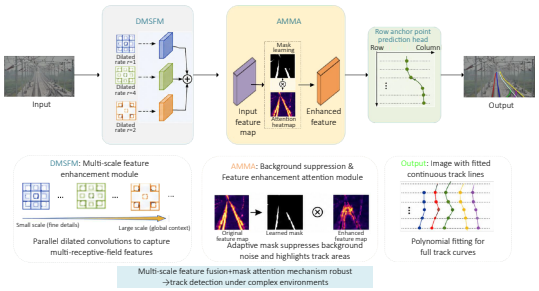
OEE | 光电工程 Opto-Electronic Engineering

Article

2026 年 6 月, 第 53 卷, 第 6 期

多尺度特征融合的铁路轨道检测



曹启平[1,2], 林珊玲[1,2], 林坚普[1,2], 吕珊红[1,2], 林志贤[1,2,3*], 郭太良[2,3]

[1] 福州大学先进制造学院, 福建 泉州 362251;
[2] 中国福建光电信息科学与技术实验室, 福建 福州 350116;
[3] 福州大学物理与信息工程学院, 福建 福州 350116

摘要: 针对复杂铁路场景中轨道线检测易受背景干扰、远尺度特征表达能力不足的问题, 本文提出一种融合多尺度特征增强的轨道检测模型。该模型在行锚点检测框架的基础上, 引入空洞多尺度融合模块以增强对弯曲与交叉轨道的几何特征捕获能力; 同时构建自适应多头掩码注意力机制, 结合 SimAM 无参注意力与 DropBlock 正则化, 有效抑制背景干扰并提升特征判别性。在 Rail-DB 数据集上的实验结果表明, 所提方法检测准确率达到 94.53%, 较基线模型提升 1.73%, 在夜间、遮挡等极端场景下亦表现出良好鲁棒性。该方法在保持轻量化的同时兼顾检测精度与泛化能力, 为铁路智能安防系统提供了一种可行技术方案。

关键词: 轨道线检测; 行锚点; 多头掩码注意力; 残差空洞卷积; Rail-DB 数据集

The English title and abstract appear at the end of the paper.

DOI: 10.12086/oee.2026.250315 | CSTR: 32245.14.oee.2026.250315 | 中图分类号: TP391.4 | 文献标识码: A

曹启平, 林珊玲, 林坚普, 等. 多尺度特征融合的铁路轨道检测 [J]. 光电工程, 2026, **53**(6): 250315

Cao Q P, Lin S L, Lin J P, et al. Railway track detection based on multi-scale feature fusion[J]. *Opto-Electron Eng*, 2026, **53**(6): 250315

1 引 言

随着轨道交通向高速化方向发展, 列车运行速度不断提升, 这对安全监控系统提出了更为严苛的要求。轨道异物侵限检测作为保障行车安全的核心环节, 其有效性直接关系到列车运行安全[1]。在铁路系统中, 列车必须严格沿轨道行驶, 因此精准的轨道线检测成为划定行车限界的前提与核心技术, 也是当前智能铁路研究的热点问题。

随着计算机视觉技术的发展, 轨道检测方法经历了由传统图像处理向深度学习方法的演进。早期研究主要依赖边缘检测、几何特征提取以及模型拟合等方法实现轨道定位, 例如 Wang 等利用 BM3D 去噪与 Zernike 矩实现轨道边缘检测[2]。此外, Xu 等对基于深度学习的轨道图像识别技术进行了系统总结, 指出深度学习方法在复杂场景中的优势[3]。传统图像处理范式通过边缘与直线特征实现轨道线检测, 其中陈明曦等[4] 以改进 Canny 算子自适应阈值搜索边缘提升精度, 却易丢失边缘细节; 林向会[5] 以 Sobel 与 Canny 级联检测后利用霍夫变换与最小二乘拟合直线轨道; 赵鸿亮等[6] 以加权灰度化、高斯滤波去噪, 再经 Sobel 边缘提取并以三次函数拟合曲线; 盛峰等[7] 则基于视觉传感器, 以帧间差分与灰度特征场景识别后, 采用自适应边缘检测提取轨道轮廓, 并结合滑动窗口与卡尔曼滤波完成精细分割。

除此之外, 国内外研究者还系统性地提出了一类面

收稿日期: 2025-10-21; 修回日期: 2026-02-27; 录用日期: 2026-02-28; 网络出版日期: 2026-06-25
基金项目: 国家重点研发计划 (2021YFB3600603)
*通信作者: 林志贤, lzx2005000@163.com
版权所有©2026 中国科学院光电技术研究所

向结构化道路的视觉检测范式，该类范式可细分为“基于特征的识别”与“基于模板的识别”两大分支[8]。其中，基于特征的识别策略核心在于充分利用轨道与枕木、道砟在纹理、灰度、方向等属性上存在的固有差异，从而实现轨道区域的可靠分割[9-10]；典型工作如居耀勇[11]所提出的“主方向迭代矫正”轨道检测算法，即通过持续估计并修正轨道主走向信息来完成轨道线的精准提取。尽管该方法在复杂背景下表现出较高的鲁棒性，但其迭代优化过程涉及大量方向计算与矫正操作，当需要对视频序列进行逐帧处理时，数据量与计算量均呈指数级增长，严重制约了处理效率，导致系统难以满足实时性要求。

基于深度学习的轨道线检测算法普遍利用卷积神经网络自动提取特征。Wang 等[12]率先将铁路检测任务形式化为语义分割问题，提出专门面向铁路场景的网络，实现像素级轨道区域分割。刘力[13]则提出 UNet 框架对轨道像素类别进行预测，并沿横向维度提取轨道位置特征，随后借助多项式函数对轨道走向进行拟合，完成轨道线提取。王辉等[14]进一步引入行锚框思想，将传统逐像素分割改为逐网格预测，同时利用轨道连续性先验实现限界区域智能识别。随着目标检测研究的深入，部分学者尝试将通用目标检测框架迁移至轨道线检测任务[15]；然而，该类方法采用矩形边界框对轨道进行定位，而轨道线通常呈细长且可能弯曲的几何形态，矩形框难以紧密贴合，导致检测精度受限。

相比之下，公路场景的车道线检测技术已相对成熟，鉴于轨道线与车道线在几何形态上的相似性——均呈现细长、连贯的线性结构，将成熟的公路检测方法迁移至铁路场景具备可行性。当前车道线检测方法主要分为四类：将车道线识别形式化为语义分割任务[13-14]；基于关键点的方法将车道线上的像素视为离散关键点[15-16]；基于参数曲线的方法通过回归多项式系数直接预测车道形状[17-18]；基于锚的方法通过回归预设锚点的位置偏移来刻画车道线形状。在四类方法中基于锚的方法精度最高，可分为行锚与线锚两类。Ko 等[19]提出一种基于关键点估计与点实例分割的车道线检测方法，通过联合关键点定位和点实例分割实现车道线实例的精确识别，有效提升了复杂场景下车道线检测的准确性和鲁棒性。Wang 等[20]提出了一种基于关键点的全局关联网络 (global association network，GANet)，通过全局关联模块建立关键点之间的空间关系，增强了车道线的连续性表达，提高了遮挡、弯道等复杂场景下的检测性能。Tabelini 等[21]提出 PolyLaneNet 模型，将车道线检测建模为多项式回归问题，通过直接预测车道线多项式参数实现端到端检测，在保证检测精度的同时兼顾了推理效率。LSTR[22]利用 Transformer结构建模全局上下文信息；UFLD[23]通过行锚点分类策略实现实时车道线检测；Tabelini 等引入注意力机制指导车道线检测，提高了复杂场景下的特征提取能力[24]；Li 等提出基于行表示的轨道检测网络，实现了高效的轨道线检测，并构建了新的轨道检测基准数据集[25]。此外为兼顾检测精度与实时性，Hou 等提出了基于自注意力蒸馏的轻量化车道线检测网络，有效提升了轻量模型的检测性能[26]。Zheng 等提出跨层细化网络 CLRNet，通过利用高层语义特征进行初始检测，并结合低层细节特征对车道线位置进行逐层细化，同时引入 ROIGather 模块增强全局上下文信息，从而显著提升了车道线的定位精度和检测性能[27]。

随着计算机视觉技术的不断发展，目标检测算法在各种应用中取得了显著进展。在轨道检测任务中，准确识别轨道线及其他小目标对于智能铁路系统至关重要。近期，栾庆磊等[28]提出基于 YOLOv8 的检测方法在目标检测领域取得了良好的成果，尤其是在红外道路目标检测方面。付明雨等[29]YOLOv8 的轻量化改进使得轨道小目标的异物检测成为可能，显著提升了检测精度。此外，张进等[30]提出了一种融合元素乘法和细节优化的道路提取算法，表明通过细节优化和多尺度特征融合可以有效提升道路检测的精度。Yu 等[31]提出了基于动态锚线的轨道检测网络 DALNet，通过动态建模轨道结构特征提升了复杂场景下的轨道检测性能。然而，现有方法在复杂场景中仍面临挑战，特别是对于弯曲轨道、交叉轨道等结构的检测。

综合分析现有方法，轨道线检测仍面临诸多挑战：铁路场景相比公路环境更为复杂，轨道线常分布于山区、林地和隧道等自然环境，背景中存在大量与轨道颜色、纹理相近的干扰信息；多条轨道密集排列容易产生遮挡；光照条件变化剧烈且存在恶劣天气影响。现有方法在特征提取能力、多尺度适应性、背景干扰抑制等方面存在不足。

针对上述问题，本文在 RailNet 框架基础上，提出了一种融合多注意力机制的轨道线检测网络 (dilated and masked attention network，DMA-Net)。主要创新包括：1) 设计空洞多尺度融合模块 (dilated multi-scale fusion module，DMSFM)，通过并行空洞卷积增强多尺度上下文感知能力，使得模型能够同时应对不同尺度的目标，尤其是轨道线的弯曲、交叉等复杂结构；2) 提出自适应多头掩码注意力 (adaptive multi-head mask attention，AMMA) 机制，能够有效抑制背景干扰，使得模型能够在复杂背景下保持高鲁棒性，并专注于轨道线区域；3) 引入

无参数注意力模块 (simple attention module，SimAM) 与 DropBlock 正则化，增强了模型在恶劣环境下的鲁棒性，使得模型能够适应不同光照、气候条件，减少过拟合，提升对新场景的泛化能力。通过综合改进，本文方法在复杂铁路场景下实现了更高精度的轨道线检测性能。

2 本文算法

2.1 系统整体框架

本文提出的轨道线检测方法基于行锚点检测框架，其整体结构如图 1 所示。首先，输入图像经过改进的 ResNet18 主干网络提取多尺度特征；随后，利用空洞卷积与多尺度融合模块 DMSFM 整合不同感受野信息，增强特征的上下文表达能力。在此基础上，引入 AMMA，通过可学习的二值掩码有效抑制背景区域的干扰，突出轨道线的显著特征。同时，结合 SimAM 空间注意力机制进一步提升像素级响应能力，并通过 DropBlock 正则化减少过拟合风险。最后，经过分类预测层输出每行轨道线的列坐标，结合多项式拟合恢复完整的轨道线。

2.2 空洞卷积与多尺度融合模块 (DMSFM)

为了克服传统卷积在提取全局信息时感受野不足的问题，同时避免单一空洞卷积产生的“栅格效应”(即感知区域的不连续性)，本文设计了 DMSFM，其结构如图 2 所示。该模块通过并行设置不同空洞率的卷积分支来提取多尺度上下文特征，从而增强模型对不同尺度信息的感知能力。

空洞卷积 (Dilated convolution) 的原理：空洞卷积是通过在卷积核中的元素间插入间隔 (空洞) 来增加感受野的大小，而无需增加额外的参数或计算量。在标准卷积中，卷积核的元素是连续的，而在空洞卷积中，元素之间通过空洞进行分隔，从而能够在每次卷积操作中访问到更大的区域。

空洞卷积的一个重要特点是通过调整空洞率 (dilation rate) 来扩展感受野。较小的空洞率 (1、2) 通常用于捕捉局部特征，而较大的空洞率 (4) 能够捕捉到远距离的上下文信息。因此，通过使用多个不同空洞率的卷积分支，DMSFM 能够同时提取局部细节与全局信息，从而提升特征的多尺度表达能力。

设输入特征图为 F，空洞率分别为 r_1, r_2, r_3，则每个空洞率对应一个独立的卷积分支:

$$F_i = Conv_{d=r_i}(F), i = 1,2,3\,. \tag{1}$$

具体来说，小的空洞率能够捕捉局部细节信息，而大的空洞率则能提取更远距离的上下文信息。通过这种方式，空洞卷积可以同时在局部和全局范围内提取特征。

多尺度特征融合：为了进一步增强对多尺度上下文的表达能力，DMSFM 采用了并行多个不同空洞率的卷积分支。每个分支通过不同空洞率的卷积操作提取不同尺度的特征信息。设每个分支的输出为 F_{out}，则最终的输出通过加权融合的方式得到：

$$F_{\text{out}} = \sum_{i=1}^{3} w_i \cdot F_i\,, \tag{2}$$

其中：w_i为可学习权重。该模块能够自适应选择合适的感受野，兼顾局部与全局信息，提高轨道线特征的完整性与鲁棒性。这种融合机制能够根据输入数据的具体情况，自适应地选择合适的感受野，从而有效地同时处理局部和全局信息。

多尺度上下文特征的提取：通过并行设置多个空洞率的卷积分支，DMSFM 模块可以同时从多个尺度提取上下文信息，从而有效增强轨道线检测模型的鲁棒性。局部特征由小空洞率的卷积负责捕获，而较大的空洞率则有助于捕捉更长程的上下文信息。这种多尺度特征融合机制能够在保证计算效率的同时，显著提升对弯曲轨道、交叉轨道等复杂场景中轨道线特征的表达能力。

通过这种设计，DMSFM 模块能够自适应地选择最合适的感受野，在不增加大量计算复杂度的前提下，有效地增强模型对不同尺度轨道线特征的感知与识别能力。

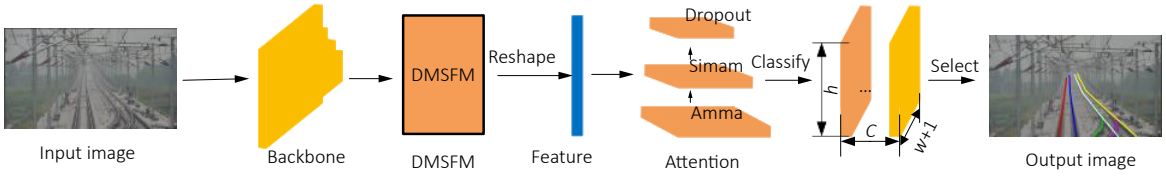


图 1 DMA-Net 整体架构图

Fig. 1 Overall architecture of DMA-Net

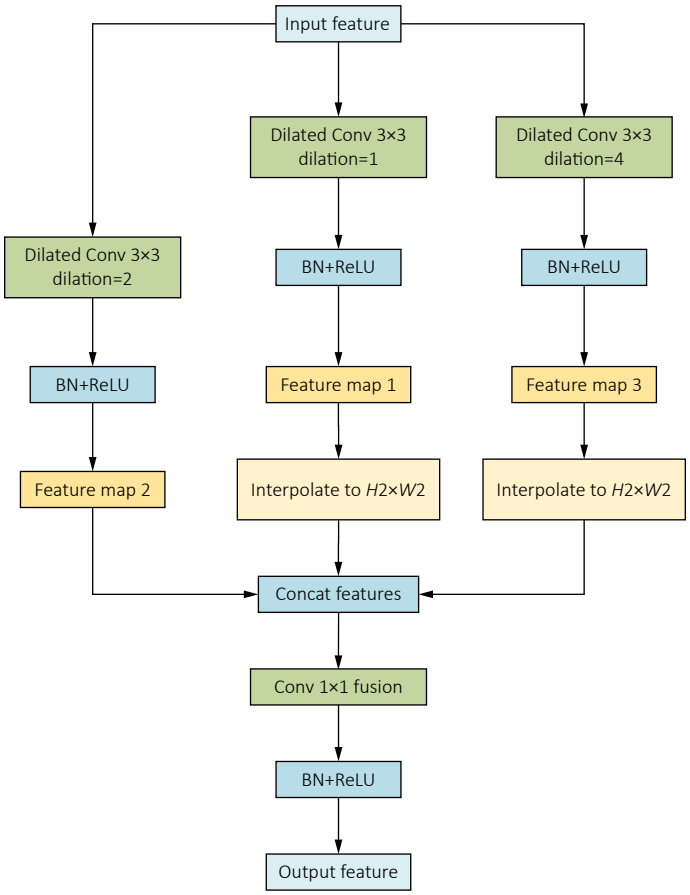


图 2 DMSFM 模块结构

Fig. 2 Architecture of the DMSFM module

2.3 自适应多头掩码注意力机制 (AMMA)

在复杂铁路场景中，轨道线常与背景中的道砟、枕木及植被等纹理特征相似，传统注意力机制在对全局特征加权时，容易将冗余区域与轨道线一并突出，从而导致误检。为此，本文提出自适应多头掩码注意力机制 AMMA，其整体结构如图 3 所示，包含核心注意力计算单元和注意力输出与特征重构单元两个主要部分。与传统注意力相比，AMMA 机制的作用有以下几个方面。

1) 区域抑制性：AMMA 通过引入可学习的掩码，对输入特征进行调整，从而动态抑制非目标区域 (如道砟、植被等) 对模型的干扰。该机制通过掩码的学习，使得模型能够自动识别并忽略与轨道线不相关的区域，减少背景干扰的影响。

2) 自适应性：AMMA 中的掩码权重是通过训练自适应学习的，可以根据不同场景的复杂度动态调整。尤其在背景纹理复杂、轨道线与背景高度相似的情况下，AMMA 能够有效聚焦轨道线的关键信息。

3) 轻量化：AMMA 机制相比传统注意力机制具有较低的计算开销，尽管保留了多头注意力的建模能力，但通过掩码机制大幅减少无关区域的计算，从而保持了较高的计算效率。

2.3.1 核心注意力计算单元

核心注意力计算单元负责对输入特征进行多头自注意力计算，流程如下。

1) 输入特征投影：将输入特征分别通过线性投影层得到查询 ($\boldsymbol{Q}$)、键 ($\boldsymbol{K}$) 和值 ($\boldsymbol{V}$) 矩阵。

2) 多头重组：将 $\boldsymbol{Q}$，$\boldsymbol{K}$，$\boldsymbol{V}$ 按头数进行切分与重组，以便并行计算多个子空间的注意力，多头机制允许模型在不同的子空间中并行地关注来自不同表示子空间的信息。

3) 注意力得分计算：计算缩放点积注意力得分，公式为

$$Attention(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = Softmax\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_{\mathrm{k}}}}\right)\boldsymbol{V}, \tag{3}$$

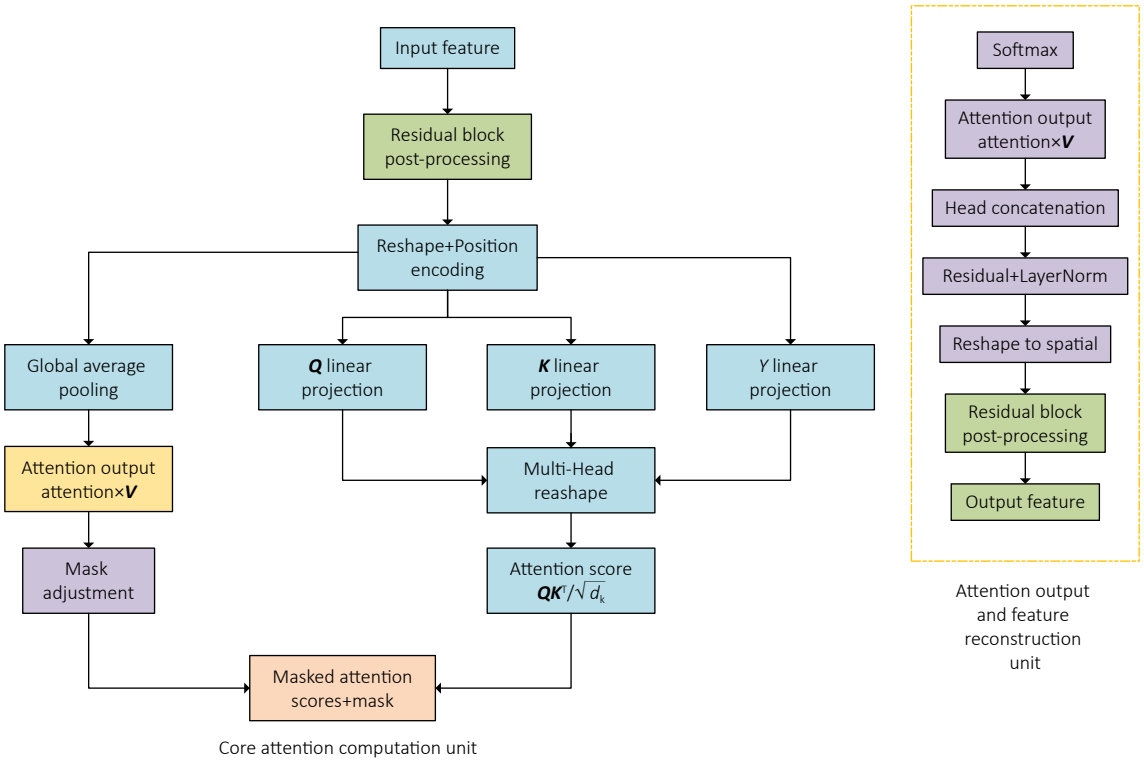


图 3 AMMA 模块结构图

Fig. 3 Architecture of the AMMA module

其中：d_k为键向量的维度，$\boldsymbol{Q}$，$\boldsymbol{K}^{\mathrm{T}}$计算了所有位置之间的相似度，形成一个注意力图。缩放因子 $\sqrt{d_k}$用于防止在维度d_k较高时点积结果过大，导致 Softmax 函数梯度消失。Softmax 函数将相似度归一化为概率分布，代表了其他所有位置对当前位置的“重要程度”。最后与 $\boldsymbol{V}$ 加权求和，实现根据全局上下文信息对每个位置的特征进行增强。

4) 掩码调节：在注意力得分上施加可学习的掩码，以自适应地增强或抑制不同位置之间的关联强度，提升模型对局部结构的感知能力。

这是 AMMA 模块的核心创新之一。在复杂的铁路场景中，背景 (如道砟、植被) 与前景 (轨道) 特征相似，传统的注意力机制会不加区分地赋予高响应。可学习掩码的作用是自适应地调节注意力。

2.3.2 注意力输出与特征重构单元

该单元负责将注意力输出重新整合为空间特征图，并进行后续处理。

1) 头拼接与线性投影：将多个注意力头的输出拼接后通过线性投影层融合信息，将多个注意力头捕捉到的不同子空间的表征信息进行集成。线性投影层负责将这些信息融合，并压缩到合适的维度，为后续处理做准备。

2) 残差连接与层归一化：将投影结果与输入特征进行残差连接，并通过层归一化稳定训练过程，遵循 ResNet 的思想，有效地缓解了深层网络中的梯度消失问题，确保了训练过程的稳定性。

3) 空间重塑：将序列形式的特征重新还原为空间维度，恢复特征图结构，便于进行自注意力计算的序列数据，转换回卷积神经网络擅长处理的空间网格数据，为后续可能存在的卷积操作以及最终的特征图输出做好准备。

4) 后处理残差块：使用包含卷积、激活函数的残差块进一步提炼特征，增强局部建模能力，注意力机制是全局的、内容相关的，但对局部细节的建模能力可能弱于卷积。此残差块作为一个强大的局部特征提炼器，对全局增强后的特征进行进一步的局部优化和非线性变换，增强模型的表示能力，并确保输出特征在局部空间上是平滑和连贯的。

5) 全局平均池化与位置编码：在部分分支中引入全局平均池化获取全局上下文，并通过位置编码保留空间位置信息。由于自注意力机制本身是置换不变的，它无

法感知特征在空间中的绝对或相对位置。加入位置编码后，模型能够利用轨道的空间连续性和几何结构这一重要先验知识，对于生成连贯、平滑的轨道线至关重要。

2.4 SimAM 空间注意力机制

引入 SimAM 空间注意力模块，用于增强像素级特征响应，提升模型对关键区域轨道线像素的关注能力。SimAM 模块基于神经元能量模型构建注意力图，其核心思想是通过计算每个空间位置的激活值与当前通道的均值之间的偏离程度，再加上其自身的平方项，来衡量该像素的重要性。这种方式能够帮助模型聚焦于关键区域并抑制背景噪声。SimAM 的作用如下。

1) 提升像素级响应能力：SimAM 模块计算每个空间位置的激活值与该通道均值的偏差，并结合偏差的平方来生成注意力权重。该方式使得模型能够更加关注轨道区域的特征，从而提高特征响应能力。

2) 增强特征判别力：通过这种方式，SimAM 能够提升模型对轨道线像素的关注，使其在复杂背景下，尤其是背景和轨道线特征相似的情况下，具备更强的判别能力。

3) 无参数设计：SimAM 不需要额外的学习参数，其计算依赖于输入特征的激活值与均值的简单运算，减少了额外的计算开销。这使得 SimAM 在实时应用中非常有效，尤其适合要求较高实时性的轨道检测任务。

所述 SimAM 空间注意力模块基于以下能量函数生成像素级注意力权重：

$$E(x)=\frac{(x-\mu)^2}{\sigma^2-\varepsilon}+\lambda x^2\,, \tag{4}$$

式中：x 表示单个像素位置的激活值；μ 表示通道内特征的均值；σ^2 表示通道内特征的方差；ε 表示用于防止分母为零的稳定性常数；λ 表示调控注意力幅度的权重项。该能量函数通过计算每个像素与通道均值的偏离程度，结合方差信息，来衡量像素的重要性，从而生成注意力图。SimAM 使得模型在进行特征选择时更加精确地聚焦于轨道线区域，提升了轨道线的检测能力和模型的整体鲁棒性。

2.5 DropBlock 正则化模块

为缓解模型过拟合，提升在遮挡等复杂场景下的鲁棒性，本文在网络深层引入 DropBlock 正则化模块。该模块并非随机丢弃单个神经元，而是随机丢弃连续的区域块 (block)，其数学表达为

$$M_{ij}\sim Bernoulli(p)\,, \tag{5}$$

其中：p 为丢弃概率，生成了一个二值掩码后，通过最大池化操作扩散丢弃区域：

$$M_{\text{drop}}=MaxPool2d(M,kernel_size=b,\ stride=1,padding=b//2)\,. \tag{6}$$

最终输出特征为

$$Y=X\odot M_{\text{drop}}\,. \tag{7}$$

其中：X 表示输入到 DropBlock 模块的特征图，M_{drop} 为经过随机采样及块扩散操作得到的二值掩码矩阵，$\odot$ 表示逐元素乘法运算。通过将输入特征图与掩码矩阵相乘，可随机屏蔽连续区域的特征响应，从而降低特征间的共适应性并提升模型的泛化能力。

训练中，丢弃概率 p 从 0.05 线性递增至 0.2，块大小 b 设置为 5，从而随训练进程逐步增强正则化强度。

3 实验结果与分析

3.1 实验数据集

本文使用的是 Rail-DB 数据集，Rail-DB 是一个专门为铁路轨道线检测任务设计的高质量数据集，它包含多样化的场景和精确的标注，适用于评估和训练深度学习模型。Rail-DB 数据集由 7432 对图像及其对应的标注组成，如图 4 所示，这些图像覆盖了多种不同的环境条件，包括不同的光照 (如晴天、雨天、夜晚)、道路结构 (如直线、曲线、交叉道口) 和视角 (如近景、远景)。这样的设计确保了数据集能够全面地覆盖铁路轨道可能出现的各种情况，从而提高模型的泛化能力。为了评估模型的性能，将 7432 对图像随机划分为训练集、验证集，比例分别为 80% 和 20%。这种划分方式确保了模型在看不见的数据上也能表现出良好的泛化能力。

通过使用 Rail-DB 数据集，可以全面评估所提出的轨道线检测方法在不同场景下的性能，包括其准确性、鲁棒性和泛化能力。

3.2 实验环境及参数配置

本研究的实验环境是 Ubuntu 20.04 操作系统，搭载 NVIDIA RTX 3090，30GB RAM， 以及 Intel (R) Xeon (R) CPU E5-2680 v4 @ 2.40GHz * 7 核 CPU。本研究是在 Python 3.8 运行环境中实现的，使用 Pytorch 1.12.1 深度学习框架，CUDA 11.3 加速库。为了增强模型泛化能

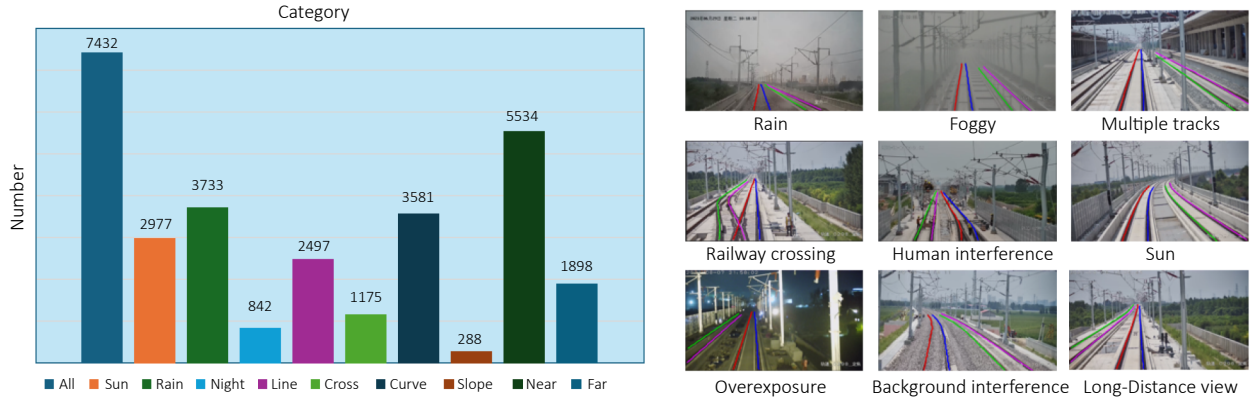


图 4 数据集类型及其数量

Fig. 4 Dataset category statistics

力，采用随机水平翻转和随机仿射变换 (旋转、平移、缩放) 等方式进行数据增强。采样点数设置为 52，每张图像中可识别轨道线数量最大为 4。采用 AdamW 优化器，初始学习率设置为 4×10^{-4}，并采用余弦退火策略衰减。训练轮次为 100，批次大小为 32。

3.3 评价指标

本文选用准确率 (*Acc*) 与 F_1 分数作为评估指标。

准确率的计算公式如下：

$$Acc=\frac{\sum_{\text{clip}}C_{\text{clip}}}{\sum_{\text{clip}}S_{\text{clip}}}, \tag{8}$$

其中：C_{clip}表示图像中正确预测的轨道点数量，S_{clip}表示图像中所有真实轨道点的总数。如果真实值与预测值之间的差异宽度小于阈值，则认为预测点是正确的。默认将 Rail-DB 的阈值设置为 6 pixel。此外，还采用 F_1 分数进行综合评价。通过统计真阳性 (*TP*)、假阳性 (*FP*) 和假阴性 (*FN*) 的数量，分别计算精确率 (Precision) 和召回率 (Recall)，进而得到 F_1 分数，其计算公式如下：

$$F_1=\frac{2\times P_{\text{precision}}\times R_{\text{Recall}}}{P_{\text{precision}}+R_{\text{Recall}}}, \tag{9}$$

$$P_{\text{precision}}=\frac{TP}{TP+FP}, \tag{10}$$

$$R_{\text{Recall}}=\frac{TP}{TP+FN}. \tag{11}$$

F_1 分数是精确率 Precision 和召回率 Recall 的调和平均数，能够综合反映模型的性能。除上述指标外，本文还比较了模型的推理速度，以每秒检测帧数 (frames per second, FPS) 表示。

3.4 对比实验

3.4.1 对比方法介绍

为便于理解表 1 中各基线方法的差异，现对其核心

表 1 对比试验结果

Table 1 Results of comparison experiments

Method	Backbone	Acc/%	F_1/%	FPS	GFLOPS
UFLD[23]	18	84.57	83.65	272	9.60
UFLD	34	84.82	85.55	160	16.32
LaneATT[24]	18	86.32	85.44	240	10.88
LaneATT[24]	34	86.61	86.03	153	17.07
ClrNet[27]	18	87.80	87.04	200	13.05
ClrNet[27]	34	88.04	88.64	151	17.30
DALNet[30]	18	93.26	93.41	255	9.75
DALNet[31]	34	93.64	93.90	160	16.58
Rail-Net[25]	18	92.80	92.20	**310**	**8.42**
Rail-Net[25]	34	93.13	92.55	165	15.82
DMA-Net (Ours)	18	**94.53**	**94.88**	268	9.80
DMA-Net (Ours)	34	94.32	94.61	155	17.38

思想进行简要说明。

1) UFLD：典型行锚 (row-anchor) 范式，将车道/轨道线检测转化为按行进行的离散位置分类问题，依赖全局特征对每一行的列坐标进行选择，速度快但在强遮挡与远距离细节上容易受限。

2) LaneATT：典型线锚 (line-anchor) 范式，将检测转化为对预设锚线的回归，并用注意力机制聚合与锚线相关的全局上下文，提高线状结构在复杂场景下的可分性。

3) CLRNet：通过跨层特征交互与逐级细化策略，对初始预测进行多阶段几何修正，改善细长目标的定位精度，但在强干扰遮挡场景下仍可能出现局部断裂/偏移。

4) DALNet：提出动态锚线 (dynamic anchor line) 建模，通过起始点热力图与方向/斜率估计生成与轨道实例更贴合的锚线表示，并引入多尺度上下文增强远距离轨道一致性；在人员穿行等局部强遮挡下，动态锚线的局部证据不足时可能出现漏检或中断风险。

5) Rail-Net：面向铁路场景的行锚式轨道检测基线，利用行级结构化先验生成轨道候选并进行位置细化，整体效率高，但在复杂干扰与长程几何连续性建模上仍有提升空间。

综上，表 1 覆盖了行锚、线锚、跨层细化与动态锚线等代表性范式，可较全面地反映本文方法在不同主流建模思路下的相对优势与代价。

3.4.2 对比实验结果分析

为了评估所提 DMA-Net 模型的有效性，本文在铁路轨道数据集上将其与多种主流方法进行了对比实验。对比方法覆盖了典型的锚点/锚线范式：包括基于行锚的 UFLD 与 Rail-Net，基于线锚的 LaneATT 与 ClrNet；进一步加入了基于动态锚线思想的 DALNet (dynamic anchor line) 作为最新对比方法。所有方法在相同评价协议下进行对比，分别在 ResNet18 与 ResNet34 主干网络配置上测试，结果如表 1 所示。

从表 1 可以看出，DMA-Net 在两种主干网络配置下均取得了最优或接近最优的综合表现。其中，DMA-Net (ResNet18) 的准确率达到 94.53%，F_1 分数达到 94.88%，在所有对比方法中均为最高；DMA-Net (ResNet34) 的准确率与 F_1 分数分别为 94.32% 和 94.61%，同样保持在较高水平。与原始 Rail-Net 相比，DMA-Net 在精度上提升显著：以 ResNet18 为例，Rail-Net (ResNet18) 的 Acc 为 92.80%、F_1 为 92.20%，而 DMA-Net (ResNet18) 分别提升至 94.53% 与 94.88%，对应带来 1.73% (pp) 的 Acc 增益与 2.68 pp 的 F_1 增益，说明本文所提出的多尺度融合与注意力增强策略能够有效提升轨道线特征表达与整体检测性能。

在计算复杂度方面，DMA-Net (ResNet18) 的推理计算量为 9.80 GFLOPs/帧，相较 Rail-Net (ResNet18) 的 8.42 GFLOPs/帧仅增加 1.38 GFLOPs/帧 (约 16.4%)。对应地，推理速度由 Rail-Net (ResNet18) 的 310 f/s 降至 268 f/s，属于以小幅计算开销换取显著精度增益的改进模式。总体来看，DMA-Net 在精度、速度与复杂度之间实现了更优折中，综合性能优于其他对比方法。

进一步与 UFLD、LaneATT、ClrNet 等方法比较可见，DMA-Net 在检测准确率与 F_1 分数上均表现出明显优势，同时计算开销仍处于合理范围内。具体而言，UFLD (ResNet18/34) 的计算量为 9.60/16.32 GFLOPs/帧，LaneATT (ResNet18/34) 为 10.88/17.07 GFLOPs/帧，ClrNet (ResNet18/34) 为 13.05/17.30 GFLOPs/帧。可以看到，DMA-Net (ResNet18) 在获得更高精度的同时，其计算复杂度并未显著高于同类 ResNet18 配置方法，且相比 ClrNet-18 仍具有更低的 GFLOPs，体现了较好的效率优势。

值得注意的是，作为最新引入的对比方法，DALNet 在表 1 中同样展示了较强竞争力：DALNet (ResNet18) 的 Acc 与 F_1 分别达到 93.26 % 与 93.41 %，同时保持 255 f/s 的实时推理速度，并具有 9.75 GFLOPs/帧的较低计算复杂度。相比之下，DMA-Net (ResNet18) 在计算量基本相当 (9.80 vs. 9.75 GFLOPs/帧) 的前提下，进一步将 Acc 与 F_1 分别提升至 94.53% 与 94.88%，说明本文提出的 DMSFM 多尺度上下文融合与 AMMA/SimAM 注意力增强机制能够在保证效率的同时，进一步强化轨道区域的判别特征表达，从而取得更优的检测精度。

此外，从不同主干网络版本的对比可见，ResNet18 配置在本任务中表现更优：DMA-Net (ResNet18) 不仅取得最高的 Acc 与 F_1，同时在推理速度 (268 f/s) 与计算复杂度 (9.80 GFLOPs/帧) 之间保持了更均衡的折中；相较之下，DMA-Net (ResNet34) 的计算量增至 17.38 GFLOPs/帧、速度降至 155 f/s，但精度与 F_1 反而略低。这一现象可能与数据规模及噪声扰动下的泛化有关：更深的网络在有限数据条件下更易过拟合，且在复杂纹理或遮挡场景中可能削弱对轨道边缘等低层细节的保留，从而对空间定位精度产生影响；而 ResNet18 更利于保持局部纹理与边缘信息，因而在该轨道线检测任务中获

得更稳定的综合性能。

综上，DMA-Net (ResNet18) 以 9.80 GFLOPs/帧的适中计算复杂度实现了最高检测精度，并在精度与实时性之间达到良好平衡，能够有效满足实际铁路场景下的在线检测需求。

对比实验的可视化结果 (图 5) 清晰地展示了各方法的性能差异。现有主流方法在复杂铁路场景下面临诸多挑战：ClrNet 在多轨道及人员干扰下拟合效果较差，难以正确检测轨道；UFLD 存在显著的漏检问题，无法完整检测所有轨道；LaneATT 虽在直线检测上表现良好，但对曲线轨道的拟合精度不足，且时有漏检；Rail-Net 虽整体性能更优，但在处理远距离弯曲轨道时仍不够准确。此处未加入 DALNet 的可视化对比图是因为在一般直线/常规弯道以及部分复杂背景场景中能够获得与本文方法相近的检测效果，但对于存在明显人员遮挡或人员穿行等强干扰情况，局部轨道线段更易出现漏检或断裂，表现出一定的不稳定性。

相较之下，本文提出的 DMA-Net 模型展现出全面而鲁棒的检测性能，这在复杂场景中体现得尤为明显。在图 5 所示的复杂道砟与植被干扰场景中，DMA-Net 成功实现了对多组轨道的精确区分，有效抑制了背景干扰。其显著优势在于对远距离弯轨的精准处理，这得益于 DMSFM 模块对长程上下文信息的捕获能力，使得模型在准确检测近处轨道的同时，能输出更平滑、连续的远方弯曲轨道。

此外，该模型在极具挑战性的道岔区段及人员干扰下，仍能准确识别并分离交叉轨道线。即使在雨、多轨道、人员干扰等恶劣天气条件下，通过 AMMA 机制对背景噪声的有效抑制，模型依然保持了稳定的检测能力。这些结果综合表明，DMA-Net 不仅相较基线模型在指标上提升了 1.73 pp，更在视觉结果上展现出更优异的泛化性和鲁棒性，为复杂铁路环境下的安全监控提供了更可靠的解决方案。

可视化实验结果充分表明，本文提出的模型在不同工况下均能保持可靠的轨道检测性能，具有良好的环境适应性、抗干扰能力和鲁棒性。

3.5 消融实验

为验证本文所提出的模型中每个部分的有效性，本研究以 ResNet18 为主干网络开展消融实验，结果如表 2 所示。以基准模型 (第一行) 为起点，通过逐次引入 DMSFM、AMMA 机制和 SimAM，系统验证了各组件

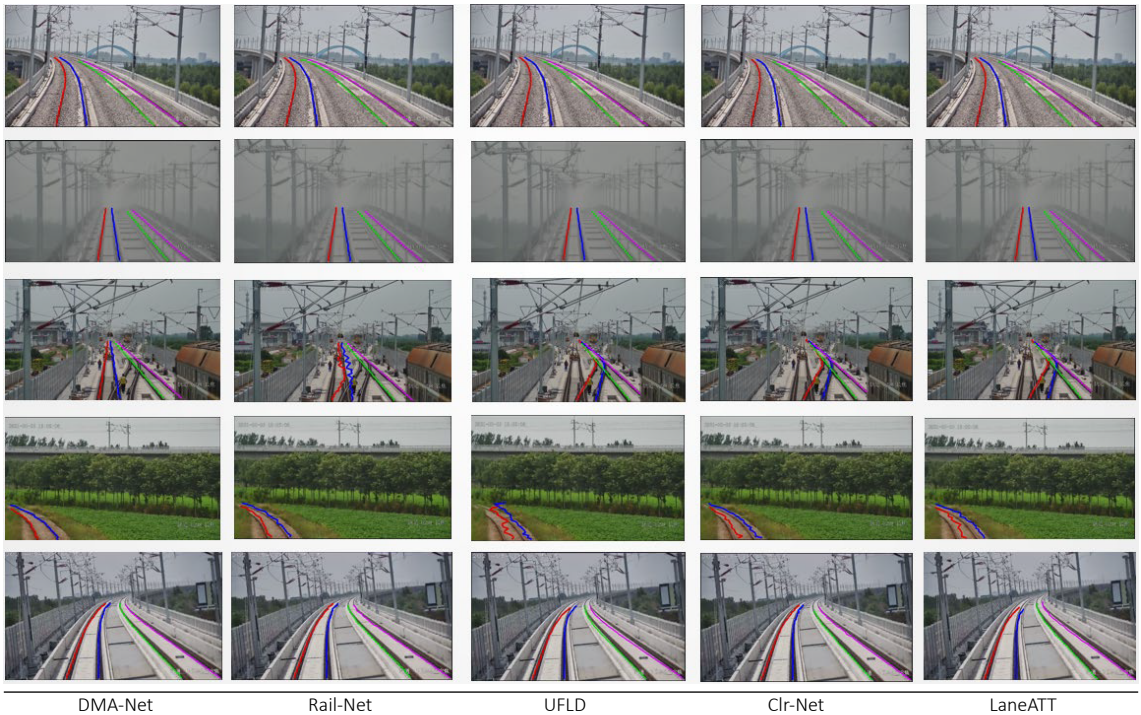


图 5 可视化结果对比

Fig. 5 Comparison of visualization results

表 2 消融实验对比结果

Table 2 Results of ablation experiments

DMSFM	AMMA	SimAM	Acc/%	F_1/%	FPS	GFlops
			92.80	92.20	**310**	**8.42**
√			93.32	93.45	280	9.06
	√		93.63	93.82	285	8.95
		√	93.18	93.31	300	8.63
√	√		94.05	94.13	273	9.59
√		√	93.82	93.95	278	9.27
	√	√	93.57	93.62	281	9.17
√	√	√	**94.53**	**94.88**	268	9.80

对模型性能的贡献。

单模块有效性验证 (表 2 第 2~4 行) 显示：

1) 引入 DMSFM 后，Acc 由 92.80% 提升至 93.32%，F_1 分数由 92.20% 提升至 93.45%。与此同时，模型推理速度从 310 f/s 降至 280 f/s，对应计算复杂度由 8.42 GFlops/帧增至约 9.06 GFlops/帧。这说明 DMSFM 能够在较小复杂度增量下有效聚合多尺度特征信息，增强模型对不同尺度轨道结构的表征能力。

2) 加入 AMMA 模块后，Acc 提升至 93.63%，F_1 达到 93.82%，推理速度为 285 f/s，对应计算复杂度约 8.95 GFlops/帧。该结果验证了 AMMA 通过可学习掩码机制抑制背景干扰、突出轨道关键区域的有效性，并保持了适中的计算开销。

3) 使用 SimAM 注意力模块时，Acc 提升至 93.18%，F_1 达到 93.31%，推理速度仍可达 300 f/s，计算复杂度仅约 8.63 GFlops/帧。这表明无参注意力机制能够以较低计算成本提升特征判别力，在效率约束下具有较高的实用价值。

表 2 第 5~7 行进一步展示了模块组合引入后的效果。DMSFM 与 AMMA 组合 (第 5 行) 使 Acc 提升至 94.05%，F_1 达到 94.13%，推理速度为 273 f/s，计算复杂度约 9.59 GFlops/帧，表明多尺度特征融合与掩码注意力具有明显协同效应。相比之下，DMSFM+SimAM (第 6 行) 在 278 FPS 下取得 93.82%/93.95% 的 Acc/F_1 (9.27 GFlops/帧)，AMMA+SimAM (第 7 行) 在 281 FPS 下取得 93.57%/93.62% (9.17 GFlops/帧)，两者虽均有提升，但增益弱于 DMSFM+AMMA 组合。

当三者同时引入 (表 2 第 8 行) 时模型达到最佳：Acc 94.53%，F_1 94.88%，较基准分别提升约 1.7 pp。此时速度由 310 f/s 降至 268 f/s，但计算复杂度仅增至 9.80 GFlops/帧 (相对基准增加 1.38 GFlops/帧)，说明在可控开销下实现了显著精度提升，仍满足实时检测需求。综上，三模块分别有效且具备协同增益，使 DMA-Net 在精度与效率间取得更优平衡。

3.6 可视化分析

如图 6 所示是 DMA-Net 模型在铁路轨道检测数据集上不同场景的可视化检测结果。在图 6(a) 所示的轨道交叉场景中，模型成功实现了多组交叉轨道的可靠检测，但在轨道线条密集交织的边缘区域仍存在少量漏检现象。针对铁路场景中具有挑战性的人员干扰区段 (图 6(b))，即使有人员闯入造成轨道局部区域遮挡，模型能够准确识别大多数有效轨道线，仅出现个别漏检情况。值得注意的是，在极端光照条件下 (图 6(c))，面对曝光突变造成的图像亮度不均，DMA-Net 仍保持稳定的轨道线检测能力。此外，模型在恶劣天气条件下也表现出优异的鲁棒性。图 6(d) 展示了雨天环境下的检测结果，图 6(e) 验证了雾天场景下的检测性能，尤为关键的是，图 6(f) 中前两张自主采集图片与后三张 RailSem19 数据集图片的组合，充分验证了 DMA-Net 模型在“常规采集-恶劣环境”跨场景下的适应性与鲁棒性：实验所用前两张图片由 6400 万像素摄像头实景采集 (非标准数据集范畴)，采集后统一降采样至 1280 pixel×720 pixel 分辨率，以此模拟实际部署中对摄像头像素规格无严苛限制的场景；后三张则直接选取 RailSem19 数据集 (非实验数据集) 中真实雨、雾、积雪情况下的轨道素材，其自带的环境退化成像效果，可直接复现恶劣天气下的实际成像状态，无需额外叠加画质劣化处理。实验结果显示，无论是在自拍的降分辨率图像中，还是在 RailSem19 的真实雨雾/积雪图像里，DMA-Net 均能实现轨道线条的精准识别；仅在 RailSem19 雾天样本中浓度最高、细节丢失最严重的局部区域，出现极个别漏检情况。这一结果充分证明，



图 6 不同环境可视化结果。(a) 轨道交叉场景；(b) 人员干扰场景；(c) 强曝光场景；(d) 雨天场景；(e) 雾天场景；(f) 非数据集场景

Fig. 6 Visualization results in different scenarios. (a) Track crossing scenario; (b) Human interference scenario; (c) Overexposure condition; (d) Rainy condition; (e) Foggy condition; (f) Non-dataset scenarios

DMA-Net 对采集设备的像素规格具有较强包容性 (高像素原图降分辨率后仍可有效检测)，同时对真实恶劣天气导致的图像质量退化具备较高容忍度，能够在跨场景条件下维持稳定的轨道检测性能。

综上，通过不同环境可视化结果表明，DMA-Net 模型在标准数据集场景、非标准自拍采集场景下均能保持可靠的轨道检测性能，对摄像头像素规格、图像分辨率及恶劣环境下的画质劣化均具备良好适应性，鲁棒性显著。

4 结 论

本文针对复杂铁路场景下轨道检测任务所面临的背景干扰强、轨道遮挡、弯曲轨道拟合难等挑战，提出了一种新颖的轨道检测网络 DMA-Net，主要工作和结论如下：

1) 提出了动态多尺度融合模块 DMSFM。该模块通过自适应权重分配机制，有效聚合了不同尺度的轨道特征，增强了模型对远近、粗细等不同尺度轨道结构的统一表征能力，为后续的精准检测奠定了坚实的基础。

2) 设计了自适应多头掩码注意力机制 AMMA。该机制通过引入可学习的掩码，使模型能够自主抑制道砟、植被等背景噪声的干扰，并聚焦于轨道关键区域，显著提升了模型在复杂纹理背景下的特征判别力和鲁棒性。

3) 构建了完整的 DMA-Net 检测模型。通过系统性地集成 DMSFM、AMMA 以及 SimAM 等组件，形成了一个优势互补的解决方案。在 Rail-DB 铁路轨道检测数据集上的综合实验表明，本文方法在检测精度 (Acc 和 F_1 分数) 上显著优于 Rail-Net、UFLD、LaneATT、CLRNet 等主流方法，同时保持了良好的实时性，在精度与速度之间实现了最佳平衡。

4) 通过充分的实验验证了模型的有效性。系统的消融实验证实了每个核心模块均为模型性能带来稳定提升；与先进方法的对比实验则证明了 DMA-Net 的综合优势；可视化结果进一步表明，该方法在各种复杂工况 (如道岔、恶劣天气、背景干扰) 下均能实现稳定、平滑且准确的轨道检测，展现出优异的实用价值和泛化能力。

尽管 DMA-Net 在复杂铁路场景下表现出了出色的检测性能和鲁棒性，但该方法也存在一些局限性。

1) 对极端环境的适应性：虽然 DMA-Net 在多种复杂场景中表现良好，但在一些极端恶劣天气条件下 (如曝光、大雾等) 或极度复杂的背景环境中，仍可能出现少量漏检或误检现象。这是因为模型在处理非常复杂的视觉特征时，可能受到背景噪声、低光照或极端天气条件下图像质量退化的影响。

2) 计算复杂度与实时性：虽然 DMA-Net 在精度和实时性之间达到了较好的平衡，但其计算复杂度仍然较高，尤其是在使用深层网络 (如 ResNet34) 时，推理速度有所下降。尽管该方法能够适应实时轨道线检测任务，但在低计算资源环境下，可能需要进一步优化以实现更高效的推理速度。

3) 对极度遮挡的处理能力：在存在较大遮挡或轨道线部分消失的情况下，DMA-Net 可能无法完全恢复完整的轨道线结构。虽然通过多尺度和注意力机制对遮挡有一定的抑制作用，但仍然有提升空间，特别是在复杂场景下轨道线严重遮挡的情况下。

综上所述，DMA-Net 有效解决了复杂铁路场景下的轨道检测难题，为铁路系统的智能化运维提供了一种高精度、高效率的技术手段，且在多种复杂场景下均能稳定检测轨道线，具有优异的鲁棒性，为铁路智能化安全防护提供实时、高精度的轨道线检测方案，以满足实际应用的需求。未来的研究将着重于提升模型在极端环境下的适应性、降低计算复杂度以及增强对遮挡和噪声的鲁棒性。

作者贡献

曹启平：改进方法的提出，论文构思和撰写，实验数据整理与分析；林珊玲：改进方法的研讨，论文审核与编辑写作；林坚普：论文审核与编辑写作；吕珊红：调查研究；林志贤：论文思路方案和实验执行进度总体指导，论文审核；郭太良：提供课题实验条件与资源。

利益冲突

所有作者声明无利益冲突

参考文献

1. Wang Q D, Yang Y, Luo Y P, et al. Review on railway intrusion detection methods[J]. *J Railw Sci Eng*, 2019, **16**(12): 3152−3159.
王泉东, 杨岳, 罗意平, 等. 铁路侵限异物检测方法综述[J]. 铁道科学与工程学报, 2019, **16**(12): 3152−3159.
2. Wang N, Hou T, Zhang T M. Research on railway track edge detection based on BM3D and Zernike moments[J]. *Arch Transp*, 2023, **68**(4): 7−20.
3. Xu X C, Shi X X, Geng C G, et al. Rail image recognition technology based on deep learning[J]. *J Railw Sci Eng*, 2024, **21**(12): 5232−5241.
徐心慈, 师秀霞, 耿晨歌, 等. 基于深度学习的铁轨图像识别技术[J]. 铁道科学与工程学报, 2024, **21**(12): 5232−5241.
4. Chen M X, Hou T, Zhao Y Z, et al. Research on rail edge detection based on new improved Canny algorithm[J]. *For Electron Meas Technol*, 2021, **40**(11): 48−52.
陈明曦, 侯涛, 赵晏章, 等. 基于新型改进 Canny 算法的铁轨边缘检测研究[J]. 国外电子测量技术, 2021, **40**(11): 48−52.
5. Lin X H. Design of a video-analysis-based railway-road safety system for detecting intrusion of foreign objects[D]. Guiyang: Guizhou University, 2021.
https://doi.org/10.27047/d.cnki.ggudu.2021.001586.
林向会. 基于视频分析的铁路异物侵限检测系统的设计[D]. 贵阳: 贵州大学, 2021. https://doi.org/10.27047/d.cnki.ggudu.2021.001586.
6. Zhao H L, Guo Y M. Research on track line detection method based on machine vision[J]. *Mech Electr Eng Technol*, 2024, **53**(5): 162−166.
赵鸿亮, 郭佑民. 基于机器视觉的轨道线检测方法[J]. 机电工程技术, 2024, **53**(5): 162−166.
7. Sheng F, Shen T, Xie Y X, et al. Track area detection algorithm for obstacle detection of fully automatic driverless train[J]. *J Tongji Univ (Nat Sci)*, 2025, **53**(3): 402−409.
盛峰, 沈拓, 谢远翔, 等. 全自动无人驾驶列车障碍物检 测的轨道区域检测算法[J]. 同济大学学报 (自然科学版), 2025, **53**(3): 402−409.
8. Xu H R, Wang X D, Fang Q. Structure road detection algorithm based on B-spline curve model[J]. *Acta Autom Sin*, 2011, **37**(3): 270−275.
许华荣, 王晓栋, 方遒. 基于 B 样条曲线模型的结构化道路检测算法[J]. 自动化学报, 2011, **37**(3): 270−275.
9. Liu L Y, Chen Q, Li Y N, et al. Feature points extraction from high

speed railway track images based on second derivative Harris operator[J]. *Sci Surv Mapp*, 2015, **40**(5): 84−88.
刘丽瑶, 陈强, 李艳娜, 等. 应用二阶 Harris 算子的高铁轨道近景影像特征点提取[J]. 绘科学, 2015, **40**(5): 84−88.

10. Li R, Wu X C. Automatic identify of linear tracks based on digital image processing[J]. *Video Eng*, 2014, **38**(3): 167−169.
李睿, 武晓春. 基于数字图像处理的直线型铁轨的自动识别[J]. 电视技术, 2014, **38**(3): 167−169.
11. Ju Y Y, Chen L. An iterative principal orientation adjustment-based algorithm for railway track detection[J]. *J Wuhan Univ Sci Technol*, 2012, **35**(6): 473−477.
居耀勇, 陈黎. 基于主方向迭代校正的铁轨检测算法[J]. 武汉科技大学学报, 2012, **35**(6): 473−477.
12. Wang Y, Wang L D, Hu Y H, et al. RailNet: a segmentation network for railroad detection[J]. *IEEE Access*, 2019, **7**: 143772−143779.
13. Liu L. Research on detection method of railway intrusion obstacles based on deep learning[D]. Lanzhou: Lanzhou Jiaotong University, 2022. https://doi.org/10.27205/d.cnki.gltec.2022.000733.
刘力. 基于深度学习的轨道侵限异物检测方法研究[D]. 兰州: 兰州交通大学, 2022. https://doi.org/10.27205/d.cnki.gltec.2022.000733.
14. Wang H, Wu Y J, Fan Z Z, et al. Fast detection algorithm of railway clearance based on deep learning[J]. *J Railw Sci Eng*, 2023, **20**(4): 1223−1231.
王辉, 吴雨杰, 范自柱, 等. 基于深度学习的铁路限界快速检测算法[J]. 铁道科学与工程学报, 2023, **20**(4): 1223−1231.
15. Jiang K, Shi J Q, Chen G W. Railway track detection method based on improved YOLOv5s[J]. *Comput Sci*, 2023, **50**(S2): 23020010.
姜珂, 石建强, 陈光武. 基于改进的 YOLOv5s 列车轨道线检测方法[J]. 计算机科学, 2023, **50**(S2): 23020010.
16. Pan X G, Shi J P, Luo P, et al. Spatial as deep: spatial CNN for traffic scene understanding[C]//*Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, New Orleans, 2018. https://doi.org/10.1609/aaai.v32i1.12301.
17. Zheng T, Fang H, Zhang Y, et al. RESA: recurrent feature-shift aggregator for lane detection[C]//*Proceedings of the 35th AAAI Conference on Artificial Intelligence*, arXiv:2008.13719. https://doi.org/10.48550/arXiv.2008.13719.
18. Wang Z, Ren W Q, Qiu Q. LaneNet: Real-time lane detection networks for autonomous driving[Z]. arXiv: 1807.01726, 2018. https://doi.org/10.48550/arXiv.1807.01726.
19. Ko Y, Lee Y, Azam S, et al. Key points estimation and point instance segmentation approach for lane detection[J]. *IEEE Trans Intell Transp Syst*, 2022, **23**(7): 8949−8958.
20. Wang J S, Ma Y C, Huang S F, et al. A keypoint-based global association network for lane detection[C]//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), New Orleans, 2022: 1382–1391. https://doi.org/10.1109/CVPR52688.2022.00145.
21. Tabelini L, Berriel R, Paixão T M, et al. PolyLaneNet: lane estimation *via* deep polynomial regression[C]//*Proceedings of 2020 25th International Conference on Pattern Recognition* (*ICPR*), Milan, 2021: 6150–6156. https://doi.org/10.1109/ICPR48806.2021.9412265.
22. Liu R J, Yuan Z J, Liu T, et al. End-to-end lane shape prediction with transformers[C]//*Proceedings of 2021 IEEE Winter Conference on Applications of Computer Vision* (*WACV*), Waikoloa, 2021: 3693–3701. https://doi.org/10.1109/WACV48630.2021.00374.
23. Qin Z Q, Wang H Y, Li X. Ultra fast structure-aware deep lane detection[C]//*Proceedings of the 16th European Conference on Computer Vision*, Glasgow, 2020: 276–291. https://doi.org/10.1007/978-3-030-58586-0_17.
24. Tabelini L, Berriel R, Paixão T M, et al. Keep your eyes on the lane: real-time attention-guided lane detection[C]//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), Nashville, 2021: 294–302. https://doi.org/10.1109/CVPR46437.2021.00036.
25. Li X P, Peng X J. Rail detection: an efficient row-based network and a new benchmark[C]//*Proceedings of the 30th ACM International Conference on Multimedia*, Lisboa, 2022: 6455–6463. https://doi.org/10.1145/3503161.3548050.
26. Hou Y N, Ma Z, Liu C X, et al. Learning lightweight lane detection CNNs by self attention distillation[C]//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*, Seoul, 2019: 1013–1021. https://doi.org/10.1109/ICCV.2019.00110.
27. Zheng T, Huang Y F, Liu Y, et al. CLRNet: cross layer refinement network for lane detection[C]//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), New Orleans, 2022: 1372–1381. https://doi.org/10.1109/CVPR52688.2022.00097.
28. Luan Q L, Mao Y D, Chen Z H, et al. Infrared road target detection based on YOLOv8-AFDP[J]. *Opto-Electron Eng*, 2025, **52**(8): 250079.
栾庆磊, 毛宜东, 陈梓华, 等. 基于 YOLOv8-AFDP 的红外道路目标检测[J]. 光电工程, 2025, **52**(8): 250079.
29. Fu M Y. Research on lightweight railway small target foreign object detection algorithm based on YOLOv8[D]. Dalian: Dalian Jiaotong University, 2025. https://doi.org/10.26990/d.cnki.gsltc.2025.000883.
付明雨. 基于 YOLOv8 的轻量化轨道小目标异物检测算法研究[D]. 大连: 大连交通大学, 2025. https://doi.org/10.26990/d.cnki.gsltc.2025.000883.
30. Zhang J, Lv M H, Feng Y A, et al. A road extraction algorithm that fuses element multiplication and detail optimization[J]. *Opto-Electron Eng*, 2024, **51**(12): 240210.
张进, 吕明海, 冯永安, 等. 融合元素乘法和细节优化的道路提取算法[J]. 光电工程, 2024, **51**(12): 240210.
31. Yu Z C, Liu Q L, Wang W, et al. DALNet: a rail detection network based on dynamic anchor line[J]. *IEEE Trans Instrum Meas*, 2024, **73**: 5018214.

作者简介



曹启平 (2000-)，男，福建三明人，硕士研究生，2023 年于东华理工大学获得学士学位，主要从事深度学习、图像处理及图像分割的研究。

E-mail：m18206056421@163.com



【通信作者】林志贤 (1975-)，男，福建泉州人，博士，教授，博士生导师，2010 年于福州大学获得博士学位，主要从事电子纸显示、平板显示器件驱动及图像处理技术的研究。

E-mail：lzx2005000@163.com

Railway track detection based on multi-scale feature fusion

Cao Qiping[1,2], Lin Shanling[1,2], Lin Jianpu[1,2], Lü Shanhong[1,2], Lin Zhixian[1,2,3*], Guo Tailiang[2,3]

[1] School of Advanced Manufacturing, Fuzhou University, Quanzhou, Fujian 362251, China;
[2] Fujian Science and Technology Innovation Laboratory for Photoelectric Information, Fuzhou, Fujian 350116, China;
[3] College of Physics and Information Engineering, Fuzhou University, Fuzhou, Fujian 350116, China

Abstract

Objective Accurate railway track detection is a critical component of intelligent railway safety systems, including obstacle intrusion warning, autonomous train navigation, and infrastructure monitoring. Reliable detection improves operational safety and enables real-time decision-making in complex railway environments. However, railway track detection remains challenging due to various environmental and structural factors. Railway tracks are often affected by strong background interference caused by ballast, vegetation, shadows, and surrounding infrastructure. These background elements share similar visual characteristics with track structures, which reduces feature discriminability. In addition, railway tracks exhibit significant scale variation across image regions, ranging from large, clearly visible structures in near-view areas to thin and indistinct structures in distant regions. This variation requires the model to maintain strong multi-scale feature representation capability. Furthermore, railway scenes frequently contain complex geometric configurations, such as curved tracks, intersections, and switches, which require the model to capture long-range spatial dependencies and preserve structural continuity. Existing deep learning-based methods, particularly those derived from lane detection frameworks, improve detection performance by leveraging convolutional neural networks and structured prediction strategies. However, these methods often suffer from insufficient receptive fields, limited multi-scale context modeling capability, and weak suppression of background interference. These limitations reduce detection accuracy and robustness in challenging railway scenarios. Therefore, developing an efficient and robust railway track detection method capable of enhancing multi-scale feature representation and suppressing background interference is essential for improving intelligent railway safety systems.

Methods A railway track detection network named dilated and masked attention network (DMA-Net) was developed to address these challenges. The proposed method was built upon a row-based anchor framework, which efficiently localizes horizontal track positions and preserves structural consistency. Several key modules were introduced to improve feature representation and discrimination capability.A dilated multi-scale fusion module (DMSFM) was designed to enhance multi-scale feature extraction. This module employed parallel dilated convolutions with different dilation rates to capture contextual information at multiple spatial scales. Dilated convolutions expanded the receptive field without reducing spatial resolution or increasing computational complexity. The parallel structure enabled effective aggregation of both local structural features and global contextual information, improving representation of curved, distant, and intersecting tracks.An adaptive multi-head masked attention (AMMA) mechanism was introduced to suppress background interference and enhance track-related feature representation. This module applied a learnable mask to the attention map, which selectively suppressed irrelevant background regions and emphasized important track areas. The multi-head attention structure enabled the model to learn diverse feature relationships and improve discrimination capability in complex environments.In addition, the SimAM attention module and DropBlock regularization were incorporated to further improve feature quality and model robustness. SimAM enhanced feature discriminability by evaluating neuron importance without introducing additional parameters. DropBlock improved generalization capability by randomly suppressing contiguous feature regions during training. These modules improved robustness while maintaining a lightweight network structure.

Results and Discussions Experiments were conducted on the public Rail-DB benchmark dataset to evaluate the effectiveness of DMA-Net. The proposed method achieved a detection accuracy of 94.53%, which exceeded the original baseline rail detection method by 1.73%. Compared with other mainstream railway track detection methods, DMA-Net improved accuracy by an average of 8.07%. These results demonstrate the effectiveness of the proposed multi-scale fusion and masked attention mechanisms.The DMSFM module significantly improved the model's ability to capture track features at different spatial scales. The enlarged receptive field enabled better representation of long-range contextual information, which improved detection performance in distant and curved track regions. The AMMA mechanism effectively suppressed background interference and enhanced attention to track-related features, which improved discrimination capability in complex environments containing ballast, vegetation, and structural interference.Visual analysis further demonstrated that DMA-Net

Foundation item: National Key Research and Development Program of China (2021YFB3600603)
*Correspondence: Lin Z X, E-mail: lzx2005000@163.com

produced smoother and more continuous detection results compared with existing methods. Performance improvements were particularly evident in challenging scenarios such as nighttime conditions, occlusion, and complex track intersections. These results confirm that the proposed modules effectively improve feature representation and detection robustness. DMA-Net maintained a lightweight design while achieving high detection accuracy. The proposed network achieved a processing speed of 268 f/s, which satisfies real-time processing requirements. The efficient architecture and lightweight attention mechanisms enable practical deployment in real-world railway safety systems.

Conclusions DMA-Net provides an effective and robust solution for railway track detection in complex environments. The DMSFM module enhances multi-scale feature representation and enlarges the receptive field, improving the model's ability to capture track geometry across different spatial scales. The AMMA mechanism improves discrimination capability by suppressing background interference and emphasizing track-related features. The integration of SimAM attention and DropBlock regularization further improves robustness while maintaining computational efficiency. Experimental results demonstrate that DMA-Net achieves superior detection accuracy, robustness, and real-time performance compared with existing methods. The proposed network provides a reliable technical solution for intelligent railway monitoring, autonomous train navigation, and railway safety applications.

Keywords: rail track detection; row anchor; multi-head mask attention; residual dilated convolution; Rail-DB dataset

Cao Q P, Lin S L, Lin J P, et al. Railway track detection based on multi-scale feature fusion[J]. *Opto-Electron Eng*, 2026, **53**(6): 250315; DOI: 10.12086/oee.2026.250315

Opto-Electronic Journals Group
www.oejournal.org

扫描二维码，获取 PDF 全文