文章编号：1007-2780(2026)08-1158-17

SimCLR-YOLO：面向低标注依赖的公路病害检测算法

向天明，燕　芳*，李嘉辉，孙　瑞
（内蒙古科技大学 自动化与电气工程学院，内蒙古 包头 014010）

摘要：针对公路病害检测中样本稀缺、弱纹理漏检率高及复杂背景干扰等难题，提出融合无监督预训练SimCLR与YOLOv8n的公路病害检测算法SimCLR-YOLO。首先，将SimCLR骨干替换为YOLOv8n的CSPDarknet，实现结构对齐与特征无缝迁移，解决网络结构不匹配导致的特征丢失问题。其次，将掩码重构融入对比学习，驱使编码器聚焦裂缝与坑槽的纹理轮廓，弥补对比学习对弱纹理特征的欠拟合缺陷。最后，嵌入CBAM注意力模块，引导模型在预训练阶段抑制路面纹理、阴影等背景噪声，增强关键区域的特征响应。在RDD2022公开数据集上的实验表明，SimCLR-YOLO算法的mAP@0.5达到89.4%，较基线模型提升1.7%，mAP@0.5:0.95提升1.1%，F1-score提升1.8%。不同标注比例下的验证结果显示，引入预训练权重在各梯度下均优于从头训练，在50%标注比例下，SimCLR-YOLO算法的mAP@0.5仍可达87%，接近全量标注下的基线模型，显著降低了对标注数据的依赖。此外，改进仅在预训练生效，推理参数(3.01M)无额外开销，为样本匮乏条件下的公路病害实时检测提供了有效技术方案。

关　键　词：公路病害；YOLOv8n；SimCLR；无监督预训练；低标注依赖；预训练
中图分类号：TP391.4　　**文献标识码：**A　　**doi：**10.3724/CJLCD.2026-0119　　**CSTR：**32172.14.CJLCD.2026-0119

SimCLR-YOLO: a road disease detection algorithm for low annotation dependence

XIANG Tianming, YAN Fang*, LI Jiahui, SUN Rui

(*School of Automation and Electrical Engineering*, *Inner Mongolia University of Science and Technology*, *Baotou* 014010, *China*)

Abstract: To address the challenges in road disease detection, such as sample scarcity, high miss rates for weakly textured targets, and complex background interference, a road disease detection algorithm named SimCLR-YOLO is proposed to integrate unsupervised pre-training SimCLR with YOLOv8n. Firstly, the SimCLR backbone is replaced with YOLOv8n's CSPDarknet to achieve structural alignment and seamless feature transfer, resolving the feature loss issue caused by network structure mismatch. Secondly, masked

收稿日期：2026-06-30；**修订日期：**2026-07-13.
基金项目：内蒙古自然科学基金（面上项目）（No. 2025MS06002）；内蒙古自治区重点研发和成果转化计划（No. 2025SYFHH0062）
Supported by Natural Science Foundation of Inner Mongolia Autonomous Region (No. 2025MS06002); Key Research and Development and Achievement Transformation Program of Inner Mongolia Autonomous Region (No. 2025SYFHH0062)
*通信联系人，E-mail：0472yanfang@163.com

reconstruction is incorporated into contrastive learning, forcing the encoder to focus on the texture contours of cracks and potholes, thereby compensating for the under-fitting of contrastive learning to weakly textured features. Finally, a CBAM attention module is embedded to guide the model in suppressing background noise such as road texture and shadows during the pre-training phase, while enhancing feature responses in critical regions. Experimental results on the public RDD2022 dataset show that the SimCLR-YOLO algorithm achieves an mAP@0.5 of 89.4%, which is 1.7% higher than the baseline model, with a 1.1% improvement in mAP@0.5:0.95 and a 1.8% increase in F1-score. Validation under different annotation ratios demonstrates that the scheme incorporating pre-trained weights outperforms training from scratch across all data gradients. At a 50% annotation ratio, the SimCLR-YOLO algorithm still achieves an mAP@0.5 of 87%, approaching the baseline model under full annotation, significantly reducing the dependence on annotated data. Moreover, the improvements are only effective during pre-training, with no additional overhead in inference parameters (3.01M), providing an effective technical solution for real-time road disease detection under sample-scarce conditions.

Key words: road disease; YOLOv8n; SimCLR; unsupervised pre-training; low annotation dependence; pre-training

1　引　　言

公路在使用过程中长期承受恶劣气候条件、自然灾害、交通荷载及重载车辆等多重因素影响，易产生裂缝、坑槽等不同类型的缺陷。这些缺陷不仅会显著缩短公路使用寿命，还会对车辆的动力学性能、制动效能及操纵稳定性产生不利影响[1]，严重时直接威胁行车安全。因此及时有效地检测和修复路面缺陷，对于保障行车安全和延长道路使用寿命至关重要。开展公路病害检测对延长公路使用寿命、保障车辆安全平稳运行具有重要现实意义[2]。

传统道路病害检测主要依赖人工现场识别，检测速度慢、工作强度大并且存在一定的安全风险，难以满足道路养护的精准需求[3]。随着计算机处理能力的提升，基于图像阈值分割、边缘检测和小波变换等数字图像处理技术逐步得到应用，然而这类方法仍存在检测精度有限、环境适应性弱、实时性不足等缺陷，制约了其实际应用效果[4-5]。随着计算机技术的持续发展和硬件性能不断提升，基于深度学习的目标检测算法已成为该领域的核心技术路径[6]。当前基于深度学习的物体检测方法主要分为单阶段和双阶段两种技术路径[7]。双阶段检测方法通过先筛选候选框再进行分类和定位，检测准确率较高但是检测速度不足，实时性较差，双阶段算法的典型算法包括R-CNN[8]、Faster R-CNN[9]、Mask R-CNN[10]等。单阶段检测方法通过端到端的网络架构同时完成目标类别识别和位置定位，具有较快的检测速度，其中YOLO系列[11]和SSD[12]等模型是这类方法的典型代表。目前YOLO算法被广泛用于公路病害检测。刘彦磊等[13]针对公路图像背景复杂的特点，通过引入Ghost模块与坐标注意力机制，构建了轻量型YOLOv5s车载检测网络，在降低计算资源消耗的同时提升了检测精度，但仍有复杂环境适应性差、小目标检测精度不足、数据质量依赖性强、所需标注样本多等缺陷。为解决此类问题，罗磊等[14]将YOLOv8的部分卷积层替换为幻影卷积(Ghostconv)，并引入全局注意力模块，有效捕捉输入特征图的全局上下文信息，提升复杂场景检测能力，使检测精度大幅提高，但改进后的模型体积相比原模型增加了8%。路成龙等[15]提出了一种改进的YOLOv8轻量级道路缺陷检测算法，该算法将YOLOv8n的主干网络替换为重参数化轻量网络RepHGNetV2以保持高分辨率，并引入无参数注意力机制SimAM增强小目标特征，有效弥补了小目标检测精度不足的缺陷。王雪秋等[16]提出了一种基于YOLOv8的道路缺陷检测算法DML-YOLO，在主干网络中引入多路径聚合坐标注意力机制(MPCA)，并设计了融合可变形卷积的C2f-MPDC模块，增强特征鲁棒性，解决了对数据质量依赖性强的不足。叶发茂[17]等人基于YOLOv8算法设计了双骨干网络，采用边

缘特征先验及双骨干特征提取，降低达到目标精度所需样本量。在现有改进YOLOv8的公路病害检测算法研究中，改进后的模型本质上依旧依赖大量的样本标注。

为解决现有公路病害检测算法存在的标注样本稀缺、细长弱纹理小目标漏检率高及复杂背景干扰强等难题，本文提出了一种基于SimCLR(Simple Framework for Contrastive Learning of Visual Representations)预训练的YOLOv8n算法，有效提高了检测精度。本文的内容主要包括：

(1) 提出了一种面向公路病害检测的自监督预训练框架。将SimCLR的骨干网络(Backbone)替换为YOLOv8的CSPDarknet，实现预训练与检测网络的结构对齐，解决了异构网络间的特征迁移损耗问题，为下游检测任务提供更优的初始化权重。

(2) 提出注意力引导的对比学习机制。在SimCLR对比学习中引入卷积块注意力模块(Convolutional Block Attention Module，CBAM)，引导模型在预训练阶段自动抑制路面纹理、阴影等背景噪声，聚焦于裂缝边缘、坑槽轮廓等关键区域，提升病害特征的判别性。

(3) 设计掩码重构辅助的预训练策略。在SimCLR引进掩码重构，迫使编码器提前学习裂缝、坑槽等“细长弱纹理”目标的高频边缘与局部纹理特征，弥补SimCLR全局对比损失对微小目标特征欠拟合的缺陷，减少漏检，提升公路病害特征提取能力。

实验结果表明，该算法在仅使用50%标注样本时，检测精度仍接近全量标注下的性能，且改进模块仅在预训练阶段生效，推理阶段无额外计算开销，可为样本匮乏条件下的公路病害实时检测提供有效解决方案。

2 SimCLR-YOLO模型结构

SimCLR-YOLO是一种融合SimCLR无监督预训练与YOLOv8n的公路病害检测算法，其结构如图1所示。该算法采用领域内预训练+YOLOv8n两阶段训练策略，使用SimCLR对比学习的方法，在大量无标注的数据集中提取公路

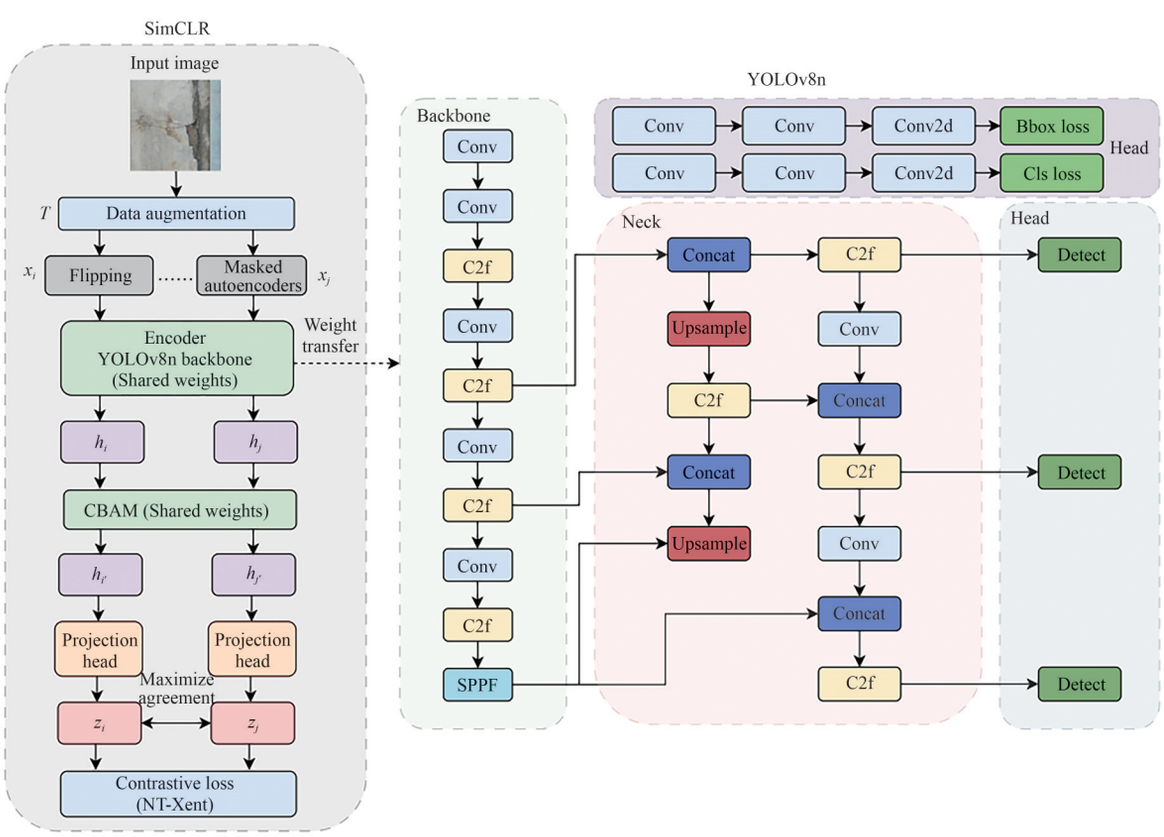


图1 SimCLR-YOLO网络框架

Fig. 1 SimCLR-YOLO network structure

病害的通用特征，并在SimCLR的骨干网络输出端引入CBAM注意力模块及掩码重构数据增强的方式，更好地聚集公路病害部位，弥补SimCLR全局对比损失对“细长弱纹理”欠拟合的缺陷。

2.1　SimCLR预训练

预训练是提升目标检测模型性能的重要手段。目前，YOLO系列模型常采用在COCO通用数据集上获取的预训练权重进行初始化，以加速模型收敛并提升精度。然而，COCO数据集以自然场景图像为主，其纹理分布、目标形态与公路病害图像存在显著差异，直接迁移易导致特征适配不佳，限制了模型在病害检测任务上的性能上限。相比之下，自监督预训练无需依赖人工标注，能够利用大规模无标注数据学习领域特有的特征表示，更契合公路病害检测中标注样本稀缺的现实需求。

SimCLR对比学习框架由Chen等[18]于2020年提出。该框架是一种典型的自监督学习方法，能够在无需人工标注的情况下，使模型自主地从数据中学习具有语义意义的特征表示。SimCLR框架的核心思想是通过最大化同一图像不同增强视图间的一致性，学习具有判别性的视觉表征。如图2所示，SimCLR通过数据增强构建正样本对，经由编码器提取特征后，利用投影头映射至低维空间计算对比损失。训练完成后，仅保留编码器用于下游任务，以剔除与检测无关的冗余信息。

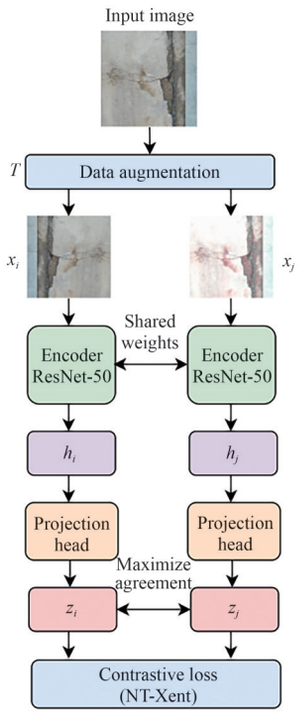


图2　SimCLR网络框架

Fig. 2　SimCLR network structure

原始SimCLR通常采用ResNet-50作为编码器，但其瓶颈层结构与YOLOv8n的CSPDarknet主干在网络层级、感受野及通道维度上存在异构性。若将ResNet-50预训练权重迁移至YOLOv8n，需对特征图进行重采样适配，易造成裂缝边缘、坑槽轮廓等高频病害特征的不可逆丢失，如图3所示。为解决上述异构网络间的特征迁移损耗问题，

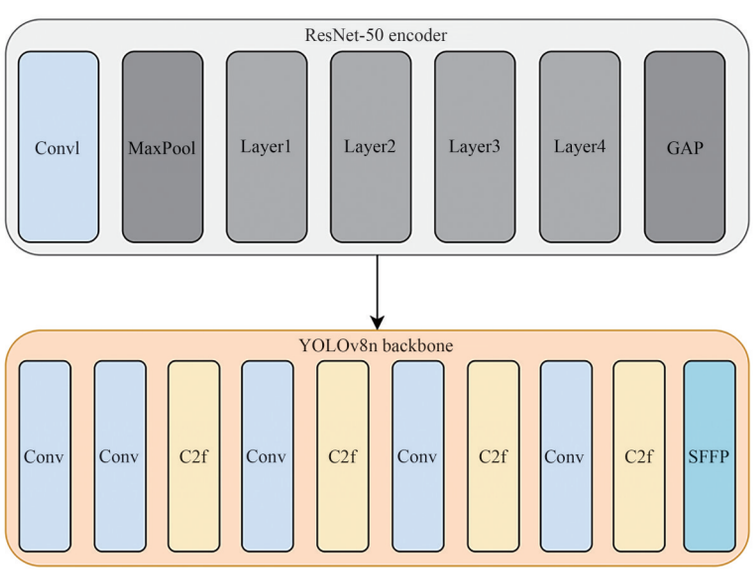


图3　结构不匹配示意图

Fig. 3　Structural mismatch illustration

本文构建了领域自适应预训练框架，将SimCLR的编码器直接替换为YOLOv8n的CSPDarknet主干。该策略实现了预训练与检测网络的结构对齐，确保预训练阶段学习到的病害特征可无缝迁移至下游检测任务，无需额外的特征适配操作，为后续引入注意力机制与掩码重构辅助任务奠定了统一的架构基础。

两个视图分别送入同一个ResNet-50编码器，将图像从像素空间映射到一个高维特征空间，参数完全共享，输出h_i、h_j，经过一个投影头（Projection Head）映射到另一个空间z_i、z_j，投影头结构是一个简单的多层感知机（MLP），将2 048维映射到128维，其网络框架如图4所示。

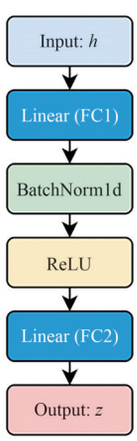


图4 MLP网络框架

Fig. 4 MLP network structure

投影头会丢弃一些对下游任务有用但对对比学习无关的信息，如颜色、方向等，训练完成后投影头会被丢弃，只保留编码器输出的h作为最终特征，参与到下游的检测任务。SimCLR的核心思想是最大化一致性（Maximize Agreement），负责让来自同一张图的z_i和z_j在特征空间中尽可能接近，让来自不同图的z在特征空间中尽可能远离。最后使用NT-Xent（Normalized Temperature-scaled Cross Entropy Loss）损失函数对比损失，如公式（1）所示：

$$\ell_{i,j}=-\log\frac{\exp(\mathrm{sim}(\boldsymbol{z}_i,\boldsymbol{z}_j)/\tau)}{\sum_{k=1}^{2N}1_{[k\neq i]}\exp(\mathrm{sim}(\boldsymbol{z}_i,\boldsymbol{z}_k)/\tau)},\quad(1)$$

其中：$\ell_{i,j}$是样本i与其正样本j之间的损失；$\boldsymbol{z}_i$、$\boldsymbol{z}_j$是同一图像两种增强后的特征向量（正样本对）；$\mathrm{sim}(\boldsymbol{z}_i,\boldsymbol{z}_j)$为余弦相似度；$\tau$为温度参数，控制分布的锐利程度；$2N$指一个batch有$N$张图，每张图2种增强，共$2N$个样本；$1_{[k\neq i]}$为指示函数，排除样本自身（$K=i$时为0，否则为1）。

2.2 CBAM注意力机制

对比学习的核心思想在于通过拉近正样本对的特征距离、推远负样本对的特征距离来学习有效的视觉表示。在公路病害检测这一特定场景下，原始的SimCLR框架在预训练过程中仍对所有图像区域赋予同等权重，缺乏对病害关键区域的针对性关注。由于公路病害图像中病害区域占比通常较低，而背景中包含大量路面纹理、阴影及杂物等干扰信息，极易导致编码器学习到冗余的背景特征，从而降低病害特征的判别性。为了引导模型聚焦裂缝边缘、坑槽轮廓等病害关键区域，本文在SimCLR框架中引入了CBAM注意力模块[19]，通过通道与空间维度的双重校准提升特征表示的质量。改进后的SimCLR+CBAM整体网络框架如图5所示。

由图5可知，从双分支数据增强后生成的样

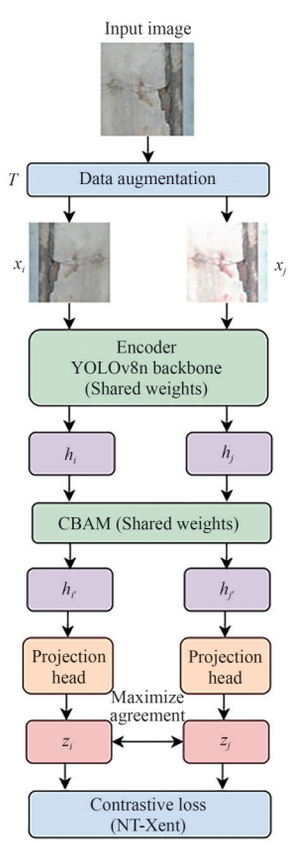


图5 SimCLR+CBAM网络框架

Fig. 5 SimCLR+CBAM network structure

本对 x_i 和 x_j 输入共享权重的编码器，其输出的特征表示 h_i 和 h_j 将进入 CBAM 模块进行自适应加权。为便于数学表达，记编码器输出的特征图为 $F\in R(C\times H\times W)$。该模块通过串联通道注意力与空间注意力两个子模块对特征进行重标定，其 CBAM 网络结构如图 6 所示。

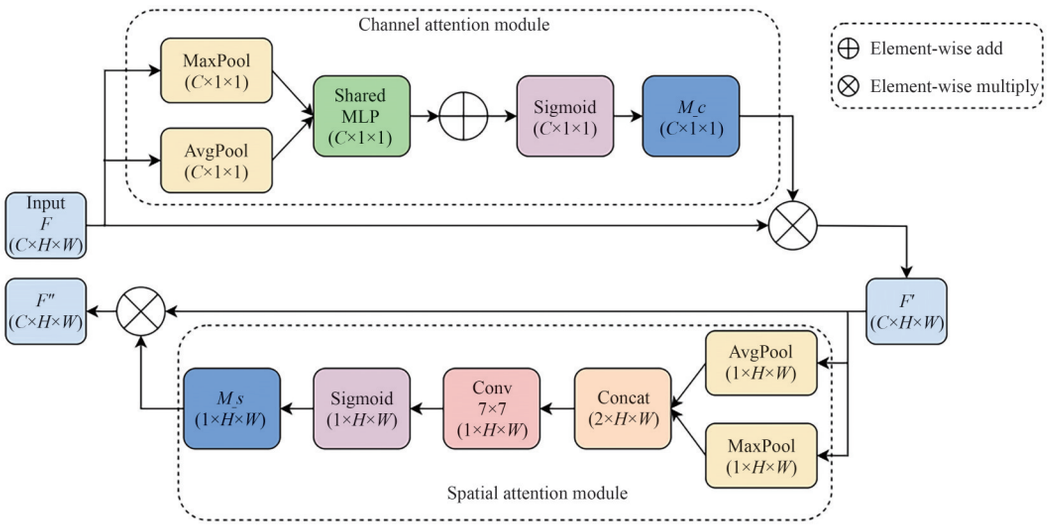


图 6　CBAM 网络框架

Fig. 6　CBAM network structure

在通道注意力计算阶段，子模块首先在空间维度上进行压缩，利用全局最大池化（MaxPool）和全局平均池化（AvgPool）提取空间背景特征，随后通过共享多层感知机（MLP）和 Sigmoid 激活函数 σ 进行映射。其通道权重 $M_C\in R(C\times 1\times 1)$ 的中间计算如公式（2）所示：

$$M_C(F)=\sigma(\mathrm{MLP}(\mathrm{AvgPool}(F))+\mathrm{MLP}(\mathrm{MaxPool}(F))). \tag{2}$$

得到通道权重 M_C 后，对原始特征图 F 进行逐通道相乘，得到通道维度重标定后的中间特征图 F'，如公式（3）所示。

$$F'=M_C(F)\otimes F, \tag{3}$$

式中，$\otimes$ 表示逐元素相乘。

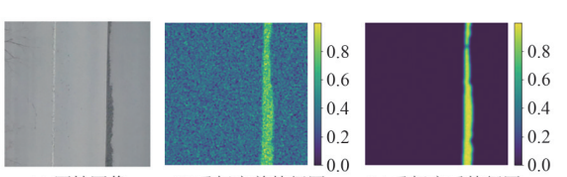
在未进行重标定的特征图 F（即图 5 的 h_i 和 h_j）中，如图 7（b）所示，由于缺乏通道筛选机制，包含路面碎石、阴影等高频背景噪声的冗余通道与包含病害信息的有效通道被同等对待，导致非病害区域存在大量散落的斑点状噪声。为了滤除这些噪声，通道注意力子模块首先计算出通道自适应权重 M_C，在该权重中，响应背景纹理的通道权重被抑制到接近 0，而响应裂缝边缘特征的通道则被赋予高权重。这在物理上实现了初步的背景过滤。在此基础上，空间注意力子模块以中间特征图 F' 作为输入，沿通道维度分别进行最大池化与平均池化并进行特征拼接，通过一个 7×7 的卷积操作 $f^{7\times 7}$ 与 Sigmoid 激活函数，生成空间权重 $M_S\in R(1\times H\times W)$，如公式（4）所示：

$$M_S(F')=\sigma(f^{7\times 7}([\mathrm{AvgPool}(F');\mathrm{MaxPool}(F')])). \tag{4}$$

最终，将中间特征 F' 与空间权重 M_S 逐元素相乘，得到重标定后的输出特征图 F''（即图 5 的 h_i' 与 h_j'），如公式（5）所示：

$$F''=M_S(F')\otimes F'. \tag{5}$$

通过引入空间权重 M_S，模型在上一步通道净化的基础上，进一步在空间维度上对非病害区域进行选择性压缩。对比重标定后的最终特征图 F'' 与重标定前的特征图，如图 7（b）及图 7（c）



（a）原始图像　（b）重标定前特征图　（c）重标定后特征图

（a）Original image　（b）Feature map before attention recaliberation　（c）Feature map after attention recaliberation

图 7　注意力重标定前后特征图可视化对比

Fig. 7　Visualization comparison of feature maps before and after attention recalibration

所示，可以直观地观察到自适应权重添加前后的显著差异。

在图 7(b)中，左侧平坦路面区域存在大面积非零的蓝绿色噪声响应，而经过双重加权计算后(图 7(c))，左侧背景区域的空间权重 M_S 被强制衰减至 0(呈现大面积深蓝色)，噪声干扰被彻底清除；同时，右侧裂缝病害区域在空间权重的高激活响应作用下，特征能量得到高度聚焦与凝聚(响应值逼近 1.0，呈现连续的明亮黄色带)，裂缝的几何轮廓被精细地提取出来。

上述对比充分证明，通过通道与空间双重自适应加权机制的协同计算，特征图完成了从“背景多噪”到“病害聚焦”的转变，不仅过滤掉了路面的干扰杂质，还增强了病害区域的特征响应。引入 CBAM 后，公路病害特征学习过程发生了本质性的改变。这种针对性的特征提取方式，使得同一病害图像在不同数据增强条件下生成的特征向量具有更高的相似度。增强后的特征 h' 送入投影头映射到低维空间并计算 NT-Xent 对比损失，通过端到端的梯度反向传播，使编码器能够有效过滤掉路面纹理、阴影、杂物等干扰因素，极大地提升了对比学习的训练效率与特征判别能力。

2.3 SimCLR 掩码重构

在对比学习框架中，数据增强策略的选择直接影响模型所学特征的质量与泛化能力。对于公路病害检测任务，病害区域通常仅占图像局部，且形态多变、边界模糊，而传统的颜色抖动、随机裁剪等增强方式难以有效引导模型聚焦于病害的结构性特征，容易导致模型对背景纹理或光照变化产生过拟合。受 MAE(Masked Autoencoders)[20] 启发，本文在 SimCLR 框架中引入掩码重构机制作为辅助任务，通过强迫模型推理被遮挡的病害信息，进一步增强编码器对局部纹理与结构特征的提取能力。MAE 网络框架如图 8 所示。

首先将输入图像划分为固定大小的图像块，随机遮挡 50% 的区域，将被遮挡位置的像素值置零后送入 YOLOv8n 编码器提取特征，随后通过轻量级解码器重构原始图像，并仅在被遮挡区域计算重构损失。该设计的核心思想在于：当图像存在大面积信息缺失时，编码器无法依赖局部纹理的简单拼接来完成重构任务，必须从可见区域推断被遮挡部分的语义内容与结构连续性。在公路病害检测场景下，这种约束促使模型深入理解裂缝的延伸趋势、坑槽的边缘形态等病害的结构连续性与结构特征，而非仅记忆特定的纹理模式。

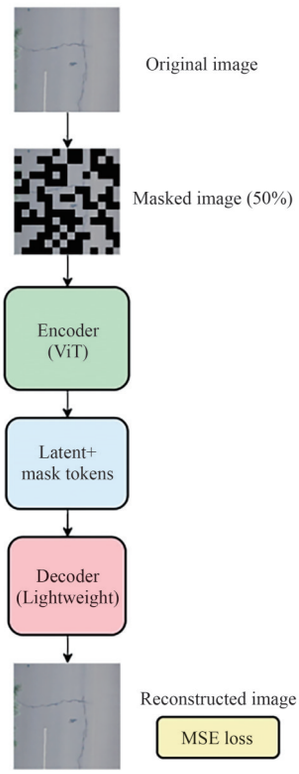


图 8 MAE 网络框架

Fig. 8 MAE network structure

与原始 MAE 不同，本方法保留了 CNN 的局部感知优势，并采用将遮挡区域置零后输入完整编码器的方式，以保持 SimCLR 对比学习分支的架构一致性。训练中，掩码重构作为辅助任务，其均方误差(MSE)损失与对比损失协同优化。通过迫使模型重建高比例遮挡的图像，编码器在具备实例判别力的同时，能从有限可见区域中捕获裂缝纹理、边缘走向等全局语义与空间结构，从而提升特征的鲁棒性，为下游检测提供优质的初始化权重。

3 实验和结果

3.1 实验平台与超参数设置

为了验证本文提出算法的有效性，建立实验平台，所有消融实验及对比实验均在相同实验环境与配置参数下进行。在实验过程中，本文使用

Windows11操作系统。实验平台的GPU配置为NVIDIA GeForce RTX 5060Ti(16 GB)，CPU配置为12th Gen Intel(R) Core(TM) i5-12600KF(16 GB)。运行库版本为CUDA 12.8，深度学习框架采用PyTorch 2.10.0，编程语言为Python 3.10.19，并选择YOLOv8n为基准网络模型。实验参数设置如表1所示。公路病害目标多为小尺度且形态细长，输入图像尺寸统一设置为640×640以保留细节特征。优化器选用SGD以提升模型泛化能力，初始学习率设为0.01，动量系数为0.937，权重衰减系数设为0.000 5以抑制过拟合。

表1　实验环境参数

Tab. 1　Experimental parameters

参数	Value
迭代次数	300
批大小	16
输入图像尺寸	640×640
工作线程数	0
优化器	SGD
学习率	0.01
动量	0.937
权重衰减系数	0.000 5

3.2　数据集与评价指标

本文选用RDD2022[21]数据集中来自中国无人机和车载摄像头拍摄的图像，作为有标注数据放入下游任务，包含纵向裂缝(D00)、横向裂缝(D10)、网状裂缝(D20)、坑槽(D40)和道路修补(Repair)五个公路病害类型共4 398张中国道路图像，其中无人机拍摄的公路病害图像2 401张，车载相机拍摄的公路病害图像1 977张。公路病害类型实例如图9所示。

数据集按照8∶1∶1的比例划分为训练集、验证集和测试集。为保证实验严谨性，本文使用划分80%的训练集及部分RDD2022数据集中日本车载摄像头拍摄的图像，共计5 626张无标注图像，放入SimCLR中进行预训练。不同公路病害类型在数据集中分布情况如表2所示。

为了客观评价模型的检测性能，本实验采用式(6)～式(10)所示的评价指标：精度(Precision，P)、召回率(Recall，R)、平均精度(mean Average Precision，mAP)，F1-score、参数量(Params)、计算量(GFLOPs)等指标。

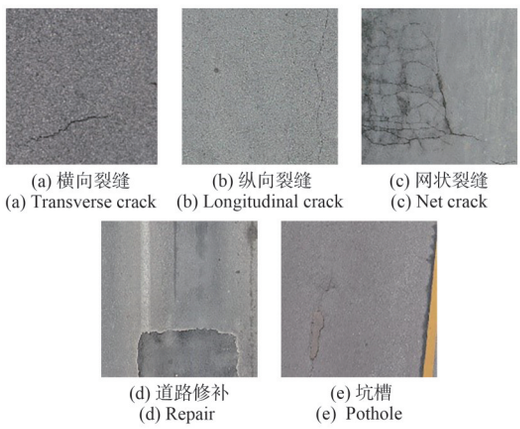


图9　数据集中不同公路病害类型示例图

Fig. 9　Example images of various road defect types in the dataset

表2　数据集中病害类型及数量

Tab. 2　Type and number of defects in dataset

缺陷类型	标签	数量
纵向裂缝	D00	3 270
横向裂缝	D10	1 895
网状裂缝	D20	756
坑槽	D40	255
道路修补	Repair	821

$$P=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}, \tag{6}$$

$$R=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \tag{7}$$

$$\mathrm{AP}=\int_0^1 P(R)\mathrm{d}R, \tag{8}$$

$$\mathrm{mAP}=\frac{\sum_{i=1}^{N}\mathrm{AP}_i}{N}, \tag{9}$$

$$\text{F1-score}=\frac{2(\mathrm{Precision}\times\mathrm{Recall})}{\mathrm{Precision}+\mathrm{Recall}}, \tag{10}$$

式中：TP代表检出的真实病害，FN代表没能检出的病害，FP则属于误报的情况，AP对应各类别P-R曲线下的面积，mAP即各类别AP的平均值。mAP@0.5是指将置信度阈值设为0.5时测得的精度，mAP@0.5∶0.95则涵盖了置信度在0.5～0.95区间内的平均表现。最后用F1-score(精确率和召回率的加权平均)判定模型的稳定性。

3.3　消融实验

为定量评估预训练策略及模块组合对

YOLOv8n算法的优化效果，以及对比学习与注意力机制结合对预训练的影响，本文共设置6组对照实验，均在RDD2022训练集上进行训练与测试。为保证实验的公平性，所有组别均采用表1所示的超参数配置，仅对预训练策略及网络结构进行变量控制。消融实验结果如表3所示。

表3 消融实验结果

Tab. 3 Ablation experiment results

组别	Scratch	Coco-pretrained	SimCLR	CBAM	MAE	mAP@0.5/%	mAP0.5:0.95/%	F1-score/%	Params/($\times10^6$)	计算量/GFLOPs
1	√					84.2	53.9	79.7	3.01	8.1
2		√				87.7	59.5	83.0	3.01	8.1
3			√			87.9	59.9	84.0	3.01	8.1
4			√	√		88.2	59.5	83.2	3.01	8.1
5			√		√	88.0	59.0	84.0	3.01	8.1
6			√	√	√	**89.4**	**60.6**	**84.8**	**3.01**	**8.1**

注：粗体数字表示最优值

组别1采用随机初始化的YOLOv8n模型进行从头训练（Scratch），由于缺乏预训练权重的先验知识，受限于有限的标注数据规模，模型难以在训练中充分泛化，导致mAP@0.5和F1-score仅分别达到84.2%和79.7%，检测性能显著低于引入预训练权重的其他对照组。组别2为基线模型，直接使用ImageNet预训练权重初始化YOLOv8n主干网络，mAP@0.5为87.7%，F1-score为83%，相对无预训练的组别1有较大的提升。ImageNet以自然场景图像为主，与路面图像的纹理分布差异较大，预训练特征的迁移效果受限。组别3将预训练方式替换为SimCLR，在路面数据集上进行对比学习。mAP@0.5提升至87.9%，F1-score从83%提升至84%。提升幅度不大，原因在于对比学习的核心目标是学习视图不变的语义特征，这类特征擅长区分是否为病害，但对病害的精确边界定位帮助有限。mAP@0.5:0.95仅从59.5%微升至59.9%印证了这一点，但小幅提升也表明，在路面数据集上进行对比学习能够获得比通用预训练更适配的特征表示。组别4在对比学习基础上引入CBAM注意力模块。mAP@0.5提升至88.2%，较组别2提高0.3个百分点，F1-score下降至83.2%。F1-score的下降并非模型退化，而是精确率与召回率的权衡变化：CBAM通过抑制背景响应提高了检测精确率，但同时也过滤掉了部分低置信度的真阳性样本，导致召回率略有下降。mAP@0.5:0.95保持在59.5%，说明单纯的注意力机制对高精度定位的贡献有限。组别5在组别3基础上加入掩码重构任务，不使用CBAM，mAP@0.5为88%，但mAP@0.5:0.95从59.9%下降至59%。这一现象揭示了单独使用掩码重构的问题：重构任务要求网络保留大量低层纹理细节用于像素预测，而对比学习要求网络提取抽象的语义特征，两个目标相互干扰。缺乏有效的特征调节机制时，网络在两个任务间摇摆，反而损害了定位精度。组别6为本文完整模型，融合SimCLR、CBAM和掩码重构。mAP@0.5达到89.4%，较基线提升1.7个百分点；mAP@0.5:0.95达到60.6%，较基线提升1.1个百分点；F1-score达到84.8%，较基线提升1.8个百分点。3项指标均达到最优，且相比任意单一改进或两两组合都有明显提升。组别6相对于组别4和组别5的提升说明了3个模块的协同机制：CBAM充当了对比学习与掩码重构之间的“调节器”。通道注意力使网络能够区分服务于语义理解的特征通道和服务于细节重构的特征通道，空间注意力引导网络在病害区域保留重构所需的纹理信息、在背景区域提取对比学习所需的语义信息。这种分工使两个预训练任务从相互竞争转为相互补充，最终体现为检测性能的全面提升。相较于从头训练的组别1，引入预训练权重的对照组在检测性能上实现了大幅提升。消融实验对比结果定量地证实了预训练策略在优化目标检测精度中的核心作用。

计算开销方面，完整模型参数量为 3.01M，GFLOPs为8.1，与基线完全一致。这是因为本文所有改进均作用于预训练阶段：SimCLR对比学习、CBAM注意力模块和掩码重构分支都只在预训练时参与计算，用于引导编码器学习更具判别性的特征表示。下游检测任务仅迁移预训练得到的编码器权重，上述模块全部移除，前向传播路径与基线模型完全相同，即预训练阶段的额外计算换来了更好的权重初始化，但推理时不引入任何开销。

3.4　CBAM注意力机制及掩码重构热力图分析

为验证CBAM注意力模块以及MAE对SimCLR的正向影响，对未引入CBAM注意力模块及MAE的SimCLR、引入CBAM注意力模块的SimCLR、引入MAE的SimCLR、引入CBAM注意力模块及MAE的SimCLR进行热力图分析。由于SimCLR属于无监督，输入图像没有任何标注，因此采用特征图激活可视化（Feature Map Activation Visualization）进行热力图分析，结果如图10所示。

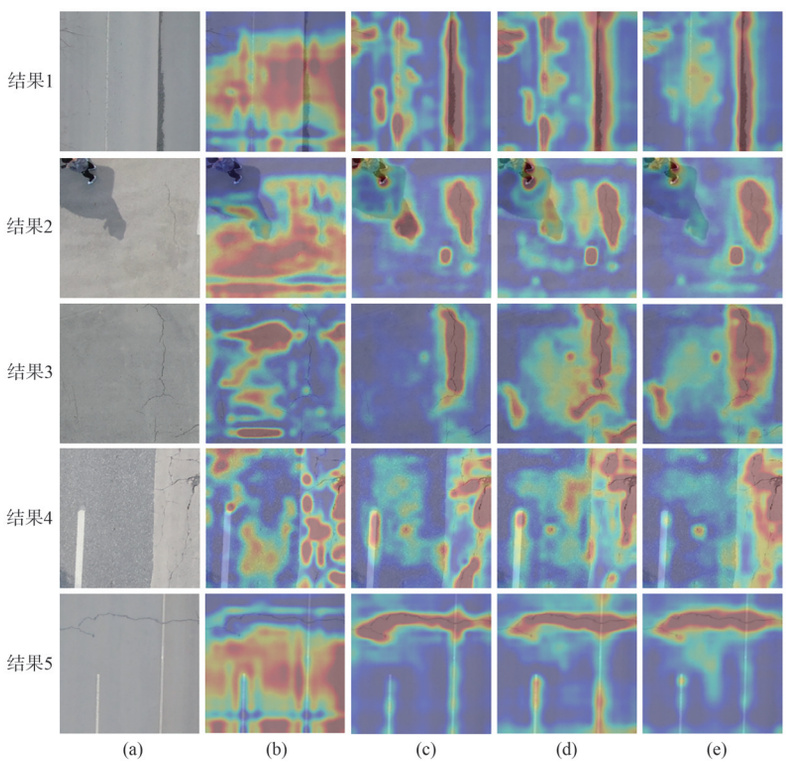


图10　不同消融配置下SimCLR的特征激活热力图对比。(a) 原图；(b) SimCLR；(c) SimCLR＋CBAM；(d) SimCLR＋MAE；(e) SimCLR＋CBAM＋MAE。

Fig. 10　Comparison of feature activation heatmaps of SimCLR under different ablation configurations. (a) Original image; (b) SimCLR; (c) SimCLR＋CBAM; (d) SimCLR＋MAE; (e) SimCLR＋CBAM＋MAE.

热力图结果直观地反映了各模块对特征表示的重塑作用。在未引入CBAM注意力模块及MAE的SimCLR（图10(b)）中，所有5组实验的高亮激活区域几乎完全集中在平整、无害的背景路面上，严重偏离了真实的病害位置。原因在于传统的实例判别任务仅要求模型区分不同的图像个体，这导致模型倾向于提取面积占绝对主导、易于区分的低频全局特征，而将对下游定位任务至关重要的局部细微裂缝视为无用噪声予以平滑和忽略。引入CBAM注意力模块有效扭转了这一偏差，如图10(c)所示。在结果1、3、5中，原本在图10(b)中被大面积激活的平整路面被成功暗化，而真正的裂缝区域则激发了显著的橙红色高亮响应。这表明CBAM通过通道和空间双重注意力机制，有效引导模型在预训练阶段抑制了无用的路面背景纹理，显著增强了关键病害区域的特征响应。然而，单独使用CBAM对背景噪声的抑制仍不够彻底，如结果2和4所示，虽然大部分

背景被成功抑制，但在阴影边缘及强干扰的白色车道线依然附着着一些激活斑块。在SimCLR仅引入掩码重构的图10(d)中，特征激活表现出极强的空间几何敏感性。在结果1、3、5中，高亮区域迅速收敛，呈现出清晰的线性形态，极度契合裂缝的物理走向。特别是在结果3中，MAE成功高亮激活了图10(c)未能有效捕获的隐蔽裂缝；同时，在结果2和4中，复杂阴影与白色车道线对模型的负面干扰被明显削弱，使模型的视线更加聚焦于裂缝本体。然而，图10(d)的背景区域整体充斥着大范围的黄绿色微弱激活。这表明单独的掩码重构任务虽然能够通过像素级重建强迫编码器精确定位裂缝与坑槽的微观纹理轮廓，但由于其缺乏对特征的选择性过滤，导致对无用背景纹理的抑制效果显著逊于CBAM。最终，融合了CBAM与MAE的图10(e)实现了两者的优势互补。如结果1和2所示，相较于单独引入CBAM或单独引入MAE时背景区域仍残留部分高亮响应的窘境，两者的有机结合显著削弱了模型对背景噪声的关注。在结果1、2、4、5中，完整模型对复杂阴影及白色车道线强干扰源的平抑效果均明显优于任何单一模块的配置；而在对病害区域的聚焦表现上，完整模型的聚焦精确度与边界拟合度明显强于单独引入CBAM的图10(c)，但稍微逊色于单独引入MAE的图10(d)，这种中和状态成功避免了MAE过度激活背景的问题，实现了定位精度与背景压制的最佳平衡。

这一对比强有力地证明了二者结合的必要性：掩码重构任务赋予了模型驱使编码器聚焦裂缝与坑槽的纹理轮廓，弥补对比学习对弱纹理特征的欠拟合缺陷；而CBAM注意力模块则充当了智能特征过滤器，引导模型在预训练阶段强力平抑普通的道路背景噪声。两者协同作用，实现了关键病害特征响应的最大化增强与背景无用噪声的完美剥离。

3.5 不同算法对比

为了验证本文算法的检测性能，将SimCLR预训练的YOLOv8n与当前主流的目标检测算法进行对比实验，对比算法包括Faster R-CNN、SSD、YOLOv5n、YOLOv8n、YOLOv8s、YOLOv9t、YOLOv10n、YOLOv11n。算法性能对比结果如表4所示。

表4 基于RDD2022数据集的算法性能对比

Tab. 4 Performance comparison of different algorithms on RDD2022 dataset

模型	Precision/%	Recall/%	mAP@0.5/%	mAP0.5:0.95/%	Params/($\times10^6$)	计算量/GFLOPs	F1-score/%	模型大小/MB
Faster R-CNN	71.2	82.7	79.4	48.5	41.45	94.4	76.5	108
SSD	76.7	61.2	62.6	43.8	26.77	60.7	67.6	33.7
YOLOv5n	84.3	81.3	87.1	57.6	2.51	7.2	82.8	5.1
YOLOv8n	**85.3**	80.7	87.7	59.5	3.01	8.1	83	6.0
YOLOv8s	83.9	82.6	87.4	60	11.1	28.7	83.3	21.5
YOLOv9t	84.9	83.7	87.6	60.5	2.62	10.7	84.3	17.1
YOLOv10n	84.3	79.9	86.5	59.3	2.71	8.4	82	5.5
YOLOv11n	83.4	84.8	88	59.9	2.9	7.6	84.1	5.8
Ours	84.7	**84.9**	**89.4**	**60.6**	3.01	8.1	**84.8**	6.0

注：粗体数字表示最优值

由表4可知，本文所提算法(Ours)在RDD2022数据集上展现出显著的综合优势。在检测精度方面，本文算法在多项核心指标上均取得最优结果。具体而言，召回率达到84.9%，较YOLOv8n提升4.2%；mAP@0.5达到89.4%，为所有对比算法中的最高值，分别较YOLOv8n、YOLOv8s、YOLOv9t、YOLOv10n和YOLOv11n提升1.7%、2.0%、1.8%、2.9%和1.4%；mAP@0.5:0.95达到60.6%，同样位列第一，较次优的YOLOv9t高出0.1%；F1-score达到84.8%，较YOLOv8n和YOLOv9t分别提升1.8%和0.5%。虽然YOLOv8n在精度上以85.3%略高于本文算法的84.7%，但本文算法在召回率上的显著优势表明其具有更强的缺陷检出能力，漏检率更低，这

对于道路缺陷检测的实际应用场景尤为重要。在实际的检测任务中，由于本文算法对特征提取能力的增强，精度的轻微下降并不会提升对公路病害的误检概率。

在计算复杂度方面，本文算法同样表现优异。参数量为3.01×10^6，计算量为8.1 GFLOPs，模型大小仅为6.0 MB。与传统检测算法相比，本文算法的参数量仅为Faster R-CNN的7.3%，计算量仅为其8.6%，模型大小仅为其5.6%，实现了数量级的压缩。与SSD相比，参数量降低88.8%，计算量降低86.6%，模型大小降低84.9%，同时mAP@0.5提升17.8%。与同系列YOLO算法对比，本文算法在保持轻量化的同时实现了精度提升。相较于YOLOv5n，本文算法参数量增加19.9%，但mAP@0.5提升2.3%；相较于YOLOv8s，参数量降低72.9%，计算量降低71.8%，而mAP@0.5仍高出2.0%；相较于YOLOv9t，参数量增加14.9%，但计算量降低24.3%，且mAP@0.5高出1.8%；相较于YOLOv10n和YOLOv11n，算法在参数量和计算量相近的情况下，mAP@0.5分别高出2.9%和1.4%。相较于YOLOv8n，YOLOv8s仅在召回率和mAP@0.5：0.95上略优，但在精度、mAP@0.5和F1-score上均低于YOLOv8n。这可能由于RDD2022数据集相对较小，更大的模型容易过拟合，道路缺陷中存在较多小目标，轻量级模型有时在小目标检测上反而表现更好。这从侧面说明，模型并非越大越好，针对特定任务选择合适规模的模型更为重要。

综上所述，本文算法在RDD2022数据集上实现了检测精度与计算效率的最优平衡。在所有对比算法中，本文算法以最高的mAP@0.5（89.4%）和mAP@0.5：0.95（60.6%）确保了检测准确性，同时以较低的参数量（3.01M）、计算量（8.1 GFLOPs）和模型大小（6.0 MB），实现零成本提升模型性能。

3.6　低标注依赖场景下的模型性能分析

自监督预训练方法的核心优势在于能够从大规模无标注数据中学习具有领域适应性的特征表示。在本研究中，SimCLR预训练阶段利用大量无标注的公路病害图像，通过对比学习机制挖掘公路病害的共性视觉特征。相较于ImageNet预训练权重所包含的通用自然图像特征以及从头训练，这种领域特定的特征表示与公路病害检测任务能够为下游检测模型提供更优的参数初始化。当预训练特征与目标任务的特征分布更为接近时，下游模型的收敛速度更快，最终性能的上限与下限均能得到提升。

为定量验证上述分析，本文设计了不同标注比例下的对比实验。将RDD2022训练集分别按100%、70%、50%、30%、10%的比例进行采样，在从头训练的YOLOv8n、COCO预训练的YOLOv8n与经SimCLR预训练的YOLOv8n上分别进行训练，实验结果如表5所示。

表5中呈现了无预训练（Scratch）、通用监督预训练（COCO）以及本文自监督预训练方法（Ours）在不同标注比例下的检测性能对比。从整体趋势来看，引入预训练权重可显著提升模型在各数据梯度下的检测精度，验证了参数初始化对公路病害检测任务的关键作用。在所有设计的标注比例下，本文方法的检测性能均一致地超越了从头训练与COCO预训练，在标注数据削减70%时，本文算法的mAP@0.5仍能达到88.7%，相比全量标注下的COCO预训练模型实现了1%的性能反超，充分证明了SimCLR预训练在特征表示层面的优越性。当标注数据缩减50%时，本文方法的mAP@0.5为87.0%，与全量标注下的COCO预训练模型仅相差0.7%，且显著优于全量标注下从头训练的性能。这表明本文方法以近乎无损的精度表现，换取50%的标注成本削减，有效缓解了公路病害检测中人工标注成本高、难度大的工程瓶颈。在10%标注样本的极端情况下，本文算法的mAP@0.5相较于COCO预训练模型和从头训练分别提升了3.9%和9.4%，证明了在新建道路或罕见新病害检测等标注样本极度匮乏的特定工程领域，本算法仍能展现出优异的特征捕捉能力，具备较高的泛化潜力与实用价值。

综上所述，预训练在公路病害检测中发挥着不可替代的基石作用。在下游病害检测任务中，针对性地引入领域内预训练不仅能为网络提供更契合病害特征的初始化权重，还能在数据极度稀疏的局限下，提升性能上限。

为验证本文算法在低标注依赖下的优势，本文选取100%标注比例及50%标注比例下3种算法的公路病害检测结果对比，如图11所示。

表 5 不同标注比例数据集的检测性能对比

Tab. 5 Detection performance comparison under different annotation ratios

预训练类型	标注比例	Precision/%	Recall/%	mAP@0.5/%	mAP0.5:0.95/%	F1-score/%
Scratch	100%	81.4	78.1	84.2	53.9	79.7
COCO		85.3	80.7	87.7	59.5	83.0
Ours		84.7	84.9	89.4	60.6	84.8
Scratch	70%	80.0	74.8	82.1	52.2	77.3
COCO		84.2	80.8	86.5	57.4	82.5
Ours		85.5	83.2	88.7	57.8	84.3
Scratch	50%	78.5	74.6	80.6	51.1	76.5
COCO		80.0	82.0	83.9	55.7	81.0
Ours		85.0	82.3	87.0	57.1	83.6
Scratch	30%	82.8	66.5	75.7	45.7	73.8
COCO		81.8	76.0	79.6	50.9	78.8
Ours		80.1	78.1	82.1	54.5	79.1
Scratch	10%	68.7	53.2	58.8	31.2	60.0
COCO		81.2	54.4	64.3	38.6	65.2
Ours		76.4	64.2	68.2	40.5	68.0

在道路修补检测场景中(第1行图),当标注量降至50%时,COCO预训练模型在目标边缘产生了一处明显的冗余检测框;而本文算法在100%与50%标注下均表现稳定,检测框定位精准且无冗余。在第2行图中,从头训练的Scratch算法在100%标注下出现漏检,在50%标注下将右侧无用的背景错误地检测为道路修补;COCO预训练模型在50%标注下发生了分类混淆,将本应属于网状裂缝的病害误检为纵向裂缝;相比之下,本文算法在50%标注下依然能够实现准确分

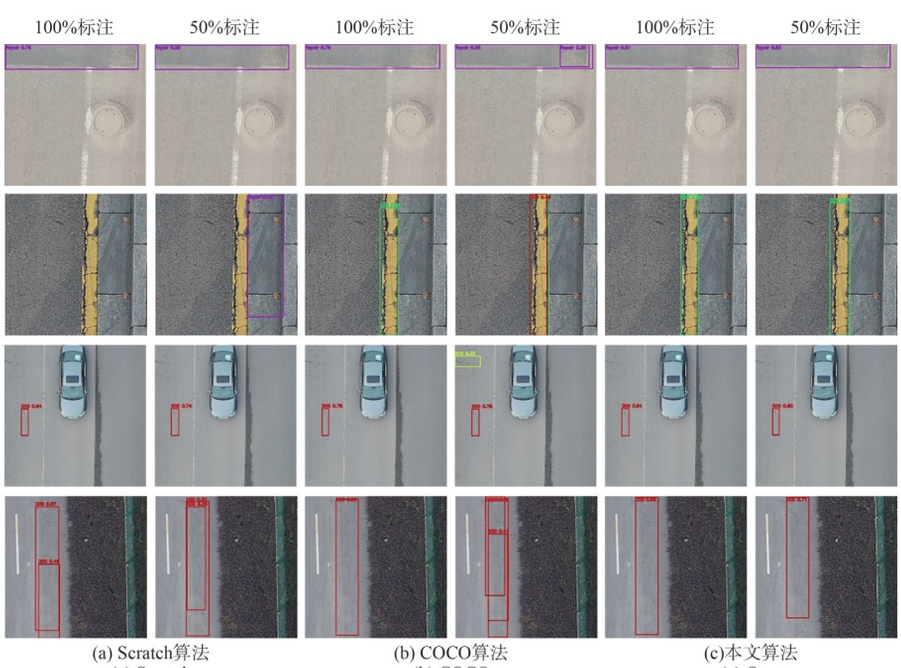


图 11 100%及50%标注比例下各算法公路病害检测结果对比

Fig. 11 Comparison of highway distress detection results under 100% and 50% annotation ratios

类与边界回归。在第3行图中，COCO预训练模型在50%标注下受到路面背景噪声干扰，将左上角的树枝背景误检为横向裂缝；而本文算法与Scratch算法均未受到该噪声干扰。第4行图中，Scratch算法在100%和50%标注下均出现了定位退化，产生多个重叠的冗余回归框；COCO预训练模型在50%标注下也退化出了重叠框；而本文算法在50%标注下依然保持了单一、精准的检测框输出，检测效能与100%全量标注下的COCO基线等效。

综上所述，在监督信息减半（50%）的严苛条件下，常规预训练与Scratch模型极易受到路面噪声、阴影及边界相似纹理的干扰，从而导致大量的误检、漏检以及多余冗余框。而本文算法得益于SimCLR自监督预训练阶段所构建的强大道路特征表征能力，不仅成功抑制了背景噪声，有效避免了误检与重叠框问题，还在减少50%标注数据依赖的情况下，实现了与全量标注下COCO预训练算法旗鼓相当的检测精度与鲁棒性。

3.7　公路病害检测可视化对比

为准确描绘SimCLR预训练的YOLOv8n算法在实际公路病害检测的效果，本文在RDD2022数据集中划分的测试集上随机选取部分图像，使用100%标注比例的训练集将本文算法、YOLOv8n算法及当前性能较优的YOLOv11n算法进行检测，检测效果对比图如图12所示。

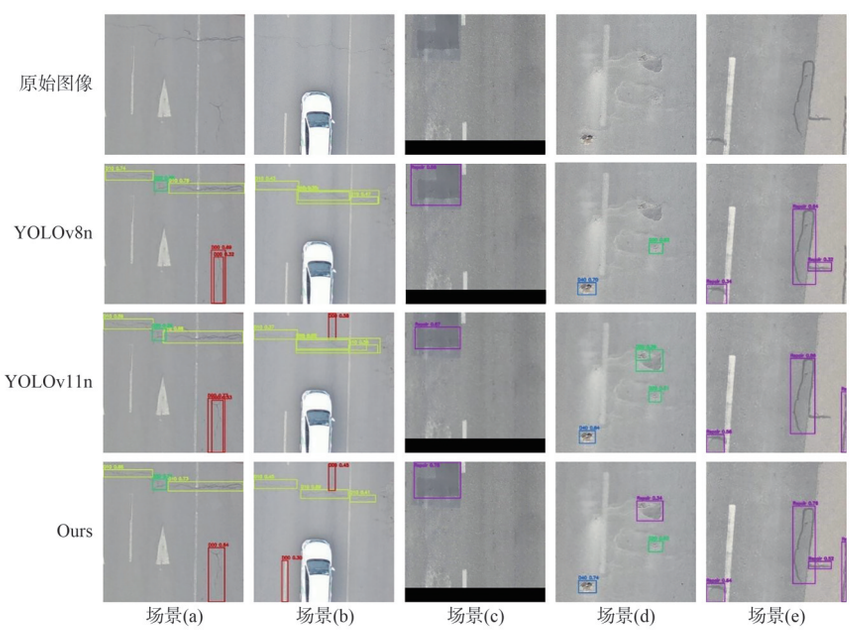


图12　检测效果对比图

Fig. 12　Comparison of detection result

从图12可以看到，在场景（a）与场景（b）中，YOLOv8n及YOLOv11n均在图像中产生了多个重叠的检测框，暴露出其在非极大值抑制（NMS）阶段未能有效过滤冗余检测的缺陷；而本文算法在两处场景中均无冗余框，实现了对目标的精准定位。场景（b）中，YOLOv8n与YOLOv11n产生了漏检，YOLOv11n较YOLOv8n相比少漏检了一个纵向裂缝，均漏检了白色车道线内的纵向裂缝；而本文算法不仅没有漏检，且在场景（c）中，并未将完好的车道线误判为纵向裂缝，这证明本文算法在场景（b）中对车道线内裂缝的检出并非随机巧合，而是实现了真正的背景与病害特征解耦。在场景（d）和场景（e）中，YOLOv8n漏检了场景（d）中的道路修补和场景（e）中右下角的修补，YOLOv11n虽定位到场景（d）中修补却将其误检为网状裂缝，并在场景（e）中遗漏了向右延伸的修补分支；而本文算法则精准无误地检出了场景（d）、（e）中的所有病害，分类无一误判，表现出极高的召回率。在低对比度的道路修补场景（c）中，3类算法虽然成功检测出目标，但YOLOv11n与本文算法的检测框贴合度不及YOLOv8n紧凑，这与表4中YOLOv8n拥有最高精确率而本文算

法及YOLOv11n精确率略低相吻合，表明模型在追求高召回率以减少漏检的同时，在定位精度上做出了一定权衡，这也为后续优化指明了方向。

综上所述，虽然本文算法在局部定位精度上做出了微弱牺牲，但在消除冗余检测、抑制病害漏检以及复杂背景解耦方面均展现出显著优势，整体检测性能与鲁棒性均优于对比算法，更适用于实际复杂路况下的公路病害检测。

4 结 论

本文提出一种融合无监督预训练SimCLR与YOLOv8n的SimCLR-YOLO算法，利用大量无标注图像提升模型的特征提取能力，使其在50%标注样本条件下也能达到接近全量标注的检测精度。具体改进包括以下3个方面：

首先，将SimCLR中原有的ResNet-50编码器替换为YOLOv8n的Backbone结构，实现了预训练与检测网络的结构对齐与特征无缝迁移，有效解决了因网络结构差异造成的特征迁移损失问题；其次，在对比学习中引入掩码重构机制，通过重构掩码区域，迫使编码器聚焦裂缝边缘、坑槽轮廓等高频局部纹理特征，弥补了对比学习对“细长弱纹理”目标特征的欠拟合缺陷。最后，在编码器后嵌入CBAM注意力模块，引导模型在预训练阶段抑制路面纹理、阴影等背景噪声，显著增强了病害关键区域的特征响应。

在RDD2022公开数据集上的实验结果表明，SimCLR-YOLO算法的mAP@0.5达到89.4%，较基线模型提升1.7%，mAP@0.5:0.95提升了1.1%，F1-score提升1.8%，同时Flops和Params均不变。尤为关键的是，本算法显著降低了检测任务对标注数据的依赖，在不同标注比例验证中，引入预训练权重的方案在所有数据梯度下均优于从头训练；在标注数据量仅为50%的少标注依赖场景下，本文算法的mAP@0.5仍可达87%，仅较全量标注基线模型的87.7%下降0.7%，证明了本文算法在标注样本稀缺条件下的实用价值与鲁棒性。综上所述，本文的改进算法在公路病害检测中具有轻量化零成本提升检测精度的优势，能够满足样本匮乏条件下公路病害实时检测的需求，为智慧交通养护提供了有效的解决方案与技术路径。

参 考 文 献：

[1] 唐湘松，杨炜，李相衡，等. 基于改进EfficientNetV2网络的路面类型及状态检测方法[J]. 汽车工程学报，2026，16(2)：294-304.
TANG X S, YANG W, LI X H, *et al*. Road-surface type and condition detection based on an improved EfficientNetV2 [J]. *Chinese Journal of Automotive Engineering*, 2026, 16(2): 294-304. (in Chinese)

[2] 马荣贵，黄训燕，董世浩. 基于改进YOLOv8的实时坑槽检测算法[J]. 计算机工程，2025，51(11)：226-234.
MA R G, HUANG X Y, DONG S H. Real-time pothole detection algorithm based on improved YOLOv8 [J]. *Computer Engineering*, 2025, 51(11): 226-234. (in Chinese)

[3] 朱世松，高虹，芦碧波，等. 基于改进YOLOv8n的城市道路病害检测算法[J]. 液晶与显示，2025，40(9)：1356-1368.
ZHU S S, GAO H, LU B B, *et al*. Urban road defect detection algorithm based on improved YOLOv8n [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(9): 1356-1368. (in Chinese)

[4] 周建新，张媛，贾梓涵，等. 改进YOLOv8的路面缺陷检测算法[J]. 电子测量技术，2024，47(19)：146-154.
ZHOU J X, ZHANG Y, JIA Z H, *et al*. The pavement defect detection algorithm of YOLOv8 was improved [J]. *Electronic Measurement Technology*, 2024, 47(19): 146-154. (in Chinese)

[5] 魏陈浩，杨睿，刘振丙，等. 具有双层路由注意力的YOLOv8道路场景目标检测方法[J]. 图学学报，2023，44(6)：1104-1111.
WEI C H, YANG R, LIU Z B, *et al*. YOLOv8 with Bi-level routing attention for road scene object detection [J]. *Journal of Graphics*, 2023, 44(6): 1104-1111. (in Chinese)

[6] 张子轶，马丽，吕帅，等. 基于RT-DETR的轻量化车辆目标检测算法[J]. 液晶与显示，2026，41(3)：373-387.
ZHANG Z Y, MA L, LÜ S, *et al*. Lightweight vehicle object detection algorithm based on RT-DETR [J]. *Chinese*

Journal of Liquid Crystals and Displays, 2026, 41(3): 373-387. (in Chinese)

[7] 张旭中,李波,贝绍轶,等. 改进YOLOv8的道路凹陷检测算法[J]. 重庆理工大学学报(自然科学),2024,38(10):79-87.
ZHANG X Z, LI B, BEI S Y, *et al*. Improved road depression detection algorithm for YOLOv8 [J]. *Journal of Chongqing University of Technology* (*Natural Science*), 2024, 38(10): 79-87. (in Chinese)

[8] YIN L, XIA G, WANG Q, *et al*. Infrared monitoring and identification of lubricating oil pipe on a gas turbine based on faster RCNN [C]//*Proceedings of SPIE* 12617, *Ninth Symposium on Novel Photoelectronic Detection Technology and Applications*. Hefei: SPIE, 2023: 126173I.

[9] VEERABHADRASWAMY K M, BANSOD P J, PATEL W, *et al*. Analysis of CT-images for COVID-19 using blockchain networks for mask RCNN and capsule [J]. *Journal of Environmental Protection and Ecology*, 2024, 25(5): 1717-1727.

[10] SYETIAWAN A, SUSETYO D B, LUMBAN-GAOL Y, *et al*. Deep learning-based palm tree detection in unmanned aerial vehicle imagery with Mask R-CNN [J]. *TELKOMNIKA*, 2025, 23(1): 156-165.

[11] WANG D, XUE Y K, LIU Y, *et al*. On the understanding and modeling of non-negligible subthreshold-state degradation (SSD) in sub-20-nm DRAM technology [J]. *IEEE Electron Device Letters*, 2024, 45(9): 1582-1585.

[12] 岳永恒,雷文朋. 基于改进的CycleGAN和YOLOv8联合雾天道路环境感知算法[J]. 华南理工大学学报(自然科学版),2025,53(2):48-57.
YUE Y H, LEI W P. Foggy road environment perception algorithm based on an improved CycleGAN and YOLOv8 [J]. *Journal of South China University of Technology* (*Natural Science Edition*), 2025, 53(2): 48-57. (in Chinese)

[13] 刘彦磊,李孟喆,王宣宣. 轻量型YOLOv5s车载红外图像目标检测[J]. 中国光学(中英文),2023,16(5):1045-1055.
LIU Y L, LI M Z, WANG X X. Lightweight YOLOv5s vehicle infrared image target detection [J]. *Chinese Optics*, 2023, 16(5): 1045-1055. (in Chinese)

[14] 罗磊,谢竹逵. 基于改进YOLOv8的交通标志检测算法[J]. 机电工程技术,2024,53(3):205-210.
LUO L, XIE Z K. Traffic sign detection algorithm based on improved YOLOv8 [J]. *Mechanical & Electrical Engineering Technology*, 2024, 53(3): 205-210. (in Chinese)

[15] 路成龙,庆光蔚,孙勇,等. 一种轻量级道路缺陷检测算法[J]. 汽车工程学报,2026,16(3):395-405.
LU C L, QING G W, SUN Y, *et al*. A lightweight algorithm for road defect detection [J]. *Chinese Journal of Automotive Engineering*, 2026, 16(3): 395-405. (in Chinese)

[16] GANGA B, LATA B T, VENUGOPAL K R. Object detection and crowd analysis using deep learning techniques: comprehensive review and future directions [J]. *Neurocomputing*, 2024, 597: 127932.

[17] 叶发茂,张立,袁燎,等. DB-YOLO:特征增强融合的双骨干YOLOv8道路缺陷检测模型[J]. 计算机工程与应用,2024,60(24):260-269.
YE F M, ZHANG L, YUAN L, *et al*. DB-YOLO: dual backbone YOLOv8 model with feature enhancement fusion for road defect detection [J]. *Computer Engineering and Applications*, 2024, 60(24): 260-269. (in Chinese)

[18] CHEN T, KORNBLITH S, NOROUZI M, *et al*. A simple framework for contrastive learning of visual representations [C]//*Proceedings of the 37th International Conference on Machine Learning*. Online: PMLR, 2020: 1597-1607.

[19] WOO S, PARK J, LEE J Y, *et al*. CBAM: convolutional block attention module [C]//*Proceedings of the 15th European Conference on Computer Vision* (*ECCV*). Munich: Springer, 2018: 3-19.

[20] HE K M, CHEN X L, XIE S N, *et al*. Masked autoencoders are scalable vision learners [C]//P*roceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans: IEEE, 2022: 15979-15988.

[21] ARYA D, MAEDA H, GHOSH S K, *et al*. RDD2022: a multi-national image dataset for automatic road damage detection[J]. *Geoscience Data Journal*, 2024, 11(4): 846-862.

作者简介:



向天明,男,硕士研究生,2023年于吉林化工大学获得学士学位,主要从事深度学习、目标检测方面的研究。E-mail: 1377950574@qq.com



燕　芳,女,教授,博士、2016年于北京科技大学获得博士学位,主要研究方向为无损检测及智能仪器仪表开发。E-mail:0472yanfang@163.com