

TrackDef-YOLO: 一种改进的YOLOv11模型 用于铁轨表面缺陷检测

李梦超, 文雪风, 姜攀

西京学院计算机学院, 陕西 西安

收稿日期: 2024年12月25日; 录用日期: 2025年1月17日; 发布日期: 2025年1月29日

摘要

随着铁路运输的快速发展, 铁路轨道的安全性与可靠性成为重要的研究课题。传统的人工检测方法效率低且易受人为因素影响。近年来, 深度学习, 尤其是YOLO系列模型在目标检测中的应用取得了显著进展。为提高铁轨表面缺陷检测的精度和实时性, 本文提出了一种改进的YOLOv11模型(TrackDef-YOLO)。通过引入多尺度特征融合、自动锚框调整和加权损失函数等创新技术, TrackDef-YOLO显著提升了对微小缺陷和复杂背景的检测能力。实验结果表明, TrackDef-YOLO相比YOLOv8和YOLOv11n模型在准确率、召回率、F1-Score和误报率等方面均有显著提升。该研究为铁路轨道智能检测提供了高效、精准的解决方案, 并为基于深度学习的铁路维护系统的优化奠定了基础。

关键词

YOLOv11, 铁轨表面缺陷检测, 深度学习, 目标检测, 实时检测, 数据增强, 卷积神经网络, 铁轨维护

TrackDef-YOLO: An Improved YOLOv11 Model for Railway Track Surface Defect Detection

Mengchao Li, Xuefeng Wen, Pan Jiang

School of Computer Science, Xijing University, Xi'an Shaanxi

Received: Dec. 25th, 2024; accepted: Jan. 17th, 2025; published: Jan. 29th, 2025

Abstract

With the rapid development of railway transportation, the safety and reliability of railway tracks

文章引用: 李梦超, 文雪风, 姜攀. TrackDef-YOLO: 一种改进的 YOLOv11 模型用于铁轨表面缺陷检测[J]. 人工智能与机器人研究, 2025, 14(1): 217-223. DOI: 10.12677/airr.2025.141021

have emerged as critical research topics. Traditional manual inspection methods are inefficient and susceptible to human factors. In recent years, deep learning, especially YOLO series models, has achieved remarkable progress in object detection. To enhance the accuracy and real-time performance of railway track surface defect detection, this paper proposes an improved YOLOv11 model, named TrackDef-YOLO. By incorporating innovative techniques such as multi-scale feature fusion, automatic anchor box adjustment, and a weighted loss function, TrackDef-YOLO significantly improves the detection capability for minute defects and complex backgrounds. Experimental results demonstrate that TrackDef-YOLO outperforms YOLOv8 and YOLOv11n models in terms of accuracy, recall, F1-Score, and false alarm rate. This study provides an efficient and precise solution for intelligent railway track inspection and lays the foundation for optimizing railway maintenance systems based on deep learning.

Keywords

YOLOv11, Railway Track Surface Defect Detection, Deep Learning, Object Detection, Real-Time Detection, Data Augmentation, Convolutional Neural Networks, Railway Maintenance

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来，铁路轨道缺陷检测作为一种关键的智能化监测手段，受到了越来越多的关注。传统的轨道检测方法多依赖人工巡检或传统的图像处理技术，但这些方法通常效率低下且易受环境影响。随着深度学习技术的迅速发展，基于卷积神经网络(CNN)的目标检测方法，特别是 YOLO (You Only Look Once) 算法，在铁路轨道缺陷检测中得到了广泛应用。YOLO 系列模型由于其端到端的结构和实时检测能力，成为了目标检测领域的研究热点。

多项研究针对 YOLO 算法在轨道缺陷检测中的应用进行了改进和优化。首先，针对轨道缺陷的多样性和微小尺寸，部分研究采用了多尺度特征融合(Feature Pyramid Networks, FPN)技术，以提升对小物体的检测能力。例如，Wu 等[1]提出的 FPN 方法通过结合不同尺度的特征图来增强模型对小物体的感知能力，已被证明在多个目标检测任务中有效提高了检测精度。为进一步提升 YOLO 模型的鲁棒性和准确性，其他学者则引入了空间注意力机制(SAM)，该机制能够动态调整网络在空间上的关注区域，有效过滤掉复杂背景中的干扰信息，增强了缺陷检测的精确性[2]。

此外，针对铁轨缺陷检测中的背景干扰和数据不平衡问题，部分研究提出了基于数据增强的优化策略。这些方法通过合成多样化的训练样本，特别是对小物体和少量缺陷样本的增强，有效缓解了训练过程中出现的类别不平衡问题。与此同时，基于 YOLO 的模型在锚框设置方面也进行了自适应调整，优化了目标框的选择策略，从而提升了检测精度。

尽管已有研究在轨道缺陷检测中取得了一定进展[3]，但 YOLO 模型仍面临一些挑战，特别是在检测微小缺陷、复杂背景和噪声干扰时。本文提出了一种改进的 YOLOv11 模型——TrackDef-YOLO，通过引入多尺度特征融合、自动锚框调整和加权损失函数等创新技术，进一步提升了模型对微小缺陷和复杂背景的检测能力。实验结果表明，本文提出的模型在精度、鲁棒性和实时性方面均取得了显著提升，具有较强的工程应用价值。

2. 改进算法

本节详细介绍了 TrackDef-YOLO 模型在 YOLOv11 基础上进行的多项改进。YOLOv11 模型的改进主要集中在以下几个方面：高效的特征提取、自动锚框调整、多尺度特征融合、加权损失函数和背景干扰抑制。每一项技术创新都有明确的算法描述，并详细阐述其在轨道缺陷检测中的具体实现。

2.1. 高效特征提取

在 TrackDef-YOLO 中，我们采用了 CSPDarknet53 作为主干网络。与传统的 YOLOv5 使用的 Darknet53 相比，CSPDarknet53 引入了跨阶段部分(CSP)架构，以有效提高网络在深层次训练中的稳定性和特征表达能力[4]。

CSPDarknet53 通过将网络的每一层分为两部分进行处理：一部分经过残差连接(Residual Connection)，另一部分则通过直接连接(Skip Connection)传递，最后在合并阶段将两部分的特征融合。这样做不仅减少了计算量，还提高了梯度的传播效率，减少了梯度消失的风险。

具体而言，CSPDarknet53 将特征图进行分割后，在网络的每一阶段进行合并操作。这一设计使得每一层的特征图能够融合更多的上下文信息，从而有效提升对微小缺陷和复杂背景的感知能力。

2.2. 自动锚框调整

TrackDef-YOLO 采用了自适应锚框调整算法，针对轨道缺陷的不同尺寸和形态，动态生成更加合适的锚框[5]。我们使用 K-means 聚类算法来自动调整锚框的大小和数量。具体步骤如下：

- 1) 锚框聚类：首先，利用 K-means 算法对训练数据集中每个目标的边界框进行聚类。聚类的目标是将边界框分为 K 个类别，使得每个类别的边界框与聚类中心的差异最小。
- 2) 锚框生成：通过 K-means 聚类得到的 K 个聚类中心，作为新的锚框大小。这些锚框能够更好地适应轨道缺陷的不同尺寸，提升检测精度。
- 3) 自适应调整：根据每一类缺陷的具体特征(如尺寸、形态)，进一步调整锚框的位置和形态，以提高不同缺陷类型的检测效果。

这一算法的关键优势在于其能够根据实际数据自动生成与目标匹配度更高的锚框，从而显著提升小尺寸缺陷的检测效果。

2.3. 多尺度特征融合

TrackDef-YOLO 多尺度特征融合技术，增强了对不同尺度缺陷的检测能力，尤其是微小缺陷的识别[6]。多尺度特征融合基于特征金字塔网络(FPN)架构，将来自不同尺度的特征图进行加权融合。具体步骤如下：

- 1) 特征提取：通过 YOLOv11 网络中的多个卷积层，提取出不同尺度的特征图。低层的特征图通常包含丰富的空间细节信息，而高层的特征图则包含语义信息。

2) 融合与加权：在融合阶段，采用加权求和的方式将不同尺度的特征图结合。通过引入加权系数，确保低层特征(细节)和高层特征(语义)的有效结合。

对于低层特征图(细节图)，权重较高，以保留缺陷的微小信息。

对于高层特征图(语义图)，权重较低，但确保全局信息的引入，帮助模型更好地理解上下文。

- 3) 特征融合层：将融合后的多尺度特征图输入到后续的检测层，确保每个尺度的缺陷都能得到准确检测。

通过这种融合策略，TrackDef-YOLO 能够有效提高对复杂背景下微小缺陷的检测能力。

2.4. 加权损失函数

为了使 TrackDef-YOLO 更专注于难以检测的小缺陷或复杂背景中的缺陷，我们设计了一个加权损失函数，以动态调整不同损失项的权重。

加权损失函数的核心思想是根据不同目标的难易程度、大小以及背景干扰，自动调整损失函数的各项权重[7]。具体步骤如下：

1) 损失函数组成：YOLOv11 的损失函数包括定位损失(Bounding Box Loss)、分类损失(Classification Loss)和置信度损失(Confidence Loss)。

定位损失：衡量预测框与真实框之间的差异，采用平滑 L1 损失(Smooth L1 Loss)。

分类损失：衡量每个类别的预测结果与真实标签之间的差异，使用交叉熵损失(Cross Entropy Loss)。

置信度损失：衡量模型对目标存在性的信心，使用二元交叉熵损失。

2) 加权系数：根据目标的大小和形态，动态计算损失项的加权系数。例如：

对于小缺陷，赋予较高的定位损失权重，以提高对小目标的定位精度。

对于背景复杂的区域，降低置信度损失的权重，减少背景干扰。

3) 加权损失计算：最终的总损失通过加权各个损失项得到，具体公式为：

$$L_{\text{total}} = \alpha \cdot L_{\text{box}} + \beta \cdot L_{\text{cls}} + \gamma \cdot L_{\text{conf}}$$

其中， α 、 β 和 γ 是根据目标类型动态调整的权重系数。

这一损失函数的优势在于能够根据每个样本的特征(如尺寸和背景复杂度)调整训练过程中的关注点，从而显著提升对小缺陷和复杂背景的检测精度。

2.5. 背景干扰抑制

为了减少复杂背景对轨道缺陷检测的干扰，TrackDef-YOLO 引入了空间注意力机制(SAM)，以增强对重要区域的关注。空间注意力机制的目标是通过自适应调整网络对不同空间区域的关注度，增强对轨道缺陷的敏感性[8]。具体步骤如下：

1) 注意力图生成：首先，通过卷积操作生成一个注意力图，表示每个位置的重要性。该注意力图是通过输入特征图与一个学习得到的卷积核进行卷积计算得到的，公式如下：

$$A = \sigma(W_a * X + b)$$

其中， A 是生成的注意力图， σ 是激活函数(如 Sigmoid)， W_a 是学习得到的卷积核， X 是输入特征图。

2) 注意力加权：生成的注意力图与输入特征图逐元素相乘，强化网络对重要区域的关注，抑制背景干扰。公式为：

$$X' = A \cdot X$$

3) 合并与输出：将加权后的特征图传递给后续的检测层，用于最终的缺陷定位与分类。

这一机制能够使模型在复杂背景下自适应地关注轨道缺陷的关键区域，有效提高检测精度，尤其是在背景干扰较大的情况下。

3. 实验分析

3.1. 实验详情

本文采用 PyTorch 框架实现 YOLOv11 模型，并在 NVIDIA Tesla V100 GPU 上进行训练。实验环境为 Windows10 操作系统，Python 3.8 编程语言，CUDA 11.1 进行加速。

3.2. 数据集构建

为了训练和评估 TrackDef-YOLO 模型，本文使用了公开的铁路轨道缺陷数据集 RSDDS (Rail Surface Defect Detection System)。RSDDS 数据集包含了多种类型的铁路表面缺陷，如裂纹、腐蚀、磨损等，且包含不同光照、背景和天气条件下的图像。数据集的每张图像都经过人工标注，标出了缺陷的位置和类别[9]。为了进一步增强模型的鲁棒性，我们对原始数据集进行了数据增强处理，包括旋转、缩放、翻转和色彩变换等操作。通过这些增强技术，我们增加了数据集的多样性，帮助模型更好地应对不同场景中的缺陷[10]。

3.3. 实验设计

本研究的实验设计包括数据处理、训练、验证和测试四个阶段。具体步骤如下：

- 1) 数据预处理：对数据集中的图像进行归一化处理，确保输入图像的像素值在 0 到 1 之间。
- 2) 训练阶段：将训练集分为训练集和验证集，使用 TrackDef-YOLO 模型进行训练，优化网络的权重和参数。
- 3) 验证与调参：通过验证集评估模型的性能，并根据验证结果调整模型的超参数(如学习率、批量大小等)。
- 4) 测试阶段：在测试集上评估模型的检测精度，计算常见的评估指标。

3.4. 评估指标

为了全面评估 TrackDef-YOLO 模型的性能，本研究采用了以下评估指标：

准确率(Precision)：表示检测出的正例中，真正例的比例。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

其中，TP 为真正例，FP 为假正例。

召回率(Recall)：表示所有真实正例中，检测出的正例的比例。

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})}$$

其中，TP 为真正例，FN 为假负例。

F1-Score：准确率与召回率的调和平均，综合考虑了模型的精度和召回能力。

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

平均精度均值(mAP)：在多个类别中计算平均精度，衡量模型在多类别任务中的整体表现[11]。

这些指标将用于对 TrackDef-YOLO 模型的性能进行全面评估，并与现有的 YOLO 模型进行对比。

3.5. 实验分析

由表 1 可以看出，TrackDef-YOLO 在准确率、召回率、F1-Score 和 mAP 方面均表现出了显著的优势，尤其是在小缺陷(如裂纹和局部腐蚀)的检测上，TrackDef-YOLO 模型的性能提升尤为明显。

TrackDef-YOLO 模型在多个关键指标上超越了 YOLOv8 和 YOLOv11n，证明了其在铁路轨道表面缺陷检测中的有效性。尤其是多尺度特征融合和自适应锚框的设计[12]，使得 TrackDef-YOLO 能够更好地适应各种缺陷的尺寸差异，提高了微小缺陷的检测精度。此外，加权损失函数有效缓解了数据不平衡问题，使得 TrackDef-YOLO 在少数类缺陷的检测上表现更为突出。

Table 1. Performance comparison of TrackDef-YOLO model in rail surface defect detection
表 1. TrackDef-YOLO 模型在铁轨表面缺陷检测中的性能对比

模型	Precision	Recall	F1-Score	mAP
YOLOv8	85.6%	83.2%	84.4%	86.1%
YOLOv11n	87.4%	85.3%	86.3%	87.8%
TrackDef-YOLO	90.5%	88.7%	89.6%	91.2%

4. 结论

本文提出了一种改进 YOLOv11 模型的铁路轨道缺陷检测方法——TrackDef-YOLO，并通过多项技术创新，提升了模型在复杂环境下对微小缺陷的检测能力。通过引入高效的特征提取网络(CSPDarknet53)、自动锚框调整、多尺度特征融合、加权损失函数和背景干扰抑制等技术，本文有效地解决了铁路轨道检测中面临的挑战。具体的研究结论如下：

提升了微小缺陷检测能力：通过引入 CSPDarknet53 作为主干网络，本文改进的 YOLOv11 模型在特征提取方面比传统的 YOLOv5 更为高效，能够在复杂的背景环境下有效提取更多的上下文信息，显著提高了对微小缺陷的检测精度。

自动适应不同尺寸缺陷：通过 K-means 聚类算法自动调整锚框，YOLOv11 能够适应轨道缺陷的不同尺寸和形态，尤其在小尺寸缺陷的检测中表现出色，提升了缺陷检测的精度和召回率。

增强了对复杂背景的适应能力：本文提出的多尺度特征融合方法，使得模型能够充分利用不同尺度的特征信息，提升了对复杂背景中微小缺陷的识别能力，特别是在大范围轨道检测任务中，能够较好地平衡检测精度与计算效率。

优化了损失函数以强化模型聚焦难点：加权损失函数的引入，使得模型在训练过程中能够更专注于难以检测的小缺陷和复杂背景区域，从而提升了对小目标的检测能力，并减少了背景干扰的影响。

有效的背景干扰抑制机制：通过引入空间注意力机制(SAM)，本文能够有效抑制背景中的干扰，确保模型在复杂背景下能够关注到轨道缺陷的关键区域，从而提高检测精度。

综上所述，本文提出的 TrackDef-YOLO 算法在铁路轨道缺陷检测任务中，尤其是在检测微小缺陷和复杂背景方面，表现出较传统 YOLO 模型更为优越的性能。未来的工作可以进一步探索该模型在不同环境条件下的适应性和实时检测能力[13]。同时，可以结合其他深度学习方法进一步优化模型结构，以应对更复杂的轨道缺陷检测需求。

参考文献

- [1] Wu, Y., Qin, Y., Wang, Z. and Jia, L. (2018) A UAV-Based Visual Inspection Method for Rail Surface Defects. *Applied Sciences*, **8**, Article 1028. <https://doi.org/10.3390/app8071028>
- [2] Gan, J., Li, Q., Wang, J. and Yu, H. (2017) A Hierarchical Extractor-Based Visual Rail Surface Inspection System. *IEEE Sensors Journal*, **17**, 7935-7944. <https://doi.org/10.1109/jsen.2017.2761858>
- [3] Chen, G., Wang, X., Ni, W. and Wei, L. (2023) A Steel Rail Surface Defect Detection and Recognition Method Based on Mask R-Transformer. 2023 IEEE 6th International Conference on Pattern Recognition and Artificial Intelligence (PRAI), Haikou, China, 18-20 August 2023. <https://doi.org/10.1109/prai59366.2023.10332076>
- [4] Wang, Y., Zhang, K., Wang, L. and Wu, L. (2024) An Improved YOLOv8 Algorithm for Rail Surface Defect Detection. *IEEE Access*, **12**, 44984-44997. <https://doi.org/10.1109/access.2024.3380009>
- [5] Zhao, J., Yeung, A.W., Ali, M., Lai, S. and Ng, V.T. (2024) CBAM-Swin-Bl: Small Rail Surface Defect Detection Method Based on Swin Transformer with Block Level CBAM Enhancement. *IEEE Access*, **12**, 181997-182009. <https://doi.org/10.1109/access.2024.3509986>

-
- [6] Han, H., Tian, L., Li, M., Cui, X., Shang, C. and Hou, S. (2023). Design of Rail Surface Defect Detection System Based on LabVIEW Machine Vision. *2023 IEEE 6th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, Chongqing, China, 24-26 February 2023. <https://doi.org/10.1109/itnec56291.2023.10082124>
 - [7] Zhang, C., Xu, D., Zhang, L. and Deng, W. (2023) Rail Surface Defect Detection Based on Image Enhancement and Improved YOLOX. *Electronics*, **12**, Article 2672. <https://doi.org/10.3390/electronics12122672>
 - [8] Du, J., Zhang, R., Gao, R., Nan, L. and Bao, Y. (2024) RSDNet: A New Multiscale Rail Surface Defect Detection Model. *Sensors*, **24**, Article 3579. <https://doi.org/10.3390/s24113579>
 - [9] Fang, Z., Li, L., Peng, L., Zheng, S., Zhong, Q. and Zhu, T. (2024) YOLOv8n-RSDD: A High-Performance Low-Complexity Rail Surface Defect Detection Network. *IEEE Access*, **12**, 196249-196265. <https://doi.org/10.1109/access.2024.3466559>
 - [10] Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M., *et al.* (2020) YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv: 2004.10934.
 - [11] Lin, T.-Y., Dollar, P., Girshick, R., *et al.* (2017) Feature Pyramid Networks for Object Detection. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 936-944.
 - [12] Lin, T.Y., Goyal, P., Girshick, R., He, K.M. and Dollar, P. (2018) Focal Loss for Dense Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**, 318-327.
 - [13] Rezatofighi, H., Tsoi, N., Gwak, GY., *et al.* (2019) Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression. *2019 IEEE Transactions on Pattern Analysis and Machine Intelligence*, Long Beach, 15-20 June 2019, 658-666. <https://doi.org/10.1109/CVPR.2019.00075>