

分类号: TP391.41

密 级: 公开

学校代码: 10363

学 号: 2210120102



安徽工程大学

Anhui Polytechnic University

硕士学位论文

题 目 基于图像融合增强的改进 YOLOX

轨道梁裂缝检测方法研究

论 文 作 者 黄伟

校 内 导 师 尚正阳 讲师

校 外 导 师 汪玉冰 高级工程师

专业学位类别 机械

研究方向(领域) 机器视觉

论文提交日期: 2024 年 5 月 31 日

分类号：TP391.41

单位代码：10363

密 级：公开

学 号：2210120102

安徽工程大学

基于图像融合增强的改进 YOLOX

轨道梁裂缝检测方法研究

Research on Improved YOLOX Track Beam Crack
Detection Method Based on Image Fusion Enhancement

论文作者：黄 伟

校内导师：尚正阳（讲师）

校外导师：汪玉冰（高级工程师）

专 业：机 械

研究方向：机器视觉

论文答辩日期：2024 年 5 月 17 日

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：黄伟

日期：2024年06月02日

安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 _____ 年解密后适用本版权书。

本学位论文属于

不保密 ☒。

学位论文作者签名：黄伟

日期：2024年06月02日

指导教师签名：[Signature]

日期：2024年06月02日

基于图像融合增强的改进 YOLOX 轨道梁裂缝检测方法研究

摘 要

PC 轨道梁作为跨座式单轨交通系统的重要组成部分，其表面的裂缝等缺陷会直接威胁整个交通线的运营安全。针对传统 PC 轨道梁面裂缝检测效率低、精度低、检测结果不准确等问题，本文提出一种基于图像融合增强的改进 YOLOX 轨道梁裂缝检测方法。该方法利用改进的 DCGAN 网络实现样本扩容图像的高分辨率输出，并对 YOLOX-s 算法部分结构进行改进，使图像增强与目标检测有效协同，提高检测精度。实验结果表明相较于传统人工检测方法和 YOLOX-s 检测算法，改进的网络模型 YOLOX-CRACK 不仅能保持高精度的轨道梁面裂缝缺陷识别，还能有效减少漏检与误判情况，特别是针对图像中难以检测的特征具有较好的检测效果。论文的主要工作如下：

（1）基于试验基地已有检测车上的工业相机和数据云台，构建了跨座式单轨 PC 轨道梁面图像采集模型。为解决原始轨道梁面图像分辨率过大与裂缝特征较小的不匹配问题，本文采用滑窗截图处理得到若干个分辨率较小有利于网络模型训练的轨道梁面图像，且每张图像包含突出的裂缝特征。之后对原始 PC 轨道梁面图像做预处理和数据增强操作，构建跨座式单轨 PC 轨道梁面裂缝图像数据集。

（2）针对原 DCGAN 网络生成轨道梁面图像中裂缝轮廓模糊的问题，采用 DCGAN 网络作为主要架构，通过引入自注意力机制来增强网络对图像局部和全局特征的捕获能力。同时，加深 DCGAN 模型的网络层数，以支持更大尺寸图像的生成；采用 Wasserstein 距离来取代传统的 JS 散度，优化了损失函数以提高模型的训练稳定性。通过对比实验验证，改进后的 DCGAN 网络基于 IS 和 FID 两个量化指

标上分别达到了 8.72 和 52.63，均优于原 DCGAN 网络，且生成的轨道梁面图像的裂缝轮廓更加明显，背景更加清晰。

(3) 为提高光照不均匀和杂质背景下的轨道梁面裂缝识别精度，减少模型参数量，以适应现场实际工程的应用。本文以 YOLOX 网络为基础提出图像融合增强的轨道梁面裂缝检测算法。通过在加强特征提取网络中使用 Ghost 轻量化卷积提取目标特征，以降低模型计算复杂度；在特征融合层中嵌入 ASSF 模块使模型自适应学习不同特征层间的联系，以充分利用每个特征层信息，加强网络对裂缝目标区域的关注度；同时引入 GIOU 损失函数，进一步加快网络收敛速度；采用双三次插值法进行上采样操作，更好的保持原始裂缝特征的清晰度和细节信息，提升网络检测性能。在自制 PC 轨道梁面裂缝数据集上的实验结果表明，改进后模型相较于原基准模型平均检测精度提升 4.40%，在参数量上比原模型下降 25.08%，每秒检测裂缝图像数量为 119frame。特别的，该方法对于光照不均匀以及杂质背景等不利识别因素的裂缝图像识别也具有较好的效果，为跨座式单轨 PC 轨道梁面裂缝的高效检测奠定了理论基础。

关键词：轨道梁面裂缝，特征融合，数据增强，DCGAN，YOLOX

RESEARCH ON IMPROVED YOLOX TRACK BEAM CRACK DETECTION METHOD BASED ON IMAGE FUSION ENHANCEMENT

ABSTRACT

PC rail beam is an important part of straddle monorail transit system, its surface cracks and other defects will directly threaten the safety of the entire traffic line. Aiming at the problems such as low efficiency, low precision and inaccurate detection results of traditional PC track beam surface crack detection, this paper proposes an improved YOLOX track beam crack detection method based on image fusion enhancement. This method uses the improved DCGAN network to realize the high-resolution output of sample expansion images, and improves the structure of part of the YOLOX-s algorithm, so that the image enhancement and target detection can effectively cooperate and improve the detection accuracy. The experimental results show that compared with traditional manual detection methods and YOLOX-s detection algorithm, the improved network model YOLOX-CRACK can not only maintain high accuracy in the identification of cracks on the surface of track beams, but also effectively reduce the cases of missed detection and misjudgment, especially for features difficult to detect in images. The main work of this paper is as follows:

(1) Based on the industrial camera and data head of the existing inspection vehicle in the test base, the image acquisition model of straddle monorail PC track beam surface is constructed. In order to solve the mismatch between the original track beam surface image resolution is too

large and the crack features are too small, this paper uses sliding window screenshot processing to obtain several track beam surface images with small resolution, which are conducive to network model training, and each image contains prominent crack features. After the original PC track beam surface image is preprocessed and data enhanced, the straddle monorail PC track beam surface crack image dataset is constructed.

(2) To solve the problem of fuzzy crack contour in track beam surface image generated by original DCGAN network, DCGAN network is adopted as the main framework, and self-attention mechanism is introduced to enhance the ability of the network to capture local and global features of image. At the same time, the network layer of DCGAN model is deepened to support the generation of larger size images. Wasserstein distance is used to replace the traditional JS divergence, and the loss function is optimized to improve the training stability of the model. Through comparative experiments, the improved DCGAN network reaches 8.72 and 52.63 respectively based on the quantitative indicators of IS and FID, both of which are better than the original DCGAN network. Moreover, the generated image of the track beam surface has more obvious crack contours and clearer background.

(3) In order to improve the crack identification accuracy of the track beam surface under the background of uneven illumination and impurities, the number of model parameters is reduced to adapt to the application of field engineering. Based on YOLOX network, an image fusion enhanced crack detection algorithm for track beam surface is proposed in this paper. By using Ghost lightweight convolution in enhanced feature extraction network to extract target features, the computational complexity of the model is reduced. By embedding ASSF module in the feature fusion layer, the model adaptively learns the relationship between different feature layers to make full use of the information of each feature layer and

strengthen the network's attention to the crack target region. At the same time, GIOU loss function is introduced to accelerate the convergence speed. Double cubic interpolation method is used for up-sampling operation to better maintain the clarity and details of the original crack features and improve the performance of network detection. The experimental results on the self-made PC track beam surface crack data set show that the average detection accuracy of the improved model is 4.40% higher than that of the original reference model, the number of parameters is 25.08% lower than that of the original model, and the number of crack images detected per second is 119frame. In particular, this method also has a good effect on crack image recognition of unfavorable identification factors such as uneven illumination and impurity background, which lays a theoretical foundation for the efficient detection of the cracks on the beam surface of straddle monorail PC track.

KEY WORDS: Track beam surface crack, Feature fusion, Data enhancement , DCGAN, YOLOX

目 录

第 1 章 绪论.....	1
1.1 研究背景及课题意义.....	1
1.2 国内外研究现状.....	2
1.2.1 基于传统图像处理的裂缝检测方法.....	2
1.2.2 基于深度学习的裂缝检测方法.....	3
1.3 主要研究内容.....	6
1.4 论文结构安排.....	7
第 2 章 PC 轨道梁裂缝缺陷检测方案设计.....	9
2.1 常见缺陷及危害.....	9
2.1.1 PC 轨道梁.....	9
2.1.2 PC 轨道梁面缺陷.....	10
2.2 缺陷特征分析.....	12
2.3 PC 轨道梁裂缝检测必要性分析.....	14
2.4 基于视觉检测的方案设计.....	16
2.5 本章小结.....	17
第 3 章 PC 轨道梁面裂缝图像数据集构建.....	18
3.1 PC 轨道梁面裂缝图像数据采集.....	18
3.2 PC 轨道梁面裂缝数据集预处理.....	20
3.2.1 滑窗截图.....	20
3.2.2 空间域低通滤波去噪.....	21
3.2.3 二值化.....	22
3.3 PC 轨道梁面裂缝数据集扩容.....	23
3.3.1 数据集扩容方法解析.....	23
3.3.2 基于生成式对抗网络的裂缝图像扩容方法.....	27
3.4 本章小结.....	28
第 4 章 改进 DCGAN 的 PC 轨道梁面裂缝图像增广方法	30

4.1 DCGAN 模型优化	30
4.1.1 自注意力机制引入.....	30
4.1.2 网络结构加深.....	32
4.1.3 损失函数改进.....	34
4.2 DCGAN 轨道梁面裂缝图像生成实验	36
4.2.1 生成式对抗网络模型评价指标.....	36
4.2.2 实验环境与模型训练.....	37
4.2.3 实验结果与分析.....	37
4.3 裂缝检测的数据增强有效性分析.....	39
4.4 本章小结.....	41
第 5 章 改进 YOLOX-S 的 PC 轨道梁面裂缝检测方法研究	42
5.1 YOLOX 网络运行原理解析.....	42
5.2 基于 YOLOX-CRACK 的 PC 轨道梁面裂缝检测算法设计.....	43
5.2.1 自适应特征融合机制.....	43
5.2.2 Ghost 卷积模块	45
5.2.3 双三次插值上采样.....	45
5.2.4 损失函数优化.....	47
5.3 YOLOX-CRACK 的效果验证与分析	47
5.3.1 目标检测评价指标.....	47
5.3.2 对比试验.....	49
5.3.3 消融试验.....	50
5.3.4 实验结果与分析.....	51
5.4 本章小结.....	53
第 6 章 结论与展望.....	54
6.1 研究总结.....	54
6.2 未来展望.....	55
参考文献.....	56
攻读硕士学位期间承担的科研任务与主要成果.....	62
致谢.....	63

第 1 章 绪论

1.1 研究背景及课题意义

随着全球城市化进程的加速，交通拥堵问题变得尤为凸显。近年来，大容量轨道交通系统迅速发展，在一定程度上缓解了大城市的交通拥堵问题。跨座式单轨^[1,2,3]是一种轻型轨道交通系统，可作为城区中低运量轨道交通的补充。跨座式单轨交通系统主要构成包括轨道梁^[4]、单轨车体及道岔^[5]，系统构成如图 1-1 所示。与传统轨道交通系统相比，单轨交通系统的造价成本更低、对地势的适应能力更强、没有多余的废气排放、具有投资和维护成本低等优点，适合人口密集的城市地区。图 1-2 为 2021 年在芜湖市已成功开通运营的无人驾驶跨座式单轨列车。

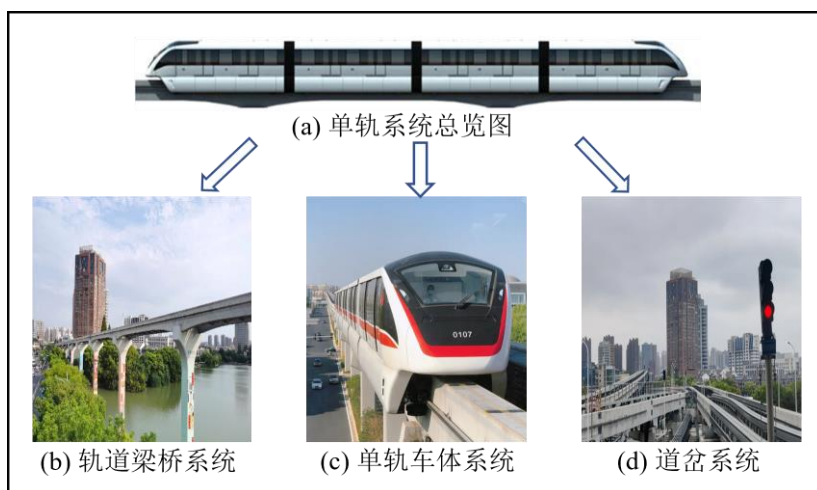


图 1-1 单轨系统构成



图 1-2 芜湖跨座式单轨列车

预应力混凝土（PC）轨道梁在列车运行时起到至关重要的作用，它不仅需要支撑车体重量，还要指引列车行驶方向。随着时间的推移，列车行驶引起的震动和持续负重会对轨道梁造成损害。国内跨座式单轨轨道梁的异常检测主要靠人眼

视觉判断,人工检测的准确性和效率受限于检测人员的经验,投入成本较高且检测精度不足。同时,轨道梁通常位于较高或隧道内部位置,附近有为列车供电的高压接触网,这些条件加剧了人工检测的困难和风险。因此设计高效率的自动化检测方案对于 PC 轨道梁面的精准有效异常检测格外重要。

1.2 国内外研究现状

1.2.1 基于传统图像处理的裂缝检测方法

在传统的图像处理技术中,检测裂缝通常采用几种方法:阈值分割^[6,7,8]、边缘检测^[9,10,11]的分割技术,以及基于频域分析的方法。

基于阈值的图像分割技术由于其较低的计算复杂度、快速的处理速度、简便的实现过程以及稳定的性能特性,已经得到了广泛应用。该方法的核心原理是通过设定特定的阈值对图像内的像素点进行比较,并基于这一比较结果来对像素点进行分类处理。Vivekananthan 等^[12]利用灰度强度调整模型,采用 Otsu 方法设定阈值范围,可以高精度检测出混凝土桥梁裂缝。Othman 等^[13]在裂缝图像的全局和局部阈值分析中实现了多分辨率级固定分区方法,结果表明该方法优于传统的 Canny 边缘检测方法。Chen 等^[14]开发了一种结合全局与局部阈值的自适应阈值裂缝检测技术,以分割图像并更准确地识别图中的道路裂缝区域。Shao 等^[15]提出了一种融合混合颗粒群优化算法的二维 Otsu 阈值分割技术,此方法不仅能高效精确的提取裂缝,而且能将检测到的线性裂缝宽度与实际测量值之间的相对误差维持在较低水平。熊玲玲等^[16]在改良后的 Otsu 算法上构建出一种新的桥梁裂缝检测算法,可自适应的识别出桥梁表面出现的裂缝缺陷。基于阈值的分割技术主要是利用像素的灰度值。当背景简单并且目标与背景的对比度较高时,这种方法可以有效地进行图像分割。在灰度值相近的图像中该方法的分割精度较低,且对噪声的抵抗力弱,分割效果显著受到阈值设置的影响。

图像的边缘分割方法依靠检测目标与背景的灰度变化差异进行图像划分,此法能精确捕捉边缘位置,执行速度快。常通过边缘检测算子来提取边缘信息,包括但不限于 Sobel^[17,18,19]、Prewitt^[20,21,22]、Canny^[23,24,25]及 Roberts^[26,27,28]等算子。郭运冲等^[29]基于边缘检测算法 Canny 结合多种改进措施,进一步增强算子对裂缝图像边缘信息的提取能力。Luo 等^[30]在 Canny 算法中引入双边滤波器和四向 Sobel 算子以增强降噪效果,更有效地提取裂缝边缘特征。Wang 等^[31]引入八向

Sobel 边缘检测算子提取裂缝边缘，边缘提取后对隧道衬砌裂缝图像进行数学形态运算，其整体裂缝识别准确率高于现有隧道衬砌裂缝识别算法，具有较高的研究意义和工程应用价值。Arya 等^[32]进行的研究涉及到 Sobel 边缘检测算子，评估了其单独及组合应用在道路裂缝识别中的效能。当图像的各个区域边界清晰可辨时，边缘检测技术可以实现较为满意的分割效果。这种方法也面临着一些限制，例如难以兼顾边缘检测的准确性与对噪声的抵抗力，存在的裂缝可能不连贯，以及可能出现假边缘等问题。这些挑战说明了在实际应用中边缘检测技术在维持高精度与高抗噪性之间需要找到一个平衡点。

采用频域检测技术，图像首先经过傅里叶或小波等正交变换被转移到频率域，依据裂缝和背景在频率域的表现差异来进行图像的区分。刘春阳等^[33]提出一种利用点云扫描线的混凝土预制箱梁表面裂缝识别方法，该方法采用小波分解来提高突变点的拾取准确性，并最终通过连接这些突变点来确定裂缝在平面上的分布。Nigam 等^[34]提出了一种利用离散小波变换检测光束裂纹的方法，光束非裂缝边缘的偏转被用作小波变换的输入信号以检测裂纹位置，该方法适用于数值模拟和实验测量的裂纹梁挠度。Vashist 等^[35]提出了一种基于主动振动控制和短时傅里叶变换相结合的开关控制策略，此技术基本上对预共振区域的信号进行操作，以此避免了裂缝扩展的机会。尽管基于频域的检测方法已经实现了某些成就，但它们对噪声高度敏感，算法复杂性较高，而且在多个领域的适用性有限。

尽管传统图像处理方法取得了进展，但面临泛化能力不足、高度依赖图像质量且对噪声敏感等问题，这些因素均限制了其检测性能。

1.2.2 基于深度学习的裂缝检测方法

AlexNet^[36]神经网络在 2012 年的 ImageNet 图像分类竞赛中夺冠引起了更多的研究者对深度学习的关注。随着时间的推移，深度学习技术在各个领域都得到了快速发展，尤其是在视频处理^[37,38]、面部表情识别^[39,40]和医疗图像处理^[41,42]等方面展现出了明显的优势。朱伟刚等^[43]针对无人机影像中裂缝目标小、图像背景复杂导致检测精度低等问题，提出了一种改进 YOLOv5 模型的深度学习道路裂缝检测方法。通过在 YOLOv5 模型骨干网络中分别加入 CBAM、SimAM、CA 注意力机制，提高模型对裂缝的识别能力及检测精度。Zhang 等^[44]提出一种改进 YOLOv3 的桥面裂缝检测算法，算法结合 MobileNets^[45,46,47]和 CBAM 卷积块注

注意力模块，实现桥面裂缝的快速准确检测。Xu 等^[48]探讨了运用深度学习技术进行道路裂缝智能识别的效果，并将 Faster R-CNN^[49,50,51]与 Mask R-CNN^[52,53,54]的性能进行了对比分析。研究发现通过采用联合训练方法，即便只利用 130 张图像进行训练，Faster R-CNN 和 Mask R-CNN 也能有效执行裂缝检测任务，并且其表现超越了 YOLOv3。尽管目标检测方法对裂缝识别和分类有效，但裂缝形态多样和尺寸不同使得仅用矩形框捕捉形状特征缺乏精确性。

采用基于语义分割的技术能够实现对裂缝缺陷的像素级别识别^[55,56]，这种方法通过精确地描绘裂缝缺陷的形状、尺寸和位置信息，为裂缝特性的深入分析提供了有力支持。杨燕泽等^[57]针对传统的沥青路面人工检测效率低、缺乏客观性的弊端，提出基于语义分割的沥青路面裂缝智能识别方法，为高速公路等场景下沥青路面裂缝的智能检测实践提供依据。Alipour 等^[58]在混凝土基础设施系统中使用深度全卷积神经网络进行像素级裂缝缺陷检测，模型能够正确检测验证集中超过 92%的裂纹像素和 99.9%的非裂纹像素。Pan 等^[59]开发了一种名为空间通道分层网络 SCHNet 的新型计算机视觉方法，该方法在像素级支持自动化和可靠的混凝土裂缝分割是可行的。Li 等^[60]提出了一种基于卷积神经网络的像素级裂缝分割网络 SCCDNet，网络使用深度可分离卷积来提高裂缝分割的准确性，同时也能降低模型的复杂性。Elhariri 等^[61]在对几种 U-Net 深度学习网络架构进行详尽研究的基础上，提出了一种自动化的深度裂缝分割方法，利用 U-Net 深度学习的像素级裂缝分割可识别对图像分类决策至关重要的像素。这类方法主要依赖于高层次的特征图来执行裂缝检测，裂缝的准确识别经常会由于遮挡以及光照条件等因素而受到影响。因此，为了实现像素级精确的识别，提升特征表示的辨识力显得尤为重要。

结合图像的多尺度信息，通过多尺度和多层级的特征融合，可以显著增强特征表示的辨识力。特征金字塔网络采用横向连接与自顶向下的架构策略，高效整合了不同层级的特征信息，从而获得了显著的性能提升。钟梅嘉等^[62]提出了一种名为 EAFNet 的道路裂缝检测方法，旨在解决复杂背景下细小裂缝难以检测和裂缝检测产生断裂的问题。通过在网络中间部分引入基于注意力机制的多尺度特征融合模块，以实现对裂缝的准确定位。该方法经过 CRACK500 数据集训练验证，并在两个道路裂缝数据集上进行测试，结果显示与目前部分检测方法相比，该算法在分割精度和泛化性能方面均有显著提升。Yan 等^[63]提出了一种新型的编码器

与解码器分割网络 CycleADC-Net, 网络采用循环生成对抗网络将弱光裂缝图像转换为亮域, 并将其输出馈送到后续的编码器-解码器分割网络进行最终检测。CycleADC-Net 在弱光裂缝和光照良好的裂缝数据库上均优于近期分割网络, 表明其在弱光照环境下具有良好的泛化能力和巨大的道路巡检潜力。Mei 等^[64]提出了一种用于像素级裂缝分割的新型深度学习架构, 其中最大池化层用于减小特征的维度, 转置卷积层用于多级特征融合, 有效的去除了孤立的像素, 提高了检测精度。Ye 等^[65]提出了一种聚合多尺度信息的像素级分割网络, 采用替代量化方法计算裂缝宽度, 该方法预测裂缝宽度的平均误差分别为 0.13 和 0.12mm, 为板坯轨道的维护和修复提供了新的方法和理论支撑。Xu 等^[66]提出一种基于卷积-反卷积特征融合的整体嵌套网络的混凝土裂缝分割网络, 通过卷积-反卷积特征融合模块, 将反卷积的深层语义信息与卷积的浅层特征有效融合, 有效改善了 VGG-16^[67,68]网络各阶段提取的语义裂缝特征信息。

近年来, 注意力机制越来越多的融入到深度学习领域中, 其已成为一种被广泛采用的改进技术。在分析裂缝图像时, 由于裂缝像素占整体的比例较低, 分配给这些区域的计算资源通常较少。这一现象削弱了模型在裂缝识别上的性能, 尤其是对于微小裂缝的检测。通过集成注意力机制可以有效地将计算资源集中于裂缝而非背景上, 从而抑制不相关特征, 减少噪声的干扰, 并降低误检和漏检的风险。Chu 等^[69]提出一种具有注意力机制的多尺度特征融合网络, 将双注意力模块被整合到架构中, 以更好地将微小的裂缝与背景分开。赵志宏等^[70]提出了一种融合 GhostUNet 的路面裂缝检测方法, 以解决现有深度学习提取裂缝特征不完整、精度低以及实时性不足的问题。在编码和解码部分, 为了增强对裂缝特征的提取能力, 引入了 ECA 注意力机制和残差连接, ECA 模块可过滤无关特征, 残差连接则可防止网络退化。与同类模型相比, 该模型具有更高的分割准确率和更少的参数。Qiao 等^[71]提出了一种基于深度特征聚合网络的新型裂缝检测算法, 其具有空间通道挤压和 SCSE 注意力机制模块, 使得网络模型在外部条件干扰下具有较好的鲁棒性和泛化能力。Yao 等^[72]提出了一种基于 YOLOv5 的路面裂缝检测方法, 利用 SCSE 模块和 CBAM 模块共同构建了 12 种不同的裂缝检测模型。相较于目前已有的路面裂缝检测技术, 该方法展现出了更高的检测精度、更快的速度以及更强的鲁棒性。

虽然基于深度学习的裂缝检测技术已经在某种程度上取得进展, 但仍面临两

大挑战。一是在收集裂缝图像的过程中，环境因素的干扰不可避免，导致获取到的图像常含有噪声，这增加了检测裂缝的难度。二是目前流行的裂缝检测算法大多设计用于识别明显的宽裂缝，而对微小裂缝的检测能力不足，使得这些小裂缝容易被遗漏。因此，解决这些挑战并提高裂缝检测的准确性和鲁棒性，成为了迫切需要解决的问题。

1.3 主要研究内容

本文针对横向裂缝、环向裂缝、斜向裂缝、锈蚀裂缝四种常见的 PC 轨道梁面表面缺陷，提出一种基于图像融合增强的改进 YOLOX 轨道梁裂缝检测方法。同时为了有效处理实地采集的轨道梁面图像数据，给出改进的 DCGAN 网络来生成质量与多样性更优异的裂缝图像；面对轨道梁面上存在光照不均匀和杂质背景干扰检测精度的挑战，提出以 YOLOX 网络为基础的图像融合增强轨道梁面裂缝检测算法，该算法旨在解决轨道梁面裂缝检测效率低、精度低、结果不准确等问题。面对与公路裂缝检测相比，单轨 PC 轨道梁面的裂缝检测面临着不同的挑战。

(1) 解决实际梁面存在的噪声和光照不均匀等图像特征干扰问题

为解决轨道梁面上存在噪声以及光照不均匀所带来的低检测精度的影响，提出一种基于改进 DCGAN 的 PC 轨道梁面裂缝图像增广方法。利用 DCGAN 网络对这些图像数据进行训练，生成逼真的合成图像。这些合成图像可以模拟出真实场景中存在的各种光照不均匀和噪声干扰的情况，使模型在面对真实数据时具有更好的泛化能力，实现在复杂环境下也能够高精度的检测和识别出裂缝。

与常规道路表面有所区别，跨座式单轨列车使用橡胶轮胎在轨道梁的走行面上按固定路径滚动，造成了轨道梁表面出现连续而明显的条纹痕迹，如图 1-3 展示。轨道梁所处在露天的外部环境条件下，长期的光照和湿度差异使得梁面颜色呈现显著的不同，以及梁面上的凹坑、水渍和油污等噪声进一步降低了梁面异常检测的准确度，这极大的干扰了梁面异常检测工作。PC 轨道梁表面上的光照不均匀现象是由车站屋顶以及周围建筑物的遮挡等因素共同导致的。车站屋顶作为一个大型遮挡物，会在一定范围内阻挡阳光直接照射到轨道梁表面上形成阴影区域如图 1-4 所示。光照不均匀会对裂缝检测产生以下几种主要影响：导致裂缝区域的亮度变化，使得某些裂缝区域过亮或过暗，干扰模型的特征提取和识别；可

能会引入阴影和光斑，使得部分裂缝被掩盖或模糊，从而影响检测的准确性；导致图像中的噪声增加，使得模型更难以区分裂缝和背景。



图 1-3 单轨梁走行面图像



图 1-4 轨道梁面光照不均匀

(2) 提高基于图像视觉的轨道梁裂缝的检测精度与速度

任何轨道梁面上的细微缺陷都会严重威胁列车的安全行驶，所以裂缝缺陷的检测对精度和速度都有着极高的要求。提出一种基于 YOLOX-CRACK 的 PC 轨道梁面裂缝检测方法，通过在加强特征提取网络中使用 Ghost 卷积提取目标特征，这种轻量化的设计有效减少了模型的计算负担，在提高检测速度的同时保持了对裂缝特征良好的提取能力；在特征融合层中嵌入 ASSF 模块使模型自适应学习不同特征层间的联系，提升模型对于轨道梁面裂缝的识别细节，从而增强检测的精度；同时引入 GIoU 损失函数，使模型在训练过程中能够更精细地调整边界框的位置和大小，减少定位误差，提升裂缝检测的精确度；采用双三次插值法进行上采样操作，有助于在最终的特征融合和分类过程中保持裂缝特征的清晰度和完整性，进一步提升检测的准确性。

1.4 论文结构安排

本文包含六个章节，每个章节的具体内容如下列所示：

第一章阐述了研究议题的背景与意义，综述了近年来国内外在裂缝检测方面的研究动态，并概述了本篇论文的主要研究内容。

第二章详述了跨座式单轨 PC 轨道梁面裂缝缺陷检测的整体规划设计。通过介绍 PC 轨道梁面常见的裂缝缺陷类型，进一步分析了进行轨道梁面裂缝缺陷检测的重要性，并提出了一个基于图像融合增强的改进 YOLOX 轨道梁裂缝检测方法。

第三章着重讨论了单轨 PC 轨道梁面裂缝图像数据集构建与数据增强方法。本章节详细描述了轨道梁面裂缝图像的采集方式及图像预处理的相关技巧，接着通过使用专门的标注工具对轨道梁面图像进行标注。随后论述了如何通过数据增强技术有效的增加轨道梁面裂缝图像数据集的规模，成功建立了符合本课题的裂缝数据集。

第四章重点提出一种针对性改进的深度卷积生成对抗网络模型。该模型结合了单轨 PC 轨道梁面裂缝图像的特征，通过引入自注意力机制、增强模型的深度和优化损失函数三个方面对 DCGAN 原模型进行了改良。这些改进措施增强了单轨 PC 轨道梁面裂缝数据集的扩增能力，为接下来章节中进行的轨道梁面裂缝目标检测任务提供了充实且高质量的样本量。

第五章阐述了对 YOLOX 模型进行改进以提升轨道梁面裂缝检测性能的内容。此章节详细介绍了 YOLOX 算法的工作原理并提出了针对性的改进措施，形成适用于单轨 PC 轨道梁面裂缝检测的改进算法 YOLOX-CRACK。构建了网络模型，并使用第三章中经过预处理和数据增强的轨道梁面图像数据集输入模型中进行训练。实验部分采用正负样本对改进的检测算法在裂缝检测任务中的适用性与效率进行验证，并对所获实验成果进行了细致的分析。

第六章中对前期工作研究成果进行了汇总，概述了文中提出的主要理论观点、关键技术及所采用的算法的改进之处。文末指出了当前研究存在的局限性及未来的改进方向，并对 PC 轨道梁面裂缝检测技术的未来发展进行了前瞻性的展望，表达了对此技术在未来取得更进一步成果的期望。

第 2 章 PC 轨道梁裂缝缺陷检测方案设计

2.1 常见缺陷及危害

2.1.1 PC 轨道梁

跨座式单轨 PC 轨道梁采用钢筋混凝土材料制成，具有矩形块状结构，其标准尺寸为高 1500 毫米，宽 850 毫米，并且长度大致为 20 米，PC 轨道梁结构如图 2-1 所示。轨道梁表面根据不同的功能需求分为走行面、导向面和稳定面。导向面设有用于安装接触网的凹槽，该接触网为列车提供必要的动力来源。走行面和导向面装配有指型板，这些指型板的作用是确保相邻轨道梁之间能够被精确地连接和对齐。整个单轨轨道是通过将若干个轨道梁与梁片支座固定在墩上，并用指型板将多个轨道梁按顺序连接起来，最终构建成一条完整的轨道线路，PC 轨道梁整体外观如图 2-2 所示。



图 2-1 PC 轨道梁结构



图 2-2 PC 轨道梁整体外观

2.1.2 PC 轨道梁面缺陷

跨座式单轨 PC 轨道梁作为现代轨道交通系统的重要组成部分，其稳定性与安全性对整个交通系统的运营至关重要。在长期使用和自然环境的双重作用下，轨道梁面可能出现多种缺陷，这些缺陷主要包括孔洞、裂缝、混凝土掉块和锈蚀等，缺陷基本统计信息如表 2-1 所示。

表 2-1 PC 轨道梁面缺陷基本信息

轨道梁面缺陷种类	主要产生原因	主要危害分析	主要数量分布
孔洞	混凝土浇筑和养护	削弱轨道梁的结构	约占轨道梁面缺陷
	不当	强度	总数量的 15%
裂缝	应力变化	降低轨道梁的整体	约占轨道梁面缺陷
		性能和耐久性	总数量的 70%
混凝土掉块	载荷振动	减弱轨道梁的承载	约占轨道梁面缺陷
		能力	总数量的 10%
锈蚀	环境中存在腐蚀性	导致混凝土裂缝和	约占轨道梁面缺陷
	介质	剥落	总数量的 5%

这些缺陷的产生基本上可以归因于以下几个方面：首先是设计与施工阶段的问题，包括材料选择不当、施工工艺控制不严格等；其次是长期荷载作用下的疲劳损伤，持续的列车通过会引起轨道梁材料疲劳，导致裂缝等缺陷的出现；再次是环境因素的影响，如温度变化、水侵蚀等，这些外界条件的变化会加速 PC 轨道梁的老化过程促使缺陷的产生。因此，为了确保跨座式单轨系统的安全可靠运行，对 PC 轨道梁进行定期的检查和维护，及时发现并处理这些缺陷，是非常有必要的。接下来是对 PC 轨道梁面裂缝缺陷详细的说明：

(1) 孔洞

跨座式单轨 PC 轨道梁面出现孔洞缺陷主要因混凝土浇筑和养护不当、原材料质量问题以及施工过程中的污染所致。不当的浇筑会导致混凝土内部孔隙形成，劣质的原材料降低了混凝土的密实度，而施工污染影响了混凝土的均匀性。孔洞的存在会削弱轨道梁的结构强度，降低承载能力，并为水分及腐蚀物质的渗透提供通道，加速内部钢筋的腐蚀过程。这不仅影响轨道梁的使用寿命，还增加维护成本。因此，防治孔洞缺陷需通过提升材料质量、优化施工工艺和加强后期养护

等措施实现。跨座式单轨 PC 轨道梁面孔洞缺陷如图 2-3 所示。

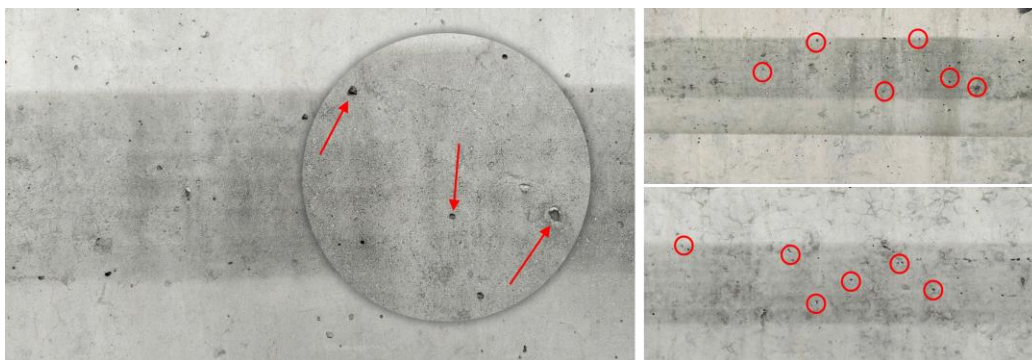


图 2-3 PC 轨道梁面孔洞缺陷

(2) 裂缝

跨座式单轨 PC 轨道梁面裂缝的产生主要受多种因素影响，包括混凝土材料特性、施工工艺、环境条件和荷载作用等。混凝土收缩、不恰当的养护方法、温度应力、超载以及施工过程中的振动都可能导致裂缝的形成。这些裂缝不仅影响轨道梁的外观，还可能降低轨道梁结构的整体性能和耐久性。因此，及时发现并修复裂缝，以及采取预防措施如优化混凝土配比、改进施工技术和加强养护管理，对于保障轨道梁的安全性和延长使用寿命至关重要。跨座式单轨 PC 轨道梁面裂缝缺陷如图 2-4 所示。

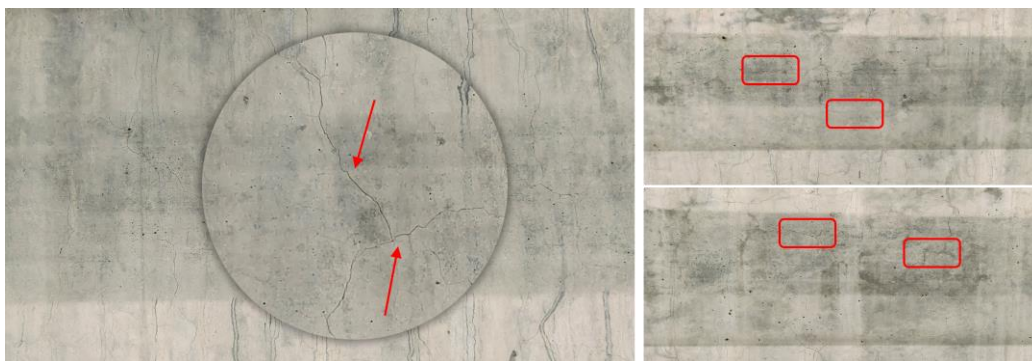


图 2-4 PC 轨道梁面裂缝缺陷

(3) 混凝土掉块

跨座式单轨 PC 轨道梁面混凝土掉块缺陷主要由混凝土本身的质量问题、不当的施工工艺、环境因素以及长期的机械疲劳等原因引起。混凝土质量不佳，如配比不当、养护不充分，会导致内部结构薄弱。此外，长期承受温度变化、湿度变化和载荷振动等环境因素，可使混凝土结构逐渐劣化，最终导致掉块。混凝土掉块不仅影响结构的完整性和美观，更重要的是会减弱轨道梁的承载力和耐久性，

增加事故风险。因此，采取合理的混凝土配比、严格施工控制及定期维护检查，对防止混凝土掉块至关重要。跨座式单轨 PC 轨道梁面混凝土掉块缺陷如图 2-5 所示。

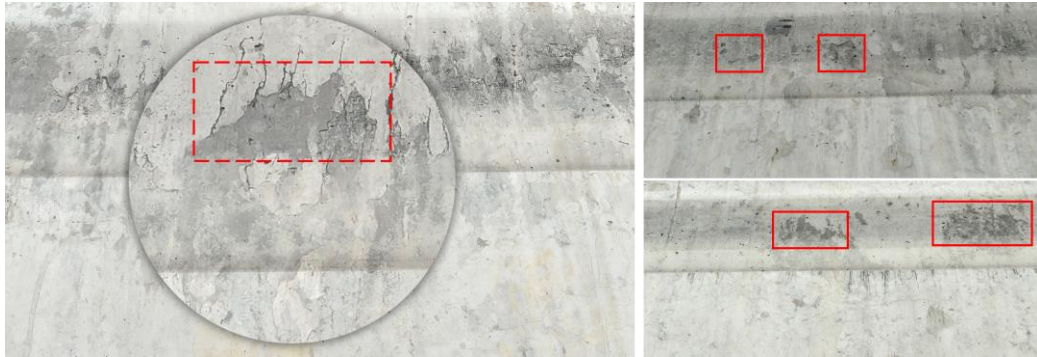


图 2-5 PC 轨道梁面混凝土掉块缺陷

（4）锈蚀

跨座式单轨 PC 轨道梁面锈蚀缺陷通常源于混凝土覆盖层内部钢筋的腐蚀，这一过程受多种因素影响，主要包括混凝土的碱性减弱、碳化、裂缝的形成、以及外界有害物质的侵入等。此外，环境中的湿度和温度变化也会加剧锈蚀速度。钢筋锈蚀会引起体积膨胀，进而导致混凝土裂缝和剥落，降低结构的整体性和承载能力，缩短服务寿命，严重时可能威胁到整个轨道系统的安全运行。因此，通过优化混凝土配方、确保足够的保护层厚度、使用防腐蚀材料和技术，以及定期进行结构健康监测和维护是防止和控制锈蚀缺陷发展的关键措施。跨座式单轨 PC 轨道梁面锈蚀缺陷如图 2-6 所示。

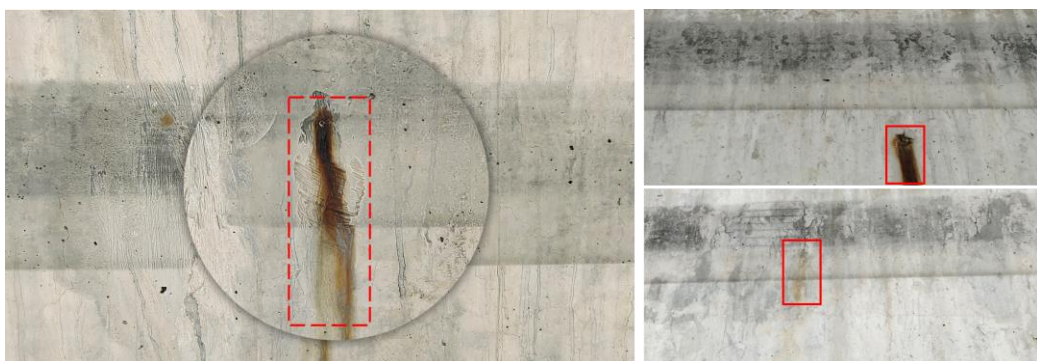


图 2-6 PC 轨道梁面锈蚀缺陷

2.2 缺陷特征分析

PC 轨道梁面的缺陷主要分为两类：结构性和功能性。结构性缺陷主要源于梁面老化问题，导致梁体本身的承重能力下降，并伴随着裂缝的出现；功能性缺

陷则是源于随着使用时间的累积，轨道梁的表面变得不平整，车辙越来越深。混凝土梁表面经常出现各种缺陷，其中最常见的是裂缝缺陷。初期结合单轨运维基地现场调研，根据裂缝特征和具体形态可将 PC 轨道梁面裂缝缺陷细分为以下四大类：横向裂缝、环向裂缝、斜向裂缝和锈蚀裂缝，它们以线状形式出现在梁的表面，如图 2-7 所示。

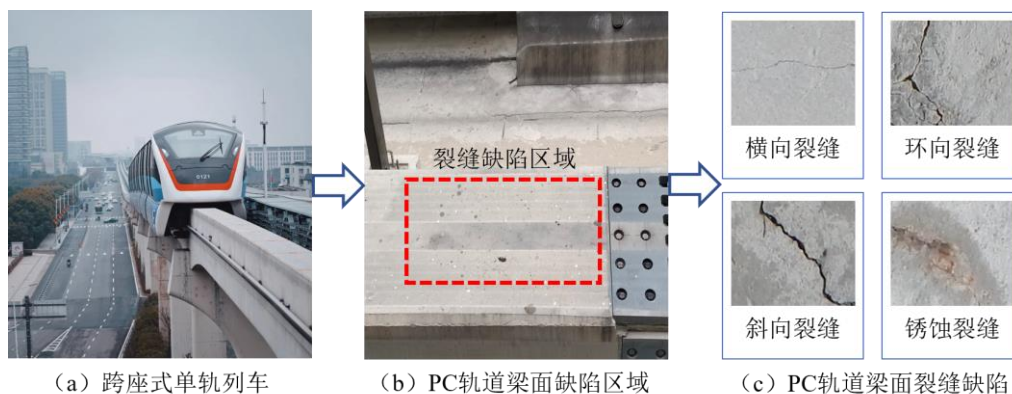


图 2-7 轨道梁面裂缝缺陷分类

不同类型的裂缝对轨道梁体结构的安全性和稳定性有着不同的影响，日常维保应根据具体情况采取不同的监测预警方法。

(1) 横向裂缝

横向裂缝缺陷主要产生原因是由于温度变化引起的应力变化。在昼夜温差较大的情况下，轨道梁受到热胀和冷缩的影响，这种热胀冷缩引起的应力可能会导致轨道梁材料的变形和开裂，从而在长时间的作用下促成横向裂缝的产生。横向裂缝缺陷如图 2-8 所示。

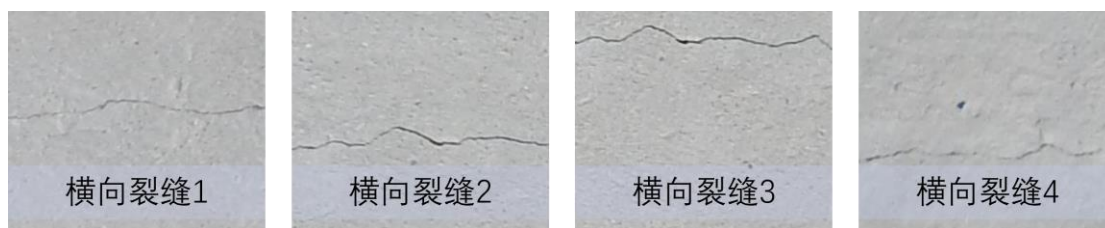


图 2-8 横向裂缝缺陷

(2) 环向裂缝

环向裂缝缺陷主要产生原因是由于混凝土收缩引起的应力集中。混凝土在固化过程中会产生收缩，在预应力混凝土轨道梁中由于预应力钢束的存在，混凝土的收缩会受到约束，导致在一些部位产生较大的收缩应力，最终引起环向裂缝的形成。环向裂缝缺陷如图 2-9 所示。

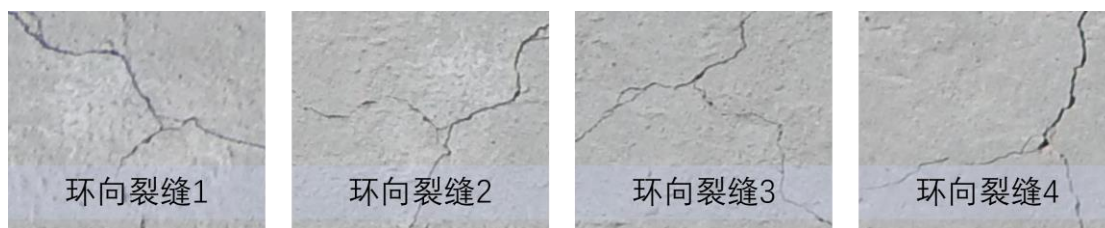


图 2-9 环向裂缝缺陷

(3) 斜向裂缝

斜向裂缝缺陷主要产生原因是由于混凝土内部受到的不均匀荷载引起的弯曲应力。在实际使用中,轨道梁可能会承受来自列车荷载、温度变化等引起的荷载作用,如果混凝土内部受到的荷载分布不均匀,就会导致混凝土承受不均匀的弯曲应力,最终引起斜向裂缝的形成。斜向裂缝缺陷如图 2-10 所示。

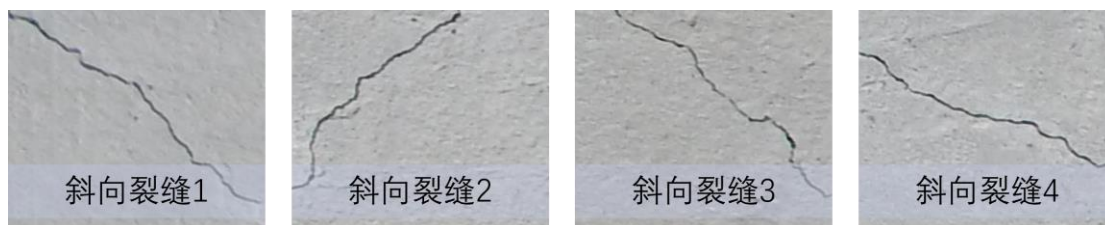


图 2-10 斜向裂缝缺陷

(4) 锈蚀裂缝

锈蚀裂缝缺陷主要产生原因是由于环境中存在腐蚀性介质。如果梁体表面的预应力混凝土被暴露在高湿度、高氯离子含量或其他腐蚀性介质的环境中,可能会导致梁体表面的保护层受到破坏,进而引发梁体内部钢筋锈蚀,最终导致锈蚀裂缝的形成。锈蚀裂缝缺陷如图 2-11 所示。

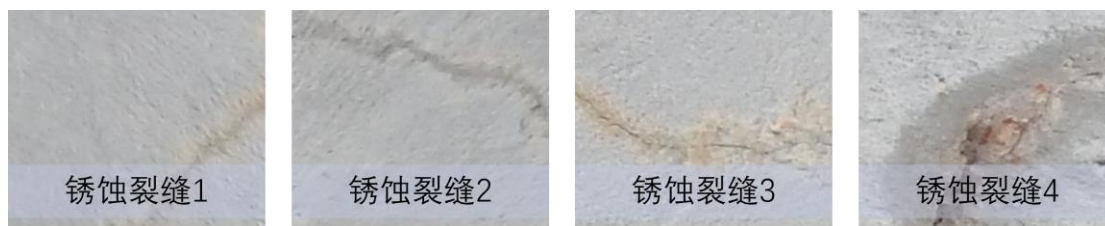


图 2-11 锈蚀裂缝缺陷

2.3 PC 轨道梁裂缝检测必要性分析

裂缝是 PC 轨道梁中一种较常遇到的问题,特指在跨座式单轨 PC 轨道梁的混凝土表层出现开裂的状况。这些裂缝最开始可能表现为单条或多条平行开裂,但如果不能及时被发现并进行修补,这些裂缝会逐步扩展,形成平行裂缝之间的

横向或斜向连接，引发锈蚀、脱落等更严重的问题，对 PC 轨道梁的稳定性构成威胁，甚至可能危及人民生命和财产安全。因此，实施跨座式单轨 PC 轨道梁面裂缝的及时检测显得尤为重要。目前，传统的裂缝缺陷检测主要依靠人工巡查，该方法通常利用装有升降装置的车辆进行作业，操作人员在升降机上通过手动调整高度，肉眼识别裂缝等缺陷，并使用测量工具记录裂缝的尺寸，将缺陷位置用粉笔标记以便后续修复。然而，这种人工检测方式效率低下，成本较高，且必须在 PC 轨道梁不承担单轨列车运营任务时进行。引入深度学习技术进行的跨座式单轨 PC 轨道梁面裂缝检测因而显得非常必要，以下从保障列车安全运行、延长轨道梁寿命、提升检测效率和准确度等方面，对基于深度学习的裂缝检测的必要性进行了细致分析。

（1）保障跨座式单轨列车运营安全

随着经济快速增长和城市化进程的加速，跨座式单轨的建设已成为提供便捷交通工具的首要任务。截至 2023 年 6 月，中国已有 9 个城市实现了跨座式单轨的覆盖，单轨交通系统承担了大多数发达城市居民日常通勤的重任。在建设和运营的不同阶段，由多种因素引起的 PC 轨道梁表面裂缝缺陷不可避免地出现。在这些缺陷中，裂缝类缺陷尤为普遍且危害性较大。因此，跨座式单轨 PC 轨道梁的裂缝检测与维护工作对于确保单轨系统稳定运行是极其关键的。

（2）延长 PC 轨道梁使用寿命

跨座式单轨 PC 轨道梁的建设是国家基建项目的重要组成部分，其初期建设需大量资金和劳力投入，具有民生工程的属性。该工程在施工初期就采纳了严格的建造流程，并有专门机构在其运行期间负责日常维护，以此确保轨道梁的安全运行与延长其服务寿命。在兴建与运行的过程中，裂缝缺陷是常见的问题。未能及时进行检测和维修，可能会使得细小的裂缝逐渐扩大和加深，增强因裂缝引起的安全隐患，也可能触发其他类型的缺陷如锈蚀和脱落等，这些都会削减 PC 轨道梁的预期寿命周期。因此，对于跨座式单轨 PC 轨道梁面裂缝的监测工作，显得尤为重要且必要。

（3）提高检测效率和精度

深度学习技术在跨座式单轨 PC 轨道梁面裂缝的检测中，能够有效克服传统人工检测的不足，显著提升检测效率。在传统的人工检查方式中，20 名检查人员用 5 个小时才能检查完毕约 1 公里的轨道梁，这种方法在效率上不仅低下，而且

过分依赖检查人员的个人经验并缺乏量化标准,因此在准确性上也存在明显不足。基于深度学习的方法主要分为两个阶段进行:第一阶段是对轨道梁面裂缝进行图像采集;第二阶段是利用深度学习网络模型处理这些图像数据,自动识别并学习图像中的裂缝特征。通常采集裂缝图像是通过配备高精度相机的自动检测车进行的,检测车能够在时速 50 公里以上的条件下获取高清的轨道梁面图像,突破了人工巡检仅在空窗期进行检测的限制,大大减少了 PC 轨道梁面检测和维护的成本。近年来,随着计算机视觉技术的进步和硬件性能的提升,图像处理技术已在众多工业领域得到应用。深度学习网络有能力学习缺陷特征,自动识别出问题并生成检测报告,从而替代大量人工检测和分析任务。当这一网络模型训练得当时,不仅可以有效解决人工检测的低效率问题,还显著提高了检测的准确度和灵活性,标志着跨座式单轨 PC 轨道梁面裂缝检测技术的一个重要发展趋势。

2.4 基于视觉检测的方案设计

图 2-12 详细展示了本文的完整研究方案,其初始步骤是对获取的跨座式单轨 PC 轨道梁面图像进行预处理,预处理步骤主要包括利用滑动窗口技术进行裁剪、图像降噪以及图像的二值化处理。由于单轨运维基地检测车拍摄获得的 PC 轨道梁面图像存在尺寸较大且包含许多非目标信息干扰的问题,故采取滑动窗口截图技术对图像进行裁剪,以生成尺寸为 800×800 像素的裂缝图像数据集,适配模型训练需求并减少冗余内容;对于图像降噪,使用均值和中值滤波方法去除轨道梁面图像的噪声,增强图像质量;将图像进行二值化处理,图像被简化为仅包含黑白两种颜色,这极大地简化了后续基于连通域的外接矩形图像提取过程,为精确标注图像上的裂缝创造了有利条件。

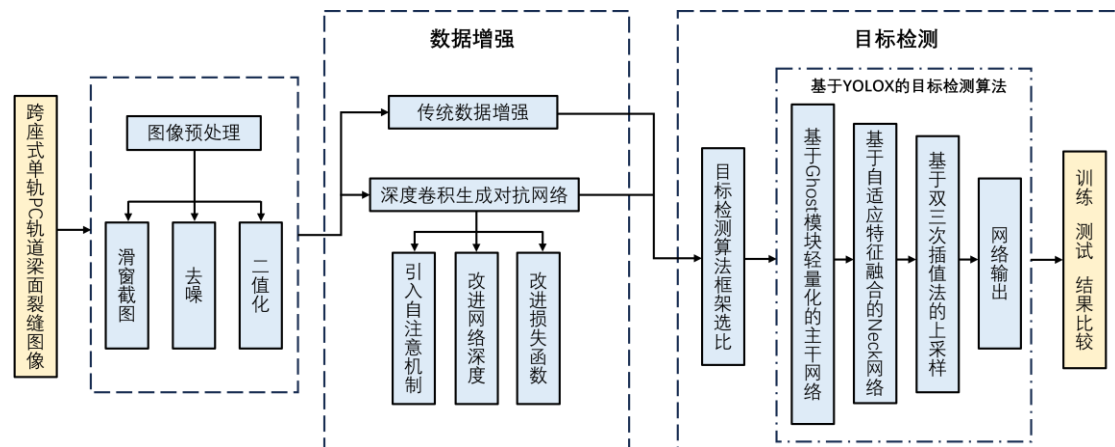


图 2-12 PC 轨道梁裂缝缺陷检测总体方案设计

完成图像预处理获得初始数据后,采用标准数据增强技巧和改进的深度卷积生成对抗网络对样本进行增强,从而为裂缝检测扩展了数据集。然而,深度卷积生成对抗网络训练困难、模式崩溃的风险高、产生的图像质量较差等问题依然存在。本文通过加深网络深度、引入注意力机制和改进损失函数来提高生成的 PC 轨道梁面裂缝图像的质量,目的是创建对裂缝识别任务更为有效的图像数据。关于目标检测框架的选择方面,本研究采用 YOLOX 目标检测框架来对原始轨道梁面裂缝数据集及其经数据增强技术产生的样本集进行分析。此外,为了进一步提高检测效率和准确性,对 YOLOX 算法进行了相应的改良和优化。首先,在加强特征提取网络中使用 Ghost 轻量化卷积提取目标特征,以降低模型计算复杂度。其次,在特征融合层中引入 ASSF 模块使模型自适应学习不同特征层间的联系,加强网络对裂缝目标区域的关注度。同时引入 GIoU 损失函数,进一步加快模型收敛速度。最后,模型采用双三次插值法进行上采样操作,更好地保持原始特征的清晰度和细节信息。最终根据对比试验的结果来进行研究分析。

2.5 本章小结

本章节详尽阐述了针对跨座式单轨 PC 轨道梁面裂缝缺陷检测的综合方案设计,并确立了算法的框架流程。分析了裂缝缺陷的细化特征,着重强调 PC 轨道梁面裂缝检测的必要性。本章就裂缝图像数据集的构建、利用生成式对抗网络增强梁面图像数据的方法、以及基于 YOLOX 模型进行轨道梁面裂缝检测的算法进行了初步说明,为接下来的图像增强和算法改进奠定了理论基础。

第 3 章 PC 轨道梁面裂缝图像数据集构建

3.1 PC 轨道梁面裂缝图像数据采集

数据集由专用检测车沿着单轨基地的试验轨道梁拍摄获取，相关图片均传送至数据台保存处理，试验基地如图 3-1 所示。工业相机被安装在检测车护栏前端，垂直距轨道梁面约 20cm 处，且镜头平行于轨道梁面，轨道梁面图像前期采集工作流程如图 3-2 所示。数据集的前期准备工作只针对单轨运维基地的指定试验线区域进行图像采集工作。考量到裂缝可能对轨道梁桥结构的安全构成潜在风险，这些裂缝在初期阶段便已得到修复和维护。最新采集的轨道梁面图像显示裂缝现象相对较少，经过筛选和处理共得到了 410 张分辨率为 800×800 的单轨 PC 轨道梁面裂缝原始图像数据集，部分数据集展示如图 3-3 所示。受到数据集的规模及其特征明显性的限制，进行数据增强操作是有必要的，详细的方法和步骤将在本章节后部分进行详细介绍。



图 3-1 单轨试验基地

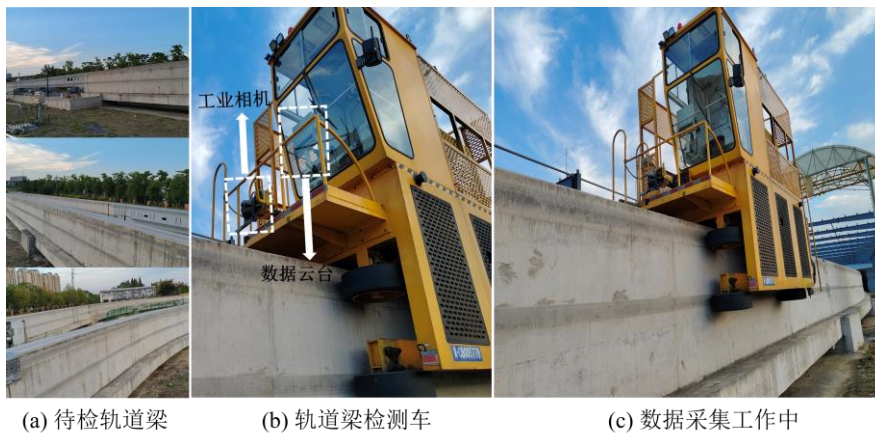


图 3-2 数据集采集

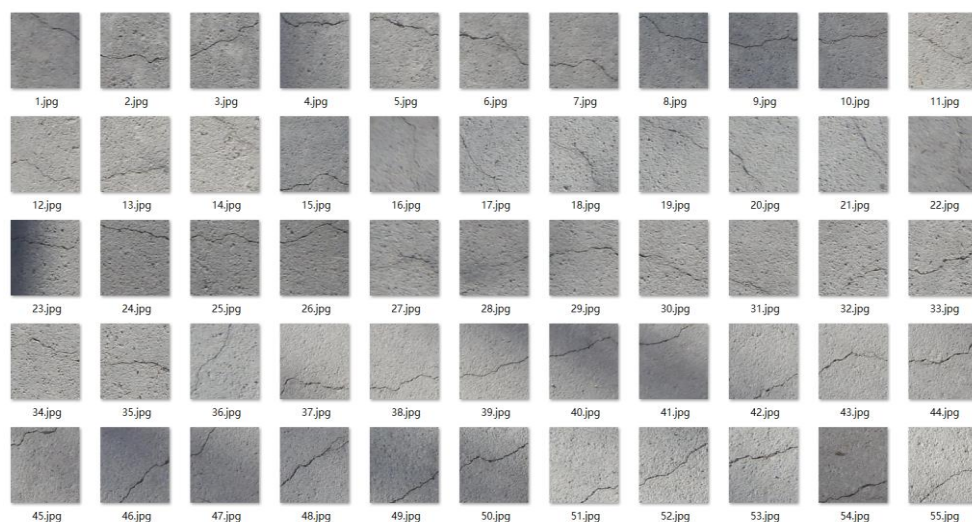


图 3-3 PC 轨道梁面裂缝数据集

在网络模型训练启动前，需对采集的裂缝图像进行精确的目标标记。本研究选用 LabelImg 这一专业图像标注工具完成裂缝的标识工作。标注的主要步骤涉及在图像中画出目标裂缝的矩形边界，并对其类别进行识别和标定，具体操作如图 3-4 所示，展现了使用 LabelImg 软件对裂缝图像进行标注的过程。

为了确保深度学习模型能够准确识别轨道梁面裂缝，标注的框架大小需适中：过大的标注框会导致模型识别效果不理想，而过小的标注框可能会使模型仅学习到图像中的噪声和其他干扰元素。基于此，本文采用了一种适度的连续多框标注方法来为目标检测算法准备数据，最大程度确保了识别精度。



图 3-4 数据集标注软件

LabelImg 工具不仅允许在原始图像上标记出裂缝的位置，而且为每张图像生成一个 txt 文件，该文件记录了边界框的坐标信息，这些坐标经过了归一化处理以适配 YOLOX 模型的输入数据格式。图 3-5 显示了这些 txt 文件的结构，其

中每一行数据对应一个目标物，包括目标类别和四个标注信息：边界框中心点的坐标以及边界框的宽度和高度。

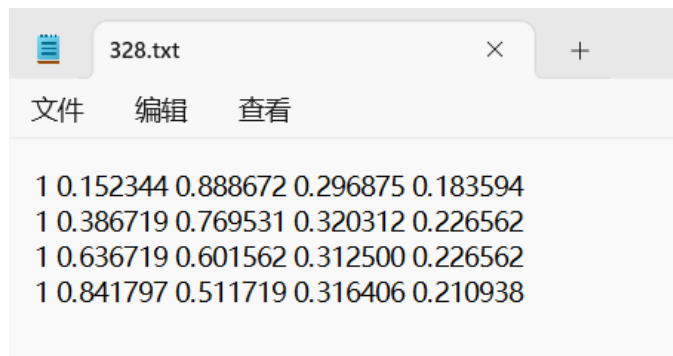


图 3-5 数据集标注文件

3.2 PC 轨道梁面裂缝数据集预处理

PC 轨道梁通常具有较长的长度，同时由检测车上的高精度工业相机获得像素高达 4608×3456 的原始轨道梁面图像，因此在连续捕获轨道梁面图像的过程中，将会产生较大量级的图像数据。直接使用这些原始采集的轨道梁面图像进行网络模型的训练，显然是不切实际的。将原始轨道梁面图像放大观察，可以注意到其混凝土表面存在大量水渍、油污和划伤等不利因素，如图 3-6 所示。在图像分析中，这些因素均表现为显著的噪声，一定程度上增加了图像处理的复杂度。

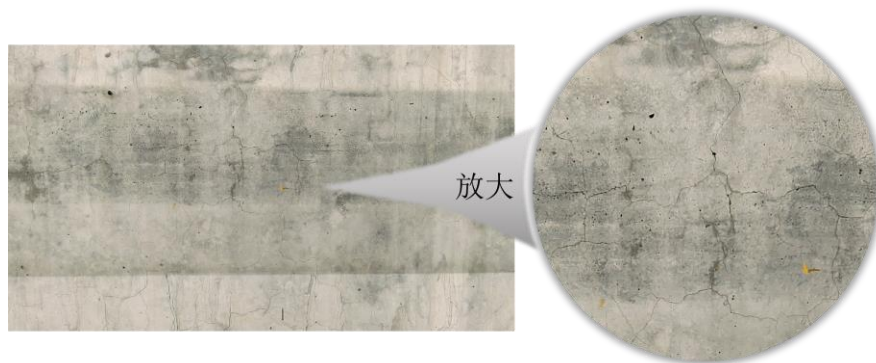


图 3-6 轨道梁面图像存在的噪声

3.2.1 滑窗截图

裂缝作为较小的目标特征，在原始 4608×3456 像素的轨道梁面图像中难以直接识别，本节采取了滑动窗口截图的方法来更好地捕捉这些特征。设置尺寸为 800×800 像素的滑动窗口，并以 100 像素的步长从原图中截取图像，部分滑窗截图所产生的结果如图 3-7 所示。因此，对初步处理获得的 410 张轨道梁面图像进行恰当的预处理，能极大降低网络所需的存储空间和计算能力，在实际环境条

件下也能提升目标检测网络的综合性能。

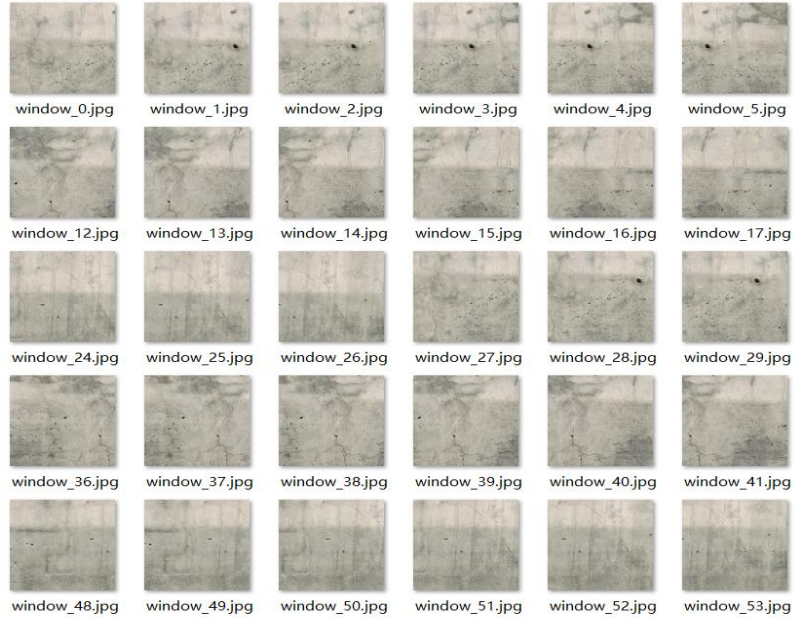


图 3-7 滑窗截图展示

由此，每张图片所获取的对应图片数量可以用以下公式表述：

$$T_x = \frac{w-d}{s} \quad (3-1)$$

$$T_y = \frac{h-d}{s} \quad (3-2)$$

$$S = T_x \times T_y \quad (3-3)$$

设定宽度为 w 、高度为 h ，而正方形滑动窗口的边长定为 d ，步长为 s 。基于这一前提，依据宽度方向上的滑动窗口操作可以确定图像的总数，记为 T_x ；同等情况下，在高度方向上进行相应操作，得到的图像总数记为 T_y 。进一步地，将每一张原始轨道梁面图像经过滑动窗口截取操作得到的图像数量总和记为 S 。本研究在处理截取图像数据集时，重点是挑选出那些梁表面上显示出明显裂缝特征的图像作为初始数据集。

3.2.2 空间域低通滤波去噪

图像的采集与数字化处理中，不可避免地会引入图像噪声，而图像去噪作用于降低或消除数字图像中的这些噪声。尤其是在 PC 轨道梁面裂缝图像的采集过程中，噪声的存在可能会导致图像质量下降，从而影响裂缝的识别与判断。去噪在图像处理的预处理阶段是一个关键步骤，以确保数据分析的准确性。

图像滤波是去噪过程中广泛应用的一种技术，它分为频率域滤波和空间域滤

波两种主要形式。频率域滤波负责对图像进行全局处理，应用诸如傅立叶变换、小波变换等方法在变换域内进行，处理完毕后再将结果转换回空间域。相较之下，空间域滤波技术着重于对图像局部区域的处理，通过直接操纵图像的像素值实现。为了达到降噪效果，这种方法采用了一个大小为 $a \times b$ 的滤波器。

在空间域滤波领域，可以区分为线性滤波和非线性滤波两大类。其中，线性滤波的代表为均值滤波，通过平滑处理有效降低噪声，但这可能会使图像的一些细节丢失。与此不同，非线性的中值滤波技术在减少噪声和平滑图像的过程中，能够更好地维持图像的细节，这是通过将滤波区内像素值排序后取中间数作为输出来实现的。考虑到跨座式单轨 PC 轨道梁面裂缝图像的特殊性，本研究采用空间滤波技术的均值与中值滤波去除噪声，目的是改善裂缝图像的预处理效果。

均值滤波作为线性滤波技术中的一种经典方法，其核心原理在于图像上移动一个固定尺寸的窗口，并计算该窗口内的像素灰度值的平均数以确定输出值，正如公式 3-4 所示：

$$g(x, y) = \frac{1}{N} \sum_{(x, y) \in M} f(x, y) \quad (3-4)$$

其中， $f(x, y)$ 代表中心像素点的值，而 $g(x, y)$ 表示经过滤波处理后的像素点的值。 N 表示窗口内的像素点数量，而 M 则表示模板内的像素点集合。

中值滤波属于非线性滤波技术，其操作过程中先对处于滤波窗口内的所有像素值进行排序，进而选取排序结果中的中值作为该点的最终输出值。符号 $\{x_{i,j}, i, j \in I\}$ 用来表示像素点的灰度值，滤波窗口 A 中的中值滤波可被定义如式 3-5 所示：

$$y_{ij} = \text{med}\{x_{ij}\} = \text{med}\{x_{i+r, j+s}, (r, s) \in A, (i, j) \in I^2\} \quad (3-5)$$

3.2.3 二值化

在本研究中，裂缝检测采取的是基于目标检测的图像识别方法，而非传统的在语义分割中常见的图像二值化技术。在图像预处理阶段，通过对轨道梁表面的裂缝图像执行二值化处理，并对裂缝的外接矩形连通区域进行标定，目的是为了获得精确的标定坐标。此步骤旨在为后续的数据增强，特别是图像旋转后的标注任务提供简化过程。二值图像是一种只包含两种颜色，即纯黑（灰度值为 0）和纯白（灰度值为 1）的图像。

二值化过程涉及设置一个界定阈值，以此阈值为基准，将灰度值超过该阈值的像素转换为 1，低于该阈值的像素转换为 0。详细的操作流程如式 3-6 所示：

$$I(x,y)=\begin{cases} 1 & g(x,y)>T \\ 0 & g(x,y)<T \end{cases} \quad (3-6)$$

其中， $I(x,y)$ 代表二值图像中坐标为 (x,y) 的像素的灰度值，而 T 则表示二值化操作所采用的阈值。

3.3 PC 轨道梁面裂缝数据集扩容

3.3.1 数据集扩容方法解析

数据增强方法能够对全部样本进行变换，有效增加样本数，尤其适合于样本量较小的数据集。针对本研究实验部分所处理的原始数据集，其仅包括数百个样本，明显属于小规模样本数据集范畴，因此，运用数据增强技术来扩展样本规模显得尤为关键。

几何变换增强不仅可以扩展数据集，还可以帮助模型更好地捕捉图像中的特征和细节，从而提高模型的性能和准确度。通过在跨座式单轨 PC 轨道梁面裂缝图像上应用平移、镜像、旋转和颜色变换等操作，可以增加数据的多样性，使模型在训练过程中能够更全面地学习裂缝的形状、大小和位置变化，进而提升裂缝检测的效果。几何变换增强还可以提高模型的泛化能力，使其在面对不同环境和条件下的裂缝图像时表现更稳定和可靠。综合而言，利用几何变换增强技术来扩充轨道梁面裂缝数据集，能够有效地提升裂缝检测模型的性能和鲁棒性，为 PC 轨道梁的健康监测和维护提供有力的支持。多种数据增强方法和模型训练策略相配合可以进一步提高裂缝检测系统的精度和稳定性，确保跨座式单轨列车运输的安全性。

(1) 平移

设置指定的距离和步进值，图像可以执行水平或垂直方向上的位移变换，此操作使图像在二维坐标系中的初始位置发生改变。某些像素区域在执行该位移过程中可能会脱离原始图像边界，因此需要用黑色像素来填补这些缺失部分，以保证视觉连贯性。该过程的数学实现遵循特定的变换如式 3-7 所示：

$$\begin{bmatrix} x_a & y_a & 1 \end{bmatrix} = \begin{bmatrix} x_b & y_b & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \Delta_x & \Delta_y & 1 \end{bmatrix} \quad (3-7)$$

其中， (x_a, y_a) 表示图像平移前的坐标点，而 (x_b, y_b) 代表图像平移后的坐标点，其中 Δx 和 Δy 分别为水平和垂直方向的偏移量。图像的平移效果如图 3-8 中所示：

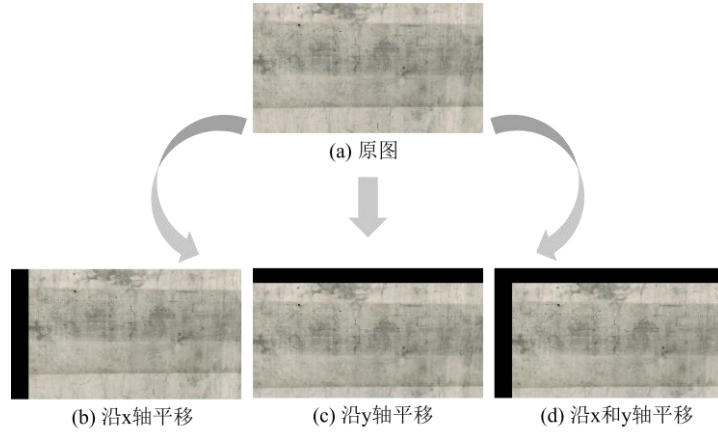


图 3-8 图像平移图

(2) 翻转

图像翻转是一种数据增强技术，它包含水平翻转、垂直翻转和对角线翻转等操作。这些操作可以使得原始数据集的规模扩大至四倍大小。应用这些不同的翻转方法可以生成具有变化的视觉效果的数据集如图 3-9 所示。对于单轨 PC 轨道梁面裂缝检测，尤其是在环向裂缝数据较为丰富的情况下，对角线翻转可用于人为增加环向裂缝的数据量，以此来改善数据集的多样性。

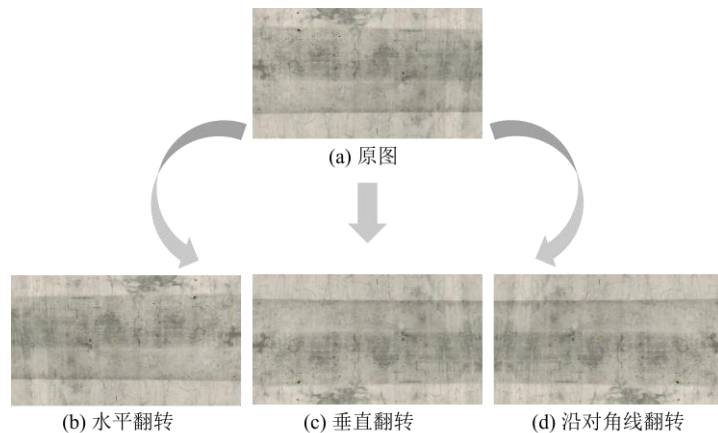


图 3-9 图像翻转图

(3) 旋转

利用图像的旋转操作，在中心点为轴进行角度调整，并对因旋转产生的空白区域进行单色填补，一方面扩大了数据集的体量，另一方面也丰富了数据集中斜向裂缝的样本数量。这种数据增强手段为深度神经网络训练提供了稳固基础，并且有效避免了网络模型过拟合的问题，同时增强了网络在识别多方向裂缝上的泛化能力。图像在顺时针方向的旋转变换可以参照公式 3-8 来说明，而其逆时针方向的旋转变换则通过公式 3-9 进行说明。

$$\begin{bmatrix} x_b & y_b & 1 \end{bmatrix} = \begin{bmatrix} x_a & y_a & 1 \end{bmatrix} \begin{bmatrix} \cos \beta & \sin \beta & 0 \\ -\sin \beta & \cos \beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3-8)$$

$$\begin{bmatrix} x_b & y_b & 1 \end{bmatrix} = \begin{bmatrix} x_a & y_a & 1 \end{bmatrix} \begin{bmatrix} \cos \beta & -\sin \beta & 0 \\ \sin \beta & \cos \beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3-9)$$

其中， (x_a, y_a) 表示图像平移前的坐标点，而 (x_b, y_b) 代表图像平移后的坐标点，旋转角度为 β 。图 3-10 展现了图像旋转的过程：

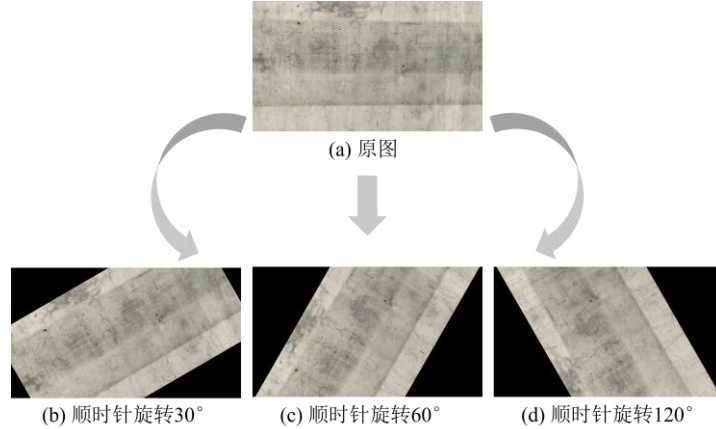


图 3-10 图像旋转图

运用目标检测标注软件 LabelImg 完成了对原始跨座式单轨 PC 轨道梁裂缝图像的标注工作，生成的标注信息遵循了 YOLO 标准，存储在.txt 格式文件中。这些文件里，每行数据都对应一个被检测的目标，包括一个类别索引和四个归一化后的位置参数，分别表示目标的中心坐标、宽度和高度。面临的挑战在于当图像经过旋转后，需要相应地调整裂缝检测框的位置参数，这种调整超出了简单仿射变换的范畴。因此，对于旋转图像的数据标签制作过程中，需要采取特定方法

来正确更新这些位置信息，以确保旋转后的图像标注的准确性，从而为后续的目标检测提供精确的训练数据。

在本研究中首先对数据集内裂缝的像素点进行了详细标注，接着对这些已标注的裂缝图像实施了二值化处理。完成二值化后，根据预设的角度 180° 旋转了图像。旋转过程之后，采用了连通域分析技术，对旋转后二值化的裂缝图像中的标注信息进行了精确提取。详细的处理流程展示于图 3-11。

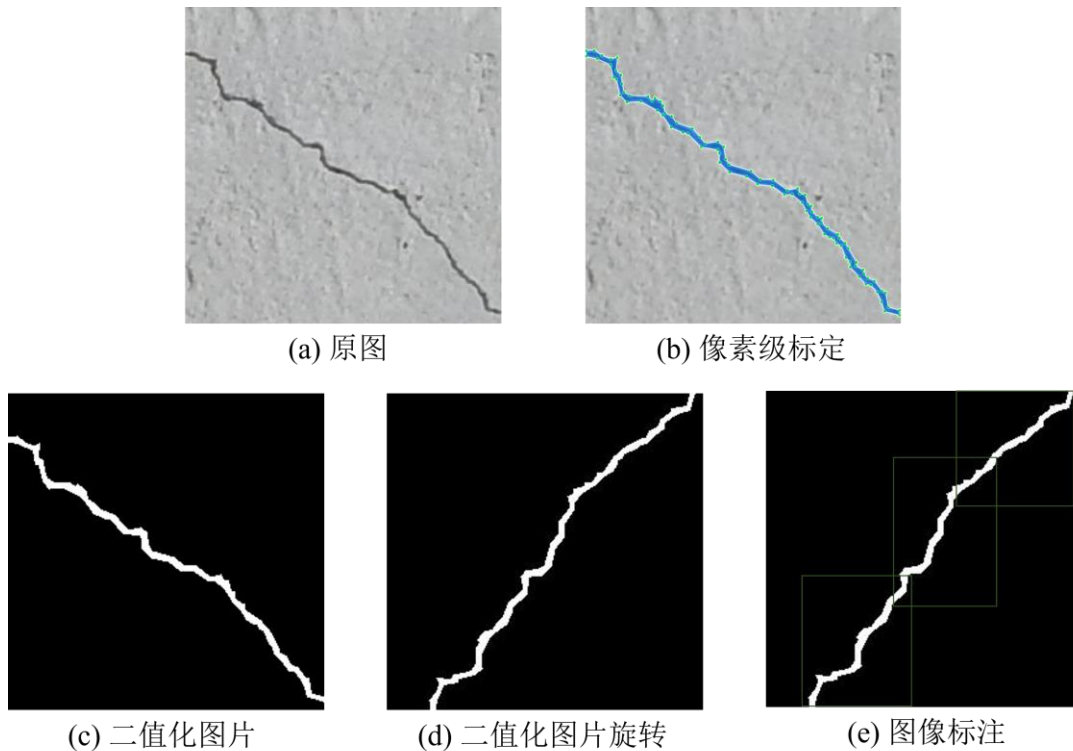


图 3-11 图像旋转图标注获取

(4) 颜色变换

在几何变换中，图像的像素仅在位置上进行调整，其像素值保持不变。相反，颜色变换类型的数据增强手段通过修改图像的像素内容来实现效果，涉及的技术包括添加噪声、应用模糊效果、进行颜色调整以及采用超像素技术等方法。

针对原始的跨座式单轨 PC 轨道梁面裂缝图像数据集的主要特点，将单通道灰度图像转换成三通道图像后进行颜色空间的变换可能并不理想。因此，本项研究采纳了不同的方法进行颜色变换增强，具体包括添加高斯噪声、采用图像模糊处理、以及调节图像的亮度和对比度，增强效果展示于图 3-12。

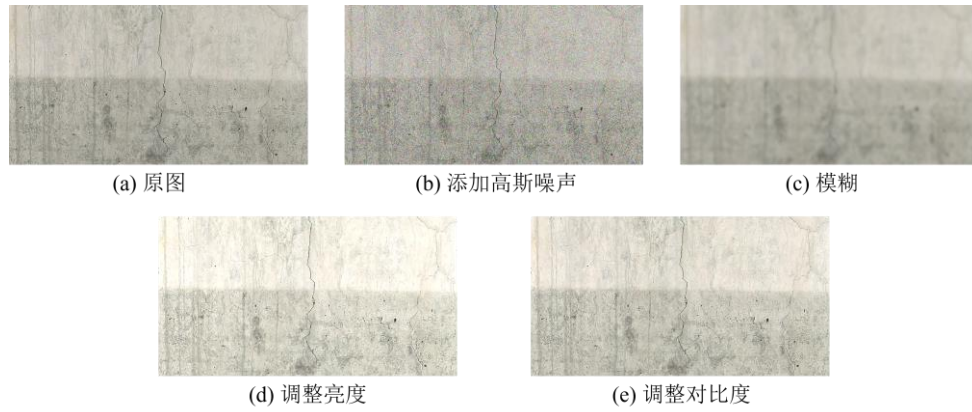


图 3-12 颜色数据变换增强

3.3.2 基于生成式对抗网络的裂缝图像扩容方法

生成对抗网络（GAN）是一种计算机网络，可以通过训练两个竞争网络（生成器和鉴别器）来生成目标解剖结构的高质量、逼真图像。GAN 的关键创新是利用生成器网络和判别器网络之间的竞争来获得高质量的合成数据。GAN 网络通过尝试欺骗鉴别器来学习生成更真实的数据，而鉴别器网络则通过尝试将样本正确分类为真实样本或合成样本来学习从人工数据中改进原始数据。整个过程中，两个网络同时得到改进，直到生成器产生与真实数据非常相似的高质量合成数据。GAN 已成功应用于各种计算机视觉任务，例如图像生成、图像到图像转换、图像超分辨率和 3D 对象生成。图 3-13 展示了 GAN 网络的核心流程。

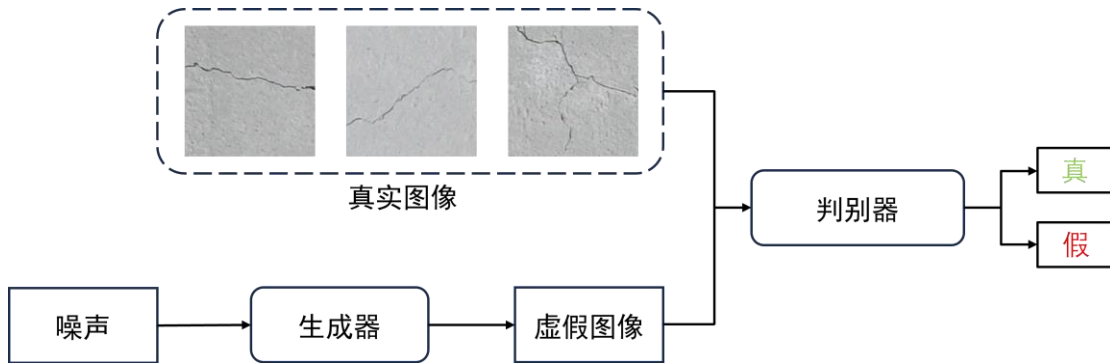


图 3-13 生成式对抗网络流程图

GAN 网络的训练具有不稳定性，这种不稳定性可能会导致训练过程出现分歧，并难以达到预期的结果。GAN 中训练的不稳定是由于模式崩溃造成的，即生成器无法探索整个数据分布，而是产生有限的样本变化，这可能会直接导致重复或低多样性输出。2016 年，研究人员引入了深度卷积生成对抗网络（DCGAN）的概念。这种方法涉及利用卷积神经网络（CNN）有效地从数据集中提取关键特征。DCGAN 旨在增强生成式对抗网络模型的训练过程，使其更易于管理以防止

收敛问题。DCGAN 的主要网络结构如图 3-14 所示：

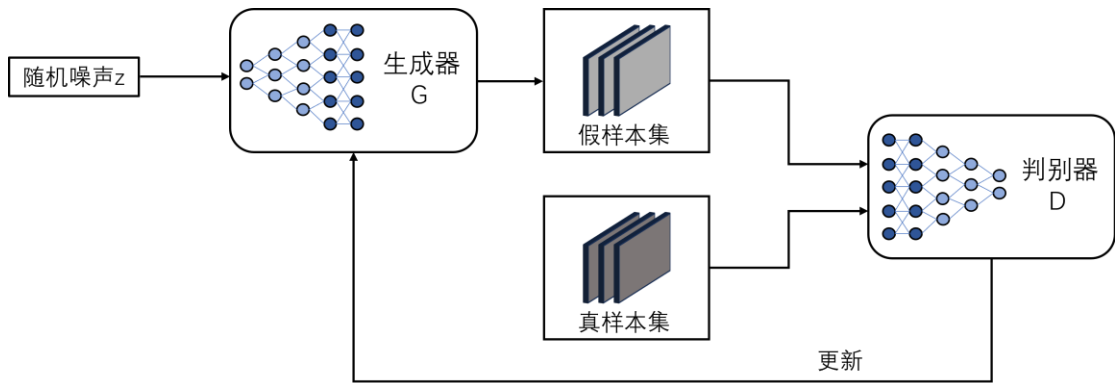


图 3-14 DCGAN 网络结构图

将经过几何变换增强的裂缝图像输入到 DCGAN 网络模型中进行简单预训练。部分由 DCGAN 网络迭代 500 轮生成的新 PC 轨道梁面裂缝图像如 3-15 所示：

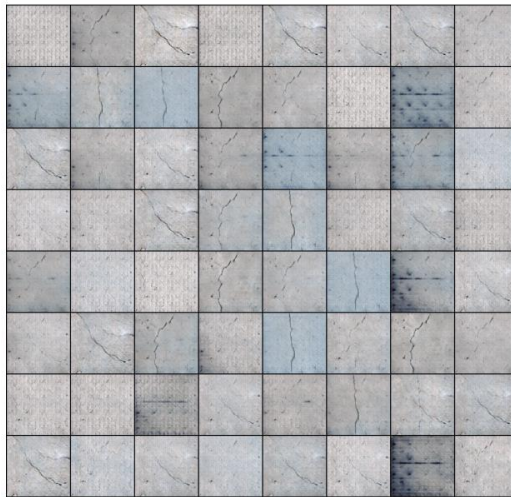


图 3-15 DCGAN 生成裂缝图像

由图 3-15 得知，DCGAN 网络生成的部分 PC 轨道梁面裂缝图像存在裂缝轮廓显示不明显以及背景模糊的现象。为了能生成裂缝轮廓与实际裂缝更为接近并且清晰度达到理想水平的裂缝图像，本文将会在第四章重点对原 DCGAN 网络进行针对性的改进。

3.4 本章小结

本章节首先阐述了获取跨座式单轨 PC 轨道梁面裂缝图像的方法及图像的特性分析。采用的图像捕获设备获取的大尺寸图像并不适合作为网络模型的输入，因裂缝数量在图像中较少，部分图像甚至不包含裂缝缺损。这些图像是在室外自然环境中获得的，采集和传输过程中难免会引入噪声。因此，原始图像的预处理

成为了后续研究工作的关键前置步骤。此阶段通过滑动窗口方法裁剪图像、应用空间域低通滤波消除噪点及执行图像二值化处理，以形成初步的轨道梁面裂缝图像数据集。详细讲述了传统的数据增强技术，随后介绍生成式对抗网络的图像增广方法。章节详细说明了用于数据标注的软件工具及标注后样本数据集的格式，这一环节为深度学习算法研究章节提供了必需的数据基础。

第4章 改进 DCGAN 的 PC 轨道梁面裂缝图像增广方法

由于利用大量数据训练能够改善求解效果,使得深度学习模型在图像识别领域取得了显著进展。数据增强是解决过拟合的一种有效手段,通过对现有数据进行多种转换或学习的方式,可以扩展数据集的规模和改善数据的质量。数据增强在深度学习的各个领域中发挥着关键作用,有助于提高模型的性能和泛化能力。

由前期数据采集和预处理所获得的 410 张跨座式单轨 PC 轨道梁面原始图像,这一数量的图片与深度学习神经网络所需的训练集数量存在明显差距。采用数据增强方法有助于优化神经网络的训练过程,提高模型的泛化能力。此方法能够基于初始较小规模的数据集,生成规模更大、更为可靠且高效的数据集,进而为 PC 轨道梁面裂缝检测的目标识别算法开发提供了充足的数据资源。应用数据增强可以增加数据样本的多样性,提高算法的准确性和稳定性。

本章内容重点讨论了对深度卷积生成对抗网络的结构进行改良后,如何利用预处理的轨道梁面裂缝图像数据集进行训练来生产出高质量的裂缝图像,在此基础上解决原始网络自身图像生成质量的问题。

4.1 DCGAN 模型优化

虽然 DCGAN 在图像生成领域表现良好,但在轨道梁面裂缝图像增广方面仍存在待解决的问题:其产生的图像分辨率通常不会超过 64×64 ,且在较大尺寸下图像清晰度受限;由于生成器和判别器是同时训练的,且它们的目标相反,这可能导致训练过程中的不稳定性。此外,模型还容易遭受模式崩塌,导致生成的图像在变化范围上过于狭窄,缺乏多样性,进而使得不同图像间呈现出高度的相似性。本研究在考虑到 DCGAN 在生成 PC 轨道梁面裂缝图像方面所存在的不足之处,提出了一个经过优化的生成对抗网络模型,旨在提升图像的高分辨率生成能力。此模型通过集成注意力机制、增加生成器与判别器的网络深度,以及采用新的损失函数等策略。

4.1.1 自注意力机制引入

DCGAN 的一大亮点是其在生成器和判别器中使用的卷积神经网络架构,这一架构虽然在提取图像特征方面效果显著,但因其依赖局部感受野,仅能分析相邻像素之间的联系,从而使得该模型在训练时偏重于局部特征的提取,而忽视了

全局特征。这导致在生成 PC 轨道梁面裂缝图像的应用中，DCGAN 倾向于产生与实际情况不符的裂缝图像。为解决这一问题，本文提出了一项改进措施，即引入自注意力机制到生成对抗网络中。这一机制使得模型在学习过程中能够更加综合地理解图像的全局特征，从而显著提升了图像生成的精确度与真实感。自注意力机制模块图如图 4-1 所示：

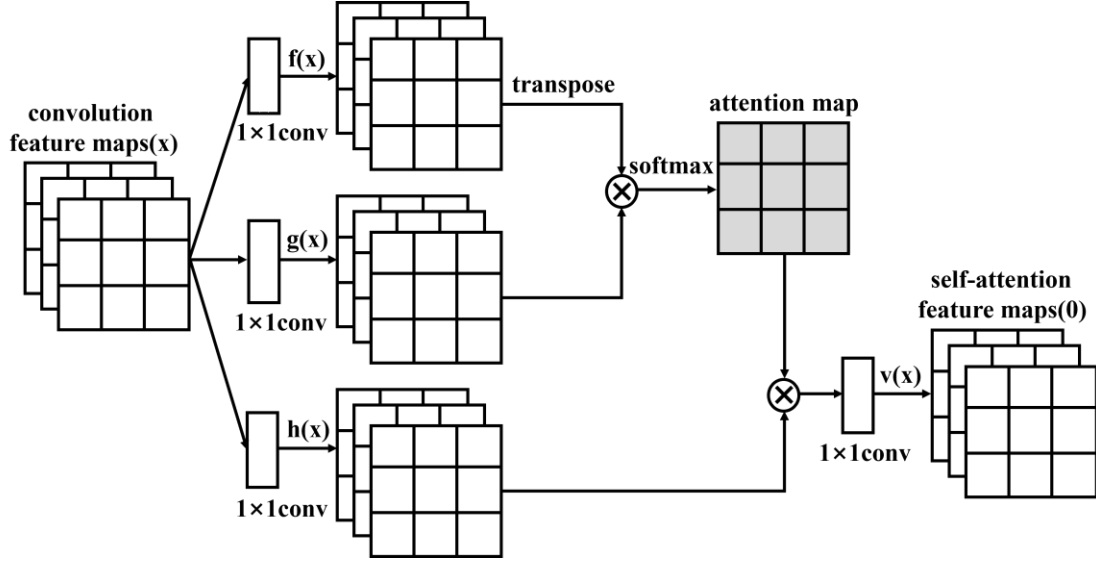


图 4-1 自注意机制模块图

自注意力模块的引入与卷积神经网络相结合，使得图像的局部和全局特征能够有效地被提取。以下是自注意力机制运算步骤的详细描述。

首先，对于卷积神经网络输出的特征图 x （尺寸为 $M \times M$ ），通过三个 1×1 卷积层将其映射到三种不同的特征表示 $f(x)$ 、 $g(x)$ 和 $h(x)$ ，这一步骤旨在传递特征图的初始信息；然后，将 $f(x)$ 执行转置操作后与 $g(x)$ 执行矩阵乘法，接着对乘积矩阵的每行实施 softmax 函数处理，以此产生注意力映射 $\delta_{b,a}$ 。该过程的计算细节见公式 4-1：

$$\delta_{b,a} = \frac{\exp(s_{ab})}{\sum_{a=1}^N \exp(s_{ab})} \quad (4-1)$$

其中 $s_{ab} = f(x_a)^T g(x_b)$ 。

其次，对第二步得到的注意力特征图 $\delta_{b,a}$ 进行后续处理。利用点乘的方式将其与 $h(x)$ 结合，接着应用 1×1 的卷积以形成最终的自注意力特征图 β_b 。此过程的详细数学描述参照公式 4-2：

$$\beta_b = v \left(\sum_{a=1}^N \delta_{b,a} h(x_a) \right) \quad (4-2)$$

其中 $h(x_a) = W_h x_a$, $v(x_a) = W_v x_a$ 。

在公式中, W_h 和 W_v 代表卷积层对各个对象的权重矩阵。

最后, 通过引入缩放因子 γ 并将之乘以自注意力特征图 β , 再将此乘积结果与原始的输入特征图相加, 便形成了最终的输出 y_a 。具体计算方法详见公式 4-3:

$$y_a = \gamma \beta_b + x_a \quad (4-3)$$

式中的缩放因子 γ 被设置为 0, 致使网络在初始训练阶段自然而然地集中处理细节特征。训练过程中随着 γ 参数的不断调整和优化, 神经网络逐步获得能力, 使其能够综合考虑图像中局部与全局的相互作用的信息, 从而逐渐捕获到图像全局层面的信息。

4.1.2 网络结构加深

为缓解目标检测中的数据集限制, 需确保通过 DCGAN 网络生成的轨道梁面裂缝图像在质量和尺寸上与原始图像一致。这要求对抗生成网络输出的图像具有 800×800 的分辨率, 并且必须与原图的分辨率相匹配, 避免因分辨率不一致而对后续目标检测模型训练产生负面影响。

当前的 DCGAN 网络模型, 因其层次较浅且深度有限, 主要适用于生成小尺寸图像, 如 MNIST 数据集和 Celeb-A Face 数据集等, 通常输出的图像仅为 28×28 或 64×64 像素, 这无法达到本研究所需要的高分辨率图像标准。因此, 本研究对原始的 DCGAN 模型结构进行了改良, 增添了三层卷积层以扩展其五层网络结构, 进而使得网络能够生成更高尺寸的裂缝图像, 以满足项目的具体需求。

扩展神经网络的层数可以提高其在抽取图像更深层次特征方面的能力, 进一步增加网络的非线性表现力, 以及在面对复杂问题时的适配性。然而仅仅通过增加层数并不无缺点, 随着网络深度的增长, 训练过程依赖于反向传播算法, 此时如果参数在反向传递中变得异常大或小, 就容易引发梯度消失或梯度爆炸问题。若在没有进行适当调节的情况下只是简单地扩展网络深度, 不仅无法实质性提高网络在特征表达上的能力, 反而可能会造成图像重建的质量下降, 以及网络性能的退化等不良后果。

本研究采用了残差模块的方法来克服增加网络层数所导致的性能不稳定的现象, 借助实施层间跳过连接来有效地解决了这一问题。残差网络通过在其基本结构中引入跳连, 使得不同层的参数能够直接传递, 从而构成了其核心机制。图

4-2 展示了残差单元的设计方案。

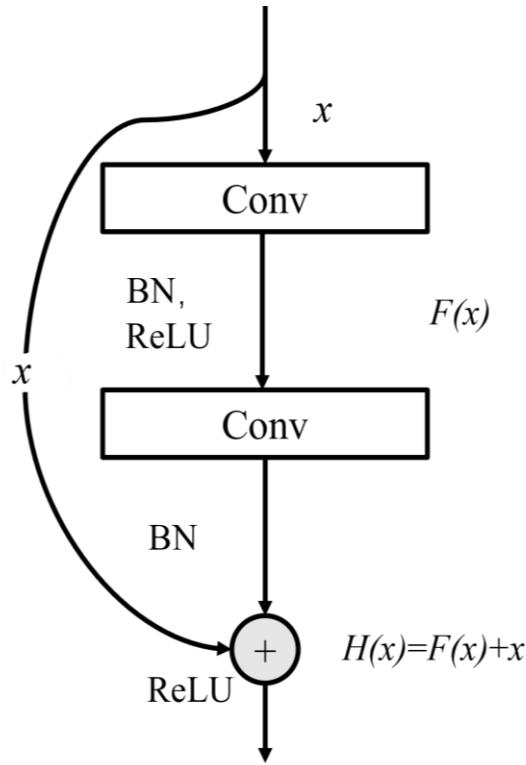
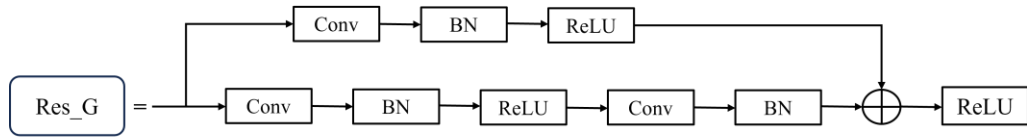


图 4-2 残差单元结构图

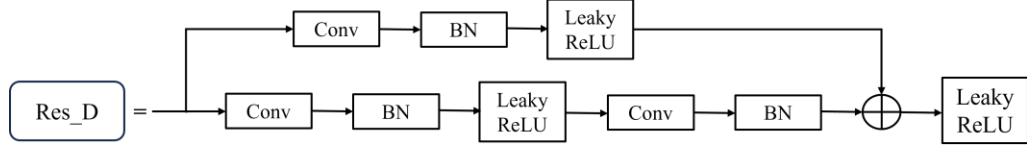
在此情形下， x 代表非线性输入，而 $F(x)$ 则表示通过卷积层处理的输出。 F 代表残差单元内的操作，包括卷积与激活等过程。左侧的曲线显示了跳过连接，实现了对输入 x 的恒等映射。最终的输出 $H(x)$ 是通过将非线性输入与其恒等映射进行线性结合得到的。整个残差架构可用公式 4-4 进行表述。

$$x_{l+1} = F(x_l, W_l) + x_l \quad (4-4)$$

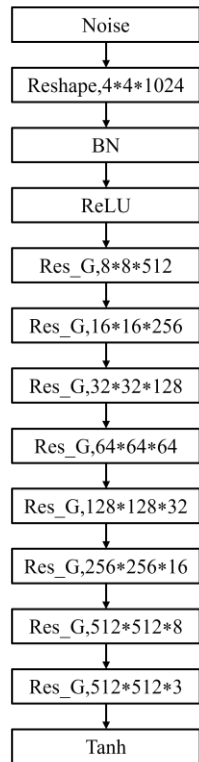
将深度 DCGAN 网络与残差单元相结合，应用残差单元来代替生成器与判别器中的步长卷积，有效地提升了网络在增加层数时的稳定性并防止了网络退化。该方法相较于传统的网络层叠策略，在调节网络架构及其层数上能够产生更优秀的图像合成效果。同时，集成残差网络不仅减轻了模型的参数负担，而且还把模型的整体性能提升了一个层次。图 4-3 描述了在增加网络深度并引入残差模块后，生成器和判别器的结构示意图。



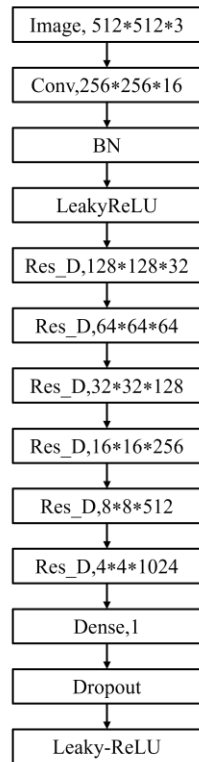
(a) 生成器残差块



(b) 判别器残差块



(c) 生成器网络结构图



(d) 判别器网络结构图

图 4-3 引入残差单元 DCGAN 结构图

4.1.3 损失函数改进

在原始的 DCGAN 网络架构中，生成器与判别器形成了一对对抗性系统，这种对抗性质可能会在训练过程中导致模型间的平衡难以维持。具体来说，如果判别器过于优秀，可能会引发生成器遇到梯度消失的问题；反之，若生成器性能较强，则判别器的损失函数可能会快速上升。为了在训练过程中保持两模型之间的有效平衡，本研究引入了 Wasserstein 距离，以取代基于 JS 散度的传统 DCGAN 网络损失函数计算方式，从而有效地生成了高质量的轨道梁面裂缝图像。

在 DCGAN 原始模型的训练流程中，随机噪声 z 被输入到生成器 G 中，用

于产生样本 $G(z)$ 。目标是使这些由 G 生成的样本能够欺骗判别器 D ，令 D 对这些样本的真实性评估概率趋向于 1，即 $D(G(z))$ 尽可能接近 1，显示出生成的样本足以让判别器将其错判为真实数据集中的样本。因此 $1-D(G(z))$ 趋近于 0，表明了将生成器 G 的模型优化问题简化为最小化问题。反观 D 网络，在接收到真实数据 x 时，其目标是让判别概率 $D(x)$ 接近于 1。当判别器 D 评估由生成器 G 创造的数据 $G(z)$ 时，其目标是降低这些生成数据的判别概率 $D(G(z))$ ，使其趋近于 0，亦即让 $1-D(G(z))$ 尽可能地靠近 1，将此目标视为 D 网络优化问题的核心。因此，在生成对抗网络框架内， G 网络和 D 网络的训练目标本质上是在 $V(G, D)$ 上执行的最大化与最小化过程。这一过程可以通过目标函数公式 4-5 来描述。

$$\min_G \max_D V(G, D) = E_{x \sim P_{data}(x)} [\log D(x)] + E_{z \sim P_z(z)} [\log(1 - D(G(z)))] \quad (4-5)$$

式中， P_{data} 代表真实的数据分布，而 P_z 代表噪声的分布。

当 G 网络的参数保持不变时，可以通过公式 4-6 来描述最优判别器 D 的目标。

$$D_G^*(x) = \frac{P_{data}(x)}{P_{data}(x) + P_g(x)} \quad (4-6)$$

一旦判别器 D 经过充分的网络训练优化，其训练目标就转变为最小化生成数据与真实数据分布之间的 JS 散度。相应的，这一目标函数在公式 4-7 中被简化表示。

$$C(G) = -\log 4 + 2 \cdot JSD(P_{data} \| P_g) \quad (4-7)$$

JS 散度的目标函数在理论上预设了两种数据分布应有交集，实际情况中，数据分布 P_{data} 与 P_g 之间往往缺乏重合，或者其重合程度极为有限。这种分布特征的差异可能会导致生成器在训练过程中遇到梯度消失的问题，从而使得网络训练变得困难。针对这一问题，本文提出了一种解决方案，将原始 DCGAN 模型的损失函数替换为 Wasserstein 距离，可以有效提升网络训练的效果。

Wasserstein 距离也被称为 Earth-Mover(EM)距离，其主要计算公式如公式 4-8 所示：

$$W(P_{data}, P_g) = \inf_{\gamma \in \Pi(P_{data}, P_g)} \mathbb{E}_{(x, y) \sim \gamma} [\|x - y\|] \quad (4-8)$$

在此框架下， $\Pi(P_{data}, P_g)$ 的集合体现了真实与生成数据分布所有潜在的联合分布。对任意给定的联合分布 γ ，通过对 γ 采样得到的样本对 $(x, y) \sim \gamma$ 来计算出一

个真实样本 x 与生成样本 y 之间的距离。定义 Wasserstein 距离为在 γ 所指定的联合分布下，真实样本与生成样本间距离的期望值，并在所有潜在联合分布上寻求此期望值的最低界。与 JS 散度相比，Wasserstein 距离的优势在于，即使真实与生成数据分布没有任何交集，它也能有效衡量两分布之间的距离，展现它们的相似性或差异性。

4.2 DCGAN 轨道梁面裂缝图像生成实验

4.2.1 生成式对抗网络模型评价指标

在评估生成式对抗网络生成的图像质量时，常见的评价方法可分为两大类：主观质量评价与客观质量评价。主观质量评价依赖于观察者对真实图像与 GAN 生成图像之间真伪的评判，但由于个体对图像美学的认知存在显著差异，这种评价方式具有较高的主观性，无法提供一致性的质量评定，因而不适合作为一个普遍适用的标准。相较之下，客观质量评价采用量化指标进行评估，能够提供更加标准化和定量的结果。本研究主要围绕两种客观评价指标分别是 Inception Score (IS) 和 Frechet Inception Distance (FID) 来对 GAN 生成的图像质量进行评估。这两种方法提供了一种定量的手段来衡量生成图像的质量，旨在减少评估过程中的主观性，提高评价结果的可靠性和准确性。

(1) IS

IS 评分指标通过分析类别概率分布来衡量生成图像的品质，具体的数学形式在公式 4-9 与 4-10 中进行了阐述。

$$IS = \exp \left[\mathbb{E}_{x \sim G(z)} D_{KL} (p(y|x) \| p(y)) \right] \quad (4-9)$$

$$p(y) = \int p(y|x = G(z)) dz \quad (4-10)$$

其中， x 表示需要进行评估的图像，而 $p(y|x)$ 代表给定图像 x 后的评价分布。 $p(y)$ 则是对生成模型制作出的图像多样性的度量。本文通过 $\mathbb{E}_{x \sim G(z)}$ 来表达在生成模型分布下图像的期望值。另外， D_{KL} 用来衡量两个概率分布间的 KL 散度。KL 散度的增大导致 IS 增高，而当 IS 升高时，表明生成对抗网络产生的样本集在多样性和真实性方面表现得更加出色，从侧面反映着模型的性能提高。

(2) FID

FID 评价通过比较真实和生成图像特征的 Frechet 距离，该过程基于两组特

征信息在高斯分布上的计算得以完成。FID 的数学计算公式如 4-11 所示：

$$FID = \|\mu_{data} - \mu_g\|^2 + tr \left(\sum_{data} + \sum_g - 2 \left(\sum_{data} \sum_g \right)^{\frac{1}{2}} \right) \quad (4-11)$$

其中， μ_{data} 表示真实样本的高斯分布均值， μ_g 表示生成样本的高斯分布均值。FID 分数的降低表明生成对抗网络的生成器所产生图像与真实样本之间的相似度提高，进而指示了图像质量的提升。

4.2.2 实验环境与模型训练

本实验采用表 4-1 中所示的硬件与软件设置，使用 Python 作为编程语言，并通过开源 Pytorch 框架搭建了深度学习训练模型的网络。实验所需的具体环境配置详见下述表格。

表 4-1 生成式对抗网络实验环境配置

系统配置	配置参数
操作系统	Ubuntu 16.04 LTS
CPU	AMD Ryzen 7 5800H with Radeon Graphics @3.20 GHz
GPU	RTX 3060
内存	16GB
CUDA	11.3
Python	3.7
Pytorch	1.11.0
开发工具	Pycharm

对预处理后的裂缝图像执行了批量 Resize 操作，制作了 410 张分辨率达到 512×512 的图像，以此作为网络模型训练的输入。模型训练过程中使用的超参数配置详细信息，请参考表 4-2。

表 4-2 训练超参数设置

超参数	img_size	epoch	batch_size	Learning_rate	optimizer
数值	512	500	32	0.0001	Adam

4.2.3 实验结果与分析

图 4-4 呈现了原 DCGAN 和经过改进的 DCGAN 网络模型在 0、100、200、

300、400、500 个 epoch 时生成的 PC 轨道梁面裂缝图像的结果。

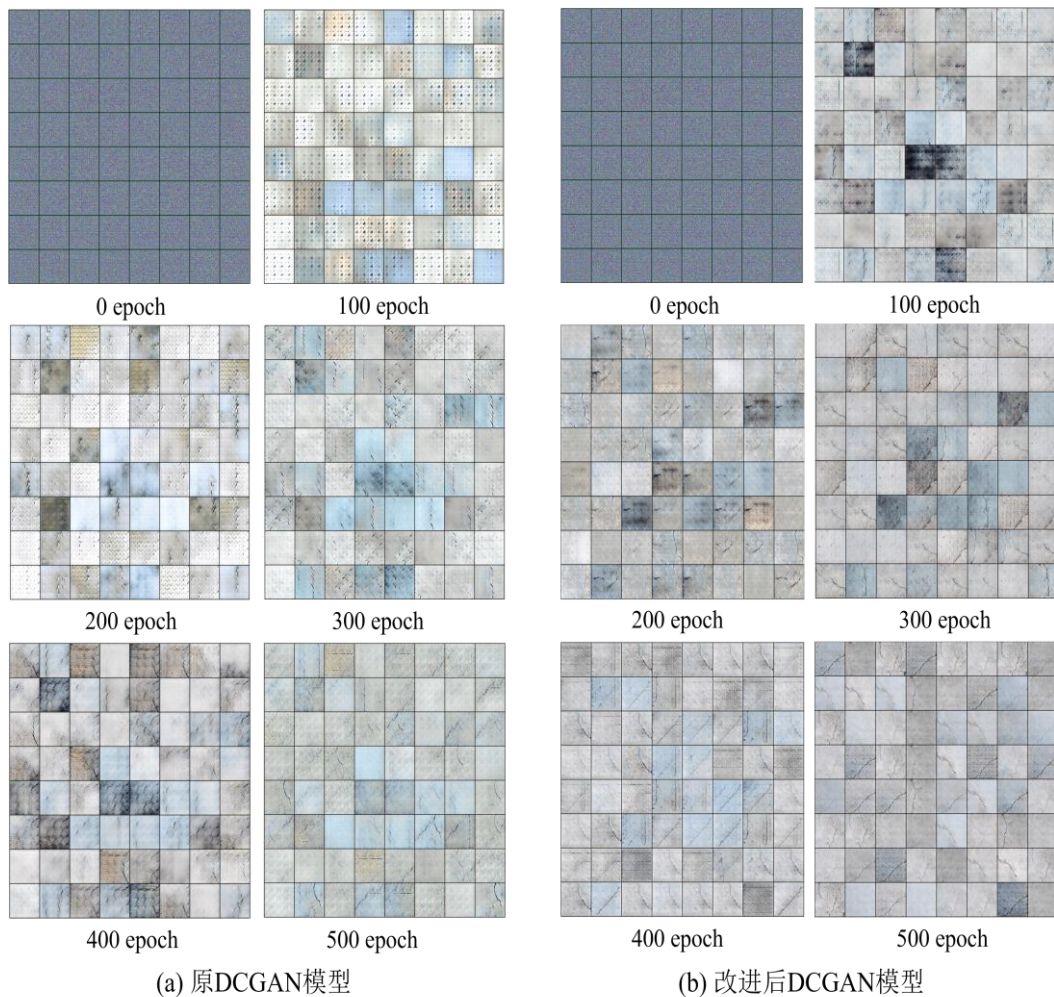


图 4-4 DCGAN 生成 PC 轨道梁面裂缝效果图

从图 4-4 的分析中可以观察到，在相同模型迭代次数下，改进的 DCGAN 模型比原 DCGAN 模型所生成的裂缝图像更为清晰逼真。改进的 DCGAN 模型最初使用随机噪声生成裂缝图像。100 次迭代训练之后，虽然模型可以产生裂缝图像，但图像呈现显著模糊和网状结构特征。经 200 次训练迭代，模型显著减少了网状结构的现象。至 300 次迭代，裂缝图像开始显示出较为明显的裂缝轮廓，但清晰度仍未达到理想水平。400 次迭代后，裂缝轮廓与实际裂缝更为接近，尽管背景仍有模糊。经 500 次迭代训练，生成的裂缝图像与真实图像的相似度显著提高，肉眼可清楚识别。这个过程体现了随着模型训练迭代次数的增加，模型在生成裂缝图像方面的性能逐步提升。

本文对所生成的轨道梁面裂缝图像进行了 IS 和 FID 指标的定量评价，相关数值见表 4-3 所示：

表 4-3 PC 轨道梁面裂缝图像 IS、FID 值对比表

模型	IS	FID
真实图像	12.26±0.16	10.32
DCGAN	6.94±0.11	66.57
改进 DCGAN	8.72±0.13	52.63

为了对裂缝生成效果进行直观评估，本研究制作了真实图像、原 DCGAN 模型和改进的 DCGAN 模型在 IS、FID 值上的比较图，具体见图 4-5。

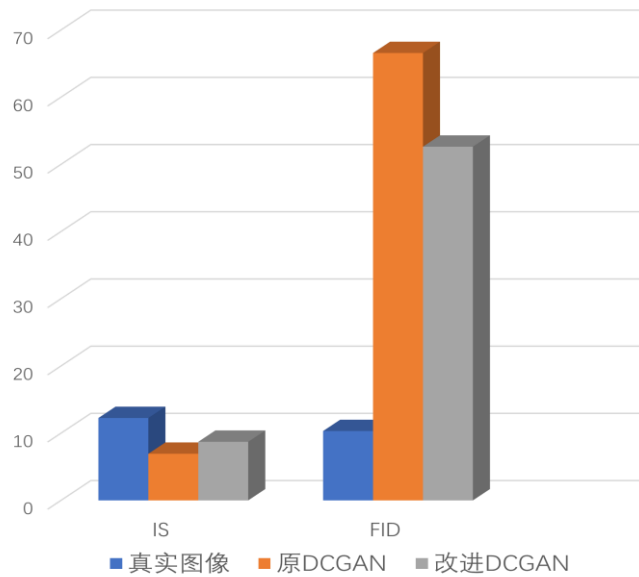


图 4-5 PC 轨道梁面裂缝 IS、FID 值对比图

尽管通过使用 DCGAN 网络来评估其产生的裂缝图像时，IS 和 FID 作为评价指标能够在一定程度上反映模型的性能，但是 IS 存在一定的局限性。IS 对于模型中权重的变化过于敏感，并且无法揭示模型是否发生了过拟合现象。此外，IS 不考虑真实样本与生成样本之间的详细差异。FID 评估了生成样本与真实样本在特征层面上的距离，比 IS 更能反映两者之间的相似度，但 FID 同样不能识别模型的过拟合问题。由于 FID 评估基于特征提取，它只能评价是否存在特定的特征而不能评估这些特征在空间中的相对位置。

4.3 裂缝检测的数据增强有效性分析

实验使用的数据集在模型的训练与测试阶段主要包括三个类别。具体而言，这三个类别包括：一是包含 410 张原始轨道梁面裂缝图像的数据集记作 A，二是经过传统方法数据增强处理扩增至 4100 张的数据集记作 B，三是通过生成式对

抗网络生成的 4000 张图像数据集记作 C。以上三类数据集的合集记作 D。详细信息见下表 4-4:

表 4-4 目标检测样本集分类

测试目标对象	数据集种类	数量
A	初始数据集	410
B	初始数据集+经传统数据增强处理后的数据集	4100
C	初始数据集+经生成式对抗网络生成的数据集	4100
D	初始数据集+经传统数据增强处理后的数据集+生成式对抗网络生成的数据集	8200

将三种数据集分别组合形成四个不同的测试实例,为确保每个样本中的图像与相应标注文件一一对应,所有通过生成式对抗网络生成且原始分辨率为 512×512 的图像,均被统一调整至 640×640 分辨率,简化了训练与测试的流程。此外,按照 8:1:1 的比例划分数据,形成训练集、测试集及验证集,以便于各阶段的数据处理和模型评估。

为了验证改进 DCGAN 网络以及一系列数据增强技术对算法目标检测精度的正面效应,本研究采用了统一的基础算法框架 YOLOX-s 对原始数据集以及多种数据增强策略处理后的数据集进行了对比实验。网络训练参数设置为:选用 Adam 优化器;训练批量大小设定为 8;初始学习率调整为 $1e-3$;优化器内部 momentum 参数设定为 0.937。模型总共训练迭代 300 轮,其中前 50 轮采用冻结训练。实验结果如下表 4-5 所示:

表 4-5 数据增强方法效果对比

算法	mAP@0.5 (数据集 A)	mAP@0.5 (数据集 B)	mAP@0.5 (数据集 C)	mAP@0.5 (数据集 D)
YOLOX-s	0.821	0.883	0.886	0.898

根据表 4-5 的分析结果显示,利用数据集 B、数据集 C 和数据集 D,这些经过数据增强处理的数据集作为目标检测网络的输入,能够有效地提高目标检测的整体性能。输入这些采用传统增强方法处理后的数据集,相较于未经增强的原始数据集,模型的平均精度 mAP 得到了 6.2%的显著提升。同样地,采用改进版 DCGAN 算法生成的裂缝数据集作输入时,相较于原始数据集,mAP 得到了 7.5%的提升。进一步,将传统的数据增强技术与改进 DCGAN 技术相结合生成的裂缝

数据集用作输入，mAP 相比于原始数据集增加了 7.7%。该实验全面验证数据增强方法在提高目标检测精度方面的显著效果。

经过这一系列的对比实验不仅深化了对于数据增强技术在图像处理领域应用的理解，同时也展示了数据预处理步骤对于提升深度学习模型性能的重要性。此外，这些实验结果也为后续研究提供了宝贵的数据支持和方法论参考，指明了通过技术创新和算法优化相结合的方式，进一步提升目标检测系统的精度和效率的可行路径。

4.4 本章小结

本部分进一步探讨了深度卷积生成对抗网络的细节，通过集成自注意力机制、调整网络结构深度以及优化损失函数，本研究对 DCGAN 模型进行了有效改良。采用这一改进模型，对预处理的 PC 轨道梁面裂缝图像数据集进行了训练与优化，成功生成了大量高质量的轨道梁面裂缝图像，有效提高了图像生成的质量。实验结果的全面分析也在本章进行了详细阐述。最后采用了统一的基础算法框架 YOLOX-s 对原始数据集以及多种数据增强策略处理后的数据集进行了对比实验验证改进 DCGAN 网络以及一系列数据增强技术对算法目标检测精度的正面效应。

第5章 改进 YOLOX-S 的 PC 轨道梁面裂缝检测方法研究

5.1 YOLOX 网络运行原理解析

为了解决 PC 轨道梁面裂缝检测中图像特征背景复杂的问题,本研究拟采用 YOLOX 作为基础模型进行算法改进。YOLOX 网络的核心优势在于其采用的 Anchor Free 机制,它摒弃了传统的依赖 K-means 算法生成先验框的做法,极大地增强了算法在处理多变背景下轨道梁面裂缝检测的适应性,使其更适配于背景繁复的轨道梁面裂缝检测场景。当前,官方 YOLOX 目标检测框架提供四种不同规模的模型版本,它们按照网络深度和宽度的不同排列,依次分为 YOLOX-s、YOLOX-m、YOLOX-l、YOLOX-x。YOLOX-s 模型以其较小的体积(几十 MB)在保持较高检测速度和准确度方面表现出色。结合 YOLOX-s 模型的特性及图像的特定需求,本文选用参数量相对较少的 YOLOX-s 作为此次 PC 轨道梁面裂缝检测任务的基础模型,模型整体结构如图 5-1 所示。

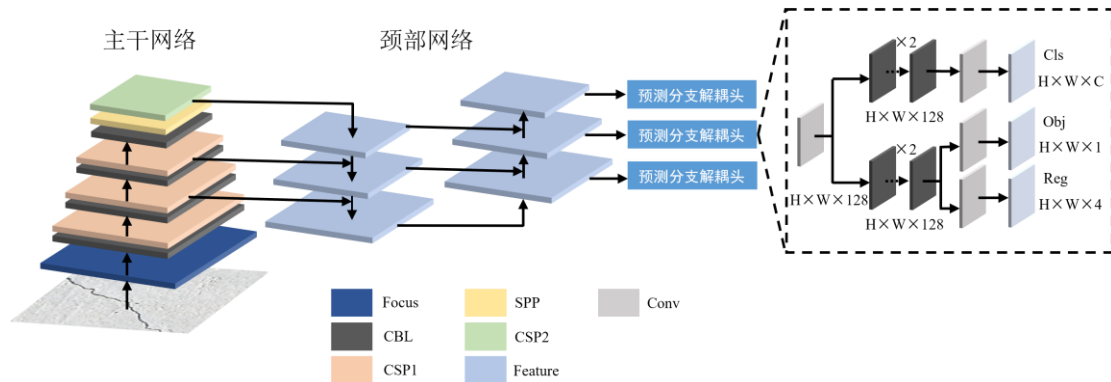


图 5-1 YOLOX-s 模型网络结构图

该模型的构架涵盖了四个主要部分:首先是输入层(Input),其次是负责特征提取的网络(Backbone),第三是颈部网络(Neck),最后是执行预测的输出层(Prediction)。YOLOX-s 主干部分中的 CSPDarknet 会对传输进来的原始图像进行特征提取,进而获得参与组建下一级网络的三个有效特征层。特征金字塔网络(FPN)与路径聚合网络(PAN)相组合共同完成特征融合的方式被应用在颈部层级上,两种特征提取器相辅相成弥补了层层卷积过程中遗漏的部分位置信息。预测部分的检测头采取了能将目标位置和类别信息导入不同网络分支学习的解耦合头(Decoupled head)结构,进一步提升了检测效果。

5.2 基于 YOLOX-CRACK 的 PC 轨道梁面裂缝检测算法设计

针对目前 PC 轨道梁面裂缝缺陷检测中存在精度低及复杂背景干扰等问题，提出一种高精度实时的跨座式单轨 PC 轨道梁面裂缝图像检测模型 YOLOX-CRACK。YOLOX-CRACK 算法主要是通过以下几种思路来改进的：全程上采样使用双三次插值法；将常规卷积换成结构更为轻量化的 Ghost 卷积；ASSF 模块嵌入特征融合层；引入 GIOU 损失函数。改进得到的 YOLOX-CRACK 轨道梁面裂缝检测算法原理如图 5-2 所示。所有输入的检测图片都被调整为 640×640 的尺寸大小，经过主干特征提取网络处理后，将生成三个有效特征层，这些特征层将用于构建加强特征提取网络。最终通过特征融合后，网络输出获得的预测结果。

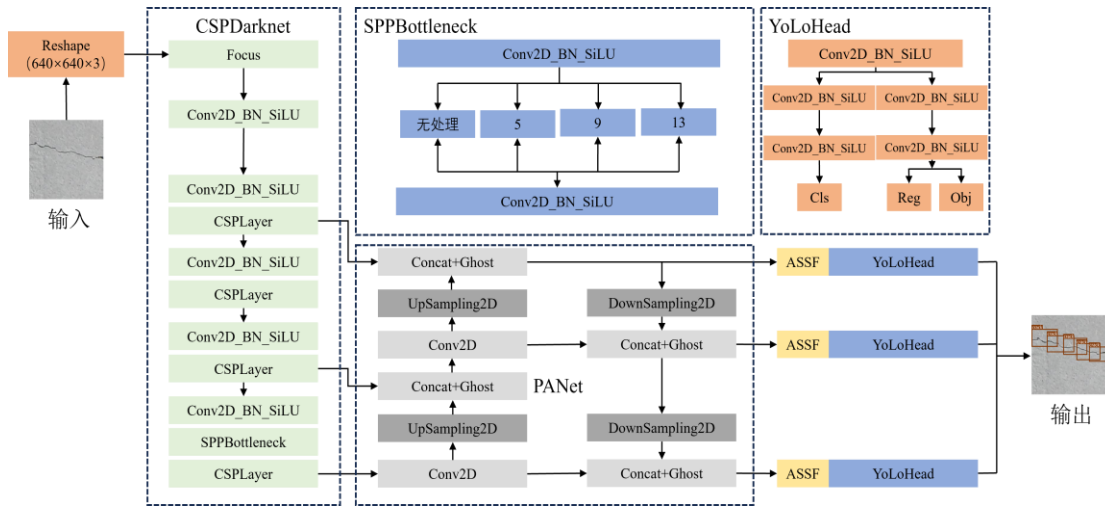


图 5-2 YOLOX-CRACK 检测算法原理

5.2.1 自适应特征融合机制

YOLOX-S 颈部层级中的 FPN 后端添加了自下而上方向的 PAN 金字塔结构，如图 5-3 所示。FPN+PAN 的结构认识到不同层级特征图之间的融合对于裂纹目标检测任务具有积极作用，所以将原有的单向多尺度特征融合转变成更为全面的双向融合。但该结构在特征融合上采取的仅是维度发生简易变化的张量拼接，而忽略了尺度不一的待检裂缝目标对于不同层级特征图所侧重的关注度也不一样。

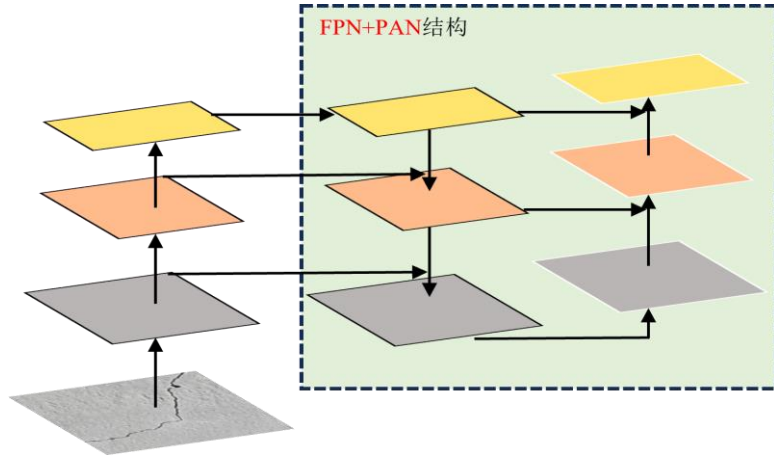


图 5-3 FPN+PAN 结构示意图

针对此问题，引入一种自适应空间特征融合的策略，实现每个空间位置上不同级别的特征自适应融合。ASSF 模块能通过自身的机制在训练规律中找到最佳的特征融合方式。若特定位置的特征包含相互冲突的信息，这些特征便会被排除。相反，如果某处的特征提供了丰富的区别性线索，这些特征则会被保留并加以强化。ASSF 模块具体实现原理如图 5-4 所示。

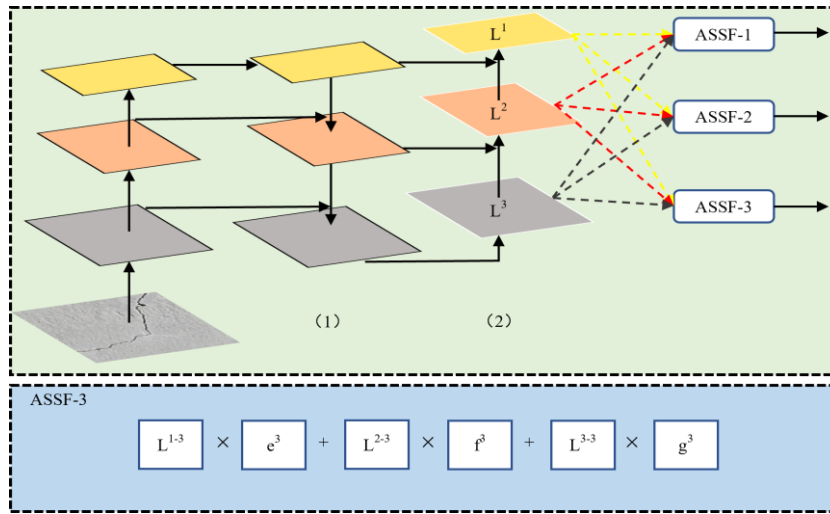


图 5-4 ASSF 模块结构

每个 ASSF 层会处理接收来自图 5-4 中 FPN+PAN 双向融合后输出的三个不同尺度的特征图，分别为 20×20 的 L^1 、 40×40 的 L^2 和 80×80 的 L^3 。例如在 ASSF-3 层级中，特征图 L^1 和 L^2 经过 1×1 的卷积以及不同倍率的上采样后，相继达到了与 L^3 相同的通道数与维度。将 L^1 和 L^2 转变后的特征图记作 L^{1-3} 和 L^{2-3} ，定义 L_{ij}^{n-3} 为特征图 L^{n-3} 在坐标位置 (i, j) 处所拥有的特征向量。即 ASSF-3 上新融合所输出的特征图 Y^3 在坐标位置 (i, j) 处的计算公式如公式 5-1 所示：

$$Y_{i,j}^3 = f_{i,j}^3 \times L_{i,j}^{1-3} + g_{i,j}^3 \times L_{i,j}^{2-3} + h_{i,j}^3 \times L_{i,j}^{3-3} \quad (5-1)$$

式中 $f_{i,j}^3$ 、 $g_{i,j}^3$ 、 $h_{i,j}^3$ 表示由网络自适应学习到的特征图的融合空间权重。定义 $f_{i,j}^3$ 的计算如公式 5-2 所示：

$$f_{i,j}^3 = \frac{e^{\varepsilon^3 f_{i,j}}}{e^{\varepsilon^3 f_{i,j}} + e^{\varepsilon^3 g_{i,j}} + e^{\varepsilon^3 h_{i,j}}} \quad (5-2)$$

其中 $\varepsilon^3 f$ 、 $\varepsilon^3 g$ 、 $\varepsilon^3 h$ 分别由特征图 L^{1-3} 、 L^{2-3} 和 L^{3-3} 通过 1×1 的卷积计算得出。

5.2.2 Ghost 卷积模块

Ghost 模块的核心设计理念并不繁琐,其初始步骤是使用一个尺寸为 1×1 的卷积来减少输入特征层的通道数,从而实现符合需要的特征浓缩。接着通过逐层卷积的方式产生多个类似于特征浓缩的特征图,根据规则把这些特征层与特征图依次累加在一起,从而构建出所需的输出特征层。Ghost 模块可以显著减少模型的计算资源,同时也使得网络训练变得更为简单。普通卷积和 Ghost 轻量化卷积的工作原理对比参见图 5-5。

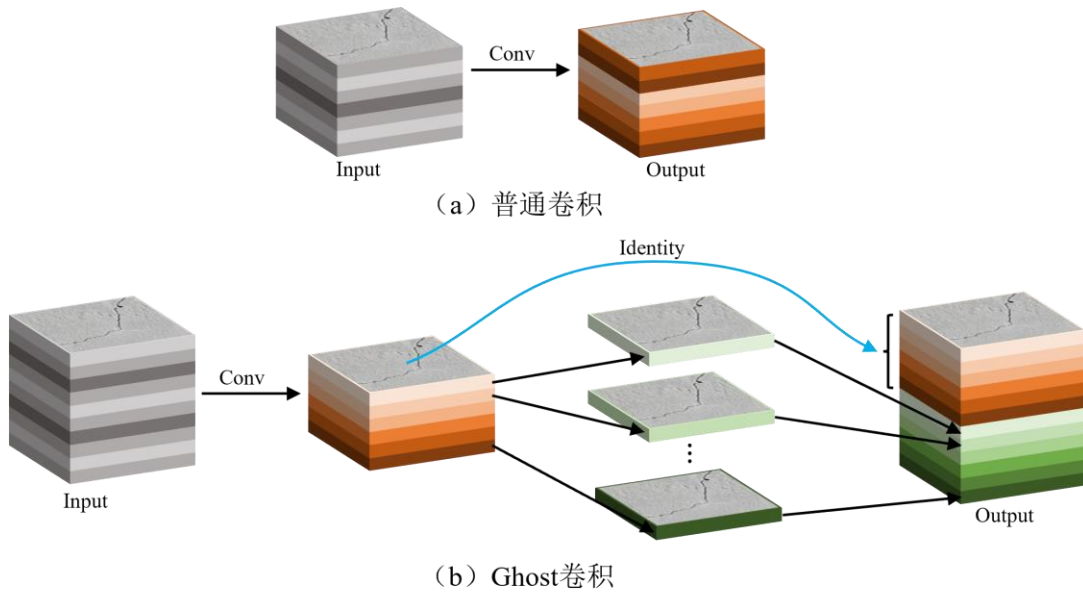


图 5-5 普通卷积和 Ghost 卷积

5.2.3 双三次插值上采样

裂缝图片在输入到网络模型前会进行相应的几何变换,以上一系列的处理依靠着插值算法来实现。由于图像中部分像素点在几何变换中并没有获得对应的赋值,插值算法利用其特定的原理计算出像素点上缺失的赋值并进行填充。在原有的 YOLOX-s 算法中采用了最邻近插值的方法,这种插值方法相对简单,只关注

与待测采样区域最接近的像素点的灰度值，而没有考虑到其他相邻区域的像素点。最邻近插值算法会导致图像中像素点灰度值之间产生一定的间断性，图像的质损程度较大的同时伴有锯齿现象。为保证上采样操作后特征图的综合图像质量，最终选择质损程度较小的双三次插值法来对图像进行上采样，图 5-6 为双三次插值法对裂纹图像进行上采样的流程图。

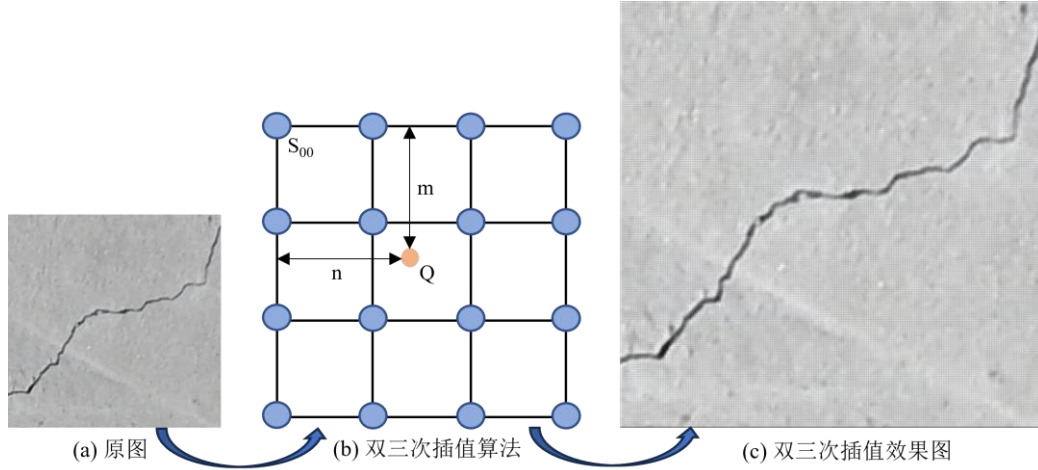


图 5-6 双三次插值上采样流程图

S_{00} 点像素的权重可以表示为 $W(m) \cdot W(n)$ ，其中 m 和 n 分别是 Q 到 S_{00} 点的纵向距离和横向距离。采用类似的方法可以计算其他 16 个整数像素位置的权重，并用这些权重对 Q 的像素值进行预估，完成对裂纹图像的上采样处理。此方法利用了距离插值点 (x, y) 最近的 16 个采样点的信息，通过加权平均的方式计算出像素灰度值 $f(x, y)$ ，从而实现对插值点的估计。每个采样点所赋有的权重并不相同，而是由该点到待求插值点的水平距离和竖直距离共同决定所得，有关权重的估算公式 5-3 为：

$$z(Q) = \sum_{i=0}^3 \sum_{j=0}^3 z(S_{ij}) \cdot W(m_{ij}) W(n_{ij}) \quad (5-3)$$

其中： S_{ij} 为原图像所提取的整数位置像素点， m_{ij} 为 Q 与 S_{ij} 之间的纵向距离， n_{ij} 为 Q 与 S_{ij} 之间的横向距离， z 表示该采样点的像素值。

双三次插值对于像素坐标点权重的计算需通过调用 Bicubic 函数来完成，相关函数表达式如公式 5-4 所示：

$$W(x) = \begin{cases} (a+2)|x|^3 - (a+3)|x|^2 + 1, & |x| \leq 1 \\ a|x|^3 - 5a|x|^2 + 8a|x| - 4a, & 1 < |x| < 2 \\ 0, & \text{其他} \end{cases} \quad (5-4)$$

其中： α 为图像边缘的锐化程度， $\alpha=-0.5$ 。

5.2.4 损失函数优化

该改进算法对预测回归边界框损失函数进行了优化，传统的 IOU 损失函数被相对更完善的 GIOU 损失函数所替代。为了更准确地评估目标之间的重叠度，GIOU 损失函数综合考虑了重叠区域和非重叠区域。此外，GIOU 损失函数还能够处理预测框和真实框没有重叠的情况，并且在网络训练时对模型的收敛起到积极的作用。相关计算公式为：

$$L_{reg} = 1 - GIOU \quad (5-5)$$

$$GIOU = IOU - \frac{|D - (P \cup R)|}{|C|} \quad (5-6)$$

$$IOU = \frac{P \cap R}{P \cup R} \quad (5-7)$$

式中： D 是指包括预测框和实际框的最小边界框， P 是指预测框， R 是指实际框， IOU 是指预测框与实际框的交叉面积与这两个框的并集面积的比值。

5.3 YOLOX-CRACK 的效果验证与分析

5.3.1 目标检测评价指标

目标检测模型的发展伴随了多样化的科学评估标准的制定，而这些评估标准的确立对于量化深度学习网络的训练成果具有关键意义，这可以有效地反映模型在训练过程和实际应用中的性能表现。

(1) 交并比

交并比，即 IOU (Intersection-over-Union)，是衡量目标检测中预测框与真实标注框交叠程度的关键指标，它体现了物体检测的准确性。在本研究的跨座式单轨 PC 轨道梁面裂缝检测任务中，将裂缝的真实边界框记为 b_{gt} ，而网络模型预测出的裂缝边界框记为 b_p ，如图 5-7 展示。两者的交并比是它们交集与并集的面积比率。理想情况下，预测框与真实框的重叠达到最大即交并比为 1，这表明预测精度非常高。交并比的增加表明预测框与真实框的吻合程度提高，是衡量检测准确度的重要指标。公式 5-8 为其具体计算公式：

$$IoU = \frac{b_{gt} \cap b_p}{b_{gt} \cup b_p} \quad (5-8)$$

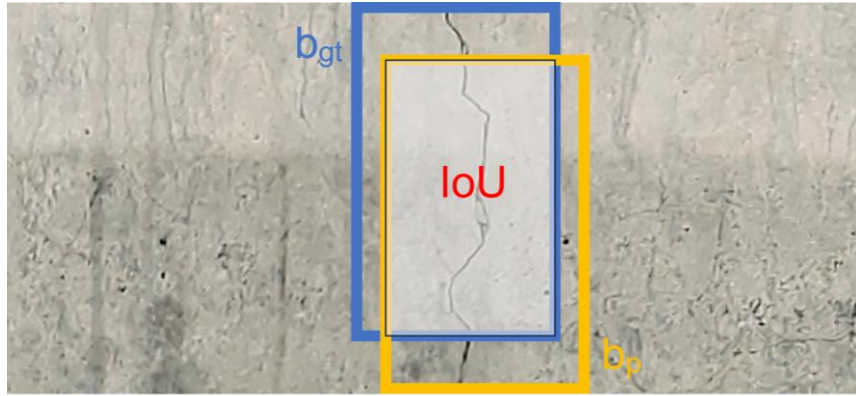


图 5-7 IOU 示意图

(2) 精确率与召回率

为了准确评估一个检测框是否成功地标记了目标物体，引入了交并比和置信度阈值的概念。只有当检测框的置信度超过特定阈值，且其与真实标注框的交并比也达到了预定标准时，该检测框才被认定为准确预测。在执行目标检测任务时，算法会分析输入图像并生成预测框，这些预测框依据它们与真实情况的符合程度，可以被分类为真正例（TP），假正例（FP），假负例（FN），以及真负例（TN）。这一分类机制为目标检测的性能评估提供了一套明确的标准，具体释义如表 5-1 所示：

表 5-1 目标检测评价参数

符号	含义
TP	准确分类为正例的数量
FP	错误分类为正例的数量
FN	错误分类为负例的数量
TN	准确分类为负数例的数量

精确率（Precision）指的是在所有被预测为裂缝的样本中，正确预测为裂缝的比例，计算方式如公式 5-9 所示：

$$Precision = \frac{TP}{(FP + TP)} \quad (5-9)$$

召回率（Recall）衡量的是在测试集中实际为裂缝的样本里，被正确判定为裂缝的样本所占的比例，其计算方法按照公式 5-10 所示：

$$Recall = \frac{TP}{(FN + TP)} \quad (5-10)$$

准确率（Accuracy），其表示的是正确预测的裂缝个数与测试集中所有待测样本个数之比，具体计算如公式 5-11 所示：

$$Accuracy = \frac{(TP + TN)}{(FP + FN + TN + TP)} \quad (5-11)$$

（3）平均精度与平均精度均值

平均精度（Average Precision, AP）反映了在所有待测样本中，正确识别出裂缝的占比，计算按照公式 5-12 所示：

$$AP = \int_0^1 Precision(t) dt \quad (5-12)$$

其中 $Precision(t)$ 表示交并比阈值取 t 时的精确率。

平均精度均值（Mean Average Precision, mAP）在多类别目标检测任务中，重要性体现于衡量模型对所有类目标的检测精准度，其具体的计算方式参照公式 5-13 所示：

$$mAP = \frac{\sum_{n=1}^N AP_n}{N} \quad (5-13)$$

由于检测类别仅涉及裂缝与非裂缝，因此本实验的 mAP 与单一类别的 AP ，即裂缝的检测精度是相等的。此处， N 代表物体类别的总数， n 代表具体某一类别， AP_n 则是该类别的检测精度。

5.3.2 对比试验

该对比实验的网络训练参数设置均与上文验证数据增强技术有效性的实验保持一致，确保了本章实验结果的真实性和重复性。为了最大限度地提高检测模型的精确度和实验总体的稳定性，本实验选取了样本量最大的数据集 D 进行实验，该数据集包含高达 8200 张图像。选择大规模数据集的目的在于提供更为丰富和多样的训练样本，从而使得模型能够学习到更加全面和细致的特征，这对于提升模型的泛化能力和识别准确度至关重要。利用大量数据进行训练还有助于减少过拟合的风险，确保模型具有良好的实际应用性能。

为展现改进后网络模型的性能，本研究在同一数据集上对 YOLOX-CRACK 算法与其他算法进行了比较实验。具体来说，比较了 YOLOX-CRACK 与原版 YOLOX-s、Faster R-CNN、SSD、YOLOv4 以及 YOLOv5-s 算法的性能。实验结果汇总于表 5-2。

表 5-2 不同算法对比实验结果

算法	mAP (%)	参数量 (MB)	FPS (frame/s)
Faster R-CNN	82.2	110.82	18
SSD	77.5	27.18	23
YOLOv4	82.6	65.56	46
YOLOv5-S	86.1	48.73	123
YOLOX-S	87.4	12.84	134
Ours	91.8	9.62	119

实验结果表明, YOLOX-CRACK 网络模型相比于 Faster R-CNN、SSD、YOLOv4、YOLOv5-s 算法和 YOLOX-s 算法的 mAP 分别提升了 9.6%、14.3%、9.2%、5.7%、4.4%, 模型参数量相比于其他五个算法分别减少 91.32%、64.61%、85.33%、80.26%、25.08%。改进后的算法在有效提升检测精度的同时, 并未显著增加时间复杂度, 证明其具备更优越的性能。

5.3.3 消融试验

为了验证所提出改进模块对裂纹缺陷检测的有效性, 该节同样在样本量最大的数据集 D 数据集上进行消融实验。在原始基线模型的基础上, 相继添加了 ASSF 模块、Ghost 模块、双三次插值法以及 GIOU 损失函数, 具体实验结果见表 5-3:

表 5-3 消融实验结果

ASSF	Ghost	Bicubic	GIOU	mAP50:95 (%)	mAP50 (%)	参数量 (MB)	FPS (frame/s)
×	×	×	×	78.4	88.6	12.84	134
√	×	×	×	79.5	90.4	13.09	132
×	√	×	×	78.7	89.5	7.57	140
×	×	√	×	79.1	90.2	13.68	125
×	×	×	√	78.4	89.8	14.84	134
√	√	√	√	80.2	91.8	9.62	119

从表 5-3 的数据来看, ASSF 模块加入后, mAP50:95 和 mAP50 分别获得了 1.1%和 1.8%的提升; 当单独添加 Ghost 模块时, mAP50:95 和 mAP50 的提升分别为 0.3%和 0.9%; 双三次插值法加入时, mAP50:95 和 mAP50 分别提高 0.7%

和 1.6%；当仅使用 GIoU 损失替代原有的损失函数时，mAP50 提升了 1.2%。同时应用 ASSF 模块、Ghost 模块、双三次插值法和 GIoU 损失函数时，可以显著提升检测效果。4 种优化方案相互协作，共同对检测结果产生积极影响。研究提出的算法在裂纹缺陷检测方面的 mAP50:95 和 mAP50 分别达到 80.2%和 91.8%，与基准模型对照，提高了 1.8%和 3.2%。YOLOX-CRACK 算法相对于基准模型，在检测速度方面下降了 15frame/s。这主要是因为引入了双三次插值法，增加了一定的计算量和网络参数量。然而，该算法的每秒图像检测数量仍能达到 119frame，为实际应用场景提供支持，并满足对实时裂纹检测的需求。

5.3.4 实验结果与分析

图 5-8 展现了原 YOLOX-s 模型的 Loss 曲线，在第 210 个 Epoch 呈现轻微的波动，而在第 250 个 Epoch 之后趋于平稳。综合对比前者，模型在第 220 个 Epoch 之后 Loss 曲线提前趋于稳定，且模型训练过程更加高效平稳。

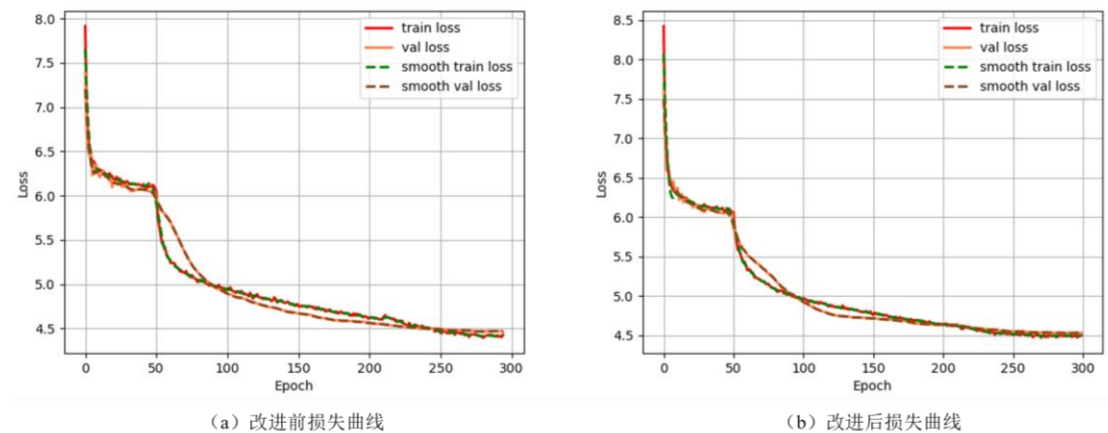


图 5-8 损失函数曲线对比

为验证跨座式单轨 PC 轨道梁面裂缝检测模型的性能，选取验证集中 820 张图像进一步分析该模型的检测结果。检测结果显示，该模型对于轨道梁面裂缝图像能够较好地识别检测，选取以下三种较典型的轨道梁面裂缝进行分析，如图 5-9 所示。

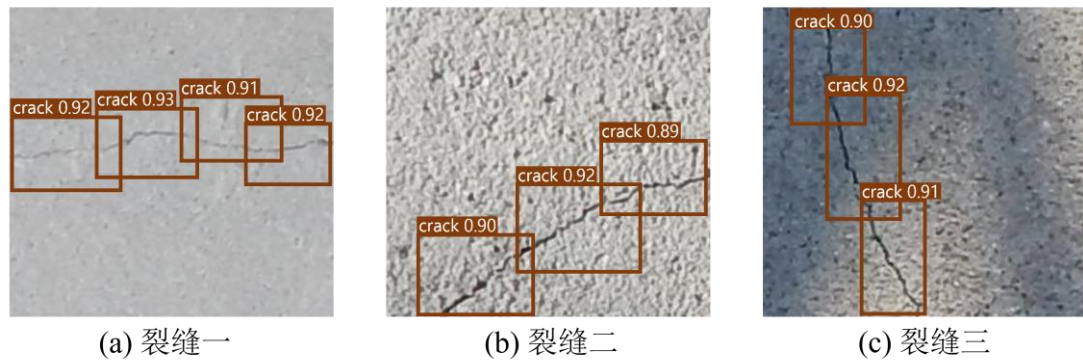


图 5-9 损失函数曲线对比

从图 5-9 中可以看出,对于背景简单和背景复杂、干扰较多的轨道梁面裂缝,该模型都能很好地识别并标记出来,且对非裂缝没有误判。在背景简单的轨道梁面图像中,可以准确地识别出细小的裂缝。在背景复杂的轨道梁面中,该模型能够很好地区分出真实裂缝与伪裂缝。在检测结果可视化对比图 5-10 中,每个检测文本框的左侧与右侧分别呈现的是预测目标的种类和置信度。为了便于识别检测结果,当预测框的置信度大于等于 0.8 时,将其标记为棕色;当置信度不足 0.8 时,则将其标记为白色。

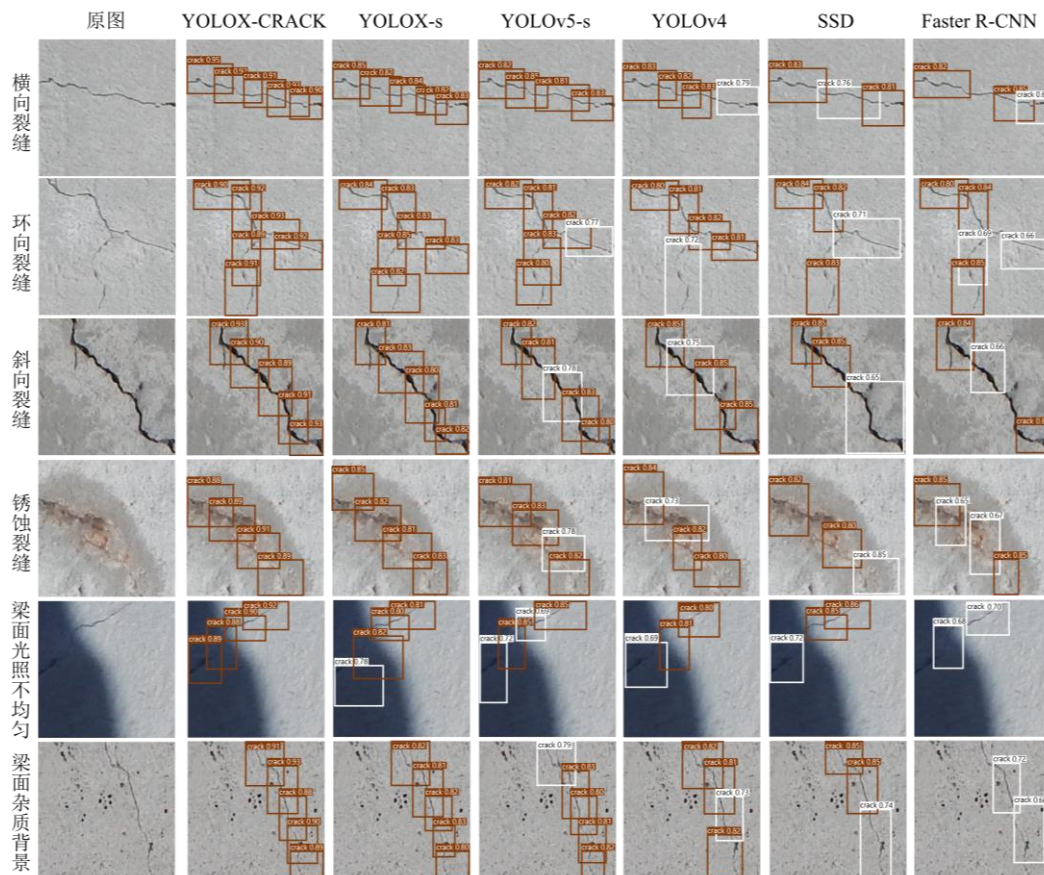


图 5-10 多种算法裂纹检测结果可视化对比

从图 5-10 中可以观察到，YOLOX-CRACK 模型在检测 PC 轨道梁面裂纹方面展现出更为精确的位置检测和更高的置信度。YOLOX-CRACK 模型能够在低照度和背景杂质的干扰下，依然可以准确地检测出裂纹的位置，并且没有出现漏检的情况。由此可见，YOLOX-CRACK 算法对 PC 轨道梁面裂纹具有更高的综合检测效率，更适合应用于工程当中。

5.4 本章小结

为实现轨道梁面裂缝的高精度检测，本章节提出了基于 YOLOX-CRACK 的裂缝检测算法。该算法在加强特征提取网络中使用 Ghost 轻量化卷积提取目标特征，以降低模型计算复杂度。在特征融合层中嵌入 ASSF 模块使模型自适应学习不同特征层间的联系，加强网络对裂缝目标区域的关注度。模型引入 GIOU 损失函数来加快模型收敛速度，并采用双三次插值法进行上采样操作，从而提升网络检测性能。通过构建网络模型并利用第四章获取的 PC 轨道梁面图像及其数据增强后的数据集进行训练，本研究进行了对比实验以探究目标检测算法在裂缝检测领域的适用性与效率。此外，通过消融实验验证了改进措施对提升 YOLOX-s 模型在检测单轨 PC 轨道梁面裂缝方面的综合性能的显著作用。在自制 PC 轨道梁面裂缝数据集上的实验结果表明，改进后模型相较于原基准模型平均检测精度提升 4.40%，在参数量上比原模型下降 25.08%，每秒检测裂缝图像数量为 119frame。特别的，该方法对于光照不均匀以及杂质背景等不利识别因素的裂缝图像识别也具有较好的效果。

第 6 章 结论与展望

6.1 研究总结

为了确保轨道交通系统能长期有效运营,轨道的维护和养护工作显得尤为关键。传统的人工检测方法已不再足以满足轨道交通发展的需求,使得发展自动化的轨道梁面裂缝检测系统显得极其重要。本文提出基于图像融合增强的改进 YOLOX 轨道梁裂缝检测方法,该方法使用深度学习目标检测算法对跨座式单轨 PC 轨道梁面裂缝进行自动化的识别与检测。采取数据增强策略来有效的拓展单轨 PC 轨道梁面裂缝的图像样本集,对生成式对抗网络进行改进,实现多样化且高质量的轨道梁面裂缝图像生成。最后,对 YOLOX 模型框架进行改进优化,并将其应用于跨座式单轨 PC 轨道梁面裂缝图像的实际场景中,实现了对轨道梁面裂缝精准快速检测。本文的主要研究内容有:

(1) 针对单轨 PC 轨道梁面图像进行了必要的图像预处理和数据增强工作,建立了较为全面的 PC 轨道梁面裂缝图像数据集。针对初始图像,实施了一系列预处理步骤:利用滑动窗口技术截取尺寸为 800×800 像素的裂缝图片;应用中值滤波与均值滤波技术以消除图像噪声;将裂缝图像处理为二值化图,并通过识别连通区域来确定外接矩形的尺寸,此步骤也同时标注了裂缝数据。在数据增强过程中,融合了传统技术与基于生成对抗网络的无监督学习策略,进而得到了一个充分准备的单轨 PC 轨道梁面裂缝图像数据集。

(2) 提出了一种优化的深卷积生成对抗网络架构,以克服传统生成式对抗网络中的模式坍塌、图像生成质量不佳以及多样性匮乏等问题。该框架结合自注意力机制,有效地捕捉了图像的全局与局部特征,同时通过加深网络层次以生成更高分辨率的 PC 轨道梁面裂缝图像。同时还优化了损失函数,引入 Wasserstein 距离取代原有的 Jensen-Shannon 散度,用以增强模型的训练稳定性。对比实验结果表明,改进的 DCGAN 模型在生成 PC 轨道梁面裂缝图像的质量及多样性方面显著优于原版 DCGAN。

(3) 针对跨座式单轨 PC 轨道梁面图像中的裂缝检测任务选定了 Pytorch 框架下的 YOLOX 算法,并且对该算法进行了相应的改良。首先,在加强特征提取网络中使用 Ghost 轻量化卷积提取目标特征,以降低模型计算复杂度。其次,在

特征融合层中嵌入 ASSF 模块使模型自适应学习不同特征层间的联系,充分利用每个特征层的信息,加强网络对裂纹目标区域的关注度。同时引入 GIOU 损失函数,进一步加快模型收敛速度。最后,模型采用双三次插值法进行上采样操作,更好地保持原始特征的清晰度和细节信息,从而提升网络检测性能。实验结果表明,YOLOX-CRACK 算法在不同场景中对 PC 轨道梁面裂纹识别表现良好。相较于原基准模型平均检测精度提升 3.40%,在参数量上比原模型下降 25.08%,每秒检测裂纹图像数量为 119frame。在保证推理速度的同时,检测准确度也具有不小的提升,说明模型在准确性和实时性方面符合实际工程需求。

6.2 未来展望

在 PC 轨道梁面裂缝缺陷的目标检测和样本生成方面,本文取得了一些研究进展,但仍存在一些需要改进的方面:

(1) 在目标检测任务中,不同应用场景对数据需求各异,且可能因多种因素导致缺乏足够的训练样本。此时,使用 GAN 来合成图像以扩展数据集是一种行之有效的方法。即使是用于生成高质量图像的 GAN 也需要大量的训练样本。当训练样本稀少至数百个甚至更少时,GAN 易发生模式塌陷,导致训练难以稳定。在此背景下,探索在小规模数据集的环境下利用生成对抗网络进行数据增强,是未来工作中亟待解决的挑战。

(2) 探索实时 PC 轨道梁面裂缝检测的新方法,特别是通过优化 YOLOX 算法框架,以便其能更有效地部署于轨道梁检测车之上,从而实现图像的即时采集和处理,将为轨道梁面的检测与维护工作带来便利,是未来研究的一个值得考虑的方向。

(3) 本文着重探讨了跨座式单轨 PC 轨道梁面裂缝的检测问题,但未对检测所得裂缝特征进行深入研究。鉴于裂缝的宽度和长度等特征对其评估至关重要,未来的工作可以考虑融合像素级别的语义分割技术,不仅实现裂缝的精准检测,同时也能提取出这些关键的特征信息。

参考文献

- [1] 杜子学,王建. 跨座式单轨列车与地铁列车通过高架线路时辐射噪声特性对比试验 [J]. 城市轨道交通研究, 2024, 27 (02): 191-194+199.
- [2] Zhou J, Huang C, Zhang L, et al. Selection of design parameters of straddle monorail system considering vehicle-bridge interaction[C]//Structures. Elsevier, 2023, 50: 1794-1808.
- [3] 李天祥,张磊. 新一代单轴跨座式单轨动力转向架构架静强度和疲劳强度分析 [J]. 铁道机车车辆, 2023, 43 (06): 167-173.
- [4] 张凯,朱尔玉,刘旭锴,等. 跨座式单轨交通轨道不平顺对梁轨振动响应的影响 [J]. 铁道建筑, 2021, 61 (07): 140-144+154.
- [5] 仵叔强. 跨座式单轨交通折线型道岔平面线形设计与研究 [J]. 都市快轨交通, 2022, 35 (02): 136-142.
- [6] Pare S, Kumar A, Singh G K, et al. Image segmentation using multilevel thresholding: a research review[J]. Iranian Journal of Science and Technology, Transactions of Electrical Engineering, 2020, 44(1): 1-29.
- [7] 王岩,胡睿甫,陈代鑫,等. 基于自适应阈值分割改进 Canny 算子的管道边缘检测方法研究 [J]. 光电子·激光, 2024, 35 (02): 164-170.
- [8] 郝琨,韩冰,李志圣,等. 基于投影阈值分割和数字序列校正的高噪声数字仪表图像识别方法 [J]. 电子科技大学学报, 2023, 52 (05): 728-738.
- [9] Sun R, Lei T, Chen Q, et al. Survey of image edge detection[J]. Frontiers in Signal Processing, 2022, 2: 826967.
- [10] 杨天敏,王晓燕. 基于强噪声条件下的激光图像边缘检测研究 [J]. 激光杂志, 2024, 45 (03): 156-160.
- [11] 陈烽,杨怀. 边缘异常识别下视频图像篡改细节检测 [J]. 计算机仿真, 2024, 41 (02): 192-195+226.
- [12] Vivekananthan V, Vignesh R, Vasanthaseelan S, et al. Concrete bridge crack detection by image processing technique by using the improved OTSU method[J]. Materials Today: Proceedings, 2023, 74: 1002-1007.
- [13] Othman Z, Abdullah A, Kasmin F, et al. Road crack detection using adaptive multi resolution thresholding techniques[J]. TELKOMNIKA (Telecommunication

- Computing Electronics and Control), 2019, 17(4): 1874-1881.
- [14] Chen C, Seo H, Jun C H, et al. A potential crack region method to detect crack using image processing of multiple thresholding[J]. Signal, Image and Video Processing, 2022, 16(6): 1673-1681.
- [15] Shao C, Chen Y, Xu F, et al. A kind of pavement crack detection method based on digital image processing[C]//2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). IEEE, 2019: 397-401.
- [16] 熊铃铃,钱庆平,张世良. 基于改进 Otsu 的桥梁裂缝检测方法 [J]. 半导体光电, 2023, 44 (06): 965-971.
- [17] Nausheen N, Seal A, Khanna P, et al. A FPGA based implementation of Sobel edge detection[J]. Microprocessors and Microsystems, 2018, 56: 84-91.
- [18] 张婧,郭风成,左泽丹,等. 融合多方向 Sobel 算子的相干斑各向异性扩散抑制 [J]. 遥感技术与应用, 2023, 38 (05): 1118-1125.
- [19] 韩红桂,甄晓玲,李方昱,等. 基于多尺度卷积神经网络的手机表面缺陷识别方法 [J]. 北京工业大学学报, 2023, 49 (11): 1150-1158.
- [20] Rahmawati S, Devita R, Zain R H, et al. Prewitt and canny methods on inversion image edge detection: an evaluation[C]//Journal of Physics: Conference Series. IOP Publishing, 2021, 1933(1): 012039.
- [21] Joshi M, Vyas A. Comparison of Canny edge detector with Sobel and Prewitt edge detector using different image formats[J]. Int. J. Eng. Res. Technol, 2020, 1.
- [22] 姚智超,楚晓亮,范筠益,等. 基于 Prewitt 算子的 X 波段雷达有效波高反演研究 [J]. 系统工程与电子技术, 2022, 44 (04): 1182-1187.
- [23] Nnolim U A. Automated crack segmentation via saturation channel thresholding, area classification and fusion of modified level set segmentation with Canny edge detection[J]. Heliyon, 2020, 6(12).
- [24] 李运堂,詹叶君,叶春泉,等. 复杂背景下航拍图像输电线提取算法研究 [J]. 计算机应用与软件, 2024, 41 (03): 220-225.
- [25] 于新善,孟祥印,金腾飞,等. 基于改进 Canny 算法的物体边缘检测算法 [J]. 激光与光电子学进展, 2023, 60 (22): 221-230.
- [26] Wanto A, Rizki S D, Andini S, et al. Combination of Sobel+ Prewitt edge detection method with Roberts+ Canny on passion flower image identification[C]//Journal

- of Physics: Conference Series. IOP Publishing, 2021, 1933(1): 012037.
- [27] 焦瑾瑾. 基于 Roberts 边缘检测的数字媒体激光重现方法 [J]. 激光杂志, 2022, 43 (05): 197-200.
- [28] 李志远,王光辉. 基于传统边缘算子的车道线识别研究 [J]. 现代电子技术, 2024, 47 (07): 61-65.
- [29] 郭运冲,李孟军,刘名果,等. 基于 Canny 算子的建筑裂缝边缘检测改进算法 [J]. 计算机仿真, 2022, 39 (11): 360-365+410.
- [30] Luo J, Lin H, Wei X, et al. Adaptive canny and semantic segmentation networks based on feature fusion for road crack detection[J]. IEEE Access, 2023.
- [31] Wang Q, Zhang N, Jiang K, et al. Tunnel lining crack recognition based on improved multiscale retinex and sobel edge detection[J]. Mathematical Problems in Engineering, 2021, 2021: 1-9.
- [32] Arya D, Ghosh S K, Toshniwal D. Automatic recognition of road cracks using sobel components in digital images[C]//Conference of Transportation Research Group of India. Singapore: Springer Nature Singapore, 2021: 139-149.
- [33] 刘春阳,曹少华,杨维双,等. 基于激光点云的公路箱梁表面裂缝识别方法研究 [J]. 应用激光, 2024, 44 (02): 77-85.
- [34] Nigam R, Singh S K. Crack detection in a beam using wavelet transform and photographic measurements[C]//Structures. Elsevier, 2020, 25: 436-447.
- [35] Vashisht R K, Peng Q. Crack detection in the rotor ball bearing system using switching control strategy and Short Time Fourier Transform[J]. Journal of Sound and Vibration, 2018, 432: 502-529.
- [36] Islam M M, Hossain M B, Akhtar M N, et al. CNN based on transfer learning models using data augmentation and transformation for detection of concrete crack[J]. Algorithms, 2022, 15(8): 287.
- [37] Sharma V, Gupta M, Kumar A, et al. Video processing using deep learning techniques: A systematic literature review[J]. IEEE Access, 2021, 9: 139489-139507.
- [38] Han S Y, Lee H W. Deep reinforcement learning based edge computing for video processing[J]. ICT Express, 2023, 9(3): 433-438.
- [39] Li S, Deng W. Deep facial expression recognition: A survey[J]. IEEE transactions

- on affective computing, 2020, 13(3): 1195-1215.
- [40] Revina I M, Emmanuel W R S. A survey on human face expression recognition techniques[J]. Journal of King Saud University-Computer and Information Sciences, 2021, 33(6): 619-628.
- [41] Abut S, Okut H, Kallail K J. Paradigm shift from Artificial Neural Networks (ANNs) to deep Convolutional Neural Networks (DCNNs) in the field of medical image processing[J]. Expert Systems with Applications, 2023: 122983.
- [42] Xu W, Fu Y L, Zhu D. ResNet and its application to medical image processing: Research progress and challenges[J]. Computer Methods and Programs in Biomedicine, 2023: 107660.
- [43] 朱伟刚,汪伦,陈田,等. 基于改进 YOLOv5 的无人机影像道路裂缝检测方法 [J]. 测绘通报, 2024, (03): 173-178.
- [44] Zhang Y, Huang J, Cai F. On bridge surface crack detection based on an improved YOLO v3 algorithm[J]. IFAC-PapersOnLine, 2020, 53(2): 8205-8210.
- [45] Bi C, Wang J, Duan Y, et al. MobileNet based apple leaf diseases identification[J]. Mobile Networks and Applications, 2022: 1-9.
- [46] 宛鹤,张金艳,屈娟萍,等. 基于改进 MobileNet V3 的矿物智能识别模型 [J]. 金属矿山, 2024, (01): 174-181.
- [47] 任惠,夏静,卢锦玲,等. 基于红外图像和改进 MobileNet-V3 的光伏组件故障诊断方法 [J]. 太阳能学报, 2023, 44 (08): 238-245.
- [48] Xu X, Zhao M, Shi P, et al. Crack detection and comparison study based on faster R-CNN and mask R-CNN[J]. Sensors, 2022, 22(3): 1215.
- [49] Cao C, Wang B, Zhang W, et al. An improved faster R-CNN for small object detection[J]. Ieee Access, 2019, 7: 106838-106846.
- [50] 于虹,龚泽威一,张海涛,等. 基于 Faster R-CNN 算法的变电站设备识别与缺陷检测技术研究 [J]. 电测与仪表, 2024, 61 (03): 153-159.
- [51] 汪西晨,彭富伦,李业勋,等. 基于改进 Faster R-CNN 的红外目标检测算法 [J]. 应用光学, 2024, 45 (02): 346-353.
- [52] He D, Qiu Y, Miao J, et al. Improved Mask R-CNN for obstacle detection of rail transit[J]. Measurement, 2022, 190: 110728.
- [53] 彭卫兵,张明见,全柳萌,等. 基于改进 Mask R-CNN 的钢桥多病害智能识别

- [J]. 中国公路学报, 2024, 37 (02): 100-109.
- [54] 张农,袁钰鑫,韩昌良,等. 基于 Mask R-CNN 的煤矿巷道掘进迎头裂隙检测与定位算法 [J]. 采矿与安全工程学报, 2023, 40 (05): 925-932.
- [55] Li D, Duan Z, Hu X, et al. Pixel-level recognition of pavement distresses based on U-Net[J]. *Advances in Materials Science and Engineering*, 2021, 2021: 1-11.
- [56] Yao H, Liu Y, Lv H, et al. Encoder - decoder with pyramid region attention for pixel - level pavement crack recognition[J]. *Computer - Aided Civil and Infrastructure Engineering*, 2023.
- [57] 杨燕泽,王萌,刘诚,等. 基于语义分割的沥青路面裂缝智能识别 [J]. 浙江大学学报(工学版), 2023, 57 (10): 2094-2105.
- [58] Alipour M, Harris D K, Miller G R. Robust pixel-level crack detection using deep fully convolutional neural networks[J]. *Journal of Computing in Civil Engineering*, 2019, 33(6): 04019040.
- [59] Pan Y, Zhang G, Zhang L. A spatial-channel hierarchical deep learning network for pixel-level automated crack detection[J]. *Automation in Construction*, 2020, 119: 103357.
- [60] Li H, Yue Z, Liu J, et al. Sccdnnet: A pixel-level crack segmentation network[J]. *Applied Sciences*, 2021, 11(11): 5074.
- [61] Elhariri E, El-Bendary N, Taie S A. Automated pixel-level deep crack segmentation on historical surfaces using U-Net models[J]. *Algorithms*, 2022, 15(8): 281.
- [62] 钟梅嘉,李宁,石林,等. 基于注意力机制和多尺度特征融合的道路裂缝检测 [J]. 计算机工程与设计, 2023, 44 (12): 3714-3721.
- [63] Yan Y, Zhu S, Ma S, et al. CycleADC-Net: A crack segmentation method based on multi-scale feature fusion[J]. *Measurement*, 2022, 204: 112107.
- [64] Mei Q, Gül M. Multi-level feature fusion in densely connected deep-learning architecture and depth-first search for crack segmentation on images collected with smartphones[J]. *Structural Health Monitoring*, 2020, 19(6): 1726-1744.
- [65] Ye W, Ren J, Zhang A A, et al. Automatic pixel - level crack detection with multi - scale feature fusion for slab tracks[J]. *Computer - Aided Civil and Infrastructure Engineering*, 2023, 38(18): 2648-2665.

- [66] Xu S, Hao M, Liu G, et al. Concrete crack segmentation based on convolution–deconvolution feature fusion with holistically nested networks[J]. Structural Control and Health Monitoring, 2022, 29(8): e2965.
- [67] Mascarenhas S, Agarwal M. A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification[C]//2021 International conference on disruptive technologies for multi-disciplinary research and applications (CENTCON). IEEE, 2021, 1: 96-99.
- [68] Tao J, Gu Y, Sun J Z, et al. Research on vgg16 convolutional neural network feature classification algorithm based on Transfer Learning[C]//2021 2nd China international SAR symposium (CISS). IEEE, 2021: 1-3.
- [69] Chu H, Wang W, Deng L. Tiny - Crack - Net: A multiscale feature fusion network with attention mechanisms for segmentation of tiny cracks[J]. Computer - Aided Civil and Infrastructure Engineering, 2022, 37(14): 1914-1931.
- [70] 赵志宏,李春秀,杨绍普. GhostConv 轻量级网络设计及故障诊断研究 [J]. 振动工程学报, 2024, 37 (01): 182-190.
- [71] Qiao W, Liu Q, Wu X, et al. Automatic pixel-level pavement crack recognition using a deep feature aggregation segmentation network with a scSE attention mechanism module[J]. Sensors, 2021, 21(9): 2902.
- [72] Yao H, Liu Y, Li X, et al. A detection method for pavement cracks combining object detection and attention mechanism[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(11): 22179-22189.

攻读硕士学位期间承担的科研任务与主要成果

一、攻读硕士学位期间论文发表情况

[1]黄伟,尚正阳,凌基等.改进 YOLOX 的跨座式单轨 PC 轨道梁面裂纹检测方法[J].邵阳学院学报.(已录用)

二、攻读硕士学位期间专利申请情况

[1]黄伟,凌基,尚正阳等.一种发动机振动幅度测试台.CN220751364U.2024-04-09.

致谢

总觉得来日方长，却不知岁月清浅，行文至此，百感交集。当敲下“致谢”二字时，意味着我的求学生涯即将落下帷幕，谨向出现在我生命中给予过我支持和帮助的人们致以深深的敬意。

春晖寸草，山高海深。首先我想感谢父母对我二十余载的培养与无微不至的关爱，是他们辛勤的打拼为我创造了最好的生活。每当我想起他俩时，我的内心就会充满力量，会感受到温暖，从而拥有克服困难的勇气和能力。我是如此平凡，又是如此幸运，希望家人身体健康，平安就是福，其余都是锦上添花。

桃李不言，下自成蹊。感谢我的导师尚正阳老师这三年来对我的悉心教导和敦敦教诲。作为您的第一个研究生弟子，我毕业论文的开题，中期，初稿的优化以及最终的定稿都离不开您严谨的学术理念和一丝不苟的工作态度，您是我终生学习的榜样。未来我也会独立踏上人生的新征程走向工作岗位，我会带着您对我的启发和美好期望去做好接下来的每一件事。漫漫求学路，感谢潘家保老师、程雷老师、李宏钊老师、王安恒老师、汧阳成明老师等给予的关心、指导和帮助。衷心祝愿各位老师工作顺利，身体健康。

山水一程，三生有幸。既相遇，皆为缘。感谢一教的师兄强哥，课题组同门基哥对我的帮助和陪伴，你们总是能给予我最有力量的鼓励。感谢曾经在男五宿舍楼门口篮球场一起打球的兄弟们，球场上有很多美好的事情值得回忆。读研的这条路上遇到了许许多多的朋友，希望我们保持那份热爱，奔赴下一场山海。

山水相逢，终有一别，北京中路8号，我们后会有期！

尚德敏学 唯实惟新

