

基于改进 YOLOv8n 的接触网螺栓识别与定位方法

吴守烨, 孟祥印, 肖世德, 卢秀杰, 迟元斌

(西南交通大学 机械工程学院, 成都 610031)

摘要: 传统的接触网螺栓识别方式效率较低, 且现有的螺栓定位方法研究较少; 针对上述问题, 提出了一种结合改进 YOLOv8n 模型和 RGB-D 相机的接触网螺栓识别与定位方法; 对目标识别模型 YOLOv8n 进行了创新改进, 增加小目标检测层以增强特征融合; 引入 SE 注意力机制突出关键特征, 提高模型检测精度; 引入 BiFPN 模块提升多尺度特征融合效率; 采用轻量级 SCDown 卷积降低模型参数量且保证检测性能; 引入 WIoU-v3 损失函数加速模型收敛并提升回归精度; 基于螺栓的对称特性和图像处理技术, 提出新的方法来实现螺栓定位, 结合目标识别模型和 RGB-D 相机深度信息, 以此获取螺栓的三维中心坐标和空间姿态; 实验结果表明, 改进后的识别模型在测试集上的 $mAP@0.5$ 达到 90.7%, 比原模型提高 2.1%, 并减少了 8.0% 的参数量, 螺栓定位方法也能够实现对接触网螺栓的有效定位, 验证了该方法的可靠性。

关键词: 电气化铁路接触网; 螺栓识别定位; 机器视觉; YOLOv8n; RGB-D 相机

Recognition and Localization Method for Catenary Bolts Based on Improved YOLOv8n

WU Shouye, MENG Xiangyin, XIAO Shide, LU Xuijie, CHI Yuanbin

(School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: Traditional recognition methods for catenary bolts are inefficient, and there is limited research on existing bolt localization approaches. To address these issues, a novel bolt recognition and localization method combining an improved YOLOv8n model with an RGB-D camera is proposed. The YOLOv8n object detection model is enhanced by adding a small target detection layer to improve feature fusion; an SE attention mechanism is introduced to highlight key features and improve detection accuracy; a BiFPN module is incorporated to enhance multi-scale feature fusion efficiency; A lightweight SCDown convolution is used to reduce the model's parameter count while maintaining detection performance; A WIoU-v3 loss function is adopted to accelerate model convergence and improve regression precision. Based on the symmetrical properties of bolts and image processing techniques, a new method is developed for bolt localization, integrating the object recognition model with depth information from the RGB-D camera to obtain the bolt's 3D center coordinates and spatial pose. Experimental results show that the improved recognition model achieves an $mAP@0.5$ of 90.7% on the test set, which is 2.1% higher than that of the original model, and the improved model reduces the parameter count by 8.0%. The bolt localization method effectively locates catenary bolts, validating the reliability of the proposed approach.

Keywords: electrified railway catenary; bolt recognition and localization; machine vision; YOLOv8n; RGB-D camera

收稿日期:2024-12-13; 修回日期:2025-01-17。

基金项目:四川省科技厅重大专项(2022ZDZX0002);四川内江高新技术产业开发区管理委员会西南交通大学产学研合作资助课题(R110223H01022)。

作者简介:吴守烨(2000-),男,硕士。

通讯作者:孟祥印(1979-),男,博士,副教授。

引用格式:吴守烨, 孟祥印, 肖世德, 等. 基于改进 YOLOv8n 的接触网螺栓识别与定位方法[J]. 计算机测量与控制, 2025, 33(12): 262-269, 341.

0 引言

电气化铁路接触网作为铁路牵引供电系统中的重要组成部分, 其可靠性对列车的稳定运行起着至关重要的作用^[1-2]。螺栓作为接触网的重要连接件, 其紧固质量直接影响着接触网的可靠性和安全性, 目前接触网螺栓紧固作业依旧采用的是人工周期性平推作业, 存在工作量大、高空作业不安全等问题^[3-4]。使用智能作业机器人代替人工对接触网螺栓进行紧固, 能够达到标准化作业的任务要求, 并且对于智能设备在铁路运维领域的全面发展具有重要意义^[5]。在利用机器人对接触网螺栓进行紧固作业之前, 需要对其进行识别和定位, 以引导机器人能够准确的到达指定位置进行作业, 因此有必要开发一种可靠的接触网螺栓识别与定位方法。

基于深度学习的目标检测算法近年来有着出色的表现^[6], 在螺栓识别与定位这一领域的研究也日渐深入。Zhao 等人^[7]提出了一种基于改进 YOLOv3 的码垛机器人定位螺栓检测方法, 实验结果表明, 检测 AP 为 86.7%, Recall 为 97%, 检测速度达到 25.47FPS, 能够实现在复杂的工业场景下完成高效准确的检测。Wang 等人^[8]提出了一种基于改进 YOLOv5 的桥梁结构健康监测螺栓定位检测方法, 采用 k-means+聚类算法和 Hypercolumn (HC) 技术对模型进行改进, precision 和 AP 比原来的 YOLOv5x 分别提升了 6.5% 和 5.9%。戴先鑫等人^[9]提出了一种基于感兴趣区域的螺栓位姿及尺寸检测方法, 将目标识别算法 YOLOv5 和传统图像处理方法结合, 实现了对螺栓位姿和尺寸的识别。除此之外, 也有相关研究结合深度相机来完成螺栓识别定位, Jiang 等人^[10]提出了一种基于深度相机的螺栓定位方法, 将深度图和彩色图像结合实现螺栓的定位, 采用该方法得到的螺栓间距绝对误差小于 2.2 mm。王向周等人^[11]将 YOLOv5s-T 网络应用于深度相机 D435i 采集的图像, 实现了角钢塔主材螺栓的实时检测、三维定位和排序等功能, 推理速度提高约 31%, 但其检测对象在同一平面且分布较为集中。虽然国内外研究学者在螺栓识别与定位这一领域取得了一定的进展, 但是目前的研究大多集中于螺栓识别, 对于定位问题的研究相对较少, 铁路接触网螺栓的识别定位领域的研究几乎为空白。螺栓三维姿态的确定能够有效引导机器人进行作业, 因此提出新的接触网螺栓识别与定位算法具有一定的意义。

针对以上问题, 本文选择 YOLO 系列中检测精度较高且模型尺寸较小的 YOLOv8n 网络作为基础, 提出了一种改进的 YOLOv8n-SBSW (SE+BiFPN+SCDown+WIoU) 接触网螺栓检测算法。在颈部网络中添加小目标检测层, 对浅层特征和深层特征进行融合, 提高网

络对小目标的检测能力; 引入 SE (Squeeze-and-Excitation) 注意力机制, 选择性地强调目标特征信息来提高模型的检测精度; 引入 BiFPN 模块, 利用双向特征融合和权重分配机制, 加强不同尺度特征融合的效率; 引入 SCDown 卷积模块替换掉颈部网络中原有的部分卷积模块, 在不降低检测精度的同时减少模型的参数量; 将 YOLOv8n 原本的边界框回归损失函数 CIoU 替换为 WIoU v3, 加快模型的收敛速度, 进一步提升模型性能。将改进后的 YOLOv8n-SBSW 模型与 RGB-D 深度相机进行结合, 提出了一种接触网螺栓的定位方法, 实验结果验证了该方法的可行性, 对智能化自动化的接触网螺栓紧固作业方式有一定参考作用。

1 螺栓识别定位系统方法

本文所提出的方法结合了深度学习目标识别模型、深度相机以及相关图像处理计算机视觉技术, 旨在实现对接触网螺栓的精准识别与三维定位。系统包括 RGB-D 相机模块、目标识别模块和三维定位模块 3 部分, 该系统方法的总体设计如图 1 所示。

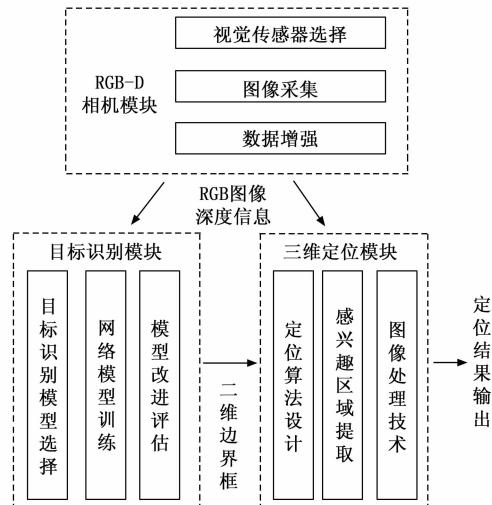


图 1 螺栓识别定位方法总体设计

1.1 RGB-D 相机

RGB-D 相机主要负责图像采集和深度信息输出, 为后续的目标识别和定位提供数据。相机选用了 Intel RealSense D435i, 它是一款集成了 RGB 相机、双目立体视觉和惯性测量单元 (IMU) 的高性能深度相机。通过双目立体视觉和主动红外投影实现高精度深度感知, 其中双红外摄像头捕获视差信息计算每个像素的深度, 红外投影器在无纹理表面和弱光环境中增强深度检测效果。RGB 摄像头与深度数据同步采集, 实现彩色与深度信息的空间对齐。内置的 IMU 提供实时姿态和运动数据, 用于动态环境中的位姿跟踪与运动补偿。同时, 用户可以使用官网提供的 Intel RealSense SDK 2.0, 对

数据进行进一步处理。借助这些功能, D435i 广泛应用于工业检测、机器人导航以及物体定位等领域。

1.2 目标识别模型

目标识别模型负责在输入的 RGB 图像中, 检测出螺栓的位置, 并给出螺栓的二维边界框和类别, 为后续的三维定位提供参考。YOLO 系列^[12-18] 网络模型在机器视觉领域上表现非凡, 尤其是近年来 Ultralytics 公司发布的 YOLOv8^[19] 模型。YOLOv8 的检测性能已经超越了现有的多数目标检测模型, 可以较好地平衡检测速度和精度, 能够满足复杂环境下的实时目标检测任务要求。YOLOv8n 是在 YOLOv8 的几个版本当中模型尺寸最小, 参数量最少的, 能够更容易嵌入到边缘设备中, 因此本文以 YOLOv8n 为基础并对网络加以改进, 进一步提高检测精度和降低计算成本, 确保模型能够快速、准确地识别出接触网螺栓。

1.3 三维定位方法

三维定位部分将结合螺栓识别模型输出的目标边界框和相机提供的深度信息, 计算出螺栓的三维位置和姿态。具体而言, 首先, 通过目标检测模型对输入图像进行处理, 识别并提取出图像中螺栓的兴趣区域, 这一过程为后续定位提供了明确的目标区域; 随后, 利用图像处理技术对提取的螺栓区域进行进一步分析和操作, 以增强目标特征; 最后, 结合从 RGB-D 相机获取的深度信息, 利用所提的定位方法计算出螺栓在三维空间中的坐标位置和空间姿态。

2 基于改进 YOLOv8n 和 RGB-D 相机的识别定位

2.1 YOLOv8n 目标识别网络

YOLOv8 的网络结构主要由 3 部分构成: Back-

bone、Neck 和 Head, 如图 2 所示。主干网络 Backbone 负责提取多层次特征, 采用 C2f 模块替代 YOLOv5 的 C3 模块, 实现梯度流优化和模型轻量化, 并通过 SPPF 模块增强模型对多尺度特征的适应能力。颈部网络 Neck 使用改进的 PAN-FPN 结构, 融合 FPN 自顶向下的深层语义信息与 PAN 自底向上的浅层位置信息, 提升了特征利用效率和检测性能。检测网络 Head 部分采用解耦头结构 (Decoupled Head), 将分类与边界框回归分离, 并引入 DFL 和 CIoU, 进一步提高检测精度和模型收敛速度。

2.2 YOLOv8n 改进策略

如图 3 所示, 该检测模型以 YOLOv8n 为基础模型, 从 5 个方面对模型进行优化改进: 引入 SE 注意力机制、增加小目标检测层、引入 BiFPN 模块、引入 SC-Down 卷积模块、引入 WIoU 损失函数。

2.2.1 SE 注意力机制

SE (Squeeze-and-Excitation)^[20] 注意力机制是一种通道类型的注意力机制, 即在通道维度上增加注意力, SE 模块主要包括两个内容: Squeeze (压缩)、Excitation (激励), 如图 4 所示。

1) Squeeze 操作, 通过全局平均池化将卷积后的特征图进行压缩, 将每个通道的二维特征 ($H \times W$) 压缩为一个实数, 从而将输入图像的特征由 $H \times W \times C$ 变为 $1 \times 1 \times C$, 计算公式如式 (1) 所示:

$$z_c = F_q(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (1)$$

式中, u 是卷积后的特征图, c 是通道数, $H \times W$ 是 u 的空间维度。

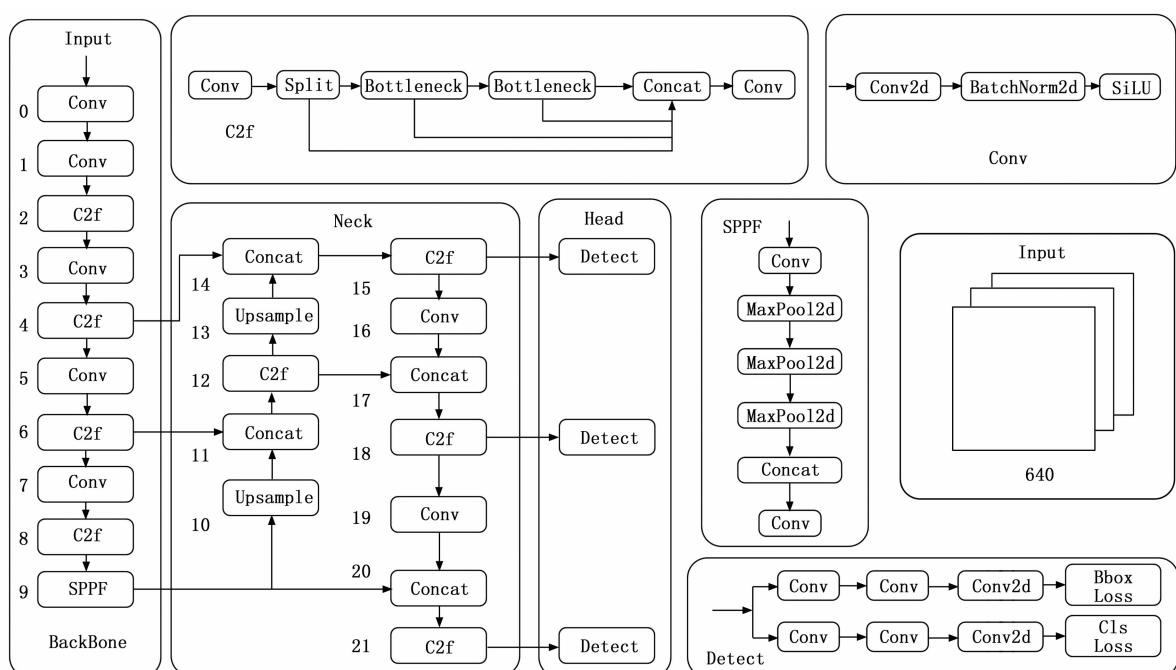


图 2 YOLOv8n 的网络结构

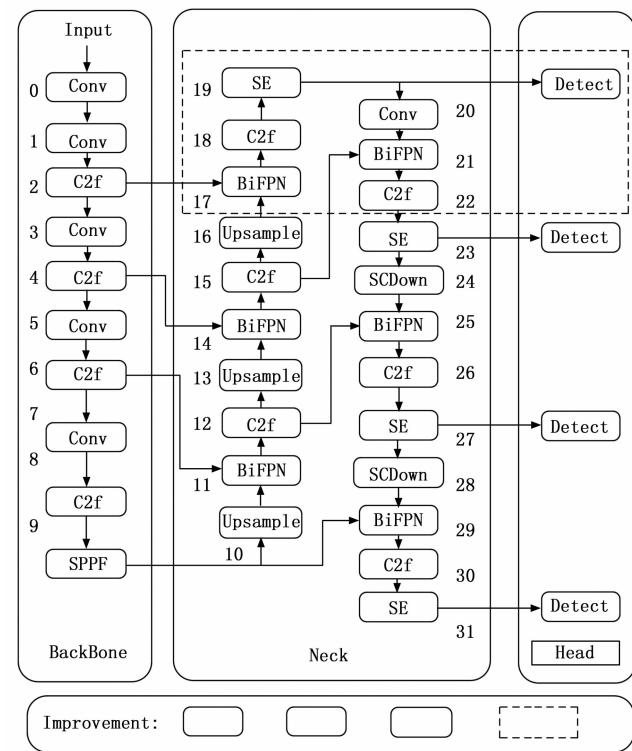


图 3 YOLOv8n-SBSW 的网络结构

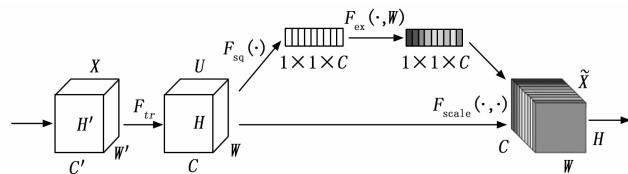


图 4 SE 模块结构

2) Excitation 操作, 将压缩得到的特征向量传到两个全连接层, 使用 ReLU 和 Sigmoid 激活函数对其计算, 生成每个特征通道的权重, 计算公式如式 (2) 所示:

$$s \equiv F_{\infty}(z, W) = \sigma^{\lceil} W_2 \delta(W_1 z) \rceil \quad (2)$$

式中, δ 是指 ReLU 激活函数, W_1 和 W_2 分别为两个全连接层的参数, σ 为 sigmoid 激活函数。

3) 进行 Scale 操作, 将各个通道的权重系数与原来的特征图相乘, 得到最终加入注意力机制后的特征图, 计算公式如式 (3) 所示:

$$\tilde{X} \equiv F_{\perp\perp}(s_\perp, y_\perp) \equiv s_\perp \times y_\perp \quad (3)$$

在网络中增加 SE 模块，可以获得特征图中各通道的重要程度，并由此为各通道赋予相应的权重，让网络更加关注这些通道，抑制对当前任务用处不大的特征，使最终的特征更符合预期要求，从而提高模型的检测性能。

2.2.2 小目标检测层

由于 YOLOv8n 的下采样倍数较大，小目标的特征信息很难由较深的特征图学习获得。YOLOv8n 的最小

检测尺度是 80×80 ，用于检测大小在 8×8 以上的目标，若原始图像中的目标高宽都小于 8 像素，模型就很难识别到其目标特征信息。为了提高 YOLOv8n 的小目标检测效果，本文在原始网络中增加一个小目标检测层，以增强模型对小目标的检测能力。

如图 3 所示，首先将主干特征提取网络中 80×80 尺度的特征层与颈部加强特征提取网络中的上采样层向上融合叠加，再经过 C2f 层和上采样处理。将获得含有小目标特征信息的深层特征层继续与主干网络的浅层位置特征层进行融合叠加，经过 C2f 层和 SE 模块后传至 Head 中新增的一个解耦头，最终得到 160×160 尺度的特征融合层。新增的解耦头检测的是分辨率高的低层次特征图，所以对小目标的检测更加准确，能够较好地解决小目标检测遗漏的问题。因此，在原网络中增加小目标检测层，提高了螺栓的检测范围，能够让网络更加准确的识别出较小的螺栓，有效的强化了网络对于接触网螺栓的检测能力。

2.2.3 SCDOWN 模块

SCDown 模块^[21]是 YOLOv10 中引入的轻量化特征提取单元，旨在提高计算效率并减少参数数量。如图 5 所示，该模块首先使用 1×1 卷积对输入张量进行降维，以减少通道数和计算量。随后，通过 3×3 深度卷积在每个通道内独立执行卷积操作，从而进一步提取特征，同时避免跨通道的计算开销。SCDown 在下采样过程中最大限度保留信息，并有效降低了推理延迟，适合实时检测应用。

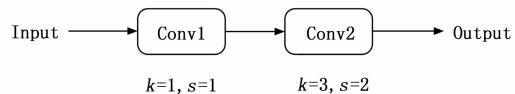


图 5 SCDown 模块

2.2.4 BiFPN 模块

BiFPN^[22]是一种改进的特征金字塔网络结构，相比传统的 FPN，BiFPN 引入了双向特征融合和权重分配机制，以增强特征图之间的信息流通。在 YOLOv8n 中的特征增强网络中引入该网络，能够加强接触网螺栓不同尺度特征融合的效率，从而提高网络目标检测的性能。BiFPN 机制如图 6 所示。

BiFPN 通过自底向上和自顶向下的双向路径对特征进行融合，确保高层特征和低层特征之间的信息可以双向流动，从而提升不同层次特征的表达能力。这种结构使得网络能够更好地捕捉不同尺度的物体。在多层特征的融合过程中，BiFPN 通过引入可学习权重，允许模型动态调整各层特征的重要性，取代了传统 FPN 中的简单平均加权方式。这一机制提高了特征融合的灵活性和准确性。此外，BiFPN 还简化了结构，删除了冗

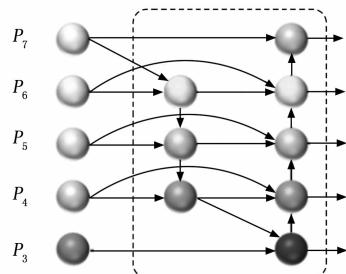


图 6 BiFPN 机制

余节点，从而在减少了计算量的同时保持了良好的特征表达能力。因此，该模块在保持模型性能的前提下实现了轻量化，非常适合实时检测任务。

2.2.5 损失函数 WIoU

针对 YOLOv8 的边界框回归损失函数 CIoU 会忽略掉高度和宽度与其置信度之间的差异以及计算开销大的问题，本文引入动态非单调聚焦机制 WIoU (Wise-IoU)^[23]，WIoU 采用了合理的梯度分配策略来对锚框进行评估，能够使其聚焦于普通的锚框，提升模型的检测性能。WIoU 有 3 个版本，其中 WIoU v3 是由离群度 β 构造的非单调聚焦系数 r 用在 WIoU v1 上得到的，离群度 β 用来描述锚框的质量， β 越小，锚框的质量越高。当离群度较大时，对质量较差的锚框分配较小的梯度增益，能够减少其产生较大的有害梯度。WIoU v3 采用合理的梯度增益分配策略，使模型能够聚焦于普通质量的锚框，提高模型的整体性能。WIoU v3 的计算公式如式 (4) ~ (6) 所示：

$$L_{\text{WIoU}_3} = r \cdot L_{\text{WIoU}_1} \quad (4)$$

$$r = \frac{\beta}{\delta \alpha^{\beta-\delta}} \quad (5)$$

$$\beta = \frac{L_{\text{IoU}}^*}{L_{\text{IoU}}} \in [0, +\infty) \quad (6)$$

式中， α 与 δ 是两个超参数，改变 α 与 δ 的值可以让模型获得不同的性能。

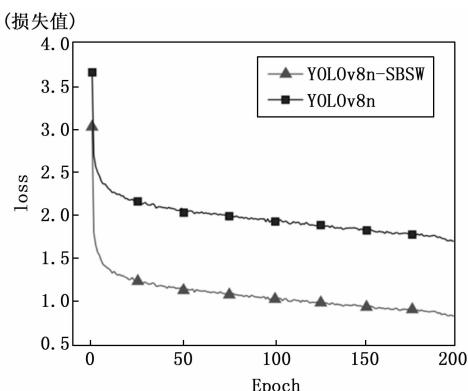


图 7 YOLOv8n 与 YOLOv8n-SBSW 损失曲线

练过程中的损失函数对比曲线，从图中可以看出，改进后的模型在训练过程中损失值更小，训练效果更加理想，验证了将边界框回归损失函数由 CIoU 替换为 WIoU 的有效性。

2.3 基于 RGB-D 相机的接触网螺栓定位方法

在使用目标检测网络 YOLOv8n-SBSW 对接触网螺栓进行识别后，将识别到的 ROI (Region of Interest) 提取出来作为输入图像，进行螺栓三维姿态的计算。螺栓定位方法的核心思路基于六角螺母的对称性特征。由于六角螺母为均匀对称体，其中心会经过任意一对中心对称的棱边所在的平面。通过识别这两条中心对称棱边，可以确定一个平面矩形，以此推算出螺母的中心位置。最后，通过结合 RGB-D 相机获取的相关深度信息，能够确定螺母中心的三维坐标及其所在矩形平面的空间位姿，从而实现对螺母的三维定位。

在对图像进行位姿计算之前，需要对图像进行预处理，以提高 ROI 特征的可读性，图 8 是图像预处理过程。首先对图像进行灰度处理，将彩色图像简化为单通道的灰度图像，为后续分析提供基础。接着通过非局部均值滤波来去除噪声，增强图像的清晰度，从而有效保留边缘和细节结构，使得关键特征在后续步骤中更加突出。在二值化阶段，将图像中符合阈值条件的像素值设置为黑白两色，以进一步增强前景与背景的对比，突出物体的轮廓。随后，通过 Canny 边缘检测来精确识别图像中的显著边缘，通过找到像素强度变化的边界，提取出物体的轮廓特征。最后，根据检测到的最左和最右的边缘线，拟合出螺母外接矩形，并计算矩形的角点和中心点，方便进一步的定位。在完成图像预处理后，即可基于上述定位方法开展后续的定位操作。

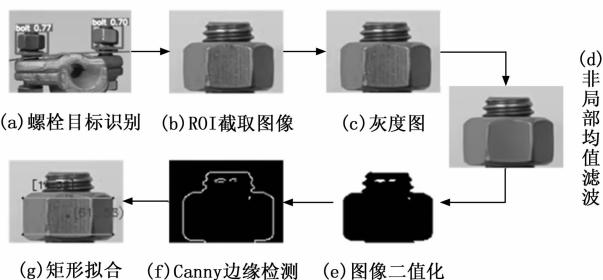


图 8 图像预处理

接触网螺栓的识别定位方法流程如图 9 所示，首先由 RGB-D 相机去采集接触网螺栓图像并获取其深度信息；然后通过深度学习网络模型去识别目标螺栓并截取出 ROI 区域，完成螺栓目标识别；在截取出的区域中进行图像预处理，再拟合出螺母的外接矩形，然后计算出该外接矩形的中心像素坐标和角点坐标，最后结合 RGB-D 相机获取的相关深度信息去计算螺母中心坐标。

图 7 是 YOLOv8n 与 YOLOv8n-SBSW 两个模型训

和三维姿态, 完成接触网螺栓的三维定位。

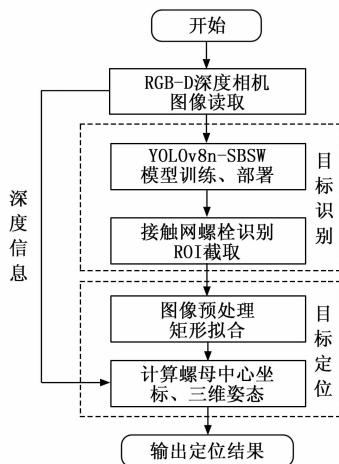


图 9 接触网螺栓识别定位方法流程图

3 实验结果与分析

3.1 数据集

本实验所使用的数据集均来源于拍摄的各类型接触网螺栓图像, 数据集中的部分代表性图像如图 10 所示。采集螺栓图像数据时, 选取不同角度进行拍摄, 共采集了 400 张接触网螺栓的图像。为了提高模型的泛化能力和鲁棒性, 本实验采用了添加噪声、图像翻转、模糊处理、亮度变换等方式进行数据增强, 最终共获得 1 922 张图像。使用标注工具 LabelImg 对图像中螺栓的位置进行标注, 将标签文件转换为训练所需要的 txt 格式进行训练。将构建好的螺栓数据集按照 8 : 2 的比例随机划分, 其中 1 537 张 (80%) 作为训练集, 385 张 (20%) 作为验证集。



图 10 数据集中代表性的图像

3.2 实验环境与训练参数

本实验用到实验平台的硬件配置为 Intel (R) Core (TM) i9-10980XE, 显卡为 NVIDIA GeForce RTX 3090 24 GB。在 win10 环境下基于 Pytorch 深度学习框架, python 版本为 3.8, pytorch 版本为 1.12.0。CUDA 版本和 cuDNN 均为 11.6。将划分好的训练集用到各模型上进行训练, 模型在训练过程中的一些相关参数设置如表 1 所示。

表 1 训练参数设置

参数	设置
Epochs	200
Initial learning rate	0.01
Final learning rate	0.01
Momentum	0.937
warmup_momentum	0.8
close_mosaic	10
Weight decay	0.000 5
Optimizer	SGD
imgsz	640
Batch size	8

3.3 实验结果

为了验证改进模型的性能, 本文选取精确率 P 、召回率 R 、平均精度均值 mAP 、模型参数数量、模型大小以及每秒帧数 FPS 作为模型的评价指标。

3.3.1 消融实验

为了验证本文提出的改进策略的有效性, 以 YOLOv8n 网络为基准模型, 依次在颈部网络中引入小目标检测层、SE 注意力机制、BiFPN 模块、SCDown 模块, 并将损失函数更换为 WIoU v3, 通过以上几种改进模块的不同组合进行了消融实验。实验结果如表 2 所示。

表 2 消融实验

小目标检测层	SE	BiFPN	WIoU UV3	SCDown	$P/\%$	$R/\%$	mAP	$mAP_{0.5} : 0.95$	参数量/ 10^6
							0.5 /%	/%	
					85.7	86.4	88.6	36.1	3.01
✓					85.9	87.3	89.0	36.3	2.92
✓	✓				86.5	87.4	89.7	36.5	2.93
✓	✓	✓			86.8	87.3	90.0	37.0	2.93
✓	✓	✓	✓		87.0	87.6	90.5	37.5	2.93
✓	✓	✓	✓	✓	87.1	88.1	90.7	37.7	2.77

从表中的数据可以看出, 在颈部网络添加了小目标检测层后, $mAP_{0.5}$ 、 $mAP_{0.5} : 0.95$ 比基准模型分别提高了 0.4%、0.2%, 说明了引入小目标检测层可以有效降低螺栓的检漏率; 添加了 SE 注意力机制后, P 、 R 、 $mAP_{0.5}$ 、 $mAP_{0.5} : 0.95$ 相对于基准模型分别提升了 0.8%、1.0%、1.1%、0.4%, 说明引入注意力机制可以让网络更加关注有效的特征信息; 引入 BiFPN 模块后, $mAP_{0.5}$ 、 $mAP_{0.5} : 0.95$ 比基准模型分别提高了 1.4%、0.9%, 说明双向特征融合机制能够有效提升网络检测性能; 将基准模型的边界框回归损失函数替换成 WIoU v3, 引入新的损失函数后, 模型的 $mAP_{0.5}$ 、 $mAP_{0.5} : 0.95$ 分别提升了 1.9%、1.4%; 最后将颈部网络中的部分 Conv 替换成 SCDown 模块

后, P 、 R 、 $mAP_{0.5}$ 、 $mAP_{0.5:0.95}$ 分别提升了 1.4%、1.7%、2.1%、1.6%，模型的参数量也减少了 8.0%，SCDown 模块的引入不仅有效的减少模型的参数量和大小，也提高了模型的检测性能。上述实验结果表明了本文所提出的算法改进策略的有效性，显著的提升了模型对接触网螺栓的检测能力。

3.3.2 对比实验

为了验证本文所提出算法的有效性和优越性，该实验将本文提出的 YOLOv8n-SBSW 模型与一些主流的目标检测模型进行了对比，实验结果如表 3 所示。

表 3 YOLOv8n-SBSW 算法与主流算法性能比较

模型	$P/\%$	$R/\%$	$mAP_{0.5}$	$mAP_{0.5:0.95}/\%$	参数量/ 10^6	模型大小/MB	FPS
Faster R-CNN ^[24]	79.6	81.5	82.9	31.5	28.39	109.7	79
YOLOv3-tiny	84.5	84.0	86.0	32.2	8.67	17.4	126
YOLOv5s	87.5	87.2	90.0	36.6	7.02	14.0	153
YOLOv7	82.4	83.3	84.2	32.5	36.48	73.03	141
YOLOv8n	85.7	86.4	88.6	36.1	3.01	6.2	158
YOLOv8n-SBSW	87.1	88.1	90.7	37.7	2.77	5.9	144
YOLOv10n	84.6	86.1	88.2	34.9	2.69	5.8	171

从表 3 可以看出，YOLOv8n-SBSW 模型的大小和参数量在这几个网络中相对较小， $mAP_{0.5}$ 和 $mAP_{0.5:0.95}$ 最高。尽管 YOLOv5s 的 Precision 表现略高，但是该模型的大小和参数量是 YOLOv8n-SBSW 的近三倍。与 YOLO 系列较新的 YOLOv10n 相比，本文所提出的模型在大小和参数量与其几乎接近的情况下，模型的检测性能均更为优越。除此之外，YOLOv8n-SBSW 的表现也要优于 Faster R-CNN。该对比实验结果表明，本文所改进后的模型与基准模型 YOLOv8n 相比，虽然 FPS 有略微下降，但是 YOLOv8n-SBSW 在减少模型大小和参数量的同时，显著地提升了检测性能，证明了本文所提算法的优越性。

3.3.3 目标识别可视化分析

为了更直观方便地展现模型改进后的优越性，描绘出其检测效果，本文进行了对比实验，从模型检测结果和热力图两方面分析了模型的检测性能。

为了直观地展示所改进模型的检测效果，本文选取了两张代表性的接触网螺栓图像作为实验对象，利用原始的 YOLOv8n 模型和改进后的 YOLOv8n-SBSW 模型对其进行检测，对比实验结果如图 11 所示。

从该对比检测效果可以看出，YOLOv8n-SBSW 具有更高的检测精度，改进后的模型在多目标检测中能够有效避免漏检，进一步说明了本文所提出改进算法的有效性。

图 12 是通过 Grad-CAM^[25] 对检测图像处理后得到

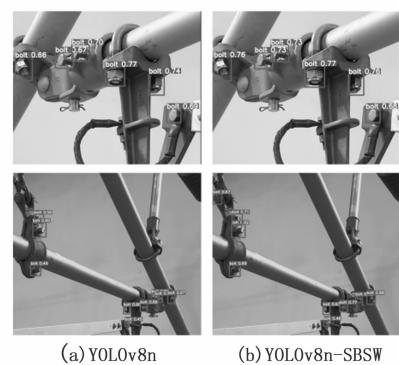


图 11 YOLOv8n 与 YOLOv8n-SBSW 对比检测结果

的热力图，热力图能够直观的展示出网络学习到的特征，容易反映出网络关注特征图的那个区域。Grad-CAM 通过反向传播对模型进行梯度计算，梯度较高的在热力图中用较亮的颜色表示，代表了网络更关注的区域。从图 12 可以看出，改进后的 YOLOv8n-SBSW 网络在热力图中能够更加准确地关注螺栓，比 YOLOv8n 的热力图更加集中。因此，该实验结果也再次表明本文所改进的模型对接触网螺栓的预测更加精准，模型的检测性能得到了有效的提升。

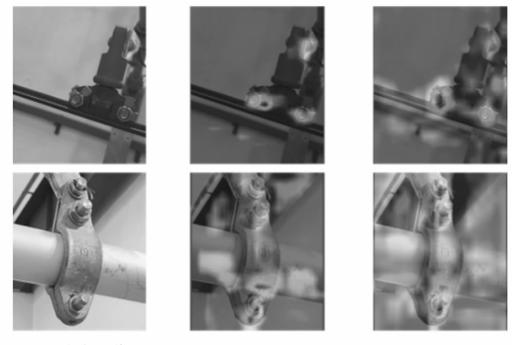


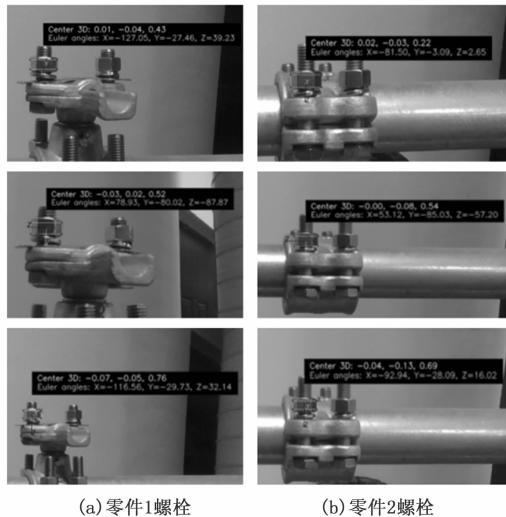
图 12 YOLOv8n 与 YOLOv8n-SBSW 热力图对比

3.3.4 螺栓定位实验与结果

为了验证所提螺栓定位方法的准确性和可靠性，本文通过对不同零部件螺栓进行定位测试，并将相机位于不同位置下进行采集。首先，使用 RGB-D 相机采集了两个零部件螺栓位于不同相机位置的图像数据，并通过训练好的 YOLOv8n-SBSW 模型对采集的 RGB 图像中的螺栓目标进行识别。对于每个识别出的螺栓，系统将裁剪出感兴趣区域 (ROI)，并结合前述的定位方法，计算出螺栓的三维坐标和姿态。

图 13 是两种接触网螺栓在不同采集位置下的定位结果，从图中可以看出该定位方法能够将图像中螺栓的位姿较为准确的描绘出来，获取到了螺母中心具体的三维坐标和空间姿态。其中，“center 3D” 表示螺母中心相对于相机坐标系的三维坐标，“Euler angles” 表示螺

母中心所在矩形平面相对于相机坐标系的偏转角。



(a) 零件1螺栓 (b) 零件2螺栓

图 13 螺栓在不同采集位置下的定位结果

本文同时还对螺母的三维中心坐标进行了实际测量，并通过对比实际测量值与计算结果，确认其误差处于可接受范围内，从而验证了该方法在定位方面的准确性。尽管螺母的空间姿态涉及角度测量，难以直接进行标准化操作，但由于其空间姿态是基于外接矩形所在平面的法向量进行计算的，因此可以通过测量矩形 4 个角点的三维坐标间接推算姿态误差。基于这一方法，通过测量对比后发现姿态的误差也是保持在合理范围内，进一步验证了该方法在姿态估计方面的可靠性。

4 结束语

1) 提出了一种基于 YOLOv8n 的接触网螺栓识别算法 YOLOv8n-SBSW。在网络中添加了小目标检测层，将浅层特征和深层特征进行融合，提升模型检测小目标的精度；在网络中加入 SE 注意力机制模块，选择性地强调信息特征并抑制不太有用的特征；在网络中引入 BiFPN 模块，增强特征图之间的信息流通，加强接触网螺栓不同尺度特征融合的效率；将颈部网路中的部分卷积模块替换为轻量级的 SCDown 卷积模块，在保证模型检测性能的同时减少了模型的大小和参数量；使用 WIoU Loss 替换 CIoU Loss 作为新的边界框回归损失函数，加快了回归损失函数的收敛速度，提高了模型的回归精度。

2) 将改进后的算法与 YOLOv8n 在自制数据集上进行了对比实验，结果表明本文所提出的 YOLOv8n-SBSW 算法显著优于原始的 YOLOv8n, P 、 R 、 $mAP0.5$ 、 $mAP0.5:0.95$ 分别提高了 1.4%，1.7%，2.1%，1.6%，模型的参数量减少 8.0%。与其它主流的目标检测算法相比，本文所改进的模型在检测精度、速度、模型大小上也更具优势，更适用于边缘设备的部署。改

进后的模型能够有效地完成接触网中螺栓的识别任务，并且对接触网其他零部件的检测具有一定的参考作用。

3) 将改进后的 YOLOv8n-SBSW 与 RGB-D 深度相机进行结合对接触网螺栓进行定位，所提出的定位方法能够有效地获取螺母中心的三维坐标和姿态，实验结果验证了该方法的可行性，为引导机器人末端进行定位和紧固作业提供了相关参考。下一步将丰富数据集，并在该模型的基础上继续研究改进网络结构，提升模型的检测精度和速度。

参考文献：

- [1] 卞建鹏, 薛秀茹, 崔跃华, 等. 基于 EfficientDet 与 Vision Transformer 的接触网吊弦故障检测 [J]. 铁道科学与工程学报, 2023, 20 (6): 2340–2349.
- [2] LIN S, SHANG C, LI N, et al. An optimization method for maintenance resource allocation in electrified railway catenary systems [J]. IEEE Transactions on Industry Applications, 2023, 59 (1): 641–651.
- [3] 蒋欣兰, 贾文博. 高铁接触网异物侵入的机器视觉检测方法 [J]. 计算机工程与应用, 2019, 55 (22): 250–257.
- [4] 程敦诚, 王倩, 吴福庆, 等. 基于深度学习的接触网顶紧螺栓状态智能检测 [J]. 铁道学报, 2021, 43 (11): 52–60.
- [5] 铁路局. 国家铁路局关于印发《“十四五”铁路科技创新规划》的通知 [EB/OL]. (2024-12-13) [2021-12-14]. https://www.gov.cn/zhengce/zhengceku/2021-12/24/content_5664357.htm.
- [6] 米增, 连哲. 面向通用目标检测的 YOLO 方法研究综述 [J]. 计算机工程与应用, 2024, 60 (21): 38–54.
- [7] ZHAO K, WANG Y, ZUO Y, et al. Palletizing robot positioning bolt detection based on improved YOLO-V3 [J]. Journal of Intelligent & Robotic Systems, 2022, 104 (3): 41.
- [8] WANG D, ZHANG M, SHENG D, et al. Bolt positioning detection based on improved YOLOv5 for bridge structural health monitoring [J]. Sensors, 2022, 23 (1): 396.
- [9] 戴先鑫, 付振山, 马栋, 等. 基于感兴趣区域的螺栓位姿及尺寸检测研究 [J]. 电子测量技术, 2024, 47 (8): 134–140.
- [10] JIANG H, YUAN B, WANG P, et al. A bolt positioning method using depth map and color image fusion from a depth camera [C] // 2023 7th International Conference on Electrical, Mechanical and Computer Engineering (ICEMCE). IEEE, 2023: 58–63.
- [11] 王向周, 杨敏巍, 郑成华, 等. 基于 YOLOv5s-T 和 RGB-D 相机的螺栓检测与定位系统 [J]. 北京理工大学学报自然版, 2022, 42 (11): 1159–1166.

(下转第 341 页)