别功能。Zhao^[24]等人以 AlphaPose 和 OpenPose 作为关键点检测框架,对构建的学生课堂行为数据集进行验证,结果表明骨骼关键点特征对多种行为的识别起到关键作用。

(3) 改进 IoU 损失函数的学生行为识别方法

损失函数作为目标检测算法中度量预测值与真实值差异的重要指标,对模型的性能起到关键作用。孙绍涵^[25]等人使用 Soft-DIoU 对 YOLOv4 网络进行优化,识别分析教室

中学生课堂行为状态，在一定程度上解决了原损失函数的局限性。胡海涛^[26]使用 DIoU 替换了 DeepSort 算法中的原始 IoU 损失函数，更加精确了预测框和真实框的相交情况，不仅提升了模型对学生课堂行为的识别效果，并且加快了网络的收敛速度。针对遮挡学生难以识别问题，Chen^[27]等人在原检测模型中引入斥力损失函数，通过优化预测框之间的损失值，使模型在多种学生行为上的平均精度超过 90%。为了有效筛选课堂场景下前后排差异过大的目标，Yang 等人^[28]将 biformer 注意力模块和 Wise-IoU 集成到 YOLOv7 网络中，对课堂中学生的举手、阅读和写作三种学习行为进行有效识别。

(4) 改进骨干网络的学生行为识别方法

骨干网络作为提取数据特征的核心结构，是深度学习模型中关键的组成部分。为了有效识别课堂环境下学生的表情变化，刘琦^[29]提出基于 YOLOv8 的改进模型，将模型的主干网络替换为 FasterNet，减少了模型的计算量，更加精准地抽取丰富的空间特征信息。闫兴亚^[30]等人以 YOLOv5 网络为基准，融合 MobileNetV3 模型设计了轻量级姿态识别模型，提高了拥挤场景下的姿态识别准确率，有效提高了模型的建模和表达能力。细粒度分类和定位由于对象类之间的相似性而相互约束，导致识别性能较差。为此，Dang^[31]等人使用自研的浅卷积神经网络模块优化了模型的骨干网络，通过并行的分支结构避免了相互约束的局限性。

1.2.3 教师课堂行为识别研究

教师作为知识的传递者，在课堂环节中扮演者重要角色，对教师课堂行为的识别与反馈是改进课堂教学的重要途径。Zhao 等人^[32]通过 3D 双线性池的行为识别网络来增强 3 维特征的融合表示，强调了教师行为和运动区域特征的深度整合，实现了答疑解惑、指向黑板、板书、擦黑板、操作互动板、提问、来回走动、操作用具等八种教师行为的识别。此外，有学者从人体骨架信息和注意力机制方面，进一步探究了课堂环境下教学行为识别。

(1) 基于骨架信息的教师行为识别方法

Wu 等人^[33]对 RGB 视频进行多模态信息整合，通过 Openpose 获取目标对象的骨架信息，并将其与 RGB 信息相结合，提高了教师行为的识别精度。张安然^[34]探究了融合空洞卷积的教师非语言行为识别方法，使用目标检测算法定位教师位置并获取教师图像，再使用人体姿态估计算法获取 RGB 骨架图像。李佳楠^[35]等人从空间和时间维度上提出多尺度卷积网络，使用多尺度语义特征融合模块在空间维度上捕捉骨架点和肢体部位的特征，有效捕获和分析教师教学行为。徐顺捷^[36]提出了基于表情和动作识别的课堂教学质量评估方法，利用 Openpose 检测骨骼关键点，通过引入高斯噪声的 ST-GCN 模型实现课堂教师关键动作的识别，为教师更细致的课堂状态研究提供了参考。

(2) 结合注意力机制的教师行为识别方法

有学者结合注意力模块的特点,对课堂环境下的相似教学行为进行有效识别。Kamiya^[37]以 CAHRNet 网络作为基准模型,提出基于协调注意力的教师姿态估计方法,使用 CA 注意力机制模块替换 HRNet 模块,提高了不同分支对特征空间位置信息的学习能力。Xu 等人^[38]结合教师的头部、手和身体状态设计了基于人体局部姿态的行为识别方法,利用特征金字塔和 CBAM 注意力模块融合局部姿态和动作行为的多尺度特征信息。针对 PoseC3D 模型对行为识别精度有限,郑益群^[39]采用 Faster-RCNN 和 HRNet 网络对教师姿态进行估计,结合结果数据和教师姿态的特点,引入 ECA 注意力机制,使得模型效果更佳。

综上所述,无论国内外,有众多的学者对课堂环境下的师生行为进行研究。研究的方法不局限于手工特征方法,有学者使用传统的机器学习去解决相应的问题,也有学者结合深度学习模型和人体骨骼关键点信息来识别多种课堂行为。无论使用何种方式,都是完成对学生课堂行为的识别。然而,当前大部分关于课堂行为的研究都侧重于学生对象,将学生的学习行为视为影响学业成绩的主要因素。但在课堂教学过程中,教师作为课堂活动的引导者和发起者,他们的教学方式和教学风格在一定程度上影响着学生听课效率。当老师面向学生的教学行为占比更高时,意味着老师和学生将会产生更多在眼神、语言、动作等方面的交流。更多的互动行为不仅能够调动学生的积极性和参与度,还提高了老师对学生的关注度,能够进一步了解学生的听课情况。

虽然国内外学者在课堂行为识别和分析方面已经取得了显著的成果,但仍然存在以下不足:

(1) 现阶段缺乏课堂环境下教师和学生的行为公开数据集,这在一定程度上制约了相关研究的深入开展。网上虽然有部分公开数据集,但行为类别较少,发布的视频数据多数以微课形式,缺乏真实课堂中的复杂情况,在进行相关研究时,需要建立符合课堂环境下的行为数据集。

(2) 真实的课堂环境较为复杂,教师和学生的行为种类繁多且动态变化,多种行为之间存在极大的相似性。现有的深度学习模型存在漏检、误检、识别困难等问题,难以满足实际的应用需求。

(3) 当前大多数学者都是针对学生的课堂行为进行研究,通过学生行为数据来反映学生的听课效率,并且与学生成绩建立某种联系,这些研究方法中的目标对象仅仅局限于学生。然而,在观察课堂教学视频时不难发现,老师的授课方式对于课堂授课效果同样起着关键作用,授课过程中老师应当加强与学生之间眼神、语言、动作等方面的交流,减少背向的行为和面向黑板的教学行为。因此,分析老师的教学行为也至关重要,从教师和学生的视角进行多方面的考虑,实现更加客观、更全面的课堂行为分析,避免仅从学生状态分析学生听课效果的片面性。

1.3 研究内容及技术路线

1.3.1 研究内容

本研究以真实课堂环境下的教师和学生为研究对象，旨在通过构建师生课堂行为识别模型，实现对教师和学生课堂行为状态的精准识别与分析，有效反馈课堂行为数据，为教师提供科学的教学改进依据，从而优化课堂管理并提升教学质量。通过对学生课堂行为识别，全面了解每个学生的课堂活动和听课状态，避免了传统课堂中教师无法关注到每位学生的情况。

结合国内外的研究现状，本文基于深度学习的师生课堂行为识别研究内容如下：

(1) 师生课堂行为数据集构建

鉴于目前没有公开的师生课堂行为数据集，本文从新疆某地区的中小学进行数据采集。通过奥威亚和大疆 DJI Pocket 2 云台设备获取大量的教师教学数据和学生听课数据，经过抽帧、筛选、打标签等数据处理，分别构建了教师课堂行为数据集和学生听课行为数据集。

(2) 面向大目标的教师课堂行为识别方法研究

本文利用目标检测算法深入挖掘潜在的行为信息，研究了基于改进 YOLOv7 的教师大目标课堂行为识别方法。通过在骨干网络中添加大目标检测层提高了查全率和查准率，使用 EMA 注意力机制增强局部特征信息的捕捉，解决类间相似行为难以识别的问题。最后引入 Inner_MPDIoU 边界框损失函数加快模型收敛的速度，提高模型的精度。改进后的模型能够有效识别多种课堂教学行为，在各项指标上的性能均显著优于其他方法。

(3) 基于改进 YOLOv8 的学生课堂行为识别方法研究

本文将当前主流的模型在学生数据集上进行实验，YOLOv8 网络凭借端到端的检测优势取得最佳的实验效果。结合学生课堂行为数据集特点，本文研究了改进 YOLOv8 的学生课堂行为识别方法。将 EfficientFormerV2 作为特征提取网络，使用 SCConv 卷积优化颈部网络中的 C2f 模块，增强模型的特征提取能力，自适应构建远程空间和通道间的依赖关系，挖掘课堂图像中的潜在特征信息，以提高学生课堂行为识别精度。

(4) 师生课堂行为识别系统研究与实现

基于改进模型构建师生课堂行为识别系统，实现师生课堂行为识别和分析，并且将结果数据可视化，便于教学人员直观了解课堂数据信息。通过教学行为分析数据了解教师的授课情况，作为教学反思和改进教学策略的重要参考，辅助教师优化教学方案，实现个性化教学设计。通过学生行为数据及时掌握学生的听课情况，重点关注听课效率低、注意力不集中的学生，有助于解决传统课堂中教师无法顾及每一位学生的问题。

1.3.2 技术路线图

本研究的技术路线如图 1-1 所示。阅读相关文献，结合现有的研究基础，确定本文的研究内容和方法。将采集的课堂视频数据进行抽帧、筛选、打标签等数据处理，分别构建出教师和学生的课堂行为数据集。同时探究不同网络模型在本文数据集上的适用性，挑选适合课堂环境下师生行为识别的基准模型，结合数据集特点优化网络结构，设计符合课堂行为识别需求的网络模型，提高对行为类别的识别精度。最后，基于改进的模型设计并实现师生课堂行为识别系统。

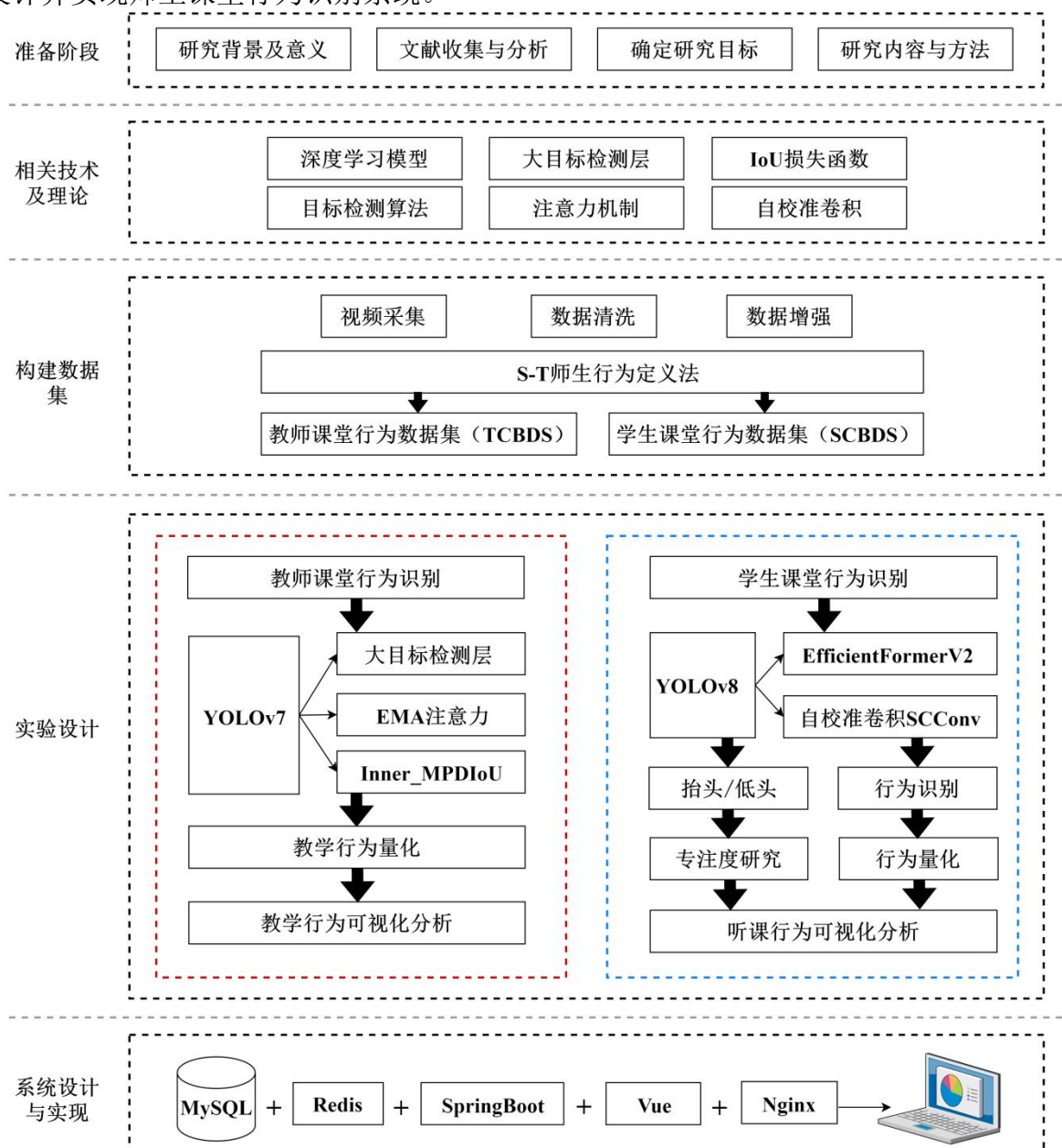


图 1-1 技术路线图

Figure 1-1 Technology roadmap