光电技术及应用

DOI：10.16818/j.issn1001－5868.20251204002

基于轻量化AWD-YOLOv8s的小目标螺栓检测算法

程　博[1,2]，王正家[1,2*]，陈安澜[1,2]，谷　峰[1,2]

（1. 湖北工业大学 机械工程学院，武汉 430068；2. 现代制造质量工程湖北省重点实验室，武汉 430068）

摘　要： 针对地铁车辆车顶图像中螺栓占比小，算法检测精度低和计算量大的问题，提出一种基于轻量化AWD-YOLOv8s的小目标螺栓检测算法。首先，设计了一种自适应权重下采样AWD模块，通过赋予局部特征信息不同的权重值，从而增强算法对小目标特征信息的提取能力；其次，引入尺度内特征交互（AIFI）模块替换SPPF模块，通过捕捉同一尺度特征间的依赖关系，实现算法的轻量化改进；最后，采用轻量化Ghost-P2结构，增强算法对小目标螺栓特征信息的提取能力。实验结果表明，改进后的算法平均精度均值达到98.7%，比原算法提高2.2%，模型参数量减少46.7%，检测速度达到115 FPS，在实现模型轻量化的同时显著提高检测精度。

关键词： 车辆车顶螺栓；轻量化YOLOv8s；小目标；AWD；AIFI；Ghost-P2

中图分类号： TP391.4　**文章编号：** 1001－5868(2026)01－0203－07

Lightweight AWD-YOLOv8s-Based Algorithm for Detecting Small Target Bolts

CHENG Bo[1,2], WANG Zhengjia[1,2], CHENG Anlan[1,2], GU Feng[1,2]

(1. School of Mechanical Engineering, Hubei University of Technology, Wuhan 430068, CHN;
2. Hubei Key Laboratory of Modern Manufacturing Quality Engineering, Wuhan 430068, CHN)

Abstract: In this study, a small-target-bolt detection algorithm based on lightweight AWD-YOLOv8s is proposed to overcome the limitations of low detection accuracy and extensive computation in the analysis of a small proportion of bolts in subway-vehicle-roof images. First, an adaptive weight downsampling module is designed to enhance the ability of the algorithm to extract small-target feature information by assigning different weight values to local feature information. Second, an intrascale feature interaction module is introduced to replace the SPPF module. A lightweight algorithm with enhanced performance is obtained, which captures the dependency between same scale features. Finally, the lightweight Ghost-P2 structure is introduced to enhance the ability of the algorithm to extract the feature information of small target bolts. The experimental results show that the mean Average Percision of the improved algorithm has reached 98.7%, which is 2.2% higher than that of the original algorithm. In addition, the number of model parameters is reduced by 46.7%, and a detection speed of 115FPS is achieved. Notably, the detection accuracy can be further improved while maintaining the "lightweight" feature of the model.

Key words: vehicle roof bolts; lightweight YOLOv8s; small targets; AWD; AIFI; Ghost-P2

0　引言

地铁车顶作为车体结构的重要组成部分，其大量的零部件是通过螺栓进行连接装配[1]。地铁车辆在长期运行中，由于振动及应力疲劳，螺栓可能出现松动的异常状态[2]。为了避免螺栓松动造成的安

收稿日期：2025－12－04.

*通信作者：王正家

全事故，地铁车辆在运行中需要频繁地对地铁车辆车顶进行人工巡检，这种检修方法存在效率低、人力成本高、漏检误检严重等问题。因此，及时检测出螺栓的异常状态对于保证车辆运行安全具有重要的研究价值[3]。随着人工智能和图像处理等技术的不断演进，许多研究致力于将这些创新技术与螺栓的异常检测相结合，以提升故障识别和监测的水平。对于自动化检测螺栓的异常状态，检测定位螺栓的具体位置是其中重要的一环。目前已有许多研究者进行了大量的研究。赵江平等[4]提出一种基于改进FasterR-CNN的铁路客车螺栓检测方法，构建了基于ResNet-50和Inception网络的骨干网络，并引入K-means^{++}聚类算法，提高了算法对小尺度螺栓检测的准确率。罗隆福等[5]提出一种基于深度学习的顶紧螺栓缺陷检测方法，通过在TDM模块中引入SSD算法，增强不同特征层之间的联系，从而提升算法对顶紧螺栓的检测精度。张洪等[6]提出一种基于改进YOLOv5s的螺栓缺陷识别方法，其通过融合GAM模块和ECA模块增强对螺栓特征的提取能力，优化了空间金字塔池化结构，同时使用MPDIoU提高边界框的回归精度。董华军等[7]提出一种基于C3F-YOLOv5的列车车底螺栓检测方法，通过改进C3模块以及引入基于LAMP分数的剪枝算法对模型进行轻量化改进，提高了模型对车底螺栓区域定位的精度与识别速度。王健等[8]提出一种基于改进YOLOv8s的PCB小目标缺陷检测算法，其通过结合不同层次的特征图信息与添加空间信息增强模块，提升模型对小目标的检测精度。

在地铁车辆车顶各子区域图像中，出风口图像像素大小为1 600×1 025、盖板图像像素大小为2 369×1 473、风机图像像素大小为1 600×385，而螺栓在各子区域图像中的像素小于30×30，因此在地铁车顶图像中螺栓像素占比小。这使得上述检测方法在应对该问题时，难以充分获取准确的目标位置信息和丰富的特征图语义信息，导致算法的检测精度低、计算量大，难以部署在工业化检测的设备上。基于此，本文提出一种基于轻量化改进的YOLOv8s算法。首先，依据RFCAConv的结构，设计了一种自适应权重下采样（Adaptive Weighted Downsampling，AWD）模块，通过赋予特征图中局部特征信息不同的权重值，增强算法对目标特征信息的获取能力，使得算力的分配更加合理。其次，引入尺度内特征交互（Adaptive Information Fusion Integration，AIFI）模块，通过有效捕捉同一尺度特征间的依赖关系，提升模型对细粒度信息的敏感度，减少模型在提取特征时算力的浪费。最后，引入轻量化Ghost-P2结构，在提高模型的小目标边缘细节信息抓取能力的同时，实现了模型轻量化。实验结果表明，改进后的算法不仅保证了检测精度，还实现了模型轻量化。

1 YOLOv8算法

YOLOv8是一阶段检测算法的代表之一，其结构如图1所示。

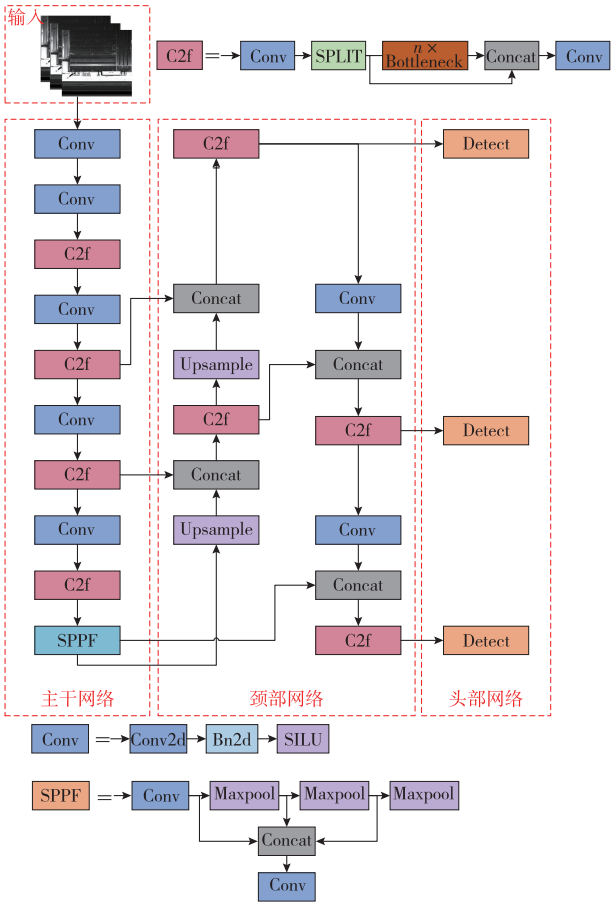


图1 原始YOLOv8s网络结构

主干网络采用CSPNet网络，通过分段和部分连接的方式，在减小计算量的同时保留了丰富的特征信息。颈部网络通过结合特征金字塔网络（FPNet）和路径聚合网络（PANet）实现对不同尺度特征的融合，显著提升了模型对不同尺寸目标的检测能力。检测头部通过多尺度设计和自适应锚框机制，有效提升了不同尺寸目标的检测能力，同时结合深度卷积层和全面的损失计算，确保在各种应

用场景中的高效和准确。此外，通过引入注意力机制，使模型在特征提取时能够更关注图像中的关键区域，增强在复杂场景中识别目标的能力，尤其是增强对小目标的识别能力。

为满足不同目标检测需求，YOLOv8提供了5个不同尺寸大小的模型版本。在工业检测部署中，大尺寸模型的性能受限。最小尺寸的YOLOv8n模型具有一定的能力，但在准确率上还有待提高。相比之下，YOLOv8s在相似参数量的情况下，提供了更高的检测准确率，并且在处理复杂任务时显示出更好的稳定性。因此，在检测地铁车顶小目标螺栓的任务中，YOLOv8s表现优秀，具有良好的通用性和适应性。因此，本研究将YOLOv8s作为基线算法。

2　改进YOLOv8模型的螺栓定位检测方法

2.1　AWD

当卷积神经网络进行卷积操作时，标准卷积核会通过共享的参数来提取信息，这会导致网络对不同位置的差分信息不敏感，使得网络的性能受限。RFCAConv[9]通过强调感受野滑块中不同特征的重要性，改善了因卷积核共享参数导致的网络性能受限的问题。然而，该模块中不同的感受野特征会进行彼此的数据交互，这会导致额外的计算开销。因此，该模块的设计者利用平均池化操作（AvgPool）来聚合每个接受野特征的全局信息，再使用1×1组卷积（Group-conv）运算来交互信息。最后，通过使用Softmax来强调感受野特征中不同特征的重要性，从而降低卷积的计算量。

为实现轻量化改进算法的同时保证算法的检测精度，本文参照RFCAConv模块的设计思路，设计了AWD模块来替换YOLOv8主干网络中的部分卷积，其结构如图2所示。首先，对输入特征图进行AvgPool以减少特征图中的不必要细节，再通过1×1的卷积提取局部特征信息。其次，使用两次Softmax注意力机制分别从通道和空间维度对特征进行权重分配，以增强后续操作中目标物特征信息的表达。然后，将这些特征权重与经过Group-conv处理的多通道特征图相乘（Mul），进一步调节特征的重要性，使得目标物的特征信息能够在模型的学习过程中占据更大权重，从而减少算力的浪费。这里的Group-conv操作是参照Focus的结构而设计的，其通过将特征图划分为多个小部分并进行空间信息的重新组合，保留了特征图中不同位置的重要局部信息，这些信息被重新分配到不同的通道中，从而在减少空间信息的同时保持了特征表达的丰富性。Group-conv通过分组卷积操作有效减少了常规Focus的切片操作时间，提升了模型的效率。最后，通过对不同特征的加权求和（Sum）操作，输出最终的特征图。

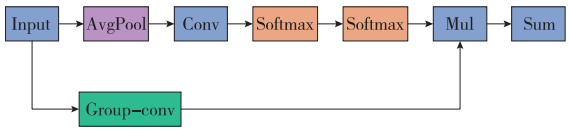


图2　AWD模块结构图

2.2　尺度内特征交互模块

AIFI[10]模块是RT-DETR模型主干网络结构中用于增强网络聚集能力的模块，其通过引入多头自注意力机制（Multi-Head Self-Attention，MSA），实现内部尺度特征的交互，以此提高特征提取的效率和有效性。该模块的核心是允许同尺度的局部特征之间进行通信，同时通过使用注意力机制捕捉并强化这些特征之间的依赖性，进而提高模型在复杂的场景中识别小目标的能力。相较于传统的特征融合所涉及的不同尺度特征，AIFI则重点放在同一尺度内部的特征融合上，这样有助于模型捕获更细粒的图像信息。SPPF模块是YOLOv8的特征提取网络层的最后一层，其通过对不同尺度特征的最大池化处理，来捕获上下文信息，从而提高网络对多尺度信息的提取。但由于SPPF的结构较为简单，在特征提取过程中无法充分利用复杂的空间信息，导致其对重要的细节特征捕捉不足，为后续的处理增加负担，使得网络的整体处理效率受限。由于AIFI模块能够根据实时输入信号的特征动态调整参数，相比SPPF模块其不仅在复杂背景下展现出更高的特征提取能力，还能使网络以较少的计算量完成对复杂数据的处理，因此本研究将SPPF模块替换成AIFI模块。为进一步提高模型对局部信息的捕捉能力，在AIFI之前额外添加一层Conv。通过结合使用AIFI和Conv，有助于捕获特征之间的依赖关系，使得模型能够更有效地处理和融合重要的特征信息，从而提高算法的整体检测性能。AIFI结构如图3所示。

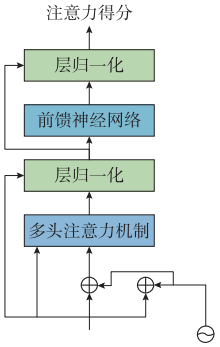


图3 AIFI结构图

2.3 轻量化Ghost-P2结构

在卷积神经网络中，浅层的特征感受野较小，更加关注图像的细节信息，而相较于P3级的特征图，P2级大小的特征图所能包含的细节信息更多，也更适合对于小目标的检测。通过引入P2级特征检测层可增强模型的对细节边缘信息的抓取能力，从而提升对小目标螺栓的检测能力。但由于P2级特征检测层的特征图是高分辨率的，直接引入会增加模型的计算量。因此，为减少引入P2级特征检测层所带来的计算增量，对颈部网络进行轻量级搭建，提出一种基于GhostNet[11]中Ghost模块的Ghost-P2结构。

2.3.1 GhostConv

GhostConv模块通过将常规卷积操作分成两步进行，以减少重复生成特征的冗余计算，并保持良好的特征提取能力。首先，使用标准卷积操作生成部分特征图；接下来，引入一种“廉价特征生成”机制，使用简单的操作（如逐点卷积、线性变换）来生成额外的特征；最后，通过逐点加法将特征图组合在一起输出完整的特征图。通过该方法使得GhostConv模块能够输出与传统的卷积相当尺寸与质量的特征图，在不降低模型性能的前提下大幅度减少模型的计算量。因此本文将GhostConv模块替换颈部网络中的所有Conv模块，如图4所示。

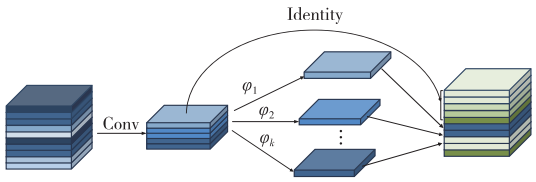


图4 GhostConv模块结构图

2.3.2 C3Ghost

C3模块通过跨阶段部分连接技术有效减少了重复的梯度信息，提升了网络的效率和性能，使其在计算量较低的情况下仍能保持较高的特征提取能力，适用于轻量化的目标检测任务。C3Ghost不仅包含C3模块的优点，还通过结合Ghost模块来减少传统卷积中的冗余计算，使得模型能够以较少的计算量生成高效的特征图，从而让C3Ghost模块在最终融合阶段获得更丰富的特征信息。因此，通过引入C3Ghost模块替换掉原有的C2f模块，可在降低模型计算量的同时，提升模型对小目标螺栓特征的提取能力，其结构如图5所示。

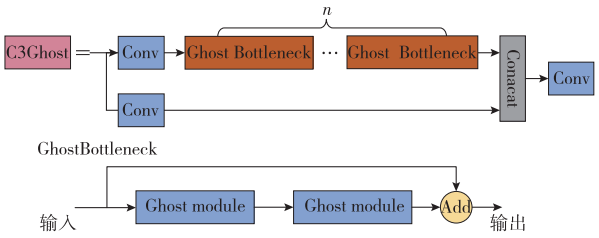


图5 C3Ghost模块结构图

最后在颈部网络中添加P2级特征检测层的同时分别引入GhostConv和C3Ghost模块，从而实现对模型的轻量化改进。改进后的AWD-YOLOv8s结构如图6所示。

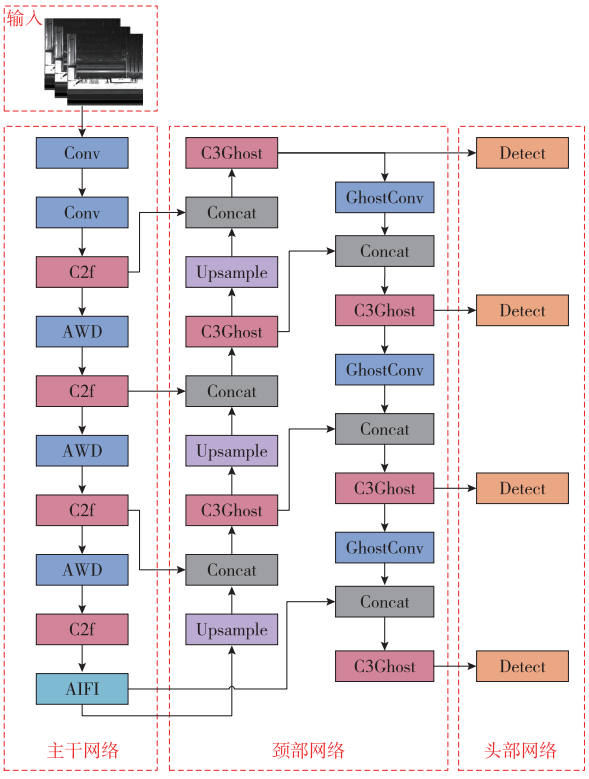


图6 AWD-YOLOv8s模型结构图

3 实验结果与分析

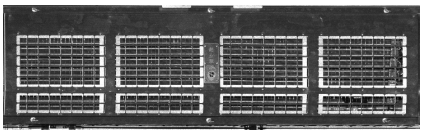
3.1 建立数据集

使用固定在地铁车辆车顶上方的线阵相机采

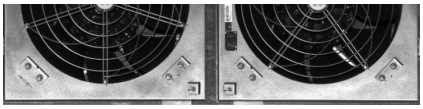
集车顶图像，并针对车顶的三个子区域：盖板、出风口和风机，建立螺栓图像数据集，共计 2 400 张图片，将数据集按 7∶1∶2 的比例分为训练集、验证集、测试集。如图 7 所示为采集到的车顶盖板、出风口和风机区域图像。



（a）车顶盖板



（b）出风口



（c）风机

图 7　车顶三处子区域图像

3.2　实验环境及评判指标

实验的软硬件环境：操作系统 Windows10，深度学习框架 Pytorch1.9，Python3.9，CUDA11.3，CPU 为 Intel i7-12900k，内存 16 G；GPU 为单块 NVIDIA GeForceRTX3080，显存 12 G。将改进的 AWD-YOLOv8s 进行模型训练，输入图像的大小为 640×640，训练时采用 SGD 优化器，学习率为 0.001，Batchsize 设为 16，epoch 设为 200。

为评价算法性能，选用每秒检测帧数 *FPS* 为模型速度衡量指标，精度率（*P*）、召回率（*R*）、平均精度均值（*mAP*）作为模型精度衡量指标。其中精确率越高表示误检率越低，召回率越高表示漏检率越低。

$$AP=\int_0^1 P(R)\mathrm{d}R \tag{1}$$

$$mAP=\frac{\sum_{i=0}^{n}AP(i)}{n} \tag{2}$$

$$P=\frac{TP}{TP+FP} \tag{3}$$

$$R=\frac{TP}{TP+FN} \tag{4}$$

式中，*TP* 表示模型正确预测的正样本数，*FP* 表示模型错误预测为正样本但实际为负样本，*FN* 表示模型错误预测为负样本但实际为正样本，*n* 表示总类别数，*i* 表示第 *i* 个区域的螺栓。

3.3　消融实验结果与分析

为了验证本文改进的有效性，在自制数据集上进行消融试验。实验在 YOLOv8s 模型的基础上首先分别单独进行 AWD 模块的添加、引入 AIFI 模块替换 SPPF 模块、对颈部网络引入轻量化 Ghost-P2 结构，其次两两组合改进，最后三种改进共同组合添加到 YOLOv8s 模型中，共设立 7 组实验，实验结果如表 1 所示。图 8 为改进后的 YOLOv8s 算法（AWD-YOLOv8s）与原始 YOLOv8s 算法在训练过程中的平均精度均值变化对比图。从图中可知，AWD-YOLOv8s 算法与原始 YOLOv8 算法的变化曲线在前 50 个 epoch 都有震荡，但在趋于平稳后，AWD-YOLOv8s 算法的平均精度均值要高于原始 YOLOv8s 算法。

表 1　消融实验

AWD	AIFI	Ghost-P2	*P*/%	*R*/%	*mAP*/%	*Parameters*	*FPS*
—	—	—	95.7	96.9	96.5	11 126 745	160
√	—	—	96.3	98.3	98.2	9 751 833	152
—	√	—	95.9	96.1	96.0	6 526 489	130
—	—	√	97.3	96.2	97.1	5 905 644	125
√	—	√	98.2	98.2	98.3	6 786 468	106
√	√	—	97.3	96.8	97.6	7 884 612	130
√	√	√	97.9	98.6	98.7	5 936 036	115

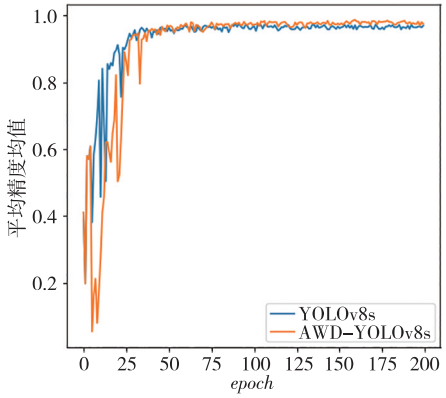


图 8　平均精度均值的变化曲线

从表 1 可知，原始 YOLOv8s 算法的精确率为 95.7%，召回率为 96.9%，*mAP* 为 96.5%，整个模型的计算参数为 11 126 745，速度为 115 FPS。在添加了 AWD 模块后，模型的精度率提升了 0.6%，召回率提升了 1.4%，*mAP* 提升了 1.7%，且模型参数量减少了 12.3%，这说明该模块能够对特征信息权重进行自适应赋值，增强模型对重要特征信息的提取，使模型对算力的分配更加集中，从而提升模型

检测精度的同时降低计算量。在单独引入AIFI模块后，模型的召回率与*mAP*略有下降，但模型参数量减少了41.3%，这说明该模块能够通过对实时输入信号的特征动态调整参数，从而以较低的计算成本完成对目标特征的提取，实现模型的轻量化改进。在单独引入轻量化Ghost-P2结构后，模型的*mAP*为97.1%，模型参数量减少了46.9%，这说明在添加进P2检测层的同时引入基于Ghost模块改进的Ghostconv与C3Ghost模块，不仅提高了模型检测小目标螺栓的检测精度，还有效降低了由于引入P2检测层所带来的计算增量。在将两种不同改进方法组合引入模型后，模型的*mAP*均有提升，且模型参数均比原始模型低。最终，将三种改进共同引入模型，与原模型相比改进后的模型精确率、召回率、*mAP*分别提高了2.2%，1.7%，2.2%，模型参数减少了46.7%，检测速度为115 FPS，提高模型检测精度的同时减少计算量，实现了对模型的轻量化改进。

3.4 对比实验

为进一步验证所提算法的有效性，设置了与SSD[12]，Faster R-CNN[13]，YOLOv3[14]和YOLOv5-tiny目标检测算法的对比实验。在相同环境下进行训练，对比实验结果如表2所示。

表2 不同检测模型的对比实验

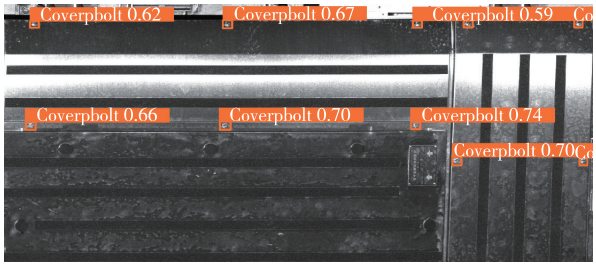
模型	*mAP*/%	*Parameters*/$\times 10^6$	*FPS*
SSD	87.42	25.67	32.6
Faster R-CNN	88.54	133.36	15.1
YOLOv3	90.65	103.75	100.8
YOLOv5-tiny	91.31	5.84	176.6
YOLOv8s	96.50	11.12	160
AWD-YOLOv8s	98.70	5.93	115

实验结果表明，AWD-YOLOv8s的*mAP*均超过其他算法，参数量相较原算法减少了46.7%，检测速度也达到115 FPS，仅比YOLOv5-tiny算法低，且能够满足实时检测需求。与原算法相比，改进后的算法虽然检测速度降低，但实现了对算法的轻量化改进，同时提高了检测精度。这进一步证明在地铁车辆车顶螺栓定位检测中，改进后的算法能够获得良好的检测结果。相比其他目标检测算法，本文算法在检测准确性与效率方面存在一定的优势。

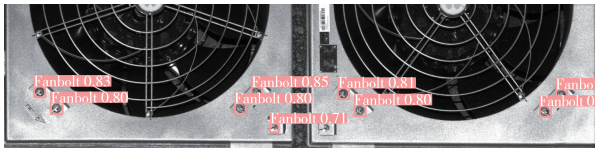
为更直观地验证AWD-YOLOv8s算法的有效性，将两种算法的螺栓定位可视化结果进行对比，对比结果如图9所示，其中(a)，(c)和(e)分别表示原YOLOv8s算法对盖板、出风口和风机的螺栓定位检测结果，而(b)，(d)和(f)分别表示改进后的AWD-YOLOv8s算法对各子区域的检测结果。从图中发现，原始算法进行对三处位置的螺栓进行检测时会存在一定量的漏检情况，而AWD-YOLOv8s能够有效检出螺栓。



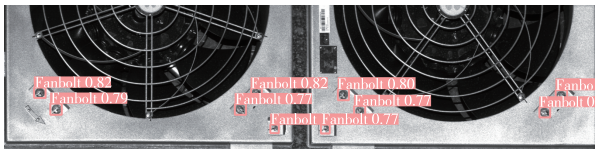
(a) YOLOv8s盖板检测结果



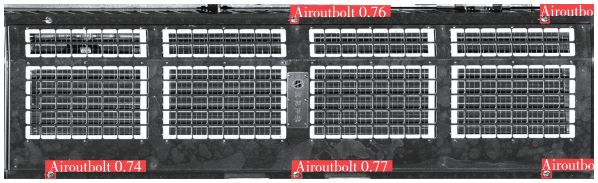
(b) AWD-YOLOv8s盖板检测结果



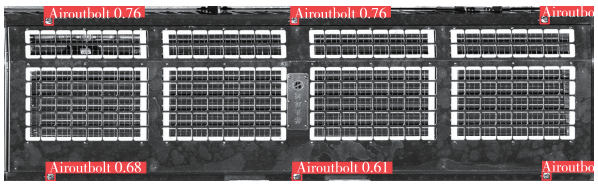
(c) YOLOv8s风机检测结果



(d) AWD-YOLOv8s风机检测结果



(e) YOLOv8s出风口检测结果



(f) AWD-YOLOv8s出风口检测结果

图9 YOLOv8s模型改进前后检测效果对比图

4　结论

本文针对地铁车辆车顶螺栓在车顶图像中占比小、难以提取目标的特征信息，导致检测精度低、算法计算量大的问题，设计了一种基于轻量化AWD-YOLOv8s的小目标螺栓检测算法。改进后的算法有效增强了对小目标的检测精度，与原始YOLOv8s相比，改进后的模型的*mAP*提高了2.2%，模型参数量减少了46.7%，检测速度可达115 FPS。该算法在实现模型轻量化同时还兼顾提升了检测精度，取得了检测精度和速度的平衡，具有一定的工程实用价值。

参考文献：

[1] 陈　旭. 地铁列车运行振动问题整治方案优化与应用[D]. 南京：东南大学，2021.
Chen Xu. Optimization and application of renovation scheme for metro train running vibration [D]. Nanjing: Southeast University, 2021.

[2] 魏云平，高崇一，魏明楠，等. 动车用连接紧固件组结构失效分析[J]. 机械设计与研究，2023，39(1)：130-134.
Wei Yunping, Gao Chongyi, Wei Mingnan, et al. Structural failure analysis of threaded fastener group in CRH train [J]. Machine Design & Research, 2023, 39(1): 130-134.

[3] 王正家，谷　峰，曾　臻. 针对低像质轨道车辆车顶螺栓松动检测方法研究[J]. 机械设计与研究，2024，40(2)：220-224.
Wang Zhengjia, Gu Feng, Zeng Zhen. Research on the detection method of loose roof bolts of low-quality rail vehicles [J]. Machine Design & Research, 2024, 40(2): 220-224.

[4] 赵江平，徐　恒，党悦悦. 基于改进Faster R-CNN的铁路客车螺栓检测研究[J]. 中国安全科学学报，2021，31(7)：82-89.
Zhao Jiangping, Xu Heng, Dang Yueyue. Research on bolt detection of railway passenger cars based on improved Faster R-CNN[J]. China Safety Science Journal, 2021, 31(7): 82-89.

[5] 罗隆福，叶　威，王　健. 基于深度学习的高铁接触网顶紧螺栓的缺陷检测[J]. 铁道科学与工程学报，2021，18(3)：605-614.
Luo Longfu, Ye Wei, Wang Jian. Defect detection of the puller bolt in high-speed railway catenary based on deep learning [J]. Journal of Railway Science and Engineering, 2021, 18(3): 605-614.

[6] 张　洪，朱志伟，胡天宇，等. 基于改进YOLOv5s的桥梁螺栓缺陷识别方法[J]. 吉林大学学报(工学版)，2024，54(3)：749-760.
Zhang Hong, Zhu Zhiwei, Hu Tianyu, et al. Bridge bolt defect identification method based on improved YOLOv5s[J]. Journal of Jilin University (Engineering and Technology Edition), 2024, 54(3): 749-760.

[7] 董华军，韩华豫，李籽骁，等. 基于C3F-YOLOv5的轻量化列车车底螺栓检测方法研究[J]. 铁道科学与工程学报，2024，21(8)：3455-3468.
Dong Huajun, Han Huayu, Li Zixiao, et al. Research on underbody bolt detection method of lightweight train based on C3F-YOLOv5 [J]. Journal of Railway Science and Engineering, 2024, 21(8): 3455-3468.

[8] 王　健，肖　迪，冯李航，等. 基于改进YOLOv8s的PCB小目标缺陷检测模型[J]. 计算机工程与应用，2025，61(15)：288-297.
Wang Jian, Xiao Di, Feng Lihang, et al. PCB small object defect detection model based on improved YOLOv8s [J]. Computer Engineering and Applications, 2025, 61(15): 288-297.

[9] Zhang X, Liu C, Yang D, et al. RFAConv: innovating spatial attention and standard convolutional operation [EB/OL]. 2023: arXiv: 2304.03198.

[10] Zhao Y, Lv W, Xu S, et al. DETRs beat YOLOs on real-time object detection [C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024: 16965-16974.

[11] Han K, Wang Y, Tian Q, et al. GhostNet: More Features from Cheap Operations[C]// Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 1577-1586,

[12] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]// Computer Vision – ECCV, 2016: 21-37.

[13] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[14] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. 2018: arXiv: 1804.02767.

作者简介：

程　博(1999—)，男，硕士研究生，研究方向为目标检测、图像处理；

王正家(1970—)，男，博士，讲师，研究方向为物联网与大数据、机器视觉。