

DOI:10.19344/j.cnki.issn1671-5276.2025.01.026

改进 FCOS 模型的细长物体检测算法研究

范佳能¹, 李开宇¹, 仲志强², 赵慧弢¹, 白千帆¹

(1. 南京航空航天大学 自动化学院, 江苏 南京 211106; 2. 南瑞集团有限公司, 江苏 南京 211106)

摘要:针对以电缆为代表的细长物体目标检测问题,以 FCOSv2 模型为基础,通过融合细长物体的“细长度”和方向角度 θ 信息改进了正样本分配策略,在主干网络和预测头加入调制可变形卷积,提升改进后 FCOS 模型的检测能力。实验结果表明:在推理速度相当的情况下,改进后的 FCOS 模型达到 83.2% 的 mAP,实现了检测精度与速度的均衡。

关键词:细长物体;目标检测;正样本分配;FCOS;特征自适应

中图分类号:TP202⁺.2 **文献标志码:**A **文章编号:**1671-5276(2025)01-0123-05

Research on Algorithm of Slender Object Detection Based on Improved FCOS

FAN Jianeng¹, LI Kaiyu¹, ZHONG Zhiqiang², ZHAO Huitao¹, BAI Qianfan¹

(1. College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China;

2. NARI Group Corporation, Nanjing 211106, China)

Abstract:To solve the problem of object detection of slender objects represented by cables, based on FCOSv2 model, proposes the fusion of "slenderness" and directional angle θ of slender objects to improves the positive sample allocation strategy, and adds modulation-deformable convolution to the backbone network and prediction head, which upgrades the detection capability of the improved FCOS model. The experimental results show that the improved FCOS model achieves 83.2% mAP at the same reasoning speed, and achieves the balance of detection accuracy and speed.

Keywords:slender object; object detection; positive sample allocation; FCOS; feature adaptation

0 引言

随着国家电网供电格局工程的推进,电缆线路维护的工作量日渐剧增。为了避免人工维护电缆的安全隐患,同时提高维护电网的效率,降低工作人员的劳动强度,机器人代替人工作业的方式逐渐成为一个趋势^[1]。通过计算机视觉,机器人可以完成回环检测、物体抓取和更换等复杂任务,而完成这些任务的基础则是探测出周围环境中存在的检测物体类别和位置。当前,机器人实现全自动维护电缆等配电设备的难点之一是如何快速高效地识别检测。

电缆是一种典型细长物体,类似的还存在许多细长物体,如剥线、滑板、棒球、滑雪板,这类物体的纵横比受物体姿态影响很大,表现为多变不固定的特征,这会对精确检测细长物体带来影响。目前,单阶段目标检测算法相比于两阶段目标检测算法,虽然丢失了小幅精度,但端到端的训练减少了提取候选框、对候选框逐一分类回归这一复杂的步骤,提高了算法运行速度^[2]。本文以单阶

段算法作进一步论述。

目前单阶段目标检测算法分为有锚框网络和无锚框网络。前者以 YOLO、RetinaNet 为代表,通常在先验框的基础上再进行坐标回归。但对于细长物体的检测,由于细长物体的纵横比受物体姿态影响变化很大,而边界框只提供不符合目标形状和姿态的粗略定位。因此,从边界框的规则单元格提取的特征容易受到背景内容或目标语义信息较少的前景区域的严重影响。这导致较低的特征质量,对细长物体的检测泛化能力较差。后者以 FCOS^[3-4]、RepPoints 为代表,此类框架摆脱了锚框的限制。针对细长物体,这类算法假设的初始正样本点很有可能落在语义信息量很小的位置,甚至是背景区域^[5],导致模型无法较好地学习到细长物体的特征。

针对上述对细长物体的检测问题,本文在 FCOSv2 模型的基础上,作出如下改进:第一,改进了正样本分配策略,融合细长物体的“细长度”和方向角度 θ 信息;第二,在主干网络和预测头加入了调制可变形卷积^[6]。

基金项目:国家自然科学基金项目(61905112);南京航空航天大学校创新计划项目(xcxjh20220335)

第一作者简介:范佳能(1999—),男,江苏张家港人,硕士研究生,研究方向为机器视觉、图像处理,1326353726@qq.com。

1 基于“细长度”和方向角度 θ 的正样本分配策略

1.1 细长度与方向角度 θ 的定义

如图 1 所示,传统的边界框纵横比(红色框)仅表示目标的具体位置,而实际细长程度无法通过边界框的纵横比精确衡量(本刊黑白印刷,相关疑问咨询作者)。因此本文提出了一个评估指标——细长度,用于评价物体的细长程度。细长度 s_l 的定义为

$$s_l = \min(W', H') / \max(W', H') \quad (1)$$

式中 W' 、 H' 分别表示最小外接矩形的宽和高。图 1 所示的橙色框为各个物体围成的最小外接矩形,同一物体当其姿态发生变化时,边界框纵横比 a 随之改变,受物体姿态扰动较大;而细长度 s_l 是保持不变的,不受物体空间姿态的影响。因此,本文用细长度 s_l 能够较为精确地反映物体的实际细长程度。此外,将 W 与 W' 的夹角记为 θ , $\theta \in [0, \pi)$ 。

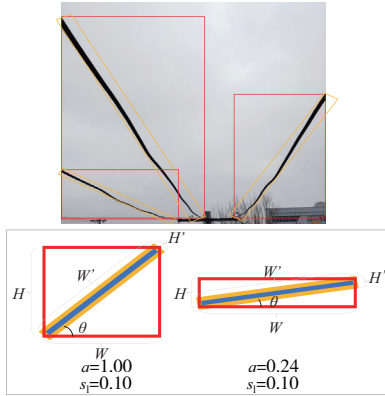


图 1 细长度和方向角度 θ 的定义

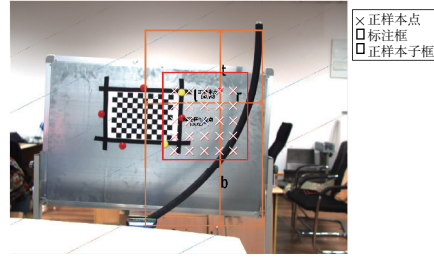
1.2 改进的正样本分配策略

FCOSv1 模型将落入标注框内的预测点视为正样本点,FCOSv2 模型则对标注框的中心部分进行采样作为正样本。

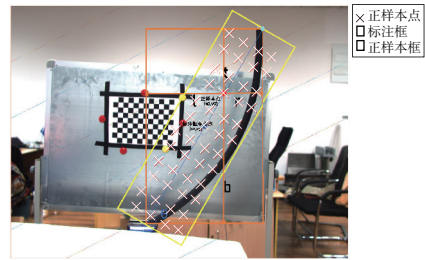
由于本文的检测对象为细长物体,此正样本分配策略的缺点如图 2(a) 所示,细长物体会导致正样本点落在语义信息量很小的位置,甚至落在背景区域,而包含更多语义信息量的样本点却被忽略,导致模型无法较好地学习到细长物体的特征。改进的正样本分配策略为:首先根据 1.1 节定义的 θ 值确定被测物体的方向,结合图 2(b) 与图 3,由于标注框中心点坐标为 (c_x, c_y) ,可计算出标注框的 4 个顶点坐标分别为 $(c_x - l_1, c_y - t_1)$ 、

$(c_x - l_1, c_y + b_1)$ 、 $(c_x + r_1, c_y + b_1)$ 、 $(c_x + r_1, c_y - t_1)$,其中 t_1 、 b_1 、 l_1 、 r_1 分别为标注框中心点到标注框上下左右边界的距离。接着根据 θ 值和细长度 s_l 可确定一个正样本候选区域,正样本候选区域是一个矩形区域。具体实现方法如图 3 所示:将标注框的对角线连线作为正样本候选区域短边中点的连线,此连线的长度等于正样本候选区域的长边,长度记为 l_{en} ,其中 l_{en} 的计算公式如式(2)所示。

$$l_{en} = \sqrt{(t_1 + b_1)^2 + (l_1 + r_1)^2} \quad (2)$$

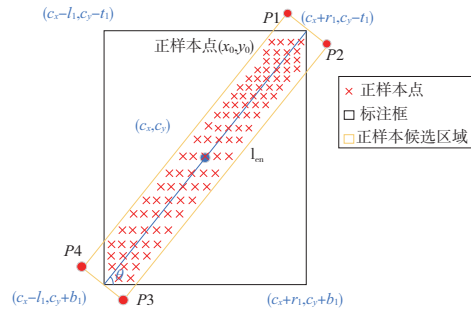


(a) FCOSv2 正样本分配策略

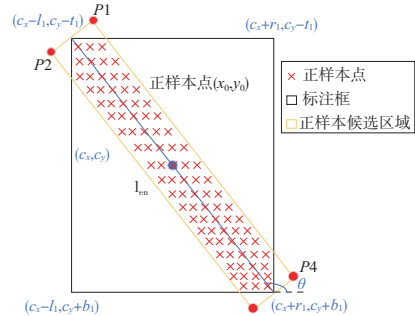


(b) 改进后的正样本分配策略

图 2 正样本分配策略



(a) $\theta \in [0, \pi/2]$



(b) $\theta \in [\pi/2, \pi)$

图 3 改进后的正样本分配示意图

结合图 3(a), 当 $\theta \in \left[0, \frac{\pi}{2}\right]$ 时, 根据 θ 值和细长度 s_1 确定正样本候选区域的顶点坐标分别为

$$\begin{cases} P1 = (c_x + r_1 + 0.5 \cdot s_1 \cdot \sin\theta \cdot l_{en}, c_y - t_1 + 0.5 \cdot s_1 \cdot \cos\theta \cdot l_{en}) \\ P2 = (c_x + r_1 - 0.5 \cdot s_1 \cdot \sin\theta \cdot l_{en}, c_y - t_1 - 0.5 \cdot s_1 \cdot \cos\theta \cdot l_{en}) \\ P3 = (c_x - l_1 - 0.5 \cdot s_1 \cdot \sin\theta \cdot l_{en}, c_y + b_1 - 0.5 \cdot s_1 \cdot \cos\theta \cdot l_{en}) \\ P4 = (c_x - l_1 + 0.5 \cdot s_1 \cdot \sin\theta \cdot l_{en}, c_y + b_1 + 0.5 \cdot s_1 \cdot \cos\theta \cdot l_{en}) \end{cases} \quad (3)$$

结合图 3(b), 当 $\theta \in \left[\frac{\pi}{2}, \pi\right]$ 时, 根据 θ 值和细长度 s_1 确定正样本候选区域的顶点坐标分别为

$$\begin{cases} P1 = (c_x - l_1 - 0.5 \cdot s_1 \cdot \sin\theta \cdot l_{en}, c_y - t_1 - 0.5 \cdot s_1 \cdot \cos\theta \cdot l_{en}) \\ P2 = (c_x - l_1 + 0.5 \cdot s_1 \cdot \sin\theta \cdot l_{en}, c_y - t_1 + 0.5 \cdot s_1 \cdot \cos\theta \cdot l_{en}) \\ P3 = (c_x + r_1 - 0.5 \cdot s_1 \cdot \sin\theta \cdot l_{en}, c_y + b_1 - 0.5 \cdot s_1 \cdot \cos\theta \cdot l_{en}) \\ P4 = (c_x + r_1 + 0.5 \cdot s_1 \cdot \sin\theta \cdot l_{en}, c_y + b_1 + 0.5 \cdot s_1 \cdot \cos\theta \cdot l_{en}) \end{cases} \quad (4)$$

本方法将 FCOSv1 的正样本区域与本文提出的正样本区域作交集, 将落入到交集的点集作为正样本点。图 3 中有详细标注, 此方法仅在训练阶段增加了少量计算量, 优点在于: 在 FCOSv1 正样本分配策略的基础上融合了物体的细长度 s_1 和方向角度 θ 的信息, 与原有的分配策略相比, 此方法使得包含更多语义信息量的样本点作为正样

本点, 理论上可有效地减少了背景区域的干扰。具体实验将在第 3 节作论述。

2 FCOS 模型的改进

2.1 调制可变形卷积

卷积神经网络中常用的标准卷积受限于模型几何变换, 无法较好地适应细长物体尺度、姿态的变化。本文将 FCOSv2 模型的部分卷积层替换为调制可变形卷积, 具体替换的层在 2.2 节作详细阐述。调制可变形卷积的定义为

$$y(p) = \sum_{k=1}^K w_k \cdot x(p + p_k + \Delta p_k) \cdot \Delta m_k \quad (5)$$

式中: w_k 表示第 k 个采样位置的权重值; $x(p)$ 和 $y(p)$ 分别表示来自输入特征图 x 和输出特征图 y 的位置 p 处的特征值; Δp_k 和 Δm_k 分别表示第 k 个位置的可学习偏移量和调制标量, 调制标量 Δm_k 属于 $[0, 1]$ 范围内, 在计算 $x(p + p_k + \Delta p_k)$ 时, 采用双线性插值的方法。

调制可变形卷积能够自适应对象几何变化的能力。偏移量分支使得对其神经特征的空间支持比标准卷积更符合对象形状结构; 调制标量则用来在学到的偏移位置上调制特征的幅度, 进一步增强了建模能力。

2.2 改进后的 FCOS 模型结构

改进后的 FCOS 模型结构如图 4 所示。本文选用的主干网络为 ResNet50。输入经过 ResNet50 的 4 个阶段分别生成 C2、C3、C4、C5, 把阶段 2、阶段 3、阶段 4 的卷积操作替换为调制可变形卷积。此外预测头经过 5 个 3×3 卷积后得到分类、回归和中心度分支。将分类和回归分支中的最后一个卷积层替换为调制可变形卷积。

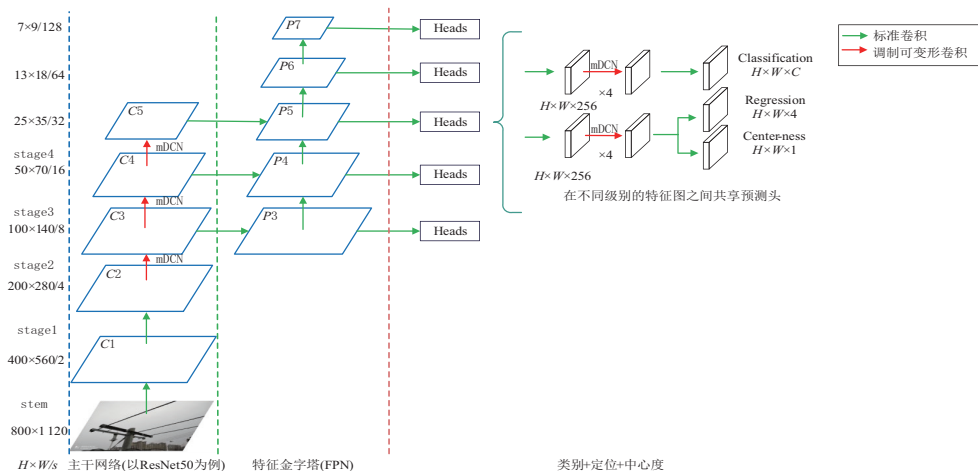


图 4 改进后的 FCOS 模型

3 实验与分析

3.1 数据集制作

本文使用配电网检修工程中的电缆(cable)、剥线(strip)样本以及 COCO2017 数据集里较为细

长的棒球(baseball)、滑板(skateboard)、滑雪板(snowboard)这5个类别的样本,共同组成了细长物体数据集。此数据集的类别定义与样本如图5所示。本数据集共10 000张样本,按照5:1的比例划分为训练集和验证集。

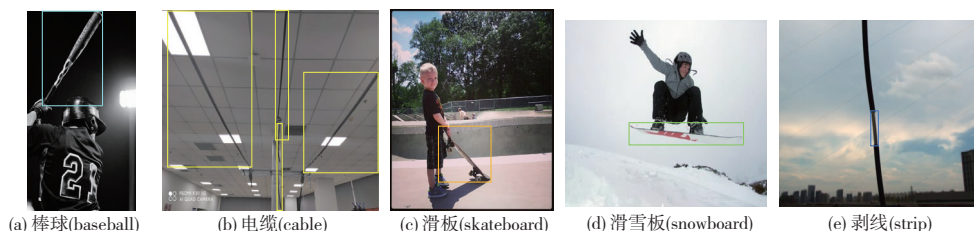


图5 细长物体数据集样本

3.2 训练细节

本次实验使用的平台为显存16 GB的Quadro RTX 5000显卡,Intel Xeon(R) Gold 5218处理器。模型算法基于pytorch1.9.1搭建,并使用CUDA11.1实现硬件加速。输入图片大小为(1 120×800)(Px)。在训练过程中,使用随机梯度下降(SGD)对模型进行50个epoch的迭代训练。训练采用迁移学习的方式统一使用ImageNet上预先训练的权重初始化主干网络。初始学习率设置为0.001,batch_size设置为6。在训练时均预热训练第1个epoch,学习率衰减因子为0.5,在训练迭代到第10、25、35个epoch时分别衰减学习率。此外将权重衰减因子设置为0.000 1,动量设置为0.9。

3.3 实验

1) 正样本分配策略对比

本实验对比了FCOSv1、FCOSv2和改进后的正样本分配策略对检测精度的影响。具体的正样本分配策略在1.2节作过论述,其他训练细节保持一致。由表1可知,改进后的正样本分配策略在检测细长物体数据集上mAP达到了0.803,比FCOSv1提升了3.5个百分点;比FCOSv2提升了2.2个百分点。

表1 正样本分配策略对比

正样本分配策略	mAP	AP50	AP75
FCOSv1的分配策略	0.768	0.938	0.883
FCOSv2的分配策略	0.781	0.942	0.907
改进后的正样本分配策略	0.803	0.954	0.920

实验结果表明,改进后的正样本分配策略融合了物体的细长度和方向角度的信息,能使得包含更多语义信息量的样本点作为正样本点,有助于模型更好地学习到细长物体的特征,检测精度得以提升。

2) 与经典单阶段检测算法的比较

本实验选择了经典的单阶段有锚框网络(RetinaNet)、YOLOv4、单阶段无锚框网络(RepPoints)与改进后的FCOS模型分别在细长物体数据集上进行了比较。训练细节与3.2节保持一致。

如表2所示,对细长物体的检测无锚框网络的表现要优于有锚框网络。此外改进后的FCOS模型的表现也优于同为无锚框的RepPoints和FCOSv2。与其他目标检测算法比,本文的算法取得了最高的检测精度,检测速度略低于其他单阶段目标检测算法。

表2 改进后的FCOS与经典单阶段检测算法的比较

网络名称	主干-颈部网络	mAP	推理时间/ms
RetinaNet	ResNet50-FPN	0.706	103
YOLOv4	CSPDarknet53-PAN	0.698	75
RepPoints	ResNet50-FPN	0.776	105
FCOSv2	ResNet50-FPN	0.781	102
FCOS+mDCN	ResNet50-FPN	0.810	112
改进后的FCOS	ResNet50-FPN	0.832	112

4 结语

针对细长物体目标检测不准确的问题,本文

在 FCOSv2 模型的基础上,通过提出融合细长物体的“细长度”和方向角度 θ 信息改进了正样本分配策略,确保了高质量的正样本用于训练,在主干网络和预测头加入了调制可变形卷积,进一步增强了主干网络的特征提取自适应能力;在细长物体数据集上达到 83.2% 的 mAP,并通过对比实验验证了算法的可行性,实现了检测精度与速度的均衡。

参考文献:

- [1] 陈曦.成立不到 2 年即填补世界空白 这个园区的企业怎么做到的[N].科技日报,2022-01-13(7).
- [2] 刘阳,谢永强,李忠博,等.基于深度学习的目标检测算法研究进展[J].通信技术,2021,54(9):2063-2073.
- [3] TIAN Z, SHEN C H, CHEN H, et al. FCOS: fully convolutional one-stage object detection [C]//2019

IEEE/CVF International Conference on Computer Vision (ICCV).Seoul,Korea (South):IEEE,2019:9626-9635.

- [4] TIAN Z, SHEN C H, CHEN H, et al. FCOS: a simple and strong anchor-free object detector[J]. IEEE Transactions on Pattern analysis and Machine Intelligence, 2022, 44(4):1922-1933.
- [5] 李昱,盖绍彦,达飞鹏,等.基于语义采样和检测框优化的目标检测算法[J].激光与光电子学进展,2022,59(18):356-363.
- [6] ZHU X Z, HU H, LIN S, et al. Deformable ConvNets V2: more deformable, better results [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019: 9300-9308.

收稿日期:2023-06-30

(上接第 122 页)

起伏多的缺点,导致判断方向次数增多,所耗时间也比改进的 A* 算法长;因此在融合改进 A* 算法和改进动态窗口算法下,既保留改进 A* 算法路径短、拐角少的特点,又解决了传统动态窗口算法在接近目标点时易扰动的问题。在目标点附近能快速朝着目标点移动,在路径长度上较传统 A* 算法减小了 5.73%,时间上节约了 40.90%。

综合来看,本文提出的融合算法能够达到全局路径最优且轨迹平滑性最好,达到了预期的目的。

4 结语

1) 为提高在复杂环境下叉车式 AGV 的路径规划和随机避障能力,本文提出了一种基于改进 A* 算法和改进动态窗口算法的融合算法。该算法通过优化搜索点的策略和启发函数、改进关键点的选择和冗余点的删除等措施,能够显著减少规划的路径长度和拐点,并节约时间。

2) 优化后的动态窗口算法改善了传统动态窗口算法在接近目标点时扰动大的问题,能够在接近目标点时迅速朝着目标点行进,动态调整路径。

3) 通过 4 种算法仿真实验结果对比,证明本文提出的融合算法在同等环境条件下有最为优秀的路径规划和避障能力。后续可根据叉车式 AGV 实际工作环境结合视觉图像^[10]处理进一步优化,以提高算法解决实际问题的能力。

参考文献:

- [1] 姚远,周兴社,张凯龙,等.基于稀疏 A* 搜索和改进人工势场的无人机动态航迹规划[J].控制理论与应用,2010,27(7):953-959.
- [2] 王凡,李铁军,刘今越,等.基于 BIM 的建筑机器人自主路径规划及避障研究[J].计算机工程与应用,2020,56(17):224-230.
- [3] 程传奇,郝向阳,李建胜,等.融合改进 A* 算法和动态窗口法的全局动态路径规划[J].西安交通大学学报,2017,51(11):137-143.
- [4] 周克帅,范平清.改进 A* 算法与人工势场算法移动机器人路径规划[J].电子器件,2021,44(2):368-374.
- [5] 邹文,韩丙辰,李鹏飞,等.融合改进 A* 算法和优化动态窗口法的路径规划[J].计算机集成制造系统,2022,55(5):145-162.
- [6] 迟旭,李花,费继友.基于改进 A* 算法与动态窗口算法融合的机器人随机避障方法研究[J].仪器仪表学报,2021,42(3):132-140.
- [7] 槐创峰,郭龙,贾雪艳,等.改进 A* 算法与动态窗口法的机器人动态路径规划[J].计算机工程与应用,2021,57(8):244-248.
- [8] 王彬,裴建军,李海洋,等.优化 A* 算法与动态窗口法的移动机器人路径规划[J].计算机集成制造系统,2022,68(9):221-237.
- [9] 绳红强,黄海英,石小锐,等.基于改进蚁群算法的自动驾驶路径规划研究[J].机械制造与自动化,2022,51(5):203-206.
- [10] MENG T C, YANG T H, HUANG J, et al. Improved hybrid A - star algorithm for path planning in autonomous parking system based on multi-stage dynamic optimization [J]. International Journal of Automotive Technology, 2023, 24(2):459-468.

收稿日期:2023-06-14