DOI：10.13890/j.issn.1000-128X.2026.06.208

基于改进GAN的地铁轨旁检测负样本生成方法

范浩彪，周宇欣，沃　野，张　俊，方　俊，田凯阳
（宁波市轨道交通集团有限公司 智慧运营分公司，浙江 宁波 315000）

摘　要：轨旁360°综合检测系统在地铁车辆故障检测中具有重要应用价值，但真实线路中故障样本的采集面临作业风险高、数量有限、类别分布不均及故障形态碎片化等问题。为此，文章提出一种基于改进生成对抗网络（generative adversarial network，GAN）的部件故障样本（负样本）生成方法，通过在给定图像外边界条件下训练生成器实现高质量故障样本的扩增。引入多种损失函数评估模型训练效果，并与DBN、Deep GSN、VAE等主流图像生成模型进行对比试验。结果表明，在给定图像外边界作为条件时，所提GAN算法生成的图像在质量上具有明显优势。以宁波地铁5号线为应用场景，该方法在仅有少量负样本的情况下生成了大量可供训练的负样本，极大地简化了传统采集流程，规避了拆装设备的风险，较好地解决了行业内轨旁检测系统小样本训练的难题。现场试验显示，在应用该文所提方法后，轨旁检测系统的真实故障检出率从初期的41.76%提升至90.84%，验证了该方法在解决轨旁检测系统小样本训练难题中的有效性与工程实用性。

关键词：宁波地铁5号线；地铁车辆；轨旁检测；GAN生成对抗网络；故障样本生成；小样本学习；城市轨道交通

中图分类号：U239.5；TP18　　**文献标志码：**A　　**文章编号：**1000-128X(2026)04-0187-09

0　引言

在轨旁检测全流程中，图像缺陷识别算法是保证轨旁检测系统故障检出率并有效控制误报率的关键。在行业实践中，训练高效稳定的轨旁图像缺陷识别算法高度依赖大规模正负样本数据集[1-3]。通过向算法模型输入大量正样本和负样本，使其自主学习正常与异常图像的特征差异，从而提升对部件缺陷识别的鲁棒性和准确性[3-4]。但在传统轨旁样本采集过程中，正样本易通过收集无故障列车图像获得，故障负样本则获取难度较大。这主要是因为地铁列车发生设备破损、紧固件脱落等真实故障的情况极少，可供采集的样本数量有限，且分布呈现碎片化特征，样本覆盖存在局限性[5]。因此，实际算法训练中经常面临负样本不足的问题，严重制约了轨旁检测系统检出率与误报率等核心性能的提升，成为轨旁应用推广的主要障碍之一。

目前行业主流的解决方案有两种：一是通过粘贴胶带等方式模拟设备破损或紧固件松脱/缺失的外观，采集模拟故障图像进行训练；二是通过拆除极少量车体设备或紧固件，采集真实故障图像进行训练。然而，这两种方案在实际应用中均存在明显短板。模拟故障方案的安全风险较低，能最大程度覆盖不同部件的样本，但胶带覆盖紧固件的模拟图像与真实脱落的故障图像差异较大，可能导致算法训练失真。真实故障方案虽然能提供真实图像，但安全风险高，且部分设备不具备设置真实故障的条件，导致故障样本数量少、覆盖不全，进一步影响训练效果。可见，无论模拟还是真实故障采集，本质上均受制于小样本困境，而轨道交通领域的小样本故障检测问题具有独特性，传统数据增强方法难以满足高精度检测需求，亟须生成式模型提供解决方案[6]。

通信作者：周宇欣，男，主要从事轨道交通车辆检修方面的研究；E-mail: 1685225342@qq.com

引用格式：范浩彪, 周宇欣, 沃野, 等. 基于改进GAN的地铁轨旁检测负样本生成方法[J]. 机车电传动, 2026(4): 187-195.DOI:10.13890/j.issn.1000-128X.2026.06.208.

Citation：FAN Haobiao, ZHOU Yuxin, WO Ye, et al. A method for generating negative samples for metro trackside detection based on improved GAN [J]. Electric drive for locomotives,2026(4): 187-195.DOI:10.13890/j.issn.1000-128X.2026.06.208.

生成对抗网络（GAN）具备强大的建模能力，能够生成高保真度图像，基于GAN的图像生成算法已广泛应用于小样本学习、数据不平衡、超分辨率、目标检测与跟踪、图像属性编辑、医学图像等领域[7]。与DBN[8]、VAE[9]等其他生成模型相比，GAN及其变体在生成图像的清晰度和真实感方面已被广泛证明具有优势，尤其适合生成高保真的故障细节[10-12]。针对上述两种采样方法的缺陷，本文提出一种基于改进条件生成对抗网络（conditional generative adversarial network，CGAN）的故障样本生成方法，将故障区域外边界作为条件信息引入生成过程，并设计边缘一致性损失与纹理类别一致性损失两个损失函数，以引导生成器在给定边界条件下补全故障区域图像。本文将首先介绍数据预处理方法，随后阐述CGAN模型的设计思路与训练过程，最后通过消融试验与横向对比来验证所提方法的有效性。

1 数据预处理

在使用GAN模型生成负样本之前，需先使用轨旁检测系统完成列车各主要设备、部件及紧固件在正常状态下的正样本采集。控制车辆以正常状态通过轨旁检测系统，由安装于龙门架四周的多组相机同步采集其360°无故障图像，经收集整理后获得对应的正样本数据。多次重复该过程，即可完成宁波地铁5号线全部车辆的正样本采集。

为提升GAN生成负样本的图像质量与准确性，同时保证算法生成的负样本尽可能逼近设备真实故障时的图像，还需采集一定数量同类型设备与紧固件的真实负样本数据。这些真实负样本将用于确定GAN的网络结构与参数优化目标，调整训练数据集的特征分布，并通过衡量生成图像与真实图像之间的差异进行迭代训练。

研究团队采用的真实负样本采集方法为：在宁波地铁5号线车辆上人为设置紧固件缺失、设备破损等可控故障，使列车在受控条件下驶过轨旁检测系统，完成对应区域的真实负样本采集，采集过程如图1所示。



图1 宁波地铁5号线真实负样本采集

Fig. 1 Collection of true negative samples from Ningbo metro line 5

通过上述方法共采集宁波地铁5号线车辆真实故障原始图像273张（紧固件缺失129张，部件外观缺陷71张，方孔锁未锁闭73张），涵盖多种实际运行中可能出现的故障类型。采集工作完成后，需对原始图像进行预处理，以便后续用于算法模型训练。预处理流程主要包括故障区域人工标注、故障区域扩边裁切及故障区域掩码3个步骤，具体预处理流程如图2所示，以下将对其进行详细分析。

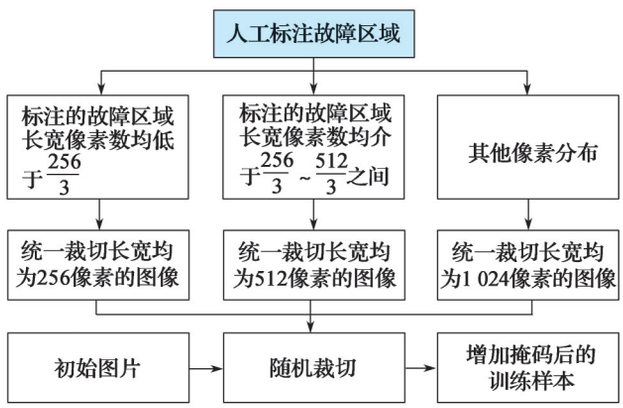


图2 原始图像预处理流程图

Fig. 2 Flowchart of raw image preprocessing

第一步，对原始图像中的故障区域进行人工标注，在标注过程中操作人员需按照统一标准精确框选出故障区域的轮廓，以保证后续扩边裁切环节的准确性和一致性。

第二步，对故障区域图像进行扩边裁切。通过扩边裁切，使人工标注的故障区域占裁剪后图像面积不超过1/9，以保留更多边缘信息。但在实际裁切过程中，各故障区域长宽尺寸存在差异，直接进行扩边裁切会导致裁切后的图像大小不统一，对后续图像训练造成不便，因此设计如下的裁切规则：

① 若标注的故障区域长和宽像素数均小于$\frac{256}{3}$，则统一裁切为长和宽均为256像素的图像；

② 若标注的故障区域长和宽像素数均介于$\frac{256}{3}$～$\frac{512}{3}$之间，则统一裁切为长和宽均为512像素的图像；

③ 其他情况下，统一裁切为长宽均为1 024像素的图像。

为丰富裁切样本的构图、背景和缺陷位置的多样性，在图像裁切过程中引入随机生成裁切框的方法：确定裁切尺寸后，在故障区域附近随机偏移裁切框位置，使每次裁切获得的图像在背景及缺陷呈现位置上有所差异，且要求裁切框内完整包含整个标注缺陷区域。在样本数量有限的情况下，该方法可有效增加训练数据的分布范围，缓解模型因样本单一而产生的过拟合问题。完成上述裁切操作后，将获得一张包含真实故障信息的训练样本，记作“I_{real}”。

第三步，对裁切后的图片中标注的真实故障区域进行掩码操作，即将故障区域像素值置为0。完成上述掩码操作后，将获得一张故障区域被遮蔽的训练样本，记作“I_{mask}”[13-14]。

类似地，对于无故障的正样本图像，只需在人工标注阶段将拟生成模拟故障的正常部件标注出来，再依次进行扩边裁切和区域掩码，即可获得用于后续图像生成的 I_{mask}。宁波地铁5号线正样本图像预处理过程见图3。

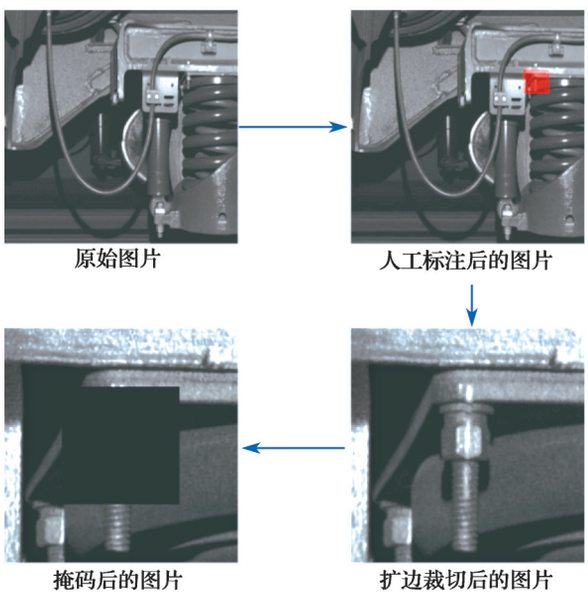


图3　宁波地铁5号线正样本图像预处理示意图

Fig. 3　Schematic diagram of image preproccessing for positive samples from Ningbo metro line 5

2　算法模型设计

2.1　模型网络框架

完成图像预处理后，将获得的 I_{real} 和 I_{mask} 分别输入用于故障负样本生成的CGAN进行训练。本文采用的CGAN是标准GAN的一种变体，其通过在训练过程中额外引入条件向量，与随机噪声向量一同输入生成器和判别器，利用条件信息指导生成过程[15]，从而能根据给定条件生成特定类型的样本。

CGAN模型主要由故障区域外边界条件编码器（Network$_1$，简称N$_1$）、生成器网络（Generator，简称G）和判别器网络（Discriminator，简称D）3个模块构成。训练时，首先由条件编码器N$_1$提取 I_{mask} 中故障区域外边界的条件信息，并将其编码为一组特征向量，作为条件向量 $\boldsymbol{C}$ 分别输入生成器网络和判别器网络[16]。随后，生成器网络根据输入的条件向量 $\boldsymbol{C}$ 和随机噪声向量 z 生成 I_{gen}，并将结果送入判别器网络。最后，判别器网络结合 I_{real}、条件向量 $\boldsymbol{C}$ 和随机噪声向量 z，判断生成的 I_{gen} 是否为真实样本。CGAN模型整体结构如图4所示。

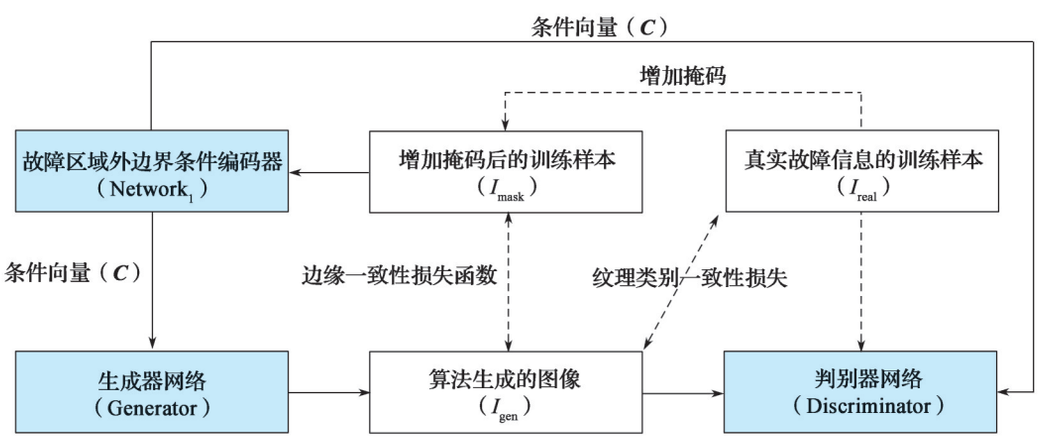


图4　CGAN模型整体结构图

Fig. 4　Overall architecture of the CGAN model

在实际训练过程中，由于负样本较少而正样本较多，为实现用正样本数据生成负样本数据，研究团队将故障区域的外边界作为条件向量，引导网络生成中间的故障部分。同理，在模型推理阶段，将正常零件区域的外边界作为条件向量，控制网络生成中间的故障部分[17]，从而利用CGAN将输入的正样本数据自动转换为近似负样本数据输出。

具体而言，在CGAN模型中，生成器网络接收2个输入：随机噪声向量 z 和条件向量 $\boldsymbol{C}$。其中，z 用于控制生成的多样性，条件向量 $\boldsymbol{C}$ 则用于约束生成故障区域的外边界。为实现条件向量的解耦，单独使用N$_1$模块提取条件向量 $\boldsymbol{C}$，该过程算法如式(1)所示。

$$\boldsymbol{C}=\mathrm{N}_1(I_{mask}) \tag{1}$$

当完成条件向量 $\boldsymbol{C}$ 的提取后，生成器网络的核心目标是在条件向量 $\boldsymbol{C}$ 和随机噪声向量 z 的共同影响下生成尽可能逼真的样本。因此在训练过程中，要求生成器网络的输出满足区域外边界条件 $\boldsymbol{C}$，并能正确补全掩码区域内的故障图像，因此最终生成器网络输出的生成样本 I_{gen} 如式(2)所示。

$$I_{gen}=\mathrm{G}(z,\ \boldsymbol{C}) \tag{2}$$

判别器网络的输入由3部分组成，分别为生成样本 I_{gen}（或真实样本 I_{real}）、条件向量 $\boldsymbol{C}$、随机噪声向量 z。判别器网络的目标是结合条件向量 $\boldsymbol{C}$、随机噪声向量 z，准确区分生成器网络生成的 I_{gen} 和 I_{real}，判断何者是真实样本，何者是生成样本。常用的判别器网络是一个简单的二分类网络，需判断生成器生成的图像是否符

合边界条件$\boldsymbol{C}$，以及补全的掩码区域图像是否接近真实情况。最终判别器输出生成样本的判别概率（PBTY gen，简称P_{gen}）如式(3)所示。

$$P_{\text{gen}}=\text{D}(I_{\text{gen}},\ \boldsymbol{C}) \tag{3}$$

判别器输出真实样本的判别概率（PBTY real，简称P_{real}）如式(4)所示。

$$P_{\text{real}}=\text{D}(I_{\text{real}},\boldsymbol{C}) \tag{4}$$

2.2 损失函数设计

GAN模型训练的本质是生成器网络与判别器网络交替训练的过程，上文已详细论述了负样本生成CGAN模型中生成器网络和判别器网络的实现方式。为使生成器网络生成的负样本更接近真实负样本[18]，研究团队参考标准GAN模型中交叉熵损失函数（cross entropy loss），引入边缘一致性损失loss_1和纹理类别一致性损失loss_2。

其中，边缘一致性损失loss_1这一参数是通过计算生成的外边界与原始外边界之间的L2损失（欧氏距离），评估生成结果与原始边界的相似程度，具体如式(5)所示。

$$\text{loss}_1=\text{L2}\left(I_{\text{mask}},\ I_{\text{gen}}\right) \tag{5}$$

纹理类别一致性损失loss_2则通过VGG网络提取生成样本与真实样本的特征，通过box函数单独取出生成区域，并去掉外边界部分，计算生成的区域和原始区域之间的差异[19]，以评估生成的图像与原始图像之间的相似程度，具体如式(6)所示。

$$\text{loss}_2=\text{VGG}(\text{box}\left(I_{\text{real}}\right),\ \text{box}(I_{\text{gen}})) \tag{6}$$

综上所述，生成器网络的最终优化目标（生成器损失函数）记为L_{G}，如式(7)所示。

$$L_{\text{G}}=-\log\left(P_{\text{gen}}\right)+\text{loss}_1+\text{loss}_2 \tag{7}$$

相反，判别器网络的目标是尽可能区分真实图像与生成图像，同理，使用交叉熵损失函数后，判别器网络的最终优化目标（判别器损失函数）L_{D}如式(8)所示。

$$L_{\text{D}}=-\log P_{\text{real}}-\log\left(1-P_{\text{gen}}\right) \tag{8}$$

3 算法训练及有效性验证

3.1 算法训练步骤与环境配置

在确定CGAN模型整体结构及判别器、生成器等主要模块的结构后，即可对CGAN模型中的判别器网络和生成器网络进行交替训练，其主要训练过程如下[20]：

步骤一：初始化生成器网络G和判别器网络D的参数。

步骤二：从训练集中选取一批真实样本I_{real}，并提取对应的条件向量$\boldsymbol{C}$。

步骤三：从标准正态分布中随机采样一批噪声向量$\boldsymbol{z}$。

步骤四：随机选取条件向量$\boldsymbol{C}$与噪声向量$\boldsymbol{z}$，并将其输入生成器网络，生成一批负样本I_{gen}。

步骤五：使用真实样本I_{real}和生成负样本I_{gen}分别更新判别器网络的参数，最小化判别器损失函数L_{D}。

步骤六：在判别器固定的情况下，使用生成负样本I_{gen}更新生成器网络的参数，最小化生成器损失函数L_{G}。

步骤七：重复步骤二~步骤六，直至完成200个epoch的训练或损失收敛。

训练在8块NVIDIA A100 GPU上进行，总批次大小（batch size）设为16，在每张GPU上并行处理2张图像。模型使用AdamW作为优化器，其权重衰减设置为10^{-4}。训练时初始学习率设置为10^{-3}，采用多步学习率衰减策略（MultiStepLR），每40个epoch将学习率降低到原来的1/10，总训练轮数设为200个epoch，训练最终学习率为10^{-7}。

3.2 算法消融试验

本试验为统一变量，所有对比试验均使用10 000张正样本与200张负样本数据，分类模型统一采用基于MMPretrain 2.0开源库的ResNet-50模型。试验将数据组1设为对照组，不额外添加生成数据；其余数据组用于验证所提基于CGAN的生成方法中各模块的有效性。试验设置中符号“√”表示该组数据训练过程中引入对应方法，相反，试验设置中符号“×”表示未引入对应方法。

完成试验设计后，为更好地量化生成样本与真实样本间的相似度，通过计算生成样本数据的FID(réchet inception distance)对图像相似度进行客观评估。FID是一种常用的评估指标，用于衡量两组图像之间的相似性，评估生成图像与真实图像的质量，其值越小，表明生成效果越好[21]。

在试验过程中，除需对生成图像的质量进行评估外，还需对CGAN算法模型在非同源数据集上固定阈值下的精确率（precision）、召回率（recall）及平均精度（average precision,AP）进行测试。为保障试验的公平性，所有试验组的精确率与召回率结果均在固定阈值0.5下测定，试验全程未调整分类阈值，以确保消融试验性能对比的客观性。在本组试验中，精确率指模型判定为真实样本的结果中实际为真实样本的比例，该指标可有效衡量模型判断的准确性；召回率指在所有真实样本中被正确判断为真实样本的比例，该指标可有效衡量模型对真实样本的覆盖程度。平均精度通过计算不同召回率下精确率的均值得到，代表PR曲线下的面积，其值越大代表模型性能越好。通过统计上述3项指标，可有效评估模型在预测真实样本和覆盖真

实样本方面的表现，不同训练参数下生成数据质量和下游模型训练结果如表1所示。

表1　不同训练参数设置生成数据质量和用于下游模型训练的结果统计表

Table 1　Statistics of generated sample quality and downstream model training performance under different training parameter settings

序号	额外使用的生成负样本					生成质量	模型性能		
	GAN	条件向量 $\boldsymbol{C}$	使用外边界条件编码器 N_1	边缘一致性损失 $loss_1$	纹理类别一致性损失 $loss_2$	FID	精确率/%	召回率/%	平均精度 AP/%
1	×	×	×	×	×	–	99.63	66.73	81.22
2	√	×	×	×	×	103.3	95.45	73.76	83.45
3	√	×	×	√	×	91.5	96.76	85.35	90.25
4	√	×	×	×	√	95.0	95.12	78.91	86.13
5	√	√	×	×	×	87.6	97.63	89.13	92.18
6	√	√	√	×	×	82.6	98.56	95.42	96.87
7	√	√	√	√	√	73.5	99.98	98.37	98.92

试验结果表明，单独使用标准GAN模型生成的样本质量不高，FID值最高为103.3，表明生成的样本与真实样本的相似度最低，但增加该生成数据后，精确率（precision）和召回率（recall）均有显著提高。此外，增加条件向量 $\boldsymbol{C}$、使用外边界条件编码器 N_1、增加 $loss_1$、增加 $loss_2$，FID值均有下降，即生成质量进一步提高，对应的精确率和召回率也进一步提升。在引入所有模块的试验组（序号7）中，精确率达到99.98%，召回率达到98.37%，平均精度达到98.92%，显著高于其他试验组。这表明，引入条件向量、使用 N_1 进行条件信息解耦、增加边缘一致性损失 $loss_1$ 和纹理类别一致性损失 $loss_2$，均对生成模型有较高的收益，应用这些方法能大幅提升缺陷零件生成图像的质量，从而显著提升模型在下游任务中的性能。

3.3　不同生成模型横向对比试验

为验证所采用的CGAN模型对生成图像质量的影响，将其与当前几种常见的条件生成模型进行了性能对比，具体包括DBN[1]、Deep GSN[1]、VAE[2]和标准GAN[11]。为保证试验公平，所有模型均在本文构建的条件生成数据集上重新训练与评测，结果统计如表2所示。

表2　不同条件下条件生成模型性能横向对比结果统计表

Table 2　Comparison of CGAN model performance under different conditions

参数项目	条件生成模型	DBN	Deep GSN	VAE	标准GAN	本文模型
参数值或说明	FID	138.6	114.3	92.4	91.5	73.5

表2统计结果显示，与目前常见的几种条件生成模型相比，本文方法在有图像外边界作为条件时，生成质量有显著提升，其FID值达到73.5，优于VAE和标准GAN。

3.4　真实故障检出对比试验

为验证算法在真实场景下对故障数据检出率的影响，研究团队使用训练后的CGAN模型生成宁波地铁5号线全部检测项点的负样本数据。其中，126类项点现场已完成正样本和少量负样本的采集，剩余269类检测项点仅完成正样本采集，需完全采用本文方法生成对应的负样本。

基于现场采集的正负样本及本文CGAN模型生成的负样本完成轨旁检测算法训练后，对训练前后的算法性能进行对比分析。鉴于精确率（现场常以“误报数量”衡量）和召回率（现场常以“检出率”衡量）是评价检测算法性能的核心指标，且现场对误报数量有严格容忍阈值，故试验设定统一约束条件：训练前后算法精确率均不低于99.5%，即在算法完成全列车7 715个项点检测后，误报数量需控制在38个以内，在此基础上重点对比召回率（检出率）的提升效果。

为严格对比模型性能，研究团队首先在同一数据集上进行了测试，在误报数量满足约束的前提下，通过对比本文方法生成负样本训练前后的模型对真实故障的检出率，评估本文方法在部件负样本生成训练中的可行性，验证结果如表3所示。

验证结果显示，使用本文方法生成的部件负样本应用于检测算法训练后，同一数据集的真实故障检出率由41.82%提升至90.00%，而平均误报数量由25.0降至22.6，基本持平，证明在几乎不增加误报的前提下检出效果大幅提升，验证了本文方法在部件负样本生成训练中的可行性。

为进一步验证算法在现场的效果，研究团队将基于本文方法生成的部件负样本训练得到的检测模型部署于宁波地铁5号线轨旁检测系统。同时，为控制试验变量，统计部署前后对同一试验列车上273个设置位置与失效方式完全一致的真实故障的检出情况，重点验证本文方法应用后检测模型在实际运维场景中的适用

性。相关验证结果如表4所示。

表4的现场真实故障对比试验验证表明，应用本文方法生成负样本并投入检测模型训练后，其真实故障检出率由原来的41.76%提升至90.84%。在现场实际应用过程中的检出率显著提升，从而一步验证了本文方法应用于故障负样本生成在提升轨旁检出率中的可行性。

表3　本文方法应用前后对同一数据集的真实故障检出情况对比统计表

Table 3　Comparison of true fault detection performance on the same dataset before and after applying the proposed method

试验类型	真实故障测试数据包名称	设置故障数量	检出故障数量	误报数量	检出率/%	平均检出率/%	平均误报数量
应用本文算法前	20230607_05004_6	24	11	17	45.83	40.92	25.0
	20230607_05001_6	19	6	22	31.58		
	20230608_05011_1	22	10	29	45.45		
	20230608_05021_6	28	13	24	46.43		
	20230608_05012_1	17	6	33	35.29		
应用本文算法后	20230607_05004_6	24	21	21	87.50	90.60	22.6
	20230607_05001_6	19	18	17	94.74		
	20230608_05011_1	22	20	25	90.91		
	20230608_05021_6	28	24	27	85.71		
	20230608_05012_1	17	16	23	94.12		

表4　宁波地铁5号线应用本文方法前后真实故障检出对比

Table 4　Comparison of true fault detection results for Ningbo metro line 5 before and after applying the proposed method

试验类型	试验序号	设置故障数量	检出故障数量	误报数量	检出率/%	检出故障总数	平均检出率/%	平均误报数量
应用本文算法前	1	21	7	23	33.33	114	42.41	23.08
	2	20	5	25	25.00			
	3	20	10	18	50.00			
	4	20	10	22	50.00			
	5	20	11	27	55.00			
	6	20	9	21	45.00			
	7	18	10	36	55.56			
	8	20	10	16	50.00			
	9	31	8	19	25.81			
	10	43	19	28	44.19			
	11	20	9	22	45.00			
	12	20	6	20	30.00			
应用本文算法后	13	14	13	17	92.86	248	91.13	21.58
	14	24	23	19	95.83			
	15	10	9	32	90.00			
	16	14	13	16	92.86			
	17	24	23	27	95.83			
	18	10	9	11	90.00			
	19	9	8	25	88.89			
	20	30	29	22	96.67			
	21	24	21	23	87.50			
	22	37	34	19	91.89			
	23	44	38	20	86.36			
	24	33	28	28	84.85			

4　算法结果可视化

按照前文所述方法完成模型训练后，CGAN模型能够较有效实现故障区域的图像补全，且生成的样本具备良好的多样性。在完成模型训练后，对于与训练数据中故障区域边界分布相近的正常零件区域，生成器网络可将其外边界作为条件输入，结合训练过程中学习的故障特征分布，生成该区域内的模拟故障部分，补全效果如图5所示。

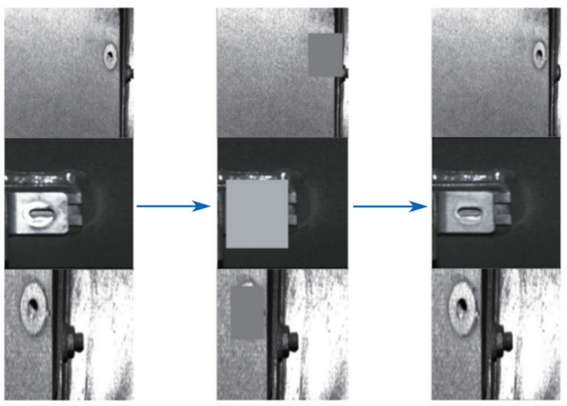


图5　训练后CGAN算法补全的效果图

Fig. 5　Completion results of the CGAN model after training

对比原始真实故障图像与CGAN补全后的图像，肉眼观察下补全区域已较为接近真实情况，模型训练基本满足负样本生成的需求[22]。此时，只需将正常零件区域的外边界条件向量作为输入，生成器网络即可生成合理的缺陷部分图像，生成的负样本效果如图6所示。

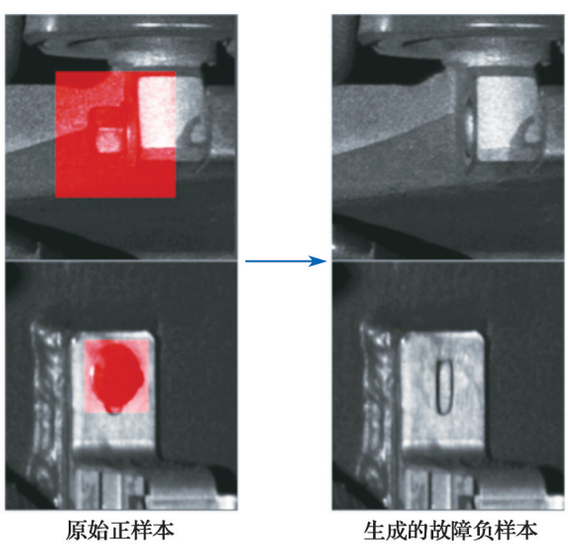


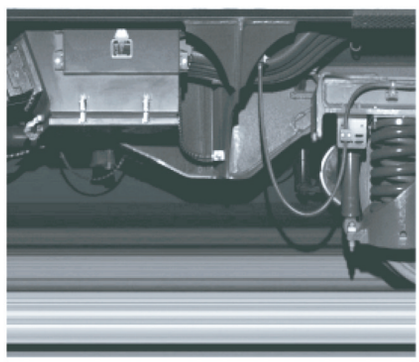
图6　生成的负样本效果图

Fig. 6　Partial results of generated negative samples

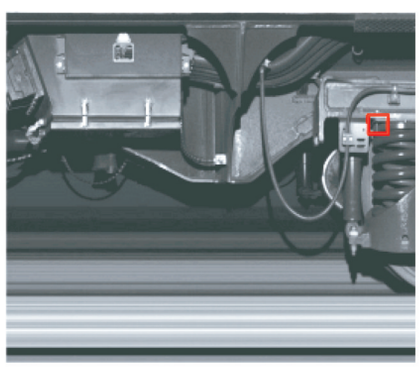
除了利用CGAN生成负样本数据，在实际应用中还可通过引入不同的随机噪声向量z，为相同区域生成具有不同细节差异的结果，从而大幅丰富样本数据。

获得CGAN生成的负样本后，需根据原始数据的裁剪坐标，将生成的故障区域粘贴回原图并对齐，最终得到一张与原始正样本图像对应的完整负样本故障图像，完整负样本的生成效果如图7所示。

上述可视化结果表明，本文提出的改进CGAN模型在给定正常零件区域外边界条件下，不仅能够准确补全故障区域，还可通过引入不同随机噪声向量生成具有多样性的样本，生成的完整负样本在视觉上与真实故障图像高度相似。这一结果与前述消融试验和对比试验的定量结论相互验证，充分验证了本文方法在地铁轨旁检测负样本生成任务中的有效性与实用性。



(a) 原始正样本数据



(b) GAN算法生成的故障负样本数据

图7　生成的完整负样本效果图

Fig. 7　Complete results of generated negative samples

5　结束语

针对地铁轨旁检测系统传统负样本采集面临的风险高、数量少、分布不均、类型碎片化等突出问题，本文提出了一种基于改进条件生成对抗网络（CGAN）的故障负样本生成方法。在技术创新层面，本研究突破了传统依赖真实故障模拟或设备拆装的采样模式，构建了生成质量高、细节保真度好的改进CGAN模型。该模型的核心创新包括两点：一是引入故障区域外边界条件编码器（N_1），将正常零件区域的外边界信息与故障生成过程进行解耦，使生成器能够根据任意给定的外边界条件精准生成对应的故障图像；二是设计边缘一致性损失（$loss_1$）与纹理类别一致性损失（$loss_2$）双损失函数，分别从边缘结构和纹理细节两个层面约束生成质量。该模型仅需少量真实负样本与一定正样本即可训练，能够精准生成符合故障特征规律、细节保真度高的负样本数据。其生成效果经FID指标验证，优于DBN、Deep GSN、VAE等主流生成模型，填补了轨旁检测领域小样本生成技术的应用空白[23]。

在宁波地铁5号线真实数据集上的试验结果表明：采用本文方法的FID值达到73.5，效果优于DBN（138.6）、Deep GSN（114.3）、VAE（92.4）及标准

GAN（91.5）。在工程部署方面，该方法不仅极大地简化了传统负样本的采集流程，还规避了设备拆装带来的安全风险，有效解决了126类已有少量负样本项点的样本扩充问题及269类无负样本项点的样本补全需求。通过生成负样本扩充训练集，系统真实故障检出率从初始的41.76%提升至90.84%，且保持99.5%以上的高精确率，验证了该方法在轨旁检测实际应用中的有效性与可行性。综上，本研究为城市轨道交通智能检测领域的技术升级提供了技术支撑与实践范式，其核心思路可进一步推广至其他轨道交通线路及铁路等相关交通领域的故障检测场景。

参考文献：

[1] 宁波轨道集团有限公司智慧运营分公司. 电客车维护规程[Z]. 宁波: 宁波轨道集团有限公司智慧运营分公司, 2023.
Smart Operation Branch, Ningbo Rail Transit Group Co., Ltd.. Maintenance regulations for electric passenger trains [Z]. Ningbo: Smart Operation Branch, Ningbo Rail Transit Group Co., Ltd., 2023.

[2] 贾建坤. 铁路客车故障轨旁图像检测系统(TVDS) 的运用研究分析[J]. 轨道交通装备与技术, 2023(4): 27-31.
Jia Jiankun. Application research and analysis of fault trackside vision detection system (TVDS) for passenger train[J]. Rail Transportation Equipment and Technology, 2023(4): 27-31.

[3] 郭燕辉. 城轨车辆多场景结合的智能列检系统研究[J]. 铁道运输与经济, 2023, 45(12): 195-202.
Guo Yanhui. Research on intelligent train inspection system of urban rail train under multi-scene combination[J]. Railway Transport and Economy, 2023, 45(12): 195-202.

[4] 石进. 基于弱监督学习的图像处理技术研究[D]. 贵阳: 贵州大学, 2023.
Shi Jin. Research on image processing technology based on weakly supervised learning[D]. Guiyang: Guizhou University, 2023.

[5] 陈瑞, 汪美玲. 改进多级网络的商品图像识别及鲁棒性研究[J]. 西安邮电大学学报, 2020, 25(2): 23-28.
Chen Rui, Wang Meiling. Research on commodity image recognition and robustness of improved multi-scale network [J]. Journal of Xi'an University of Posts and Telecommunications, 2020, 25(2): 23-28.

[6] Pei Z, Jiang H, Li X, et al. Data augmentation for rolling bearing fault diagnosis using an enhanced few-shot Wasserstein auto-encoder with meta-learning[J]. Measurement Science and Technology, 2021, 32(8): 1-22.

[7] 徐金鹏, 郭新峰, 王瑞波, 等. 基于GAN数据增强的软件缺陷预测聚合模型[J]. 计算机科学, 2023, 50(12): 24-31.
Xu Jinpeng, Guo Xinfeng, Wang Ruibo, et al. Aggregation model for software defect prediction based on data enhancement by GAN[J]. Computer Science, 2023, 50(12): 24-31.

[8] Cao H Y, Guo X, Wang G. Meta-learning with GANs for anomaly detection, with deployment in high-speed rail inspection system[EB/OL]. (2022-02-11) [2026-03-09]. https://arXiv.org/abs/2202.05795.

[9] Raza A, Sehar R, Moiz A, et al. Novel conditional tabular generative adversarial network based image augmentation for railway track fault detection[J]. PeerJ Computer Science, 2025, 11: e2898.

[10] 陈佛计, 朱枫, 吴清潇, 等. 生成对抗网络及其在图像生成中的应用研究综述[J]. 计算机学报, 2021, 44(2): 347-369.
Chen Foji, Zhu Feng, Wu Qingxiao, et al. A survey about image generation with generative adversarial nets[J]. Chinese Journal of Computers, 2021, 44(2): 347-369.

[11] Goodfellow I J, Pouget A J, Mirza M, et al. Generative adversarial networks[EB/OL]. (2014-06-10) [2026-03-09]. https://arXiv.org/abs/1406.2661.

[12] 王宥翔. 基于生成对抗网络的图像超分辨率重建[J]. 河南教育学院学报(自然科学版), 2023, 32(4): 47-52.
Wang Youxiang. Reconstruction of super resolution image based on GAN[J]. Journal of Henan Institute of Education (Natural Science Edition), 2023, 32(4): 47-52.

[13] 骆迪, 张乾, 柏武贰. 基于掩码自动编码器的图像修复研究[J]. 现代信息科技, 2024, 8(3): 69-72.
Luo Di, Zhang Qian, Bai Wuer. Research on image restoration based on mask autoencoder[J]. Modern Information Technology, 2024, 8(3): 69-72.

[14] Zhang G, Cui K, Hung T Y, et al. Defect-GAN: high-fidelity defect synthesis for automated defect inspection[C]//2021 IEEE Winter Conference on Applications of Computer Vision (WACV), 2021: 2523-2533.

[15] 晁小朋. 基于L2范数约束的生成对抗网络研究及应用[D]. 广州: 广东工业大学, 2020.
Chao Xiaopeng. Research and application of generative adversarial networks based on L2 norm constraint[D]. Guangzhou: Guangdong University of Technology, 2020.

[16] 张昀, 黄橙, 施健, 等. 一种基于生成对抗架构的目标检测增强算法[J]. 计算机学报, 2024, 47(3): 647-661.
Zhang Yun, Huang Cheng, Shi Jian, et al. An enhanced algorithm for object detection based on generative adversarial structure[J]. Chinese Journal of Computers, 2024, 47(3): 647-661.

[17] Isola P, Zhu J Y, Zhou T, et al. Image-to-image translation with conditional adversarial networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 5235-5245.

[18] 谢一博, 刘卫国, 周顺, 等. 基于显著性的双鉴别器GAN图像融合算法[J]. 应用光学, 2024, 45(1): 107-117.
Xie Yibo, Liu Weiguo, Zhou Shun, et al. Saliency-based dual discriminator GAN image fusion algorithm[J]. Journal of Applied Optics, 2024, 45(1): 107-117.

[19] Johnson J, Alahi A, Li F. Perceptual losses for real-time style transfer and super-resolution[C]//European Conference on Computer Vision, 2016: 694-711.

[20] 中车株洲电力机车有限公司. 宁波轨道交通5号线轨旁综合检测设备维护手册[Z]. 株洲: 中车株洲电力机车有限公司, 2022.
CRRC Zhuzhou Locomotive Co., Ltd.. Maintenance manual for trackside comprehensive inspection equipment of Ningbo rail transit line 5[Z]. Zhuzhou: CRRC Zhuzhou Locomotive Co., Ltd., 2022.

[21] Heusel M, Ramsauer H, Unterthiner T, et al. GANs trained by a two time-scale update rule converge to a local nash equilibrium[EB/OL]. (2017-06-26) [2026-05-18]. https://arxiv.org/abs/1706.08500.

[22] 宁波轨道集团有限公司智慧运营分公司. 5号线轨旁车辆综合检测系统作业指导书[Z]. 宁波: 宁波轨道集团有限公司智慧运营分公司, 2022.
Smart Operation Branch, Ningbo Rail Transit Group Co., Ltd.. Operation instruction manual for trackside vehicle comprehensive inspection system of line 5[Z]. Ningbo: Smart Operation Branch, Ningbo Rail Transit Group Co., Ltd., 2022.

[23] Lin H, Liu Y, Li S, et al. How generative adversarial networks promote the development of intelligent transportation systems: a survey[J]. IEEE/CAA Journal of Automatica Sinica, 2023, 10(9): 1781-1796.

A method for generating negative samples for metro trackside detection based on improved GAN

FAN Haobiao, ZHOU Yuxin, WO Ye, ZHANG Jun, FANG Jun, TIAN Kaiyang

(Smart Operation Branch, Ningbo Rail Transit Group Co., Ltd., Ningbo, Zhejiang 315000, China)

Abstract: The 360° trackside comprehensive detection system has demonstrated significant value for detecting faults in metro vehicles. However, collecting fault samples from real railway lines faces challenges such as high operational risks, limited sample sizes, imbalanced category distributions, and fragmented fault patterns. To address these, this paper proposes a component fault sample (negative sample) generation method based on an improved generative adversarial network (GAN). The method augments high-quality fault samples by training a generator conditionally under prescribed outer boundaries of images. Multiple loss functions are introduced to evaluate model training performance, and comparative experiments were conducted against mainstream image generation models such as DBN, Deep GSN, and VAE. The results show that the proposed algorithm exhibits distinct image quality advantages conditionally under the prescribed outer boundaries of images. In the Ningbo metro line 5 scenario, the method generated a large number of negative samples for training despite the initially scarce dataset, providing an effective solution to few-shot training for trackside detection systems by greatly simplifying the traditional collection process and avoiding risks associated with equipment disassembly and installation. In on-track experiments, the application of the proposed method increased the true fault detection rate of the trackside detection system from 41.76% to 90.84%, confirming the method's effectiveness and engineering practicality in addressing few-shot training challenges in trackside detection systems.

Keywords: Ningbo metro line 5; metro vehicle; trackside detection; generative adversarial network (GAN); fault sample generation; few-shot learning; urban rail transit

收稿日期：2025-02-22；修回日期：2026-05-16
责任编辑：詹　婧