

计算机深度学习在视觉目标检测中的应用

赵世杰

宁夏医科大学总医院 宁夏 银川 750001

摘要:深度学习视觉目标检测是计算机视觉的核心技术, 主要包含两阶段、单阶段与Transformer三类主流算法, 依托特征融合、模型轻量化等关键技术实现性能优化。本文系统梳理该技术在智能交通、工业制造、安防智慧城市及遥感无人机等领域的落地应用, 分析复杂场景检测精度不足、速度精度难以兼顾、跨场景泛化性弱等核心问题, 并总结优化策略, 展望轻量化、多模态融合的行业融合发展趋势。

关键词:计算机深度学习; 视觉目标检测; 应用

引言:随着人工智能与机器视觉技术快速迭代, 传统视觉检测方法已无法适配复杂场景的检测需求。深度学习目标检测算法凭借自动特征提取、高精度与高适配性优势, 逐步替代传统检测技术, 成为各行业智能化升级的核心支撑。为明晰技术发展与应用现状, 本文深入研究主流检测算法原理与关键技术, 结合多领域落地案例剖析应用痛点, 针对性提出优化方案, 并预判行业未来发展方向。

1 深度学习视觉目标检测核心技术基础

1.1 深度学习目标检测核心算法体系

(1) 两阶段检测算法(Faster R-CNN等)原理与应用适配性。Faster R-CNN通过区域建议网络(RPN)生成候选区域, 再经ROI池化与分类回归完成检测, 两阶段设计赋予其精细调整能力。在卫星图像船舶检测中, Faster R-CNN的mAP达55.1%, 显著优于SSD512(47.5%)和YOLOv3-608(45.5%), 尤其在小目标与杂乱场景中优势突出。然而其推理速度较慢(约63.2ms), 更适合对精度要求高、实时性要求宽松的工业质检、遥感分析等场景。(2) 单阶段检测算法(YOLO、SSD等)原理与落地优势。单阶段检测器摒弃区域建议阶段, 通过一次前向传播直接预测边界框与类别概率, 速度优势显著。YOLOv8s在道路缺陷检测中mAP50达70%, FPS达312.5; 在COCO数据集上YOLOv8可达165 FPS(T4 GPU)。SSD检测速度约26.5ms, 显著快于Faster R-CNN。单阶段算法凭借速度优势, 广泛部署于安防监控、自动驾驶等实时场景^[1]。(3) Transformer端到端检测算法的应用特性。以DETR为代表的Transformer检测器通过全局注意力机制捕捉图像长程依赖, 去除锚框与NMS后处理, 简化检测流程。在DOTA数据集上DETR的mAP@0.5达72%, 超越CNN检测器最高达13%。但其计算开销较大(86.3 GFLOPs), 速度仅约12 FPS, 当前更

适合对精度要求高的离线分析场景, 轻量化改进是重要发展方向。

1.2 目标检测关键技术模块

(1) 图像预处理与数据增强技术(适配不同应用场景)。数据增强通过图像变换、GAN生成、扩散模型等方法扩充训练样本多样性。针对自动驾驶、遥感等场景需定制化增强策略, 如光照适应、遮挡模拟等, 以提升模型泛化能力。(2) 特征提取与多尺度融合技术。FPN通过自顶向下路径融合多尺度特征, 使低层特征获得更强语义信息; PANet进一步引入双向特征融合机制, 实现更全面的多尺度交互。多尺度融合是提升小目标检测精度的关键。(3) 目标定位、分类与非极大值抑制技术。NMS作为后处理核心模块用于剔除冗余候选框。改进型DIoU-NMS通过手动设置阈值提升密集目标检测准确率; 引入高斯衰减的软性抑制策略可进一步优化传统贪婪NMS的不足。

1.3 模型轻量化与优化技术

(1) 网络结构简化与参数裁剪技术。剪枝可大幅压缩模型规模, 如YOLOv8n改进后参数量和计算量分别减少68.4%与56.8%; 蒸馏—剪枝—量化一体化策略可将模型体积压缩至10MB以下。(2) 注意力机制嵌入的检测精度优化方法。在YOLOv11n中嵌入高效多尺度注意力模块(EMA), 构建通道-空间协同架构, 可有效增强复杂场景下的空间感知与目标表征能力。(3) 硬件适配的模型推理加速技术。模型量化将浮点参数转为低比特(如8/4 bit)表示, 可大幅减少内存占用并提升推理速度; TensorRT加速可使YOLOv5推理时间降至约23ms, 有力支撑边缘端实时部署。

2 计算机深度学习在视觉目标检测中的应用

2.1 智能交通与自动驾驶领域应用

(1) 道路交通目标检测(车辆、行人、交通标识识

别)。蒋嘉璇等提出基于改进YOLOv8的行人及车辆检测方法,在KITTI数据集上准确率、召回率、平均准确率分别提高0.7%、2.1%、2.1%,帧率达24.64 frame/s; DGSS-YOLOv8s在KITTI数据集上所有类别mAP@50和mAP@50:95分别提升4.6%和9.8%;在BDD100K数据集上行人mAP@50从45.6%提升至49.6%。(2)自动驾驶环境感知与障碍物实时检测应用。AWD-YOLO在DAWN数据集上mAP@0.5达67.7%,较YOLOv8n提升15.4%;在RTTS和ACDC数据集上mAP@0.5分别超越YOLOv8n达6.9%和12.5%;改进YOLOv5s算法准确率和mAP50分别提升0.035和0.009,检测速度达54.1帧/s;YOLOv11-SEG在ADAS应用中mAP@0.5达97.4%^[2]。(3)交通违章识别与路况智能监测落地实践。基于YOLOX与RNN的智能识别检测系统在某省高速集团3条主干线(总里程120公里)试点部署,违规识别准确率达96.7%,响应时间0.6秒内,次生事故减少71%;YOLOv8n在交通违章检测中准确率86.7%、召回率82.1%、mAP达72.5%。

2.2 工业智能制造领域应用

(1)工业产品表面缺陷检测(PCB瑕疵、零部件损伤识别)。YOLOv9在裸板PCB六类缺陷检测中mAP达98.4%(IoU阈值0.5);优化YOLO v9架构在专用PCB数据集上mAP@50达99.2%;改进YOLOv5在NEU-DET钢材表面缺陷数据集上mAP达79.1%,较原模型提升3.4%。(2)工业生产线物料分拣与目标定位应用。基于YOLOv11的智能分拣控制系统结合机械臂、传送带执行分拣任务,显著提升物料识别的精度与速度;基于YOLOv5的喷涂产线工件识别方法实现复杂环境下目标工件的精确定位与抓取。(3)工业设备运行状态智能监测与异常识别。基于YOLOv8的旋转机械实时异常检测系统针对裂纹、渗油、螺栓松动等微小异常设计轻量化方案;电解车间轨道式机器人巡检方案将极板故障检出率提升至90%以上。

2.3 安防监控与智慧城市领域应用

(1)公共场所人员、异常行为智能检测识别。改进YOLO框架异常行为识别准确率达99.46%;YOLOv5在校园监控数据集上准确率95.2%、召回率93.8%,单帧处理仅20 ms;YOLOv10引入可变形卷积与注意力模块后mAP@0.5达89.7%。(2)安防视频实时目标追踪与入侵检测应用。YOLOv8结合DeepSORT实现多目标行人追踪;基于YOLOv8与DeepSORT的智能边缘计算系统实现实时行人入侵检测与跨帧目标跟踪;BCD-YOLO结合轨迹与停留时间判断入侵告警,有效降低复杂环境误报率^[3]。(3)城市环境占道、违规停放等场景智能巡检。

YOLOv8+ResNet方案对乱停车、乱倒垃圾等识别精度达98%,响应延迟低于1秒;YOLOv8在违规停放检测中准确率达81.86%;YOLOv8在交通信号异常检测中mAP达90.5,推理速度32 FPS。

2.4 遥感与无人机视觉检测应用

(1)无人机航拍小目标、遮挡目标精准检测。跨层动态检测网络在VisDrone2019上mAP0.5和mAP0.5:0.95分别达53.3%和33.2%,较YOLOv10s提升12.7%和9%,参数量减少23.7%;MSE-YOLOv11在VisDrone2019上mAP@0.5:0.95较基线提升20.25%,小目标AP提升31.88%,推理速度59 FPS^[4]。(2)遥感图像地物识别、灾害区域目标检测应用。改进YOLOv7的InSAR遥感检测方法自动识别大气相位与地表形变目标;YOLOv8-based六类灾害检测模型Precision达0.956。(3)国土测绘与资源巡查智能视觉检测实践。辽宁省“遥感+AI”智能监管体系已在全省14个地市实现核心场景全覆盖,累计自动识别并推送土地、林草、海岸带等领域异常图斑超100万个,问题核实查全率达90%以上。

3 计算机深度学习在视觉目标检测中应用现存问题、优化策略与未来展望

3.1 当前各领域应用落地核心问题

(1)复杂场景下小目标、遮挡目标检测精度不足。无人机航拍场景中因目标尺寸小、遮挡严重、尺度变化大等因素,常出现漏检与误检;红外小目标因成像面积小、分辨率低、易被遮挡导致检测精度低。现有YOLOv8算法因特征提取不足导致小目标检测精度受限。(2)轻量化模型检测精度与推理速度难以兼顾。轻量化目标检测模型在无人机平台部署中面临模型小目标检测效果不佳、模型复杂度与硬件资源约束存在根本性矛盾。受限于无人机平台计算性能,实现高精度与轻量化的实时目标检测具有较大挑战。(3)跨场景泛化能力弱、定制化应用成本高。目标检测算法在不同数据域间迁移时存在性能退化问题;工业表面缺陷检测中面临定位不准确、特征提取能力弱和跨域适应性差等挑战;数据采集与标注耗时耗力,制约了深度学习模型在实际应用中的推广。

3.2 面向实际应用的技术优化策略

(1)多尺度特征增强与小目标检测优化方案。可通过增加小目标检测层、引入多尺度通道与空间注意力模块提升多尺度特征提取能力;设计动态多尺度特征增强机制结合自适应特征融合提升模型对小目标的检测能力。(2)模型轻量化改进与硬件适配优化方法。CMD-YOLO算法与YOLOv10-n相比参数量降低30.5%,计算量下降19%,权重文件缩小29.3%;LW-DETR实现参数减少

71.7%、计算量降低73.7%；Lite-Edge Det模型采用蒸馏—剪枝—量化一体化策略将体积压缩至10MB以下，推理速度达30–35 FPS^[5]。（3）数据集优化与迁移学习提升跨场景应用能力。YOLOv11-EMD集成多阶段迁移学习框架，在NEU-DET和Severstal数据集上mAP@50达94.9%，较原模型提升1.6%，多阶段迁移策略可缩短训练时间3.2%的同时提升mAP 8.8%。

3.3 深度学习目标检测应用未来发展趋势

（1）轻量化、实时化智能检测应用普及。轻量化网络设计、边缘计算部署将成为研究重点，推动机器视觉向高效方向演进；基于Transformer架构的检测模型检测精度平均提升4.8个百分点，轻量化模型在嵌入式设备推理速度突破300帧/秒。（2）多模态融合的高精度复杂场景应用拓展。多模态融合及环境自适应技术将成为未来研究重点；通过融合红外与可见光等多源信息，可解决单一RGB模态在弱光、遮挡等复杂环境下的感知瓶颈。（3）人工智能与各行业深度融合的智能化应用升级。目标检测技术将持续向通用模型构建、多模态信息融合以及感知-决策一体化方向演进；随着轻量化部署与特定场景优化的推进，目标检测将在工业、交通、安防等行业实现更深层次智能化应用升级。

结束语

本文全面探究了深度学习视觉目标检测的核心技术、多领域应用成效与现存技术难题，总结了特征增强、模型轻量化、迁移学习等有效优化手段。当前目标检测技术仍存在复杂场景检测效果差、落地成本高等问题。未来需聚焦轻量化实时检测、多模态融合技术研发，持续推进算法与硬件适配、跨行业场景适配，进一步推动视觉目标检测技术的智能化、规模化落地应用。

参考文献

- [1]肇恒鑫,王超,李建龙.上下文感知多感受野融合网络的定向遥感目标检测[J].电子与信息学报,2024,46(11):3452-3460.
- [2]刘杰,张宇,王浩.复杂环境下无人机航拍小目标深度学习检测算法[J].电子与信息学报,2025,47(06):1890-1897.
- [3]王瑜,毕玉,石健彤.基于注意力与多级特征融合的改进YOLOv5目标检测算法[J].郑州大学学报(工学版),2024,45(02):78-85.
- [4]陈曦.深度学习计算机视觉目标检测技术原理与应用场景探析[J].信息技术与信息化,2024,21(09):156-159.
- [5]李泽宇.深度学习实时目标检测算法性能优化与工程应用[J].计算机技术与发展,2023,33(12):98-103.