

纪桐, 王子涵. 基于 YOLOv7 的石墨化配电室员工违规行为检测[J]. 智能计算机与应用, 2024, 14(12): 118-123. DOI: 10.20169/j. issn. 2095-2163. 241216

基于 YOLOv7 的石墨化配电室员工违规行为检测

纪桐, 王子涵

(河北工程大学 信息与电气工程学院, 河北 邯郸 056038)

摘要: 石墨化工艺是一种需要对送电功率、过程参数、变压器温度等严格把控的过程, 配电室员工违规行为可能导致安全事故及产品质量问题。针对员工打瞌睡、擅自离岗、玩手机等违规行为, 本文提出一种改进的 YOLOv7 员工行为检测算法, 构建了轻量化网络模型, 并引入 *SIoU* (Scylla Intersection over Union) 损失函数提高检测的准确度。通过数据实验验证了本文算法的有效性。改进后的 YOLOv7 算法模型相较于原始版本, 检测速度更快, 检测精度提高了 3.7%。具有较高应用价值。

关键词: 行为检测; YOLOv7; 轻量级神经网络; 损失函数; 改进策略

中图分类号: TP391

文献标志码: A

文章编号: 2095-2163(2024)12-0118-06

Detection of violations of the staff in graphitization distribution room based on YOLOv7

Ji Tong, Wang Zihan

(School of Information and Electrical Engineering, Hebei University of Engineering, Handan 056038, Hebei, China)

Abstract: Graphitization process needs strict management of power delivery, process parameters, and transformer temperature. It is indicated that violations of the staff in the distribution room may threaten safety and product quality standard. Targeted on such violations as drowsiness, departure from the post without permission, and unbridled use of mobile phones, this paper proposes an improved YOLOv7 algorithm for detecting staff behaviors. Furthermore, a lightweight network model is constructed, and the *SIoU* (Scylla Intersection over Union) is introduced to enhance the detection accuracy of the paper. Then, the effectiveness of the proposed algorithm is validated through data experiments. Compared to the original version, the improved YOLOv7 algorithm model demonstrates faster detection speed and a 3.7% increase in detection accuracy, boasting significant potential for practical applications.

Key words: behavior detection; YOLOv7; lightweight neural network; loss function; improvement strategy

0 引言

随着现代工业发展, 电炉钢产量不断攀升, 石墨电极需求量与日俱增^[1]。石墨化配电室作为石墨化工艺生产过程中的中枢神经系统, 要对送电功率、过程参数、变压器温度等进行严格把控。石墨化配电室员工的违规行为可能导致安全事故及产品质量问题。长期以来, 主要依赖人工监审监控视频的方式约束配电室员工行为。这种方式不仅浪费人力与时间, 且不具有实时性, 无法满足石墨化工艺生产中对配电室员工行为要求。本文旨在使用图像识别检测网络模型对石墨化配电室员工的违规行为进行检测, 以确保石墨化产品生产质量与生产安全。

近年来, 在行为识别检测方面主要分为 2 类算

法模型。一类为 SSD 系列算法与 YOLO 系列算法为代表的单阶段 (one-stage) 目标检测算法模型, 另一类是以 R-CNN 系列为代表的两阶段 (two-stage) 目标检测算法模型。这 2 类算法中, 单阶段目标检测算法模型相比于两阶段目标检测算法模型在检测速度上更快, 检测精度也不低, 更符合实时检测需求。目前行为识别检测算法主要以单阶段检测算法的 YOLO 系列算法为主。文献[2]使用 YOLOv3 检测算法对员工玩手机、睡觉、擅自离岗行为进行检测, 结合图像相似度和明暗度等特征进行深度分析, 增加了检测精度, 但网络模型结构复杂度较高。文献[3]基于 YOLOv5 检测算法对学生课堂行为进行识别。提出多尺度特征融合的方法, 有效提高了学生课堂行为识别精度, 但检测速度相对较慢。文献

作者简介: 纪桐 (1998—), 男, 硕士研究生, 主要研究方向: 计算机视觉。Email: 278876111@qq.com; 王子涵 (1999—), 女, 硕士研究生, 主要研究方向: 计算机视觉。

收稿日期: 2023-07-04

哈尔滨工业大学主办 ◆ 系统开发与应用

[4]对工人玩手机行为进行了检测,使用 YOLOv5 算法为基础并加以改进,通过优化主干网络增加对浅层特征的提取,加入注意力机制并修改原有损失函数,在不过多增加模型复杂度的同时提高检测性能,但检测对象较为单一,无法满足石墨化配电室人员违规行为检测的实际需要。

针对石墨化配电室人员违规行为检测的实际需要,本文以 YOLO 算法系列中较新版本的 YOLOv7 算法模型为基础。通过对 YOLOv7 算法模型网络结构的修改,替换原有损失函数,添加相关模块,在使网络轻量化的同时,有效提高对员工行为的检测精度和检测速度,具有较高的应用价值。

1 YOLOv7 网络

YOLOv7 算法是 YOLO 算法系列中的较新版本。相比之前的 YOLO 系列检测算法模型,在检测速度和检测精度上均有一定提升。YOLOv7 网络模型主要包含了输入 (Input)、骨干网络 (Backbone)、

颈部 (Neck)、头部 (Head) 4 部分^[5]。其中,Input 提供待检测的图像数据并进行预处理。Backbone 提取图像的特征,并生成高层次的特征表示。Neck 进一步处理和融合特征,Head 生成最终的目标检测结果。这 4 个部分相互配合,共同完成目标检测任务。

2 改进策略

2.1 使用 GE 结构

GE 结构是一种轻量级的梯度路径网络结构,其结构参考了 YOLOv7 中的 ELAN (Efficient Layer Aggregation Networks) 梯度路径网络结构,通过减少 ELAN 结构中的部分梯度路径,并使用 Ghost 模块 (Ghost Module)^[6]来替代传统的卷积操作,有效缓解了特征图冗余问题。显著减少了网络的参数数量和计算量,最后使用 ECA (Efficient Channel Attention) 通道注意力机制提高重要特征的权重。使网络可以学习到更多特征,其结构如图 1 所示。

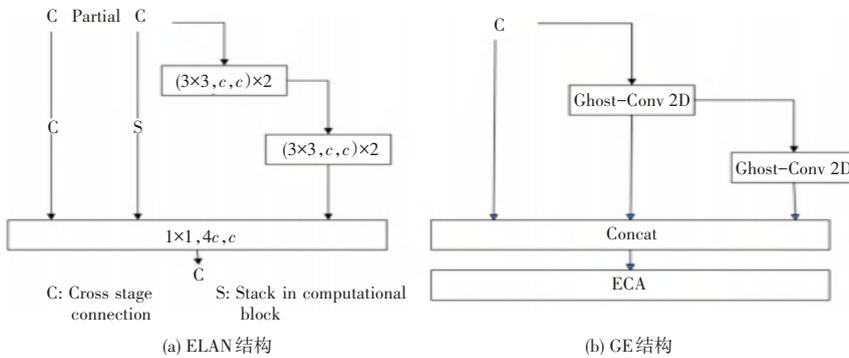


图 1 GE 结构图

Fig. 1 GE structure diagram

Ghost 模块是一种用于卷积神经网络中的模块。通过一系列的线性变换以很小的代价生成许多能从原始特征发掘所需信息的特征图,有效减少特征图冗余。该模块即插即用,旨在减少计算成本的同时保持较高的表示能力。其结构如图 2 所示。

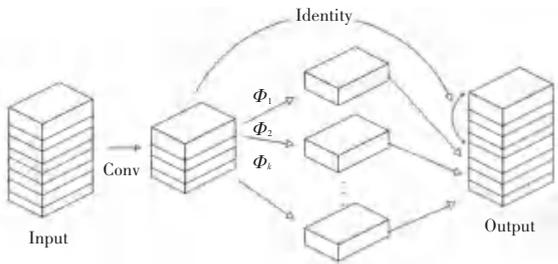


图 2 Ghost 模块

Fig. 2 Ghost module

注意力机制的核心思想是确定特征图中不同特征信息的重要性,以便将更多的计算资源用于关注更重要的信息。例如最具代表性的 SE (Squeeze-and-Excitation)^[7]通道注意力机制。考虑到 SE 注意力机制采用的降维操作可能对通道注意力的预测产生负面影响,并且其获取依赖关系的效率较低。因此,选择了一种高效的 ECA 通道注意力模块,有效避免了降维操作,实现跨通道的交互^[8]。ECA 通道注意力模块的结构如图 3 所示。

首先将原始特征图进行全局平均池化操作,特征图从 $[h, w, c]$ 的矩阵变成 $[1, 1, c]$ 的向量,再使用一维卷积(卷积核大小为 K)和激活函数得到各通道的权重,用 W_i 来表示学习到的通道注意力,对于权重 y_i ,只考虑 y_i 和相关的 k 个邻居之间的信息交互,计算公式如下:

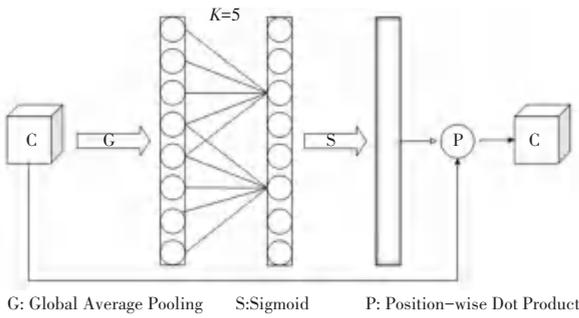


图3 ECA通道注意力模块结构图

Fig. 3 ECA Channel attention module structure diagram

$$\omega_i = \sigma\left(\sum_{j=1}^k w_i^j y_i^j\right), y_i^j \in \Omega_i^k \quad (1)$$

最后,将得到的权重与特征图中对应的元素相乘,得到新的特征图。本文 GE 结构在特征融合后加入 ECA 通道注意力模块,使模型可以有效提取特征图中的重要信息,在不增加计算量的同时,提高了模型性能。

2.2 修改损失函数

原始的 YOLOv7 算法使用 *CIoU* 损失函数来处理不同尺度框的判定问题,但其对于纵横比权重的

设计并不准确,收敛速度较慢且效率较低,最终可能产生更差的模型^[9]。为了解决这个问题,Gevorgyan^[10] 在研究中提出了一种新的损失函数 *SIoU*, 该函数充分考虑了所需回归之间的向量角度,并重新定义了惩罚指标,从而有效提高了模型训练的速度和推理的准确性。改进的损失函数在目标检测任务中表现出更好的性能和优化效果,能够更准确地处理不同目标的尺度和比例差异,加快模型的收敛速度。*SIoU* 损失函数定义公式如下:

$$Loss_{SIoU} = 1 - IoU + \frac{\Delta + \Omega}{2} \quad (2)$$

SIoU 损失函数由 4 部分组成:回归之间的向量角度(Angle cost)、中心点距离 Δ (Distance cost),长宽比 Ω (Shape cost),重叠面积(IoU cost)。其中,Angle cost 是用于解决预测框与真实框随意匹配问题,最大限度减少与距离相关变量的数量^[11]。相比于 *CIoU*,*SIoU* 的最大误差更小,且误差表面更加平滑,有效提高模型的训练速度以及准确性。改进后 YOLOv7 算法模型整体结构如图 4 所示。

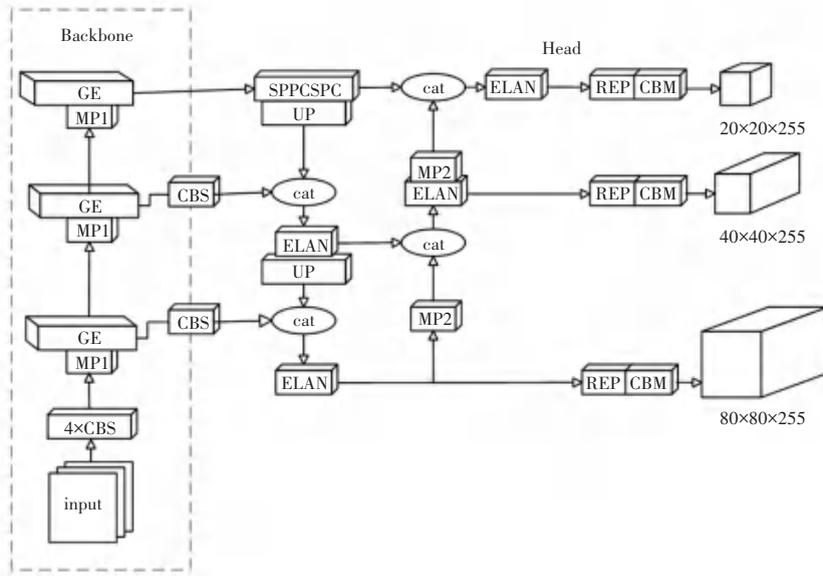


图4 改进后 YOLOv7 算法模型整体结构图

Fig. 4 The overall structure diagram of the improved YOLOv7 algorithm model

3 实验结果与分析

3.1 实验环境与数据集

本文使用的实验环境为 Ubuntu16.04 操作系统,专业级图形计算显卡 NVIDIA Quadro P2000,以及 PyTorch 1.7 和 CUDA 10.1 框架,同时使用 Python 3.8 环境。

本实验的数据来源于邯郸市成安县某碳素企业

的石墨化配电室值班监控视频,总时长为 240 h。研究中,从这些视频中截取图片 7 200 张,并对这些图片进行了归类整理,包括打瞌睡、玩手机、擅自离岗这 3 种违规行为和正常工作中行为(如图 5 所示)。每种行为分别选取了 500 张图片,使用 LabelImg 标注工具对图片进行标注,从而得到了包含 2 000 张图片的自建数据集。此数据集将用于进行石墨化配电室值班员工行为识别检测研究。



图 5 员工需要检测的 4 种行为

Fig. 5 Four behaviors of employees required to test

3.2 评价指标

本实验采用精确度 (P)、召回率 (R)、平均精确度 (AP)、平均精度均值 (mAP)、参数量、FPS 作为模型性能的评估指标。其中, P 表示模型预测出来的正样本中有多少是真正正样本; R 表示真正正

样本中有多少正样本被模型挑选出来; AP 表示精确度在不同召回率下的均值; mAP 则表示各类别的 AP 的平均值, 通常用来评价模型的整体性能。FPS 表示每秒钟模型能检测的图片数量^[12]。具体计算公式为:

$$P = \frac{TP}{TP + FP} \times 100\% \quad (3)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (4)$$

$$AP = \int_0^1 P(R) dR \quad (5)$$

$$mAP = \frac{\sum_{i=1}^n AP_i}{n} \quad (6)$$

其中, TP 表示被正确预测的正样本; FP 表示被误判为正样本的负样本; FN 表示被误判为负样本的正样本; n 表示总样本数。

3.3 实验分析

3.3.1 注意力模块实验

为了验证在 GE 结构中使用 ECA 注意力模块的有效性, 本文通过在 GE 模块同一位置分别进行 ECA、SE、CBAM^[13] 三种注意力机制实验, 实验结果见表 1。实验结果表明, 使用 ECA 注意力机制的效果最好。

表 1 融合不同注意力模块对比实验

Table 1 Comparative experiments on integrating different attention modules

算法	总体平均准确率 (Mean Average Precision) / %	召回率 (Recall) / %	FPS
YOLOv7+ECA	86.9	84.5	277
YOLOv7+SE	85.8	82.2	275
YOLOv7+CBAM	87.4	83.1	265

3.3.2 损失函数实验

为验证融合 $SIoU$ 损失函数的有效性, 本文以 YOLOv7 为基础, 通过替换原始网络损失函数为本文引入的 $SIoU$ 损失函数以及 $EIoU$ ^[14]、 $DIoU$ ^[15] 和

$GIoU$ 这 3 种损失函数进行对比实验。其实验结果见表 2。通过实验可以看出相较于其它损失函数, $SIoU$ 损失函数替换 YOLOv7 原有 $CIoU$ 损失函数的效果最好。

表 2 损失函数对比实验

Table 2 Comparative experiments of loss function

算法	总体平均准确率 (Mean Average Precision) / %	召回率 (Recall) / %	FPS
YOLOv7 ($CIoU$)	86.2	84.2	242
YOLOv7+ $SIoU$	88.2	87.3	243
YOLOv7+ $EIoU$	87.8	85.3	241
YOLOv7+ $DIoU$	85.4	83.1	242
YOLOv7+ $GIoU$	85.3	82.6	240

3.3.3 消融实验

为进一步验证本文改进算法的有效性,通过使用不同的改进策略进行消融实验,得到实验结果对

比见表3。表3中,模型1使用了GE结构,模型2更改算法原有损失函数为 *SIoU* 损失函数;模型3在模型1的基础上更改损失函数为 *SIoU*。

表3 消融实验

Table 3 Ablation experiment

算法	权重文件大小 / MB	总体平均准确率(Mean Average Precision) / %	召回率 (Recall) / %	FPS
YOLOv7	18.6	86.2	84.2	242
模型1	10.3	86.9	89.9	277
模型2	14.2	88.2	87.3	249
模型3	8.4	89.9	87.5	276

3.3.4 改进前后检测效果对比

本文将改进后的 YOLOv7 算法模型训练过程中各行为的检测效果与原始 YOLOv7 模型进行对比,发现改进后的模型在各类员工行为的检测精度上都有不同程度的提升,其中玩手机行为的检测提升效

果较为明显, $F1_curve$ 对比参见图6。且其召回率 (Recall) 也有明显提升 (R_curve 对比参见图7)。由此可见改进后的模型对于玩手机的识别能力明显提高,对配电室员工行为识别具备更好的检测能力。

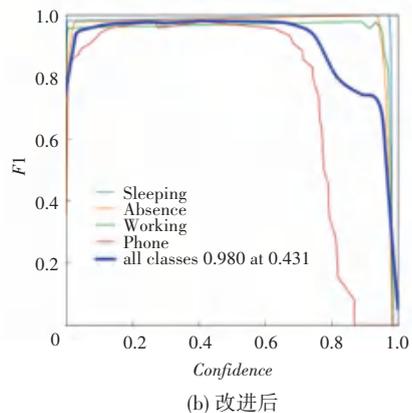
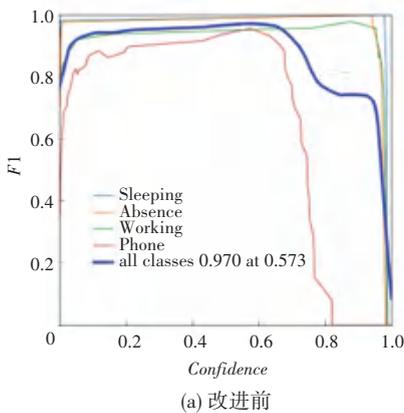


图6 改进前后对比 $F1_curve$ 图

Fig. 6 Comparison of $F1_curve$ before and after improvement

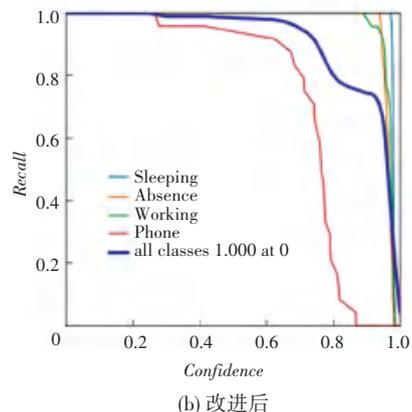
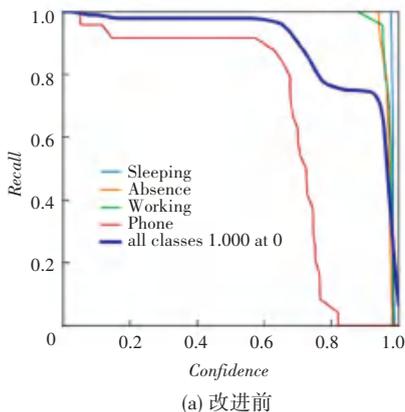


图7 改进前后对比 R_curve 图

Fig. 7 Comparison of R_curve before and after improvement

3.3.5 不同算法模型性能对比

本实验将改进后的 YOLOv7 算法与目前主流目

标检测算法进行对比,实验结果见表4。

表 4 检测算法对比

Table 4 Comparison of detection algorithms

算法	权重文件大小 / MB	总体平均准确率 (Mean Average Precision) / %	召回率 (Recall) / %	FPS
SSD	18.6	85.5	83.2	112
YOLOv3	20.3	84.5	83.3	89
YOLOv5n	10.0	85.1	84.2	228
YOLOv8n	11.8	85.5	84.3	254
改进 YOLOv7	8.4	89.9	87.5	276

由表 4 可知,改进后的 YOLOv7 相对于其他主流算法如 YOLOv8n、YOLOv5s、YOLOv3、SSD 而言,检测精准度明显提高,能够对员工行为进行更好的识别。同时,改进后的模型检测速度高于其他算法,满足实际检测的要求。

4 结束语

针对石墨化配电室员工行为识别实际需求,本文提出一种改进的 YOLOv7 员工行为检测算法。使用 GE 梯度路径网络结构替换原有 ELAN 梯度路径网络结构,采用 Ghost 模块替换传统卷积操作使得网络模型轻量化的同时,引入 ECA 通道注意力机制来提高重要特征的权重。修改原有损失函数为 $SIoU$ 损失函数,进一步考虑真实框和预测框间的向量角度,提高模型的训练速度以及推理的准确度。经过一系列对比实验,证明改进后 YOLOv7 算法模型相较于原始版本,在保持轻量化的同时,检测速度更快,检测精度提高了 3.7%,更符合石墨化企业实际需求,具有较高应用价值。后续会继续改进算法,进一步提升对配电室员工行为的检测效果。

参考文献

[1] 康进才, 吴洋, 胡春玉. 超高功率石墨电极原料针状焦的评价及应用[J]. 河南冶金, 2019, 27(1): 1-3.

[2] 云旭, 宋焕生, 梁浩翔, 等. 基于深度学习的关键岗位人员行为分析系统[J]. 计算机工程与应用, 2021, 57(6): 225-231.

[3] 张小妮, 张真真. 基于 YOLOv5 和多尺度特征融合的学生行为研究[J]. 现代信息科技, 2023, 7(8): 96-98.

[4] 林宝华, 刘坤, 朱一帆, 等. 基于 YOLOv5 的工人玩手机行为

检测方法研究[J]. 南京工程学院学报(自然科学版), 2023, 21(1): 39-44.

[5] 戚玲珑, 高建瓴. 基于改进 YOLOv7 的小目标检测[J]. 计算机工程, 2023, 49(1): 41-48.

[6] HAN Kai, WANG Yunhe, TIAN Qi, et al. GhostNet: More features from cheap operations[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2020: 1580-1589.

[7] HU Jie, SHEN Li, SUN G. Squeeze-and-excitation networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.

[8] WANG Qilong, WU Banggu, ZHU Pengfei, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 11534-11542.

[9] 李功, 赵巍, 刘鹏, 等. 一种用于目标跟踪边界框回归的光滑 IoU 损失[J]. 自动化学报, 2023, 49(2): 288-306.

[10] GEVORGYAN Z. SIoU loss: More powerful learning for bounding box regression[J]. arXiv preprint arXiv, 2205. 12740, 2022.

[11] 朱瑞鑫, 杨福兴. 运动场景下改进 YOLOv5 小目标检测算法 [J]. 计算机工程与应用, 2023, 59(10): 196-203.

[12] 王彦生, 曹雪虹, 焦良葆, 等. 基于改进 YOLOv5 的电厂人员吸烟检测[J]. 计算机测量与控制, 2023, 31(5): 48-55.

[13] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module [M]// FERRARI V, HEVERT M, SMINCHISESCU C. Computer Vision-ECCV2018, Lecture Notes in Computer Science. Cham: Springer, 2018: 3-19.

[14] ZHANG Yifan, REN Weiqiang, ZHANG Zhang, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146-157.

[15] ZHENG Zhaohui, WANG Ping, LIU Wei, et al. Distance-IoU loss: Faster and better learning for bounding box regression [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.