

基于多粒度特征提取的校园吸烟监测与预警系统设计与实现

倪帆 姚小冬 张晨 邱娟 章闯

南京机电职业技术学院, 江苏南京, 211306;

摘要: 一直以来, 校园吸烟行为违反无烟校园规定, 危害师生呼吸系统健康, 污染校园环境并干扰正常教育教学秩序。当前传统校园吸烟监测依赖人工巡查, 存在效率低、漏检率高、预警滞后等问题, 如何实现吸烟行为的实时监测与精准干预, 成为无烟校园建设领域亟待解决的现实难题。本文针对这一痛点, 探索构建基于多粒度行为识别的无烟校园行为监测与预警系统: 通过提取吸烟相关动作的多维度特征、优化智能识别模型、设计一体化监测系统, 提升校园吸烟行为的实时识别率与干预效率。研究不仅为无烟校园建设提供了智能化解决方案, 助力校园卫生安全数字化转型, 其技术框架也可为公共场所禁烟监测等领域提供参考, 具有较强的实践价值。

关键词: 多粒度行为识别; 校园安全; 监测预警系统; 深度学习

DOI: 10. 64216/3080-1516. 25. 10. 035

引言

随着城市化与教育规模化发展, 无烟校园建设成为“健康校园”建设与教育数字化转型的核心议题。据教育部校园卫生安全报告, 部分校园仍存在吸烟行为, 不仅违反《无烟学校建设指南》, 还对青少年身心健康造成负面影响, 增加校园卫生管理难度。

国内学者已展开相关探索: 石佳等指出校园行为监测在新型风险应对与智慧技术应用方面待加强^[5]; 王庆刚等人的流量预警系统未覆盖吸烟识别, 存在监测盲区^[3]。行为识别领域, 闫兴亚等的 BetaPose 模型缺吸烟专项预警能力^[9]; 曾钰琦等的 VWE-YOLOv8 算法受场景普适性限制^[2]; 王萌等的融合模型无校园吸烟专项识别机制^[4]。当前主流 YOLO 类算法在多人交互、物品遮挡(手部持烟被书本遮挡)场景性能下降, 难区分“接打电话”与“抽烟”, 吸烟漏报风险高^[8]; 虽有团队探索轻量化姿态提取^[6], 但方案仍存多模态融合实时性不足、缺校园吸烟场景差异化部署等问题。

为响应《“十四五”教育发展规划》“校园卫生智能化管理”要求, 本文设计基于多粒度特征的监测预警系统, 核心内容包括:

- 以 AlphaPose 骨架序列为基础, 融合全身与头部细节特征, 通过并行分支网络与自适应加权机制增强高危动作区分能力(见图 1);

- 开发支持 RTMP 协议与 SRS 服务器的全链路系统, 实现多摄像头并行分析, 日均处理 800GB 视频流。

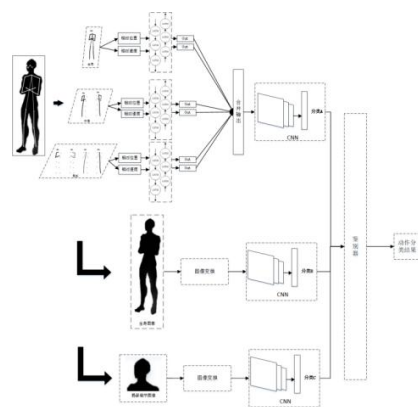


图 1. 多粒度特征提取与行为分类流程

本研究构建多粒度特征融合框架, 丰富无烟校园智能监测理论体系, 为无烟校园建设提供可落地方案, 其思路可为公共场所禁烟监测提供参考。

1 系统方法与设计

1.1 系统总体架构

本系统针对校园传统吸烟监测人工依赖度高、预警滞后、重点区域覆盖难问题, 构建前端采集、远程传输、智能分析、终端交互的分布式架构。系统由吸烟高发区域摄像头、客户端与服务器端组成, 通过流媒体协议形成数据采集→特征提取→吸烟行为识别→预警输出闭环^[7]。该架构适配校园监控分散、人员密度波动大的特性, 集中式服务器+轻量化客户端模式可平衡实时性、精度、成本, 支持多摄像头并行接入, 实现重点区域全覆盖。

1.2 核心技术方案

1.2.1 远程视频流传输方案

系统采用“RTSP 采集+SRS 转发+FFmpeg 推流”方案：网络摄像头经 RTSP 输出 H.264 视频流，局域网客户端通过 OpenCV 读取后，由 FFmpeg 推至公网 SRS 服务器；客户端从 SRS 拉流，多线程实现播放与分析。选择该方案因校园网络环境复杂（局域网隔离、带宽波动），SRS 多协议转发，FFmpeg 高效转码，确保跨网络传输稳定实时，确保跨网络传输稳定实时。

1.2.2 人体姿态估计与多粒度行为识别模型

针对校园吸烟行为动作短暂、易遮挡、误报率高的问题，系统基于 AlphaPose 人体姿态估计模型^[1]，结合多粒度特征融合方案，动态评估连续帧克服遮挡，提供 18 个关键关节点（重点关注手部、头部）的骨架特征。核心模型设计包括：骨架时序识别采用 CNN-LSTM 架构，设时刻 t 关节点 i 坐标为 P_{ij}^t ，通过相对位置与速率建模时空特征：

$$P_{ij}^t = P_i^t - P_j^t \quad (1)$$

$$S_{ij}^t = \frac{P_i^t - P_i^{t-1}}{\Delta t} \quad (2)$$

双向 LSTM 捕捉掏烟、持烟、含烟动作，CNN 提取空间特征，输出分类结果 r_1 ，适配吸烟行为短暂、动态特点。多粒度融合补充全身图像 CNN 模型，区分持烟与持书，与头部轻量化 CNN 模型，截取头部图像，识别含烟、打电话动作；融合策略采用“误差加权判决”，结果一致直接取用，差异时选误差最小类别，降低光照等因素误判。

1.2.3 客户端交互与多线程设计

为适配安保人员操作并保障高分辨率视频实时性，客户端用 PyQt5 构建界面，含视频播放器、功能控制区、日志与预警模块，支持参数配置与数据回溯。采用“视频加载线程 + 姿态识别线程 + 行为识别线程”架构，避免单线程卡顿。系统运行流程如下：

- （1）网络摄像头采集视频流，经 RTSP 协议传输至客户端后经 FFmpeg 推流至远程 SRS 服务器；
- （2）服务器端调用人体姿态检测网络，输出骨架点坐标与可视化画线结果回显；
- （3）服务器端提取多粒度特征，计算分类与误差；
- （4）误差加权输出结果，吸烟行为实时预警；
- （5）日志记录操作与结果，支持溯源。

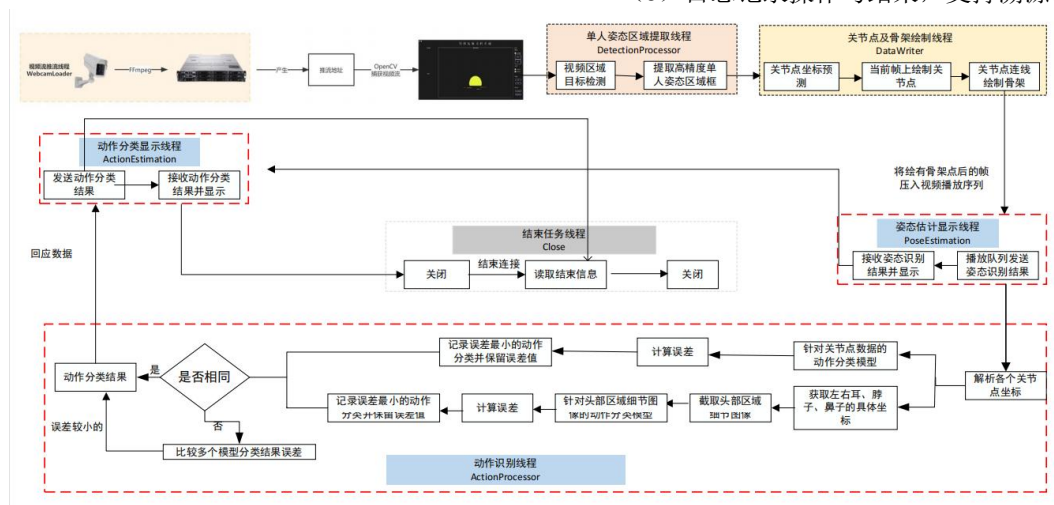


图 2. 客户端多线程工作处理流程

2 实验结果与分析

本实验基于 PyTorch 框架搭建模型，硬件环境为 Intel Core i7-12700K CPU、NVIDIA RTX 4090 GPU (24 GB 显存)，操作系统 Ubuntu 20.04，保障训练与推理高效性。

2.1 实验参数设置

核心参数配置如下：

骨架点行为识别采用 CNN-LSTM 架构，其中 LSTM 单

元数为 256，dropout 率 0.5 以防过拟合，CNN 为 AlexNet 结构，含 5 个卷积层与 3 个全连接层。用于全身图像识别的卷积神经网络包含 4 个卷积层，核大小由 16×16 递减至 1×1 ，每层接步长 2 的最大池化层。头部区域识别模型为轻量化 CNN。

初始学习率为 0.001，采用 Adam 优化器，批量大小为 32，迭代 50 轮，每 10 轮学习率衰减至原 1/10。

融合策略采用误差加权判决，结果一致直接取用，

差异时选误差小类别；交叉熵损失函数约束如公式（3）所示：

$$L_r = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^t 1\{y = j\} \log \left(\frac{\exp(\phi_j^T x)}{\sum_{1 \leq l \leq k} (\phi_l^T x)} \right) \right]$$

(3)

其中 m 为样本总数，t 为动作类别数，y 为真实标签， $\phi_j^T x$ 为预测分布。实验结合公开吸烟行为数据集与校园自建无烟监测数据集，覆盖不同场景复杂度，具体信息如表 1 所示：

2.2 数据集设置

表 1. 选用数据集及其特点

数据集	UCF101（吸烟子集）	2020 中国华录杯·数据湖算法大赛抽烟打电话数据集	校园自建
视频数量	800+	1200+	300
动作类别	吸烟、持烟等	持烟、吸烟、递烟、持烟打电话	抽烟、接打电话、摔倒等
场景特点	单人/多人、室内/室外	复杂背景、多姿态、物品遮挡	校园真实场景

校园自建数据集通过校园监控摄像头实地采集，涵盖早中晚时段，时长 5-15 秒，每段均标注动作类别，用于验证真实校园场景适用性。

2.3 剥离实验与结果分析

为验证多粒度特征融合的有效性，通过移除不同模型组件进行实验，结果如表 2 所示（以平均识别准确率为指标）。

表 2. 实验结果与剥离分析

实验设置	完整模型（骨架+全身图像+头部图像）	移除头部图像	移除全身图像	仅保留骨架点
UCF101	92.3%	84.7%	83.2%	78.8%
2020 中国华录杯数据集（吸烟类）	90.5%	81.2%	80.6%	72.9%
校园自建数据集	88.5%	80.1%	79.3%	74.6%

结果表明，完整模型在两类数据集上均取得最高的准确率，验证了多粒度特征融合策略的有效性。仅保留骨架模型时准确率最低，表明单一骨架特征在复杂场景下鲁棒性不足，需与其他特征融合提升性能。移除头部图像特征后，准确率下降 6.4%-8.4%，验证了头部细节图片特征的细粒度影响显著，说明头部细节特征能有效增强语义区分能力。移除全身图像特征后，准确率下降明显，说明行为识别需结合环境背景的动作中表现明显，证明全身图像对场景上下文信息的补充作用。

综上，多粒度融合提升吸烟识别精度，模型校园场景表现稳定，为无烟校园建设提供可靠支撑。

3 结语

本文针对校园吸烟行为“人工监测效率低、细粒度识别精度不足”问题，设计基于多粒度特征融合的无烟校园监测预警系统。构建前端采集、远程传输、智能分析、终端预警架构，实现吸烟高发区域全覆盖；提出融合骨架时序+全身图像+头部细节的识别模型，误差加权策略降低误判。实验显示，吸烟行为识别在 UCF101（吸

烟子集）准确率 92.3%，校园真实场景准确率 89.1%，可为无烟校园建设提供智能化方案。

参考文献

[1]H.-S. Fang, J. Li, H. Tang, C. Xu, H. Zhu, Y. Xiu, Y.-L. Li, and C. Lu. Alphapose: Whole-body re- gional multi-person pose estimation and tracking in real-time. IEEE transactions on pattern analysis and machine intelligence, 45 (6):7157 - 7173, 2022. 2

[2]曾钰琦, 刘博, 钟柏昌, and 钟瑾. 智慧教育下基于改进 yolov8 的学生课堂行为检测算法. 计算机工程, 50 (09):344 - 355, 2024. 1

[3]王庆刚, 顾峰, 张雪梅, 余润东, 张壹, and 王宇凡. 基于校园网流量分析的安全预警系统. 网络安全技术与应用, (7):73 - 76, 2022. 1

[4]王萌 and 郑奋. 课堂行为识别算法研究及智慧教室设计. 信息技术, (11):13-39, 2024. 1

[5]石佳, 刘娜娜, and 尹畅. 基于可视化分析的我国校

园安全研究述评(2000—2023 年). 安全, 46(1): 64 - 73, 2025. 1

[6] 裴涛, 王向阳, 谢慧志, and 赵佳辉. 基于端到端的轻量化人体姿态跟踪算法研究. 工业控制计算机, 38(03): 28 - 29+31, 2025. 1

[7] 赵欢. 基于智慧校园的高职院校安全管理体系. 中国自动识别技术, (02): 71 - 74, 2025. 2

[8] 郭伏正, 孔军, 蒋敏, et al. 自适应融合 rgb 和骨骼特征的行为识别. Laser&OptoelectronicsProgress, 57

(20): 201506, 2020. 1

[9] 闫兴亚, 匡娅茜, 白光睿, and 李月. 基于深度学习的学生课堂行为识别方法. 计算机工程, 49(07): 251 - 258, 2023. 1

作者简介: 倪帆(1997.12—), 女, 汉族, 江苏省南京市人, 助教, 硕士, 南京机电职业技术学院, 计算机视觉。