

doi: 10.7690/bgzd.2024.05.007

基于改进 YOLO 的不规范佩戴安全帽检测

郭威¹, 樊彦国¹, 栗晓政¹, 张兴富², 王满意²

(1. 国网河南省电力公司, 郑州 450018; 2. 北京中电普华信息技术有限公司, 北京 100089)

摘要: 为改善现有变电站巡检人员不规范佩戴安全帽检测时效率、精度低的问题, 提出一种基于改进 YOLO 的轻量化变电站人员不规范行为检测模型。该模型由特征提取网络、ECA-SPP 和 ECA-PANet 网络以及预测网络组成; 特征提取网络中使用 MobileNetV3; 提取 4 个尺度的特征图并将其输入到 SPP 和 PANet 网络中, 并基于注意力机制进行优化; 以建立的变电站人员不规范佩戴安全帽检测数据集为例, 验证所提模型有效性。实验结果表明: 所提模型 mAP 为 0.824 4, FPS 为 38.06, 明显优于 Faster RCNN、YOLOv4、YOLOx 等模型, 具有较高精度和更快的检测速度, 可为变电站人员不规范佩戴安全帽的实时检测提供参考。

关键词: 电力系统; 异常检测; 负荷预测; 支持向量机

中图分类号: TP393 **文献标志码:** A

Detection of Nonstandard Wearing of Safety Helmet Based on Improved YOLO

Guo Wei¹, Fan Yan'guo¹, Li Xiaozheng¹, Zhang Xingfu², Wang Manyi²

(1. State Grid He'nan Electric Power Company, Zhengzhou 450018, China;
2. Beijing China-power Information Technology Co., Ltd., Beijing 100089, China)

Abstract: In order to solve the problem of low efficiency and accuracy in the detection of non-standard safety helmet worn by the existing substation patrol personnel, a lightweight substation personnel non-standard behavior detection model based on improved YOLO is proposed. The model consists of a feature extraction network, an ECA-SPP network, an ECA-PANet network and a prediction network; MobileNet V3 is used in the feature extraction network; feature maps of four scales are extracted and input into the SPP and PANet networks, and are optimized based on an attention mechanism; The effectiveness of the proposed model is verified by the data set of the detection of non-standard wearing of safety helmets in substations. The experiment results show that the proposed model mAP is a 0.8244 and FPS is a 38.06, which is obviously better than other models such as Faster RCNN, YOLOv4 and YOLOx, and has higher accuracy and faster detection speed. It can provide a reference for real-time detection of substation personnel wearing non-standard safety helmet.

Keywords: power system; anomaly detection; load forecasting; support vector machine

0 引言

变电站^[1]是电力系统中电压变换、电流变换和配电的主要场所, 其安全稳定运行是减少停电事故发生的基础。由于变电站人员长期处于同一环境中, 心态逐渐松懈, 安全意识不可避免地减弱, 这类变电站人员在实际操作中经常违反操作规程, 导致各种事故^[2]。在实际现场监测常规操作、安全人员和个人安全防护设备已成为减少工作场所事故和工人伤害的关键问题之一。最常见的不规范佩戴安全帽行为, 对操作员存在极大的安全隐患。随着网络、大数据、物联网、通信技术的不断发展^[3-5], 可基于视频监控设备实时采集巡检人员工作过程数据, 并基于深度学习技术预警人员防护设备的使用情况, 从而减少因防护设备缺少导致的安全问题。

为此, 大量学者对变电站人员不规则行为的检

测方法进行研究, 并取得了丰硕成果。文献[6]提出了基于 PSPNet 语义分割模型的变电站人员危险行为检测模型, 可提高变电站安全管控的智能化水平。然而, 语义分割检测速度低, 实时性能低, 可能导致检查人员无法及时预防安全事故的发生。文献[7]提出了一种基于 OpenPose 多人姿态估计算法的人员危险行为检测模型。该模型通过对变电站安全区域的标定和人体关键姿态点的识别, 以人员的双脚是否在标定的安全区域内判断人员行为是否是危险行为。然而, 姿态估计算法实现困难, 且算法实时性较低。文献[8]提出了基于视频识别和 UWB 定位的变电站作业现场安全识别方法。该模型以 YOLOv3 为视频检测基础算法, 然而 YOLOv3 性能较低, 面对复杂背景识别能力较弱。

为改善上述问题, 笔者提出一种基于改进

收稿日期: 2024-01-23; 修回日期: 2024-02-25

第一作者: 郭威(1989—), 男, 河南人, 硕士。

YOLO 的轻量化变电站人员不规范佩戴安全帽检测模型。该模型主要包括特征提取网络、ECA-SPP 和 ECA-PANet 网络以及预测网络,可有效实现变电站巡检人员不规范佩戴安全帽检测。

1 基础网络结构和原理

接下来将详细介绍轻量化改进 YOLO 模型中涉及的基础网络结构和原理。

1.1 YOLO 模型

YOLO 模型是一种端到端的目标检测算法,该模型吸收了许多优秀的优化策略,如特征提取网络、模型训练、损失函数等,并且在准确性和速度方面有了显著的改进。笔者使用 YOLOv4 作为基础网络来实现轻量化网络模型,从而促进变电站巡检中不规范佩戴安全帽的实时检测。YOLOv4 模型主要由特征提取网络(CSPDarknet53)、特征增强网络(SPP+PANet)和预测网络(YOLOHead) 3 部分组成。YOLOv4 模型结构如图 1 所示。



图 1 YOLOv4 模型结构

首先,将变电站人巡检员行为数据集的图像输入到 CSPDarkNet53 中。CSPDarkNet53 由卷积-批量归一化-Mish 激活函数(convolution-batch normalization-Mish, CBM)模块和 5 个 CSP 模块组成。提取了 3 个尺度的特征,包括 52×52、26×26、13×13,并使用 SPP 和 PANet 进行融合,以增强特征信息。通过 YOLOHead 中的 3×3 和 1×1 卷积来预测目标的类别和位置。

其次,对类别和位置的预测结果进行解码,以便直观地显示检测结果。通常,预测结果在 YOLOHead 中通过 1×1 卷积处理成 S×S×M 维特征信息。其中:S 为具有一定比例的特征图的大小;M 为特征信息,包括位置和类别,其含义定义如下:

$$M=N_a \times (N_c + L_{x,y,w,h} + C) \quad (1)$$

式中: N_a 为在每个比例特征图上分配的锚点; N_c 为类别的总数; $L_{x,y,w,h}$ 为锚点的 4 个位置参数; C 为置信分数,定义如下:

$$C=P_r(o) \times I_{ou}(t_{rth}, p_{red}) \quad (2)$$

式中: $I_{ou}(t_{rth}, p_{red})$ 用于测量真值边界框和预测边界框之间的一致性,其值是相交区域和并集区域的比率; $P_r(o)$ 反映锚框是否包含目标,如果有目标,其值为 1,否则,其值为 0。

1.2 MobileNetV3

MobileNetV3 引入了深度可分离卷积、反向残差、线性瓶颈、挤压和激励网络(squeeze-and-excitation networks, SENet)、注意机制和 hard-Swish 激活函数等机制,可有效提高模型的特征提取能力。

MobileNetV3 主要由 1 个 CBH 模块和 15 个 bneck 模块组成。CBH 模块包括卷积层(convolutional layer, Conv)、批处理归一化(batch normalization, BN)层和 hard-Swish 激活功能。hard-Swish 激活功能是 MobileNetV3 中独有的激活功能,其定义如式(3)所示。通过将 swish 中的 sigmoid 函数替换为 ReLU6 函数,提高了模型效率,加快了推理速度。

$$h_{swish}(x) = x \cdot \left(\frac{\text{ReLU6}(x+3)+6}{6} \right) \quad (3)$$

$$\text{ReLU6}(x) = \min[6, \max(0, x)]$$

bneck 模块主要基于深度可分离卷积和具有线性瓶颈的反向残差设计而成。同时, bneck 模块中引入轻量级 SENet 模块,以增加模型对重要特征的学习能力。bneck 模块结构如图 2 所示。

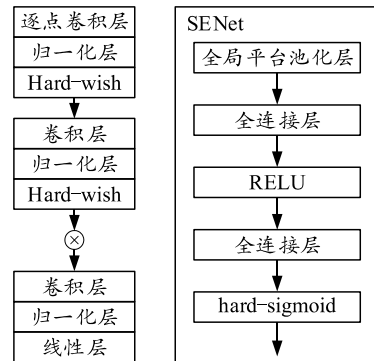


图 2 bneck 模块结构

根据上图,首先将输入带入逐点卷积层(pointwise convolution, PC)以增加维度,然后通过 BN 层和 hard-Swish 激活函数。其次,使用 3×3 卷积的深度卷积提取每个输入通道的特征,并通过 SENet 关注机制生成每个通道的相应权重系数,该机制包括全局平均池化(global average pooling, GAP)、全连接层(fully connected, FC)、ReLU 激活函数和 hard-sigmoid 激活函数。接着,权重系数立即与对应通道的所有元素相乘,以增强重要特征并抑制不重要特征。最后,使用最后的逐点卷积进行降维,并进行线性映射以减少特征的损失,从而获得更好的检测结果。

1.3 ECA 模块

由于视频监控获得的图像的背景非常复杂,需

要解决环境造成的干扰；为此，通过添加注意力机制来提高检测精度。为平衡精度和计算，笔者使用 ECA 模块来改进 SPP 和 PANet 网络，从而最小程度地增加计算量，并可以显著提高模型检测性能。ECA 模块的原理和结构如图 3 所示。

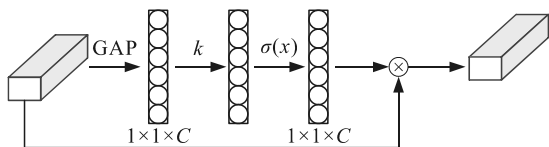


图 3 ECA 模块的原理和结构

在保持通道维度的前提下，ECA 模块对特征通道执行 GAP，并执行 1 维卷积以生成通道权重，其

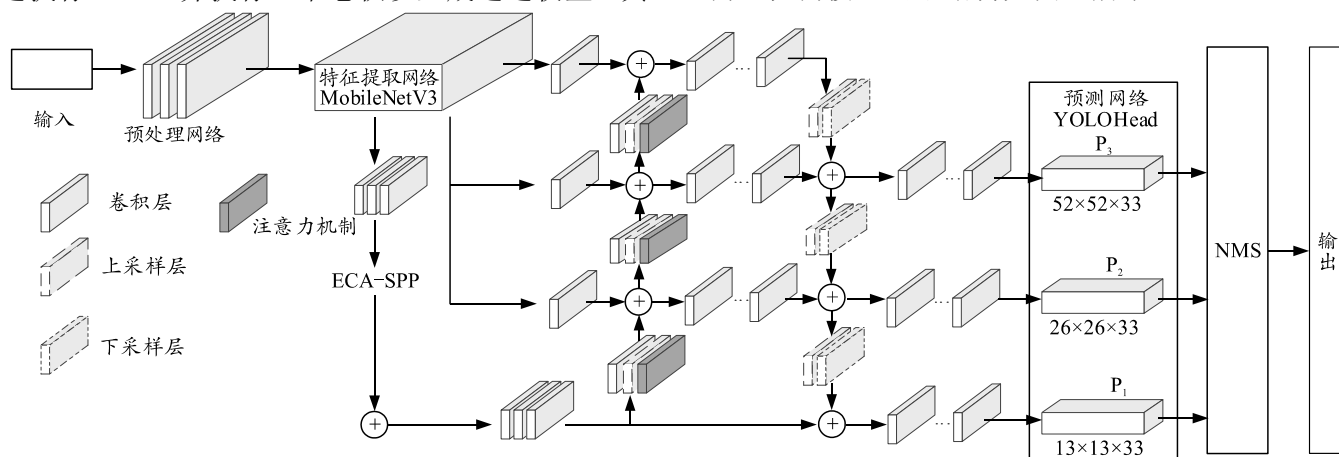


图 4 检测模型网络结构

首先，使用伽马校正图像预处理网络来减少来自无关信息的干扰并提高图像对比度。然后，通过 MobileNetV3 提取 4 个尺度特征，并与高效通道注意力-空间金字塔池化 (efficiency channel attention-spatial pyramid pooling, ECA-SPP) 和高效通道注意力-路径聚合网络 (efficiency channel attention-path aggregation network, ECA-PANet) 相结合，以融合和增强特征信息，检测层利用融合的特征来预测位置和类别。最后，采用非极大抑制 (non-maximum suppression, NMS) 消除冗余预测边界框，实现了变电站人员不规范佩戴安全帽的检测。

该模型主要包括 3 个关键网络：特征提取网络、ECA-SPP 和 ECA-PANet 网络、预测网络。接下来对各个子网络进行详细介绍。

2.1 特征提取网络

与基础 YOLO 网络不同，首先改进的 YOLO 轻量级网络中特征提取网络由 CSPDarknet53 取代为 MobileNetV3 轻量级卷积神经网络。其次，令初始 MobileNetV3 所需的输入大小从 224×224 扩展为

中卷积核大小可自适应确定，且核大小与通道维度成比例。因此，卷积核大小可以基于给定通道维度计算：

$$k = \lceil \log(C)/a + b/a \rceil_{\text{odd}} \quad (4)$$

式中： k 为卷积核大小； C 为通道维度； $\lceil \cdot \rceil_{\text{odd}}$ 为最接近的奇数； a 和 b 为线性函数系数，通常分别设置为 2 和 1。

2 改进 YOLO 模型

笔者提出一种改进的 YOLO 轻量级模型，从而实现变电站人员不规范佩戴安全帽的快速准确检测。检测模型网络结构如图 4 所示。

416×416。接着，将最初的 3 个尺度特征扩展为 4 个尺度，且 bneck3、bneck3、bneck6 和 bneck3 模块之后的特征分别用作特征 1、2、3 和 4，其中 bneck 的数量表示模块的数量。此外，SENet 作用于第 2 个 bneck3 模块、第 3 个 bneck6 模块和最后 1 个 bneck3 模块。然后，将改进 YOLO 轻量化模型中的特征 1、2、3 和 4 的大小重塑为 104×104、52×52、26×26 和 13×13，并用于 SPP 和 PANet 网络的特征融合和增强。

2.2 ECA-SPP 和 ECA-PANet 网络

为进一步提高检测精度，在提取的 4 个尺度特征之后添加了 ECA 注意机制，即将基础 SPP 和 PANet 中的卷积模块替换为 ECA 模块，从而构成 ECA-SPP 和 ECA-PANet。需注意，本文中 ECA 注意机制添加到具有 4 种池化核大小：1×1、5×5、9×9、13×13。此外，在每次上采样和下采样之后都添加了 ECA 注意机制。

2.3 预测网络

同时，将 ECA-SPP 和 ECA-PANet 输出特征图

输入到预测网络 (YOLOHead)。为进一步减少模型的计算量,笔者将基础 YOLO 中 3 个尺度的融合特征重塑为 $52 \times 52 \times 33$ 、 $26 \times 26 \times 33$ 和 $13 \times 13 \times 33$, 其中第 3 维度 33 可以被拆分为 $3 \times (6+5)$ 。其中, 3 表示分配给每个比例特征图的锚点; 6 表示变电站人员的不规则行为数据集中的类别总数; 5 可分为 $1+1+1+1+1$, 分别表示预测框的中心点、高度和宽度的横坐标和纵坐标以及置信度。

3 仿真与分析

为验证改进 YOLO 轻量化检测模型的优越性, 使用 Pycharm 2019、Tensorflow1.13.0、Opencv 4.2.0.34、Cuda 10.0.130、cudnn7.6.5.32 的软件环境和具有 11 GB 内存的 Nvidia GeForce RTX 2080TI GPU 的硬件环境, 对变电站人员的不规范佩戴安全帽进行了检测。

3.1 实验设置

3.1.1 数据集

笔者使用的变电站人员不规范佩戴安全帽检测数据集部分来自变电站视频监控系统捕获的视频数据, 部分来自变电站工人工作时拍摄的照片, 还有一些来自网络爬虫爬取的照片。

首先, 对视频数据进行抽帧并提取出关键帧, 与其他图像构成原始数据集。接着删除遮挡严重、严重失真图像, 数据集中共包含 13 046 幅图像。其次, 利用 LabelImg 软件对图像进行标注, 共包含安全头盔和安全人员 2 类。检测网络对变电站人员工作时应穿戴物品的检测可以反映变电站人员是否有不正常行为。如果控制中心发现这些不规则行为, 可以及时制止, 避免发生人身安全事故。

此外, 将数据集按 8:1:1 分为训练集、验证集和测试集。

3.1.2 参数设置

实验时选取平均精度 (mean average precision, mAP)、FPS 以及 F1 分数指标来衡量模型性能。其中, F1 分数是一种提供分类器整体性能的平衡评估度量。F1 分数为精度 p 和召回率 r 的调和平均值, 定义如下:

$$\left. \begin{aligned} p &= T_p / (F_p + T_p) \\ r &= T_p / (F_N + T_p) \\ F_1 &= 2pr / (p + r) \end{aligned} \right\} \quad (5)$$

式中: T_p 为正常数据被正确分类的数量; F_p 为异常

数据被分类为正常的数量; T_N 为异常数据被分类为异常的数量; F_N 为正常数据被分类为异常的数量。

网络训练时采用 SGD 优化器训练模型, 实验时部分参数定义如下: 批量大小设置为 16, 初始学习率为 10^{-2} , 学习率衰减率为 10^{-1} , 学习率衰减周期设置为 10, 最大迭代次数设置为 150, NMS 阈值为 0.3。

3.2 结果与分析

对比 Faster RCNN、YOLOv4、YOLOx 以及所提 YOLO 轻量化模型的训练性能。

图 5 所示为不同模型训练损失对比结果。可以看出: 所提模型训练损失收敛更低, 性能明显优于 Faster RCNN、YOLOv4、YOLOx 等模型。最终, 改进 YOLO 轻量化模型训练最小损失约为 0.103 2。

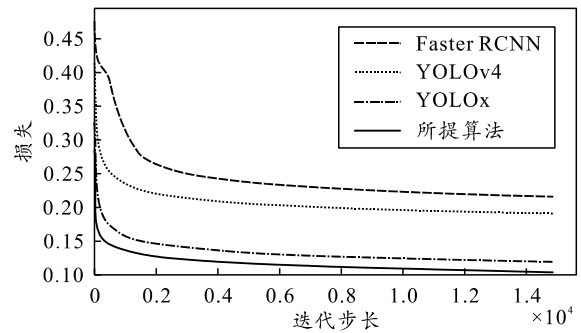


图 5 不同模型训练损失对比结果

进一步, 实验时对比了所提模型与 Faster RCNN、YOLOv4、YOLOx 的交叉对比性能。对比指标分别选取 mAP、FPS 和 F1 分数。表 1 所示为不同模型交叉对比结果。

表 1 不同模型交叉对比结果

模型	mAP	FPS/(帧/s)	F1 分数
Faster RCNN	0.609 9	12.98	0.568 3
YOLOv4	0.764 7	28.30	0.645 4
YOLOx	0.803 1	22.08	0.795 2
所提模型	0.824 4	38.06	0.792 7

根据上表, 可以看出所提改进 YOLO 轻量化模型的 mAP 优于 YOLOx 模型, F1 分数略低于 YOLOx; 因此, 所提改进 YOLO 轻量化模型识别性能与 YOLOx 性能相当, 然而 FPS 指标, 所提模型为 38.06 帧/s, 明显优于 YOLOx 的 22.08 帧/s。此外, Faster RCNN 与基础 YOLOv4 模型性能较低。分析原因, 上述模型在面对变电站复杂背景时, 特征学习能力有限, 因此识别性能整体较低。