

# 基于改进 YOLOv7 的钢轨缺陷检测方法<sup>\*</sup>

赵亚凤 宋文华 刘晓璐 胡峻峰

(东北林业大学计算机与控制工程学院 哈尔滨 150040)

**摘要:** 针对铁路轨道缺陷检测精度低,漏检率高,实时性不足的问题,本文提出了一种基于 YOLO-FCA 的钢轨缺陷检测算法。首先,将 YOLOv7 的主干网络替换成 FasterNet 轻量网络,并加入 CloAttention 注意力模块,减少参数量和计算负载的同时提高缺陷检测的精度。其次,提出 MS-ASFF,获取高层语义信息和保留低层详细特征,增强模型检测的准确性和鲁棒性。最后,在不影响精度的情况下进行网络剪枝,使模型更加轻量化,极大地提升了模型的检测速度。在公共数据集上进行实验,结果表明,YOLO-FCA 相比原始模型 YOLOv7 模型的 mAP 提高了 4.1%,达到 80.7%,同时检测速度提升了 38.5%,达到 212.5 FPS。实验结果表明,YOLO-FCA 能够高效且准确地定位检测钢轨缺陷。

**关键词:** YOLOv7 算法;钢轨缺陷;自适应的空间特征融合;注意力机制;轻量化模型

中图分类号: U216.3; TP391.41; TN249.2 文献标识码: A 国家标准学科分类代码: 520.6040

## Improved rail defect detection algorithm of YOLOv7

Zhao Yafeng Song Wenhua Liu Xiaolu Hu Junfeng

(College of Computer and Control Engineering, Northeast Forestry University, Harbin 150040, China)

**Abstract:** Aiming at the problems of low accuracy, high missed detection rate and insufficient real-time performance of railway track defect detection, this paper proposes a rail defect detection algorithm based on YOLO-FCA. First, the backbone network of YOLOv7 was replaced with the lightweight network of FasterNet, and the attention module of CloAttention was added to reduce the number of parameters and calculation load while improving the accuracy of defect detection. Secondly, a multi-scale adaptive feature fusion network (MS-ASFF) is proposed to obtain high-level semantic information and retain low-level detailed features to enhance the accuracy and robustness of model detection. Finally, the network pruning is carried out without affecting the accuracy, which makes the model more lightweight and greatly improves the detection speed of the model. Experiments on public data sets show that compared with the original YOLOv7 model, the mAP of YOLO-FCA is increased by 4.1%, reaching 80.7%, and the detection speed is increased by 38.5%, reaching 212.5 FPS. The experimental results show that YOLO-FCA can locate and detect rail defects efficiently and accurately.

**Keywords:** YOLOv7 algorithm; railway track defects; adaptive spatial feature fusion; attention mechanism; lightweight model

## 0 引言

近年来,提升轨道交通的安全性和效率已成为现代交通系统的重要目标<sup>[1]</sup>。作为轨道交通运行关键设施的钢轨,准确检测其剥落、创伤和裂纹等各种缺陷类型,是保持轨道交通安全稳定运行的核心手段。

传统铁路缺陷检测方法主要包括超声波<sup>[2]</sup>、磁粉<sup>[3]</sup>以及红外热成像<sup>[4]</sup>等,但传统方法普遍存在检测时间长、检测效率低等局限性。为提升检测效率,科研人员引入了机器视觉技术,通过成像系统获取轨道表面图像,并进行人工分

析设计或预定义特征来检测缺陷,这些方法将提取的特征输入至分类算法进行识别。在机器视觉方向,科研人员做出了大量的研究,Hua 等<sup>[5]</sup>提出了一种基于光学激光测量和时空分析的方法用于缺陷的 3D 重建。Galan 等<sup>[6]</sup>指出通过对灰度图的亮区和暗区进行阈值处理得到二值图像,然后处理这些图像的连通部分以检测由缺陷引起的阴影。

随着人工智能迅猛发展,深度学习技术特别是目标检测<sup>[7-11]</sup>在铁路缺陷检测领域取得了显著进展,学者们相应地在精度和检测速度以及模型尺寸方面提出了多种改进。熊聪等<sup>[12]</sup>结合 Swin Transformer 模块与加权双向特征金

塔网络 (bidirectional feature pyramid network, BiFPN) 提取更多差异化特征并进行跨尺度融合, 大幅度提高了缺陷检测精度。Panigrahi 等<sup>[13]</sup> 使用轻量级 SqueezeNet 替代 YOLOv3 主干网络, 显著减少计算和参数负载。Han 等<sup>[14]</sup> 对 YOLOv3 中的特征金字塔网络 (feature pyramid network, FPN) 进行剪枝处理以降低计算复杂度并提高检测速度。赵春华等<sup>[15]</sup> 提出基于 YOLOv7-tiny 网络改进的钢材表面缺陷 PC-YOLOv7 检测算法, 能在 NEU-DET 数据集上实现 78.5% 的平均精度, 提升了检测精度并缩减了模型尺寸, 但对于多尺度缺陷的检测效果不佳。卢开喜等<sup>[16]</sup> 提出 KThin-YOLOv7, 通过引入 EMA-BasicRFBC 模块, FEIOU 损失函数和 ThinNeck 结构改进 YOLOv7 算法, 显著提升了检测精度并减少了模型参数量, 但依然无法缺陷的实时检测。

尽管国内外学者们在优化检测精度和模型尺寸上已有进展, 但仍有进一步改进的空间。为了进一步平衡钢轨表面缺陷检测精度与模型尺寸的问题, 本文提出了一种基于 YOLOv7<sup>[17]</sup> 网络改进的轻量化钢轨缺陷检测模型 YOLO-FCA 主要贡献如下:

1) 结合轻量级网络 FasterNet 和 CloAttention 注意力机制, 增强了模型对小目标和细粒度缺陷的检测能力, 同时减少模型的参数量。

2) 引入新颖的多尺度自适应空间特征融合 (multi-scale adaptive spatial feature fusion, MS-ASFF) 方法, 解决了多尺度特征融合时忽视微小缺陷详细特征的问题。

3) 使用高效的网络剪枝技术, 有效降低了模型复杂性

和大小, 提高了分类效率和检测速度。

4) 在公共数据集上与其他算法进行对比实验, 结果显示本文的模型性能优于其他算法, 进一步验证了其在铁路缺陷检测领域的优异性。

这些创新使得本文方法对铁路缺陷检测具有重要的创新意义, 并为相关研究和应用提供了有益的启示。

## 1 改进的 YOLOv7 模型: YOLO-FCA

YOLOv7 模型在物体检测中表现出色, 其结构优化使得在保持高精度的同时, 能够快速处理大量数据。在针对钢轨缺陷检测时, YOLOv7 的高效性和准确性能够有效应对缺陷检测中的精度低和漏检率高的问题。通过在 YOLOv7 模型基础上进行改进, 可以提升模型对细小缺陷的检测能力, 同时保持快速响应, 从而改进检测精度和实时性。针对钢轨细粒度缺陷存在特征信息难以提取、缺乏特征表达能力等问题, 为了实现钢轨缺陷的高效检测, 本文基于 YOLOv7 模型进行了改进: 引入轻量级网络 FasterNet<sup>[18]</sup> 作为骨干网络, 缩减模型参数和计算量, 提升实时性能; 增加 CloAttention 注意力模块<sup>[19]</sup>, 增强骨干网络对全局和局部信息的特征提取能力; 提出 MS-ASFF 结构优化多尺度信息融合, 提升细粒度特征捕捉能力; 最后, 通过高效网络剪枝减少计算量和内存消耗, 进一步增强实时性能。

经过以上改进, YOLO-FCA 模型在精度和速度上均有显著提升, 能准确分类和检测钢轨缺陷图片。YOLO-FCA 模型结构如图 1 所示。

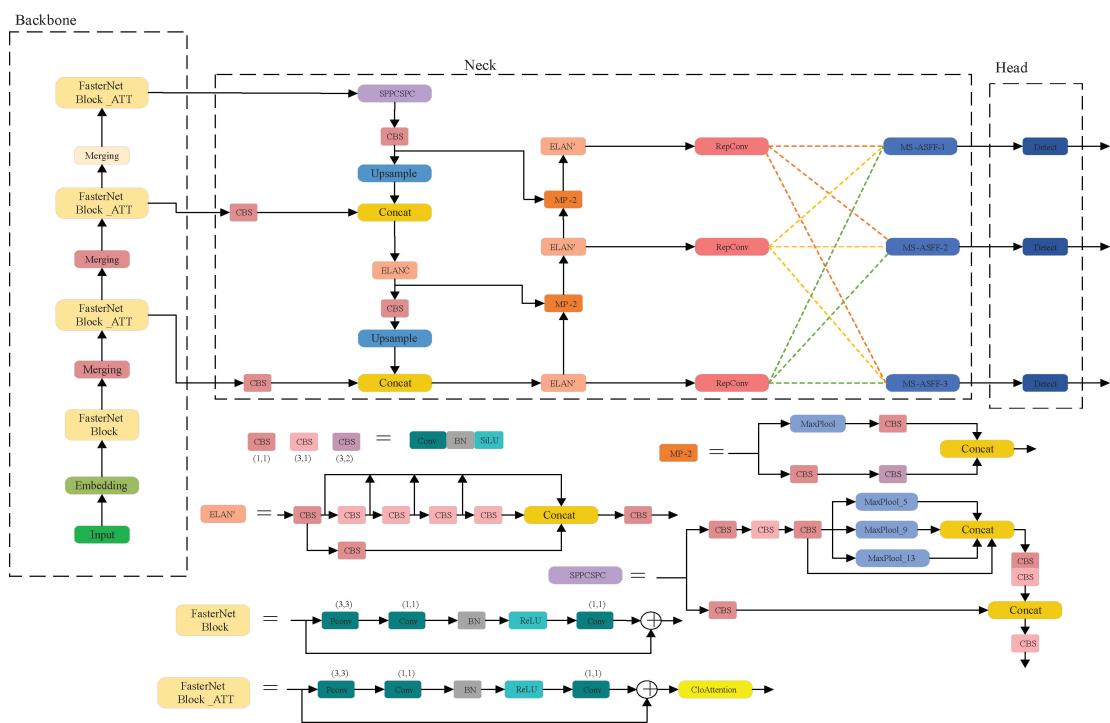


图 1 YOLO-FCA 模型结构图

Fig. 1 Structure diagram of YOLO-FCA model

## 1.1 轻量化网络 FasterNet

本文将原始 YOLOv7 模型中计算资源密集型的 Darknet53 基础网络替换为 FasterNet, 以降低模型的计算复杂度。FasterNet 利用部分卷积 (partial convolution, PConv) 在输入特征图的局部区域执行卷积操作, 有效减少冗余计算和内存访问, 提高空间特征提取效率。网络分为四个阶段, 每个阶段都包含一组 FasterNet Block, 其中包括 PConv 层和逐点卷积 (pointwise convolution, PWConv) 层, 并添加了标准化和激活层以保持特征多样性并降低延迟。这种设计显著提升了前向传播速度, 满足实时性要求, 并减轻了边缘设备上的计算负担。FasterNet 网络结构如图 2 所示。

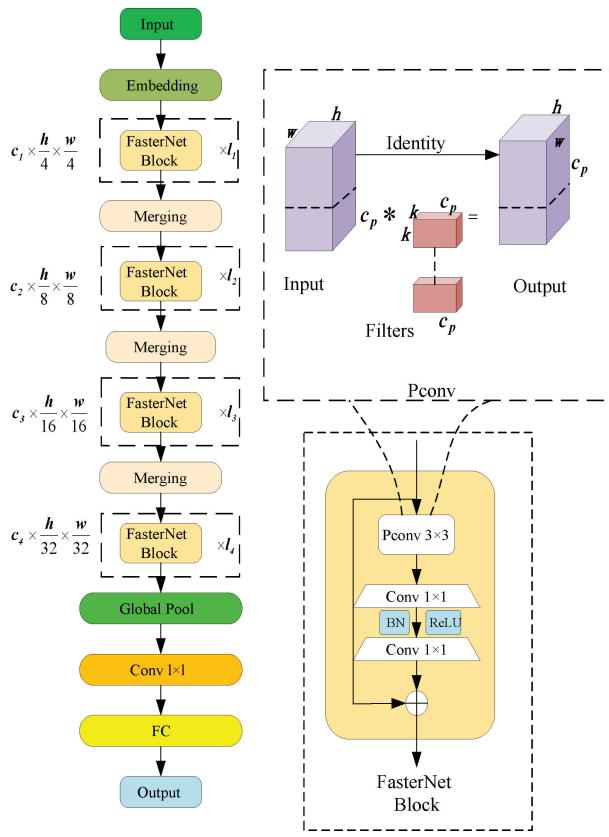


图 2 FasterNet 网络结构图

Fig. 2 FasterNet network structure diagram

## 1.2 CloAttention 注意力机制

在目标检测任务中, YOLOv7 能够实现大部分选定目标的准确检测, 但在处理划痕和硌伤等细粒度钢轨缺陷时存在不足, 因为这些缺陷与周围环境纹理、颜色和灰度值相似, 如图 3 所示。针对这些难以区分的缺陷, 本文引入了 CloAttention 机制。

CloAttention 机制通过双分支设计结构, 提升了图像分类性能。一个分支利用传统的注意力机制捕获低频全局信息, 另一个分支则采用 AttnConv 捕获高频局部信息。AttnConv 卷积算子利用注意力风格和共享权值, 结合了非

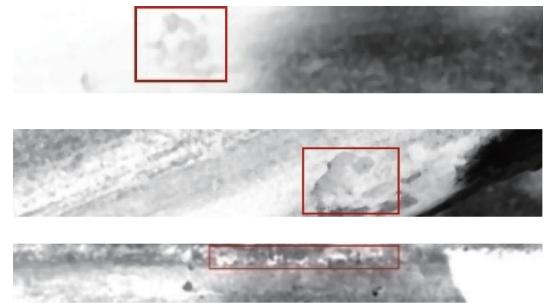


图 3 钢轨缺陷灰度图

Fig. 3 Gray level of rail defects

线性方法生成上下文感知权重, 有效地结合了全局和局部信息优势。因此, 该模块能够在钢轨缺陷检测过程中提升模型对目标细节和纹理的关注能力。注意力结构如图 4 所示。

全局分支采用传统注意力机制的风格, 但又在此上面有所改进: 对  $K$ 、 $V$  进行下采样以减少计算量, 从而捕捉低频全局信息:

$$X_{global} = \text{Attention}(Q_g, \text{Pool}(K_g), \text{Pool}(V_g)) \quad (1)$$

AttnConv 引入了局部分支, 结合了共享权重和上下文感知权重, 以有效捕捉高频局部信息。首先, 通过线性变换得到  $Q$ 、 $K$  和  $V$ 。接着, 对  $V$  进行共享权重的局部特征聚合。最后, 利用处理后的  $V$ 、 $Q$  和  $K$  进行上下文感知的局部增强:

$$V_s = \text{DWconv}(V), Q_l = \text{DWconv}(Q), K_l = \text{DWconv}(K) \quad (2)$$

AttnConv 使用深度卷积 (depthwise convolution, DWconv) 进行局部信息聚合, 其中 DWconv 共享权重。在聚合  $Q$  和  $K$  的局部信息后, 计算它们的 Hadamard 乘积, 并经过一系列转换得到在  $-1 \sim 1$  之间的上下文感知权重。

$$Attn = FC(\text{Swish}(FC(Q_l \odot K_l))) \quad (3)$$

$$Attn = \text{Tanh}(\frac{Attn}{\sqrt{d}}), X_{local} = Attn \odot V_s \quad (4)$$

最后, 将全局特征和局部特征合并起来, 并使用一个 MLP 得到最终的输出。公式如下:

$$X_t = \text{Concat}(X_{global}, X_{local}), X_{out} = FC(X_t) \quad (5)$$

从上述公式可知, AttnConv 与局部自注意力相比, 共享权重的引入可以使模型更好地处理高频信息。与传统卷积相比, AttnConv 使用共享权重和上下文感知权重, 能够更好地处理图像中不同位置之间的关系。

## 1.3 多尺度自适应空间特征融合 MS-ASFF

相比 FPN, YOLOv7 模型使用 PANet 进行特征融合有助于整合多尺度特征, 但特征融合可能导致不同层间的矛盾信息, 从而丢失来自低尺度特征的关键细节。在铁路轨道缺陷检测中, 如图 5 所示, 钢轨缺陷呈现出广泛的尺寸范围, 从  $329 \times 54$  像素的大缺陷到  $14 \times 17$  像素的小缺陷不等。

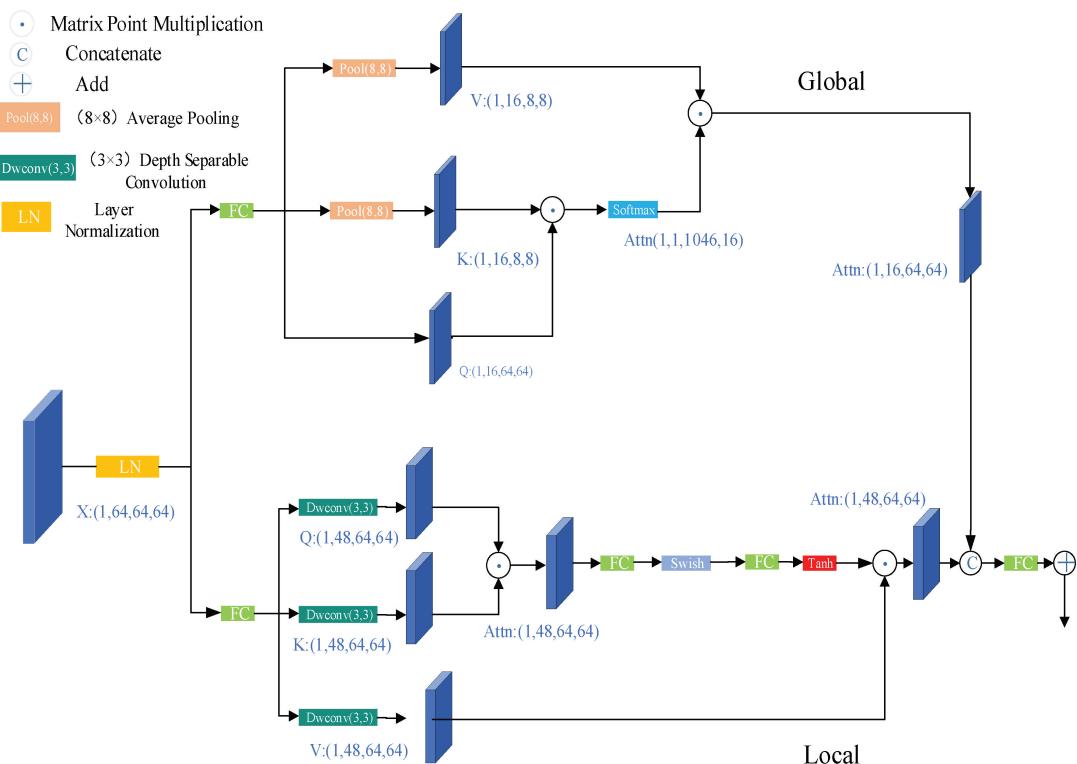


图 4 CloAttention 注意力结构图

Fig. 4 CloAttention attention structure diagram



图 5 多尺度缺陷

Fig. 5 Multi-scale defects

YOLOv7 采用了 3 个不同尺度的检测器头来实现尺度不变的目标检测，然而面临着大型目标缩小和小型目标特征学习不足等挑战。为了解决这些问题，Cheng 等<sup>[20]</sup>提

出了自适应空间特征融合 (adaptive spatial feature fusion, ASFF) 过滤冲突信息并改善不同尺度目标的特征融合。然而，ASFF 中的逐元素求和操作限制了其在适应不同尺度方面的能力。为了更好地保留细节并利用语义信息，在本文中提出了多尺度自适应空间特征融合策略 MS-ASFF。MS-ASFF 结合了对低级特征进行逐元素求和的操作，以保留其中的细粒度信息，同时对高级特征进行卷积操作，以提取其中的丰富语义内容；逐元素求和可以将这些详细信息传递给融合特征，而卷积操作则能更好地利用高级语义信息。具体结构如图 6 所示。

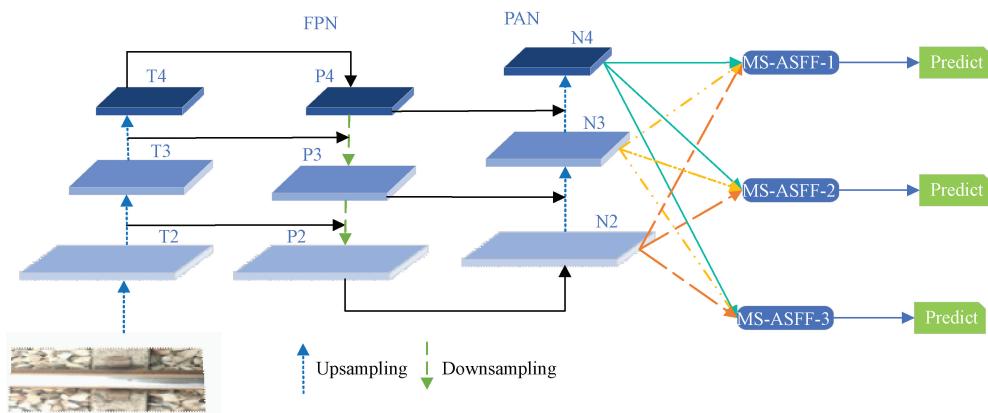


图 6 MS-ASFF 结构图

Fig. 6 MS-ASFF structure diagram

低级特征的特征融合:对于第 1 级( $i=0$ ),本文执行逐元素加法以进行特征融合:

$$\text{fused\_low} = x_i + F.\text{interpolate}(x_{i+1}, \text{size} = x_i.\text{size}()[2:], \text{mode} = \text{'nearest'}) \quad (6)$$

高级特征的特征融合:对于第 2 层和第 3 层( $i=1,2$ ),本文使用卷积进行特征融合:

$$\text{fused\_high} = nn.\text{Conv2d}(x_{i+1}, \text{size}(1), x_i, \text{size}(1), \text{kernel\_size} = 1)(x_{i+1}) \quad (7)$$

其中, $x_i$  表示第  $i+1$  层的输入特征图,通过压缩高级特征图的通道以匹配低级特征图的通道。计算每个级别的融合特征图权重,连接压缩特征图,并应用卷积计算权重。函数  $F.\text{interpolate}()$  用于改变张量的大小; $x_i.\text{size}()$  是 1 个函数调用,它返回张量  $x_i$  的形状,使用[2:]表示获取从第 3 个维度开始到最后的维度;参数 mode 值为字符串'nearest',用于指定插值的模式,此处,它表示使用最近邻插值。 $x_{i+1}.\text{size}()$ :这表示获取  $x_{i+1}$  张量的第 2 个维度(通常是通道数),用作卷积层的输入通道数。 $x_i.\text{size}(1)$ :这表示获取  $x_i$  张量的第 2 个维度,用作卷积层的输出通道数; $\text{kernel\_size}=1$ :表示使用  $1 \times 1$  的卷积核。

MS-ASFF 模块通过逐元素求和以及卷积融合,自适应地整合不同层次的特征图。这种设计增强了特征融合

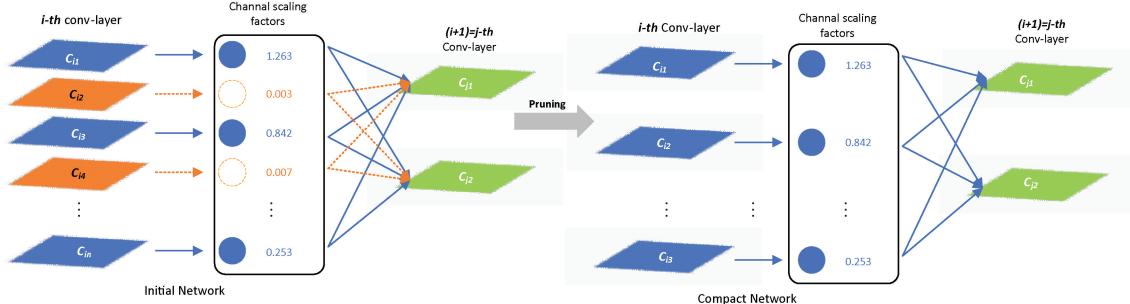


图 7 通道修剪原理图

Fig. 7 Schematic diagram of channel pruning

在修剪之前要先判断通道的重要程度,通过 BatchNormalization 层的  $\gamma$  因子评判通道的重要程度:

$$y = \frac{x - \text{mean}(x)}{\sqrt{\text{Var}(x)} + \epsilon} \times \gamma + \beta \quad (9)$$

其中, $\epsilon$  是一个极小的正数,通常设置为避免除以零的情况; $\beta$  是用于线性模型中调整截距的偏置项。在训练过程中,BN 层中的  $\gamma$  参数通过学习来评估每个通道对模型性能的贡献。当  $\gamma$  较小时,归一化后的激活值会相应较小。通过裁剪这些微弱响应,可以实现 BN 层的通道剪枝,从而减少模型存储和计算需求。

通过引入基于通道级稀疏诱导正则化训练后,在接近零缩放因子下进行通道修剪即删除相关连接和权重,并设定全局阈值为所有缩放因子某个百分位数来进行修剪。然而实验表明,过度剪枝对准确率有显著影响,根据 4 次实验取平均值得出,在保证适当准确率下选择 4% 作为合

模块的灵活性,更有效地处理多尺度特征,从而显著提升了特征金字塔网络在目标检测任务中的性能。

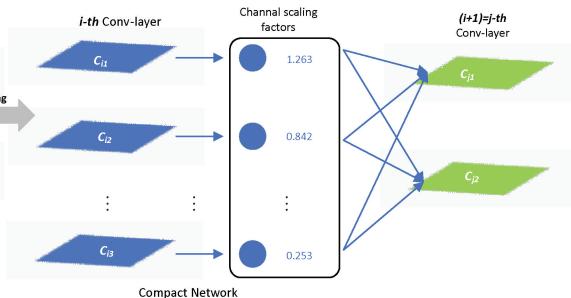
#### 1.4 网络剪枝(Pruning)

由于 CloAttention 机制以及 MS-ASFF 结构较为复杂,包含许多冗余信息,可能导致模型参数量和计算量大大增加,不利于钢轨缺陷的实时检测。

因此,为降低模型的计算需求、提升检测速度,本文采用剪枝算法对模型进行轻量化处理。Liu 等<sup>[21]</sup>提出了一种创新的剪枝策略,通过引入每个通道的缩放因子  $\gamma$  来降低深度学习模型的计算需求。在训练过程中,通过稀疏正则化优化缩放因子  $\gamma$ ,并同时更新网络权重。接着,他们修剪掉缩放因子较小的通道,然后对修剪后的网络进行微调。这种方法实现了通过稀疏化批量一化(bath normalization, BN)层中的  $\gamma$  来剪枝,其核心是删除通道的所有传入和传出连接,如图 7 所示。

$$L = \sum_{(x, y)} l(f(x, W), y) + \lambda \sum_{\gamma \in \Gamma} g(\gamma) \quad (8)$$

其中, $(x, y)$  表示训练输入和目标, $W$  表示可训练权重,第一个求和项对应 CNN 的正常训练损失, $g(\cdot)$  是稀疏性引起的尺度因子惩罚, $\lambda$  的作用则是平衡这两个项。在该实验中,选择  $g(s) = |s|$ ,即 L1 范数。



适的剪枝率(Amount),具体降低的精确度  $\nabla \text{Acc}$  和提高的 FPS 值  $\triangle \text{FPS}$ ,如表 1 所示。

表 1 Amount 与  $\triangle \text{Acc}$  以及  $\triangle \text{FPS}$  对应关系

Table 1 Correspondence between Amount and

$\triangle \text{Acc}$  and  $\triangle \text{FPS}$

Amount/%	2	4	5	10
$\nabla \text{Acc}/\%$	0.027	<b>0.8</b>	2.375	15.921
$\triangle \text{FPS}$	23.800	<b>45.7</b>	89.200	170.400

## 2 实验及结果分析

### 2.1 创建数据集

本文使用了“Railway Track Fault Detection”数据集,其他类型的缺陷出现较少,因此筛选出其中 7 种常见缺陷

类型作为实验数据:划痕、伤口、裂纹,脱落块、缺失的扣件、缺失的螺栓和缺失的夹子,如图 8 所示。筛选后的图片数据集共 873 张图片,为增强模型的泛化性能,数据集经过标记和多种数据增强<sup>[22]</sup>技术处理:亮度调整、cutout、旋转、平移和镜像操作,以扩展样本数量至 4 365 张图片。数据集按照 6:1:3 比例划分,训练集 2 619 张,验证集 440 张,测试集 1 306 张。

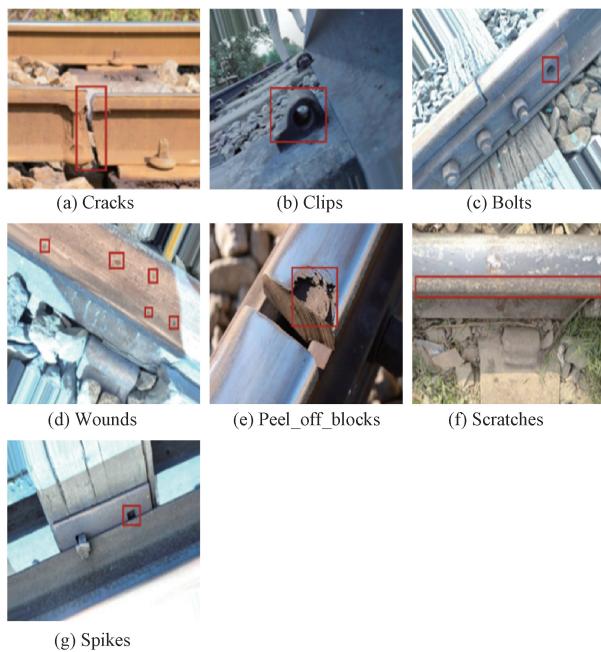


图 8 部分数据集图像

Fig. 8 Partial dataset images

## 2.2 测试参数配置

实验基于 Pytorch1.10 框架,CUDA11.2 版本进行训练,基准模型为 YOLOv7, batch size 设置为 8,共迭代 100 轮,采用 SGD 优化器,学习率为 0.005,输入图片尺寸设置为 640×640,具体实验配置如表 2 所示。

## 2.3 评价指标

为准确评价模型性能,本文采用了目标检测算法常用的 4 个性能评价指标:精确度(Precision),召回率(Recall),平均精度 (mean average\_precision, mAP) 和每秒帧数 (frame per second, FPS) 来检验本文提出的模型的性能。

表 2 实验环境配置

Table 2 Experimental environment configuration

操作系统	Ubuntu 22.04.2
CPU	Intel(R) Core(TM) i7-13700K
GPU	GeForce RTX 4090
内存	128 GB
Python 版本	Python 3.10.10

在精确度和召回率方面,预测测试样本后有 4 种状态:真阳性(TP)、假阳性(FP)、真阴性(TN)和假阴性(FN)。定义如下:

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

mAP 值和 FPS 都可以用来评价目标检测模型的性能。mAP, FPS 评分定义如下:

$$AP = \int_0^1 Precision \times Recall \, dRecall \quad (12)$$

$$mAP = \frac{1}{K} \sum_{k \in K} AP(k) \quad (13)$$

$$FPS = \frac{1}{Processing \, time \, per \, frame} \quad (14)$$

其中,K 代表数据集中类别的总数,k 表示单个类别,Processing time per frame 代表每帧处理时间。AP 和 mAP 是评估检测模型准确性的重要指标,分别用于单目标和多目标检测模型。较高的 AP 和 mAP 值反映了模型在识别目标时的优越表现。检测速度则以 FPS 来衡量,直接影响模型处理图像的速度。

## 2.4 实验结果分析

### 1) 不同注意力模块的影响

本节研究了 CloAttention 注意力机制在铁路缺陷检测任务中的影响。为了评估其有效性,本文在骨干网络中仅添加了一个 CloAttention 模块,并将其与原模型以及其他注意力机制(如 GAM、ECA 和 CBAM)进行了比较。注意力比较结果如表 3 所示,结果显示,CloAttention 在所有情况下表现最佳,尤其是在 Scratches 和 Wounds 缺陷的平均精度(mAP)方面。

表 3 YOLOv7 与不同注意机制结合的比较

Table 3 Comparison of YOLOv7 combined with different attention mechanisms

注意机制	Cracks	Scratches	Bolts	Clips	Peel_off_blocks	Wounds	Spikes	mAP/%	FPS/fps	模型尺寸/M
—	82.5	36.7	94.2	86.3	82.7	68.8	85.3	76.6	153.4	71.4
ECA	85.7	30.6	95.5	87.0	82.5	62.4	85.7	75.6	152.6	71.4
GAM	79.3	36.7	90.6	84.7	81.8	67.7	84.6	75.0	125.7	84.0
CBAM	86.7	35.4	95.4	86.8	83.2	68.8	87.9	77.7	138.8	72.0
CloAttention	87.4	38.7	95.9	87.6	84.4	69.2	89.5	<b>78.9</b>	142.3	74.6

## 2) 消融实验

为了验证各种改进策略对钢轨缺陷检测性能的贡献,在本文的数据集上进行了消融实验,结果如表 4 所示。

表 4 YOLO-FCA 模型消融实验结果

Table 4 Ablation results of YOLO-FCA model

YOLOv7	FasterNet	CloAttention	MS-ASFF	Pruning	mAP/%	FPS/fps	模型尺寸/M
√					76.6	153.4	71.4
√	√				75.7	242.0	43.8
√		√			78.9	142.3	74.6
√			√		78.6	124.0	93.0
√	√	√			77.7	216.5	50.2
√	√		√		78.3	184.0	75.5
√		√	√		82.4	105.7	118.0
√	√	√	√		81.5	166.8	90.3
√	√	√	√	√	<b>80.7</b>	<b>212.5</b>	<b>76.7</b>

FasterNet 结构效果:使用 FasterNet 替换了原始的 CSPDarknet53 后, mAP 略微降低,但是模型尺寸减小了 38.7%, FPS 提高了 88.6 fps。

CloAttention 注意力结构效果:相较于原始模型,通过在 backbone 与 neck 连接处引入 CloAttention 机制, mAP 提升了 2.2%;与 FasterNet 结合后, mAP 比 YOLOv7 提高了 1.1%,模型尺寸减小了 29.7%,保持较高精度的同时降低模型尺寸,提升了模型的检测速度。

MS-ASFF 结构效果:在 FasterNet 和 CloAttention 机制的基础上,继续引入 MS-ASFF 模块, mAP 增高了 3.8%;改进后的模型相比于原始模型, mAP 提升了 4.9%,FPS 提高了 13.4 fps。这表明 MS-ASFF 模块学习了多尺度特征,提升了对中小目标的检测识别能力。

表 4 中展示了在原始模型上分别添加各模块后的效果以及各模块相加合后对模型整体性能的影响。

## 3) 不同算法的性能比较

Pruning 的效果:继续采用剪枝策略改进模型,删除不重要的通道,过滤冗余信息,进一步提升检测速度。剪枝后的模型尺寸减小了 15.1%,FPS 提高了 27.4%。

综合对比本文模型和原始模型,模型改进后 mAP 提高了 4.1%,检测速度提升了 38.5%,使其更有利于钢轨缺陷的实时检测。

## 3) 不同算法的性能比较

为了验证本文提出的 YOLO-FCA 模型的性能以及实用性,选取了 YOLOv7、PC-YOLOv7,以及 KThin-YOLOv7,4 种模型保持同一数据集在本文 2.2 节实验环境下进行训练,采用 mAP(平均精度均值)、FPS(检测速度)、模型尺寸作为评价指标来评价模型的性能以及实用性。实验结果如表 5 所示。

表 5 与其他算法的性能比较

Table 5 Performance comparison with other algorithms

Methods	Cracks	Scratches	Bolts	Clips	Peel_off_blocks	Wounds	Spikes	mAP/%	FPS/fps	模型尺寸/M
YOLOv7	82.5	36.7	94.2	86.3	82.7	68.8	85.3	76.6	153.4	71.4
PC-YOLOv7	83.4	38.5	94.8	85.0	83.4	70.4	88.0	78.4	146.7	75.2
KThin-YOLOv7	84.2	41.3	95.0	86.8	84.2	70.8	92.3	79.8	177.6	74.8
YOLO-FCA	<b>85.6</b>	<b>43.6</b>	<b>95.3</b>	<b>88.2</b>	<b>85.0</b>	<b>72.8</b>	<b>95.0</b>	<b>80.7</b>	<b>212.5</b>	<b>76.7</b>

表 5 中的数据展示了 3 次试验的平均结果。研究表明,本文提出的 YOLO-FCA 算法在检测精度和速度方面都表现出色,特别是在 Scratches、Wounds 等多尺度缺陷的检测中表现优异,并且模型的复杂度较低。因此,可以认为 YOLO-FCA 在这些方面具有一定的优势。

为了直观地展示本研究提出方法的性能优势,对模型 YOLOv7、PC-YOLOv7、KThin-YOLOv7 以及本文提出

的模型 YOLO-FCA 的检测效果进行可视化分析。如图 9 所示,这些图像展示了不同模型在铁路钢轨缺陷数据集上的实际检测结果。

根据图 9 结果可以看出, YOLO-FCA 相比于原始模型及其几种改进算法网络,模型的检测精度更高,且在细粒度缺陷的检测效果也有大幅度提升。在复杂缺陷的模型表现上, YOLO-FCA 的漏检情况远远少于其他模型,能

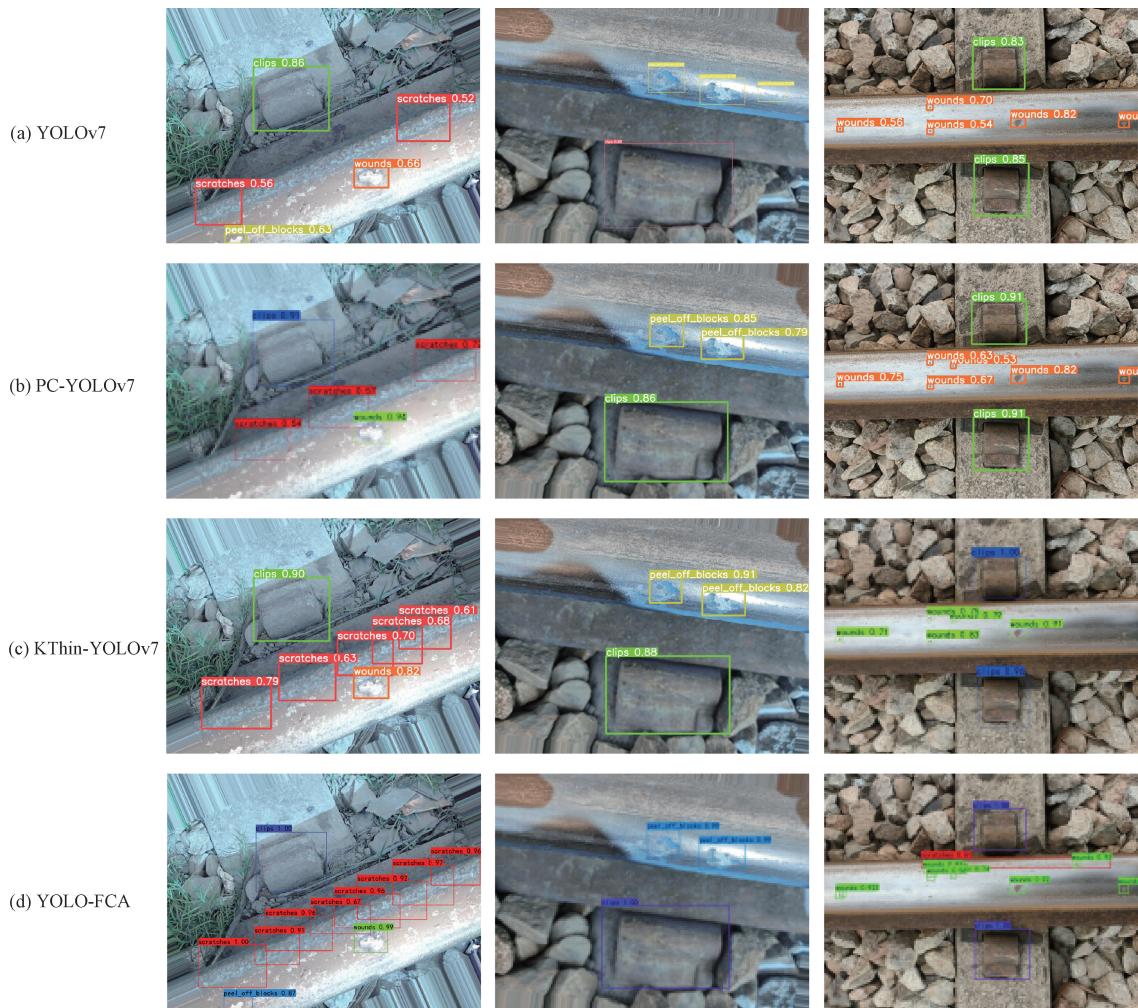


图 9 算法检测结果对比  
Fig. 9 Comparison of algorithm detection results

够更全面地检测钢轨缺陷。可见最终 3 个改进方法都使用之后,即本文提出的 YOLO-FCA 模型在工业检测场景下对钢轨缺陷检测拥有最佳的检测性能。

### 3 结 论

本研究针对钢轨细粒度缺陷难以检测等问题,提出了一种基于 YOLO-FCA 的深度学习模型解决方案,该算法能够快速且准确的发现钢轨上的潜在缺陷,减少因钢轨问题引发的事故风险。在原始的 YOLOv7 模型基础上采用轻量化的 FasterNet 替换原有骨干网络,并整合 CloAttention 注意力机制进行特征提取,提高了准确性和检测速度。引入多尺度自适应空间特征融合 MS-ASFF 进一步增强模型尺度不变性和语义特征利用率,在多尺度特征提取方面有大幅度提升。通过剪枝算法有效减小模型尺寸,保持高精度和检测速度,达到了钢轨缺陷检测的准确性和实时性要求。

实验结果显示, YOLO-FCA 在钢轨缺陷检测中的平

均检测精度达到 80.7%,检测速度为 212.5 fps,比 YOLOv7 分别提高了 4.1% 和 59.1 fps。与主流的深度学习检测模型相比,本文提出的方法不仅显著提高了缺陷检测的精确度和全面性,还大幅提升了检测速度,为钢轨缺陷检测在实际工业环境中的应用提供了新的可能性。未来的研究将重点针对钢轨缺陷和铁路复杂环境的特性,进一步优化网络设计,以增强模型的泛化能力,并开发更为精确的钢轨缺陷检测模型,以推动钢轨缺陷自动检测技术的发展。

### 参考文献

- [1] AYDIN I, AKIN E, KARAKOSE M. Defect classification based on deep features for railway tracks in sustainable transportation [J]. Applied Soft Computing, 2021, 111(7):107706.
- [2] GE H, HUAT D C K, KOH C G, et al. Guided wave-based rail flaw detection technologies: State-of-the-art review [J]. Structural Health Monitoring, 2022, 21(3):1287-1308.

- [3] YE J H, NI R H, HSU Q C. Image feature analysis for magnetic particle inspection of forging defects[J]. Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 2022, 236 ( 14 ): 1923-1929.
- [4] RODRÍGUEZ-MARTÍN M, LAGÜELA S, GONZÁLEZ-AGUILERA D, et al. Prediction of depth model for cracks in steel using infrared thermography[J]. Infrared Physics & Technology, 2015, 71: 492-500.
- [5] HUA L X, LU Y, DENG J H, et al. 3D reconstruction of concrete defects using optical laser triangulation and modified spacetime analysis [ J ]. Automation in Construction, 2022, 142:104469.
- [6] GALAN U, ORTA P, KURFESS T, et al. Surface defect identification and measurement for metal castings by vision system[J]. Manufacturing Letters, 2018, 15:5-8.
- [7] 刘子洋,徐慧英,朱信忠,等. Bi-YOLO:一种基于YOLOv8n改进的轻量化目标检测算法[J].计算机工程与科学,2024,46(8):1444-1454.  
LIU Z Y, XU H Y, ZHU X ZH, et al. Bi-YOLO: An improved lightweight object detection algorithm based on YOLOv8n[J]. Computer Engineering and Science, 2024,46(8):1444-1454.
- [8] 刘彦清. 基于 YOLO 系列的目标检测改进算法[D]. 长春:吉林大学,2021.  
LIU Y Q. An improved object detection algorithm based on YOLO series [ D ]. Changchun: Jilin University, 2021.
- [9] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv preprint arXiv: 1804.02767, 2018.
- [10] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [ J ]. ArXiv preprint arXiv: 2004.10934, 2020.
- [11] ZHAO Y, SHI Y B, WANG Z L. The improved YOLOv5 algorithm and its application in small target detection[C]. Intelligent Robotics and Applications, 2022: 679-688.
- [12] 熊聪,于安宁,高兴华,等. 基于改进 YOLOX 的钢材表面缺陷检测算法[J]. 电子测量技术,2023,46(9): 151-157.  
XIONG C, YU AN N, GAO X H, et al. Steel surface defect detection algorithm based on improved YOLOX[J]. Electronic Measurement Technology, 2023, 46 ( 9 ): 151-157.
- [13] PANIGRAHI S, RAJU U S N. MS-ML-SNYOLOv3: A robust lightweight modification of SqueezeNet based YOLOv3 for pedestrian detection[J]. Optik, 2022, 260:169061.
- [14] HAN CH Y, ZHU J L, LI F, et al. Design of lightweight pedestrian detection network in railway scenes[J]. Journal of Physics: Conference Series, 2020, 1544(1):012053.
- [15] 赵春华,罗顺,谭金铃,等. 基于 PC-YOLOv7 算法钢材表面缺陷检测[J]. 国外电子测量技术,2023,42(9): 137-145.  
ZHAO CH H, LUO SH, TAN J L, et al. Steel surface defect detection based on PC-YOLOv7 algorithm [ J ]. Foreign Electronic Measurement Technology, 2023, 42(9):137-145.
- [16] 卢开喜,段先华,陶宇诚,等. KThin-YOLOV7:轻量级的焊接件表面缺陷检测[J]. 电子测量技术,2024, 47(7):9-18.  
LU K X, DUAN X H, TAO Y CH, et al. KThin-YOLOV7: Lightweight surface defect detection of welded parts [ J ]. Electronic Measurement Technology, 2024, 47(7): 9-18.
- [17] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2023.
- [18] CHEN J, KAO S H, HE H, et al. Run, don't walk: Chasing higher flops for faster neural networks[C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2023.
- [19] FAN Q H, HUANG H B, GUAN J Y, et al. Rethinking local perception in lightweight vision transformer [ J ]. ArXiv preprint arXiv: 2303.17803, 2023.
- [20] CHENG X, YU J B. RetinaNet with difference channel attention and adaptively spatial feature fusion for steel surface defect detection [ J ]. IEEE Transactions on Instrumentation and Measurement, 2021,70:1-11.
- [21] LIU ZH, LI J G, SHEN ZH Q, et al. Learning efficient convolutional networks through network slimming[C]. 2017 IEEE International Conference on Computer Vision(ICCV), 2017.
- [22] BARGOTI S, UNDERWOOD J. Deep fruit detection in orchards[C]. 2017 IEEE International Conference on Robotics and Automation ( ICRA ), 2017: 3626-3633.

## 作者简介

胡峻峰(通信作者),博士,副教授,主要研究方向为模式识别、智能控制。

E-mail:nefuhujunfeng@126.com

赵亚凤,博士,副教授,主要研究方向为机器学习及其在林业领域的应用。

E-mail:zyf@nefu.edu.cn

宋文华,硕士研究生,主要研究方向为深度学习、目标检测。

E-mail:15272656701@163.com

刘晓璐,硕士研究生,主要研究方向为深度学习、语义分割。

E-mail:liuxiaolu9817@163.com