文章编号：2097-5899(2025)03-0015-10　　**中图分类号**：U294.2　　**文献标识码**：A
DOI：10.16669/j.cnki.issn.2097-5899.202410120001
引用格式：葛悦，王志敬，柯向喜，等. 基于场景约束的铁路货车装载状态图像样本生成技术研究[J]. 铁路物流，2025，43(3)：15-24.
GE Yue，WANG Zhijing，KE Xiangxi，et al. Scene-Constrained Image Sample Generation Technology for Railway Wagon Loading Status[J]. Railway Logistics，2025，43(3)：15-24.

基于场景约束的铁路货车装载状态图像样本生成技术研究

葛　悦[1]，王志敬[2]，柯向喜[3]，孙文桥[1]，王小卫[4]

(1. 中国铁道科学研究院集团有限公司 运输及经济研究所，北京 100081；2. 中国铁道科学研究院集团有限公司 北京中铁科客货运输技术有限公司，北京 100081；3. 中国铁路南宁局集团有限公司 柳州车站，广西 柳州 545001；4. 中国铁路昆明局集团有限公司 客运部，云南 昆明 650011)

摘　要：铁路货车装载状态图像是反映货物列车运行状态的重要载体，其智能化水平直接影响到行车安全和货运安全。受货物类型、货车类型、装载加固方案等因素影响，部分场景下铁路货车装载状态图像较少，难以满足该场景下识别项点目标检测模型构建需要。面向铁路货运检查作业需要，考虑铁路货车装载状态不良图像特征，采取基于生成-对抗的生成约束、前景项点目标提取、启发式项点目标放置等措施，构建基于生成对抗网络、实例分割等多技术融合的铁路货车装载状态图像生成技术模型，为丰富特殊场景下铁路货车装载状态图像库、构建可靠的智能识别模型提供技术支撑。

关键词：铁路货运；货检；货车装载状态；图像生成；目标检测；生成对抗网络

0　引言

铁路货运检查是铁路货运安全管理的重要组成部分，当前部分货检站搭设铁路货车装载状态视频监控设备，采集到包含货车车体、货物和货物装载加固状态的无色彩异常、完整、无畸变的清晰图像(以下简称为“标准图像”)[1]。随着基于深度学习的目标检测、实例分割等技术的快速发展与成熟应用，铁路货车装载状态标准图像智能识别装置能够对货运检查项点进行智能分析，辅助货检人员对问题车进行排查，有效提升货检作业质量和效率[2]。标准图像智能识别技术的准确检测，依托于大规模和多样化样本数据的积累。但是，受部分特殊货物运量较小、装载状态复杂等因素影响，标准图像正样本数量不足，难以确保该类型货物装载异常状态准确识别与报警。图像生成是一种增加模型训练样本以避免模型过拟合的手段，当前图像生成技术主要分为传统图像生成技术、基于风格迁移的图像生成技术、基于生成对抗网络的图像生成技术、基于大模型的多模态图像生成技术等[3]。实验发现，以上方法对于具有少样本特点项点的智能识别效果提升有限，因此，在充分分析少样本特征的标准图像智能识别需求基础上，创新基于场景约束的标准图像样本生成技术，并试验验证方法的可行性和有效性，为铁路特殊货物和特殊项点智能化识别和报警提供有效的数据支撑。

1　铁路特殊货物及特殊项点货车装载状态标准图像现状分析

1.1　业务分析

立足《铁路货运检查管理规则》《铁路货物装载加

收稿日期：2024-10-12
基金项目：中国国家铁路集团有限公司科技研究开发计划课题(J2024X004)；中国铁道科学研究院集团有限公司科研项目(2024YJ152)；中国铁道科学研究院集团有限公司运输及经济研究所自主创新研究基金项目(2023YJS06)

固规则》《铁路超限超重货物运输规则》等既有铁路货运规章制度，铁路货检作业检查的对象包含货车车体、货车承运的货物和货物装载加固状态3项内容[1]。从不同铁路货车、货物装载加固状态的检查项点来看，主要检查的对象以常见的棚车、敞车、平车等车型和散装堆货物、集装箱货物等货物为主，还有以铁路轮式、履带式货物、设备类货物、预应力梁货物为主的特殊货物，虽运量较小，但仍是铁路运输的重点货物类型，也是铁路货运安全关注的重点和难点。以轮式、履带式货物为例，从近年来铁路货运安全月度通报可以看出，铁路货运检查中出现的轮式、履带式货物相关险兆事件和事故时有发生，主要集中在掩挡失效、活动部位旋转等问题。

结合《铁路货车装载状态标准图像智能识别装置暂行技术条件》(TJ/KH041—2022)、《铁路货车装载状态视频监控设备运用管理办法(试行)》(铁货〔2023〕147号)，通用货车装载状态检查项点中，有关车体、货物的车侧和车顶“异物”的项点问题(以下简称为“特殊项点”)共14条，主要问题集中在车体、货物部位的异物和物体撒漏两方面。特殊货物检查项点如表1所示，在生产环境中的异物和撒漏问题遍布所有车种车型，且类型样式多，异物可能是闸瓦、编织袋、瓶子、货物零件等物体。通常，视频监控设备采集到的标准图像的异物对象尺寸较小(相比于全图占比0.01%～0.02%)，但在生产环境中，一些难以采集到的特征特殊、明显的异物项点图像，往往是危及行车和货运安全的核心因素[1]。异物图像数据样本如图1所示。

结合以上铁路货车装载状态业务现状，问题车图像样本收集难点如下。

(1) 特殊货物运量少，图像样本收集困难。以设备类货物、预应力梁货物等为代表的特殊货物，一是年装车量较少、类型较多，正类样本难以收集和完善，二是特殊货物安全管控措施严格，按照铁路货运安全管理要求应配备押运人，极少出现装载不良情况，因而较少收集到装载不良图像。

(2) 货物装载质量高，问题车图像收集困难。一是装卸车前后检查质量不断提升，有效增强装车源头安全管控水平，因源头装车造成货车不良装载状态的场景日益减少，难以采集到大量有明显货检项点问

表1　特殊货物检查项点

Tab.1　Special cargo inspection items

货车车种	货物类型	检查重点	检查内容	设备智能识别范围及报警标准
棚车	通用	车体异物	棚车顶部，车门、车窗、扶梯等车体外部，车辆连接处无异物	异物长度≥20 mm，且宽度或直径≥4 mm
棚车	通用	货物撒漏	车内无物体撒漏、露出	异物长度≥20 mm，且宽度或直径≥4 mm
敞车/平车	通用	车体异物	篷布顶部，未苫盖篷布的货物顶部、敞顶箱内部，货车内部、上侧梁、上端梁、车门、丁字铁、车辆连接处无异物	① 篷布顶部，未苫盖篷布的货物顶部、敞顶箱内部，货车内部、上侧梁、上端梁、车门、丁字铁、车辆连接处无异物长度≥20 mm，且宽度或直径≥4 mm；② 敞车内异物长度≥50 mm，宽度或直径≥10 mm
敞车	通用	货物撒漏	无车内物体撒漏、露出	异物长度≥20 mm，且宽度或直径≥4 mm
敞车	散堆装货物	货物装载加固	无货物撒漏造成货物表面形成锥形坑或偏载偏重	异物长度≥20 mm，且宽度或直径≥4 mm
平车	通用	车体异物	篷布顶部，未苫盖篷布的货物顶部、敞顶箱内部，车体、车辆连接处无异物	异物长度≥20 mm，且宽度或直径≥4 mm
平车	通用	货物撒漏	集装箱无货物撒漏、滴漏	异物长度≥20 mm，且宽度或直径≥4 mm
罐车	通用	车体异物	罐车顶部，罐体、侧梯及走台装配，车辆连接处无异物	异物长度≥20 mm，且宽度或直径≥4 mm
漏斗车	通用	车体异物	货车上侧梁、上端梁、檐板、扶梯、底架、底门及车辆连接处无异物	异物长度≥20 mm，且宽度或直径≥4 mm
漏斗车	通用	货物状态	无货物撒漏造成货物表面形成锥形坑	异物长度≥20 mm，且宽度或直径≥4 mm
JSQ小汽车运输专用车	通用	车体异物	车体、车辆连接处无异物	异物长度≥20 mm，且宽度或直径≥4 mm
检衡车	通用	车体异物	车体、车辆连接处无异物	异物长度≥20 mm，且宽度或直径≥4 mm
长大货物车	通用	车体异物	车体、车辆连接处无异物	异物长度≥20 mm，且宽度或直径≥4 mm

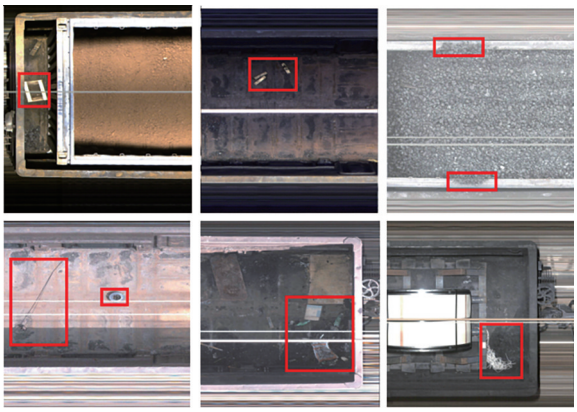


图1　异物图像数据样本

Fig.1　Foreign object image samples

题的样本图像。二是铁路货检作业质量不断提升，实现机检辅助人检协同作业，依靠货检人员专业素质能够有效识别和处置货车装载不良状态安全风险，较难收集到有明显货检项点问题的样本图像。

1.2　技术分析

铁路货车装载状态标准图像智能识别技术的核心方法是基于深度学习的目标检测模型，而大规模图像样本数据是保证模型能够稳定构建的基础；多样化项点异常特征的正样本数据，是提升项点智能识别准确性、避免漏报的重要因素[4]。图像样本收集困难导致智能识别模型难以学习到足量、有效的目标特征，使得模型难以捕捉到货物整体的旋转、局部的异常等项点问题，以及将阴影、车体印记等无关的信息判别为异常情况。图像正样本收集困难，导致依托单一性的数据训练学习的模型，面向不同天气、设备、光线变化场景下的图像，难以识别大尺寸、特殊位置、特殊变化的项点目标。因此，需要研究铁路货检标准图像样本生成技术，以标准图像为基础，增加有效样本及正样本特征的类型与规模，实际提升目标检测模型的泛化能力、鲁棒性和识别效果，以真正实现将机器智能判图的货检模式替代人眼逐张判图的技术革新。不同图像生成方法优缺点比较如表2所示。

对表2中9种图像生成方法进行归类和实验。

（1）手动使用PS工具抽取前景货物、项点目标，再手动粘贴到背景图像。由于生成的图像元素均来源于真实货检图像，保证生成图像能够直接应用于目标检测的数据集，但是人工编辑需要一定的时间和学习成本。

（2）基于图像增强方法生成图像。一种是对整张标准图像经过传统的几何、色彩、模糊等图像变换，另一种是基于强化学习的RandAugment和AutoAugment获取图像增强的最佳组合策略，能够生成并模拟不同厂家在不同线路采集到不同对比度、清晰度、视角的图像，虽然能够快速增加图像数量，提升目标检测评

表2　不同图像生成方法优缺点比较

Tab.2　Comparison of advantages and disadvantages of different image generation methods

数据生成类别	代表方法	优点	缺点
人工合成	Photoshop，Photopea	素材真实使得图像逼真，有现有的工具	人工成本较高，耗时长
几何变换	旋转，镜像，平移，裁剪，缩放等	易操作，增加数据集的空间几何信息，模拟不同视角、采集角度	数据重复记忆，可能会改变图像原本语义特征
色彩变换	色域变换，亮度变换，对比度变换	易操作，增加数据集的色彩信息，模拟不同环境、天气场景	数据重复记忆，可能改变图像的重要颜色信息
模糊变换	高斯模糊，中值模糊，边缘锐化	模拟相机虚焦情况，可突出物体细节信息	滤波与卷积网络的原理一致，不能增加新的有效信息
噪声注入	高斯噪声，伽马噪声，均匀噪声，椒盐噪声	引入冗余和干扰信息模拟出不同成像质量的图像	对模型精度提升效果不明显
生成对抗网络	基于GAN模型的众多样本生成方法	可无限量生成样本	需要一定数量的训练样本，训练难度大；生成图像的质量难以保证
图像风格迁移	基于条件生成对抗网络，基于图像翻译保留图像内容并生成不同风格图像	可对相同内容不同模态的图像之间转换	构建2个不同领域的数据集才能训练风格迁移模型
基于强化学习	AutoAugment，RandAugment	组合现有的数据增广方法，搜索最优策略	策略搜索空间大，训练复杂度高，计算开销大
基于大模型文生图	基于扩散模型StableDiffusion通过文本生成图像	能根据文本提示词控制生成的图像内容和风格，图像质量高、细节丰富	依赖提示词的质量，生成奇怪、不符合预期的图像，不理解特定文化、场景的背景概念

价指标的量化值，但是该方案生成图像的特殊项点和异物的物体特征与实际采集图像存在畸变、色彩异常等差异问题，导致生产环境中的智能识别模型仍出现较多误报、漏报的情况[3]。

（3）研究基于图像翻译TPSeNCE模型(基于GAN思想)天气场景的图像迁移技术[5]，将采集到的晴天图像与雨、雪天气的图像进行相互转化。输入晴天图像能批量生成模拟的雨、雪效果的图像，但是特殊货物和异物的正样本特征图像没有增加。

（4）使用基于对抗生成网络(GAN)的PGGAN模型生成的货物和项点问题图像存在畸形的情况，难以作为目标检测的训练数据[6]。

（5）基于Stable Diffusion大模型[7]依据提示文本和基准图像的图生图、局部重绘等生成方法。实验证实，基于扩散模型及其改进的模型，难以一次性生成符合预期的特殊货物项点目标内容，需要大量时间调优提示词和参数信息。Stable Diffusion局部重绘轮式、履带式货物图像如图2所示，使用局部重绘模式，使用“白色画笔”标识生成特殊货物的位置，以空平车为基准图像，给定例如“添加旋转的火箭炮”的正向提示词和负向提示词，使用多种采样器的Stable Diffusion1.5模型生成的图像与实际轮式、履带式货物特征差异较大，将该方法生成图像用作训练集会误导目标检测模型，导致识别效果的降低。

2　标准图像生成模型构建

对上述图像生成方法通过实验进行比较，发现以上方法着重于从数量上对整张图像进行扩充，对货物和项点特征只能手动调整，与快速生成多样化、逼真的标准图像需求之间存在较大差距。因此，为了自动化、智能化地生成更多有效的正样本图像数据，以分割模型提取特殊货物和特殊项点作为前景目标库为基础，以启发式地将已有目标置换到新的货车背景图像为思路，通过项点分割提取、启发式放置、生成-对抗约束等多模块融合，构建基于场景约束的图像生成模型，能够快速、有效地增加正样本训练集，提升智能识别模型的泛化能力。基于场景约束的铁路货车装载状态图像样本生成模型如图3所示，下面对各个模块逐一展开陈述。

2.1　特殊货物及特殊项点前景库提取模块

该模块首先从海量采集的货检图像中，筛选出包含特殊货物和特殊项点的图像，然后对货物及项点目标进行标注，使用YOLOv5-seg实例分割模型对数据学习，获得能快速从已有数据集中提取大量的特殊货物和特殊项点前景目标的模型，即将原图中的特殊货物或异物的目标区域精确地分离出来。为了避免人工花大量的时间以多边形的方式逐点、精细地标注目标对象的轮廓区域，搭建了一套自研的、基于SAM模型[8]的辅助标注软件R-ISAT(Railway-Image Segmentation Annotation Tool)。特殊货物图像目标标注如图4所示，R-ISAT软件可选择使用模型标注模式，通过鼠标交互式点击轮式、履带式货物的区域(红色区域)，以快速获取目标对象的分割区域的掩码、边界坐标和项点类别信息，以及能够快速修改超过、少于项点目标对象的局部分割范围，自动将信息保存于txt文件中，由此快速收集特殊货物和异物目标对象前景数据集。该模块使用



图2　Stable Diffusion局部重绘轮式、履带式货物图像

Fig.2　Images of wheeled and tracked cargo locally reconstructed using Stable Diffusion

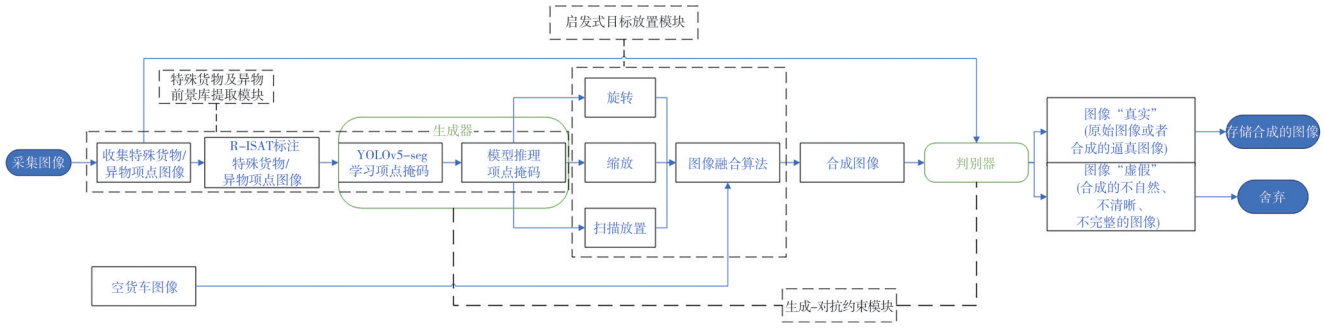


图3　基于场景约束的铁路货车装载状态图像样本生成模型

Fig.3　Scene-constrained image sample generation model for railway wagon loading status

YOLOv5-seg模型，模型的主干网络为CSPDarknet53，包含CSP结构、Focus结构和FPN多尺度特征融合技术[9]，经主干网络学习提取前景目标的特征后，再经过上采样(通过双线性插值法逐步扩大特征图尺寸)将掩码分辨率调整为与输入图像一致，避免图像信息的丢失。

2.2　生成-对抗约束模块

直接使用基于生成对抗网络及其优化算法所生成的图像，具有低分辨率、不真实、可用性较差的问题，转而考虑将生成-对抗的思路应用于本文框架，以约束和提升图像的真实性[10]。

在生成对抗网络中，生成器G以训练出能准确、自然地识别并分割出前景物体为目标，向模型分配低损失Loss(代表较少的惩罚)；反之，对于提取前景目标不真实的分割模型分配高损失Loss，该模块使用损失函数调优模型架构，以使得生成器G在前景提取模块生成能够欺骗判别器D的完整、细致、精确的掩码M。公式(1)表示前景提取模块所合成的新图像I，将生成的图像I输入到判别器D中。

$$I=M\times P_B+(k-M)\times P'_B \tag{1}$$

式中：M和B分别为获取的目标物体的掩码和边界框；P_B为原图像中前景边界框的区域；P'_B为新生成图像中前景边界框的区域；为了描述原图像和新合成图像之间的亮度、结构、像素信息的差异，优化使用结构相似性指数(SSIM)计算2张图像差异，用参数k表示。

判别器由卷积层、全连接层和Softmax激活函数组成，期望训练判别器能够正确区分图像是“虚假的”(合成效果不自然、不清晰、不正常的图像)或者是“真实的”(原始图像和合成逼真的图像)，以损失函数对齐“真实”图像和“虚假”图像的对应目标，其损失函数如公式(2)所示。

$$L_{CP}=\mathrm{E}[\log D(P_B)+\log(1-D(I))] \tag{2}$$

式中：$\mathrm{E}[\log D(P_B)]$为判别器正确识别出“真实”图像的期望；$\mathrm{E}[\log(1-D(I))]$为判别器正确识别出“虚假”图像的期望。

期望判别器能够尽可能最大化损失，生成器尽可

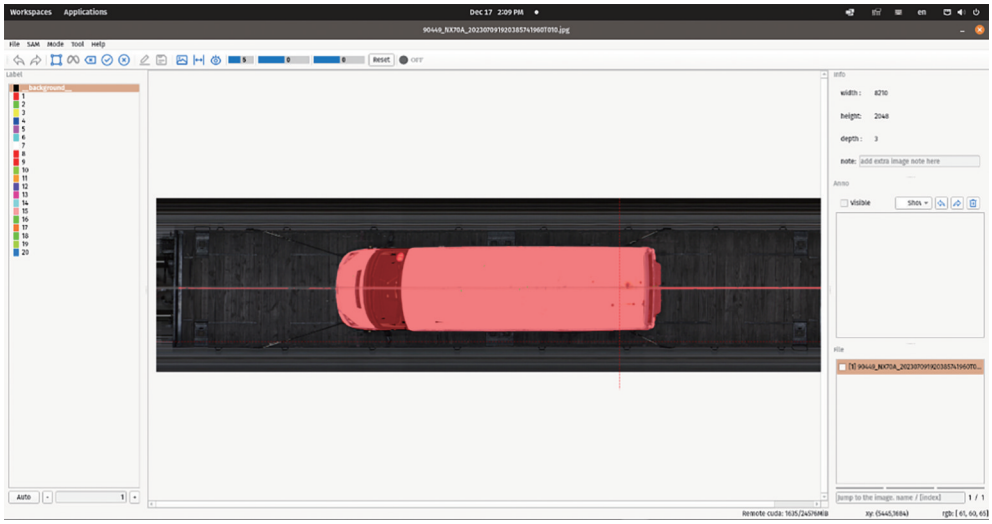


图4　特殊货物图像目标标注

Fig.4　Object annotation for special cargo images

能最小化损失。随着生成器和判别器的约束平衡，当生成器能够精准地分割出特殊货物和特殊项点对象的掩码M时，所生成的图像越逼真、自然。

但是，生成对抗网络存在退化问题，即模型可能未真正学习到准确的特殊货物和异物前景目标的分割信息，体现在判别器判别图像“真实”，合成的图像却不包含特殊货物或项点目标对象，或只包含部分目标对象2种问题。为解决以上2个问题，添加了一个额外的损失如公式(3)所示，以及一个分割损失L_{CUT}，其损失函数如公式(4)所示。

$$L_{CLS}=\mathrm{E}(\log(1-D_{CLS}(I))) \tag{3}$$

式中：L_{CLS}为分类损失；D_{CLS}为额外的分类网络。

$$L_{CUT}=E[\log(D(1-M)P_B)+\log(1-D((1-M)P_{B'}))] \tag{4}$$

式中：$(1-M)P_B$为从原图像移除前景目标对象后剩下的图像部分(理想情况是不包含目标对象的干净背景)；$(1-M)P_{B'}$为模拟在不包含对象区域进行同样的切割操作的图像部分。

如果出现只包含部分目标对象的问题，那么再次做切割会使得判别器认定剩下的2张图像相似，否则会有明显差异，以促使生成器获取完整的前景目标。将以上3种损失函数相加，并使用Hinge损失函数分别对生成器和判别器进行优化。

2.3　启发式目标放置模块

构建一套适应于特殊货物和特殊项点目标的启发式目标放置方法，分别选择不同目标对象在背景图像中的粘贴位置，依托生成－对抗约束网络确保目标分布合理、自然。

对特殊货物掩码，做以30度为旋转步长、旋转范围为[－90，90]度的旋转；依据不同设备厂家、车站的镜头的位置和设备情况，计算符合实际的缩放尺度范围，对特殊货物做以0.2为乘积参数、缩放比例范围为[3，7]，即目标缩放比例为0.6，0.8，1.0，1.2，1.4的缩放；在图像的中心点位置以标准偏差W_{box}进行随机偏移的(移动距离会有W_{box}范围内的随机变化)水平放置，W_{box}为掩码边界框宽度，在图像的中心点位置以标准偏差H_{box}进行随机偏移的垂直放置，H_{box}为掩码边界框高度。每张图像放置1～2个特殊货物掩码，背景图像选用空平车图像。

对于特殊项点掩码，做以30度为旋转步长，旋转范围为[0,90]度的旋转；依据多种车型和实际特殊项点货物的比例尺寸信息，计算符合实际的缩放尺度范围，对特殊项点目标做以2为乘积参数、缩放比例范围为[1,4]的缩放，在背景图像中以平均间隔$5\times W'_{box}$的不重叠的水平扫描移动、允许在标准偏差W'_{box}的随机偏移，W'_{box}为掩码边界框宽度，以垂直方向平均间隔$2\times H'_{box}$的不重叠的垂直方向扫描移动、允许在标准偏差H'_{box}的随机偏移，H'_{box}为掩码边界框高度。在每张背景图像中放置数量范围为[2,8]的特殊项点掩码，背景图像选用空平车、空敞车、罐车、漏斗车等无项点问题货车图像。

特殊项点图像生成流程如图5所示，由包含闸瓦的特殊项点的图像，使用特殊项点前景库提取模块提取特殊项点目标及其掩码；依据启发式目标放置模块的规则，生成多张图像，在第一张生成的图像中，随机放置2个闸瓦，其中，第一个闸瓦缩放比例为2，旋转角度为0度，第二个闸瓦缩放比例为4，旋转角度为顺时针90度，按照规则在水平和垂直方向上放置闸瓦；基于生成－对抗约束模块将符合图像放置相对位置，依据生成图像的真实性、自然效果的图像进行自动化筛选，筛选图5中的图a为可用的图像样本数据。在特殊货物和特殊项点的前景目标确定位置、变换信息后，使用泊松图像编辑算法[11]，将前景目标图像的特征与背景图像的梯度进行融合，能有效避免直接粘贴造成的边界、光线、颜色过渡不自然问题。

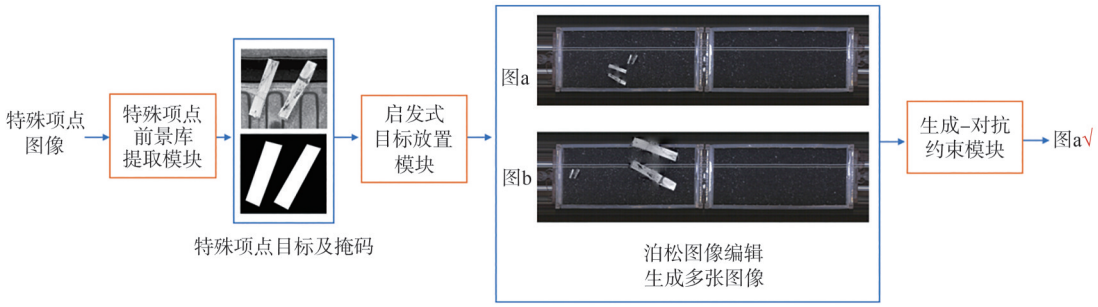


图5　特殊项点图像生成流程

Fig.5　Generation process for image at special inspection items

3　实验分析

3.1　实验数据集

图像生成模型的数据集来源于中国铁路太原局集团有限公司、中国铁路西安局集团有限公司、中国铁路南宁局集团有限公司等9个铁路局集团公司的包含了全天不同时间段和不同天气条件下的标准图像数据。实验环境中，特殊异物的图像选用的是近一年的数据，特殊项点的图像选用近一个月的数据，标注前景目标对象以构建数据集。视频监控设备采集到特殊货物图像，共计138张图，特殊货物实例掩码共141个。特殊货物原始图像统计数量如表3所示。

依据视频监控设备采集到特殊项点标准图像，依据特征主要分为21种类型，共计5 085张图像，异物实例掩码7 141个。特殊项点原始图像统计数量如表4所示。

3.2　评价指标

为了衡量生成图像的有效性和可用性，从生成图像本身的质量，以及生成图像对目标检测模型训练效果的提升效果2个角度进行分析。

生成图像的客观指标选用IS(Inception Score)，该标准常用于评价基于生成对抗网络所生成的图片质量[12]。IS指标将Inception Net－v3网络作为工具，从图像的清晰度和多样性两方面进行评估。清晰度的原理是依据熵(代表图像信息的混乱度)以及图像分类(通过预训练的Inception Net－v3网络提取高级语义特征，该网络接收图像张量作为输入，输出1 000维的概率向量，每一维度的值表示图片属于分类类别的后验概率，从而获得精确的类别概率分布)，生成的图像越清晰，分类正确的几率越高，则类别条件概率分布越高。如果一种方法生成足够多样的图像样本，生成图像在各类别分布应当是平均分布的，在所有类别的边缘分布概率较大。对于每张图像，计算其条件分布和边缘分布之间的KL散度，对所有图像的KL散度取指数平均，IS计算方法如公式(5)所示。

$$IS(img)=\exp\left(\frac{1}{N}\sum_{i=1}^{N}D_{KL}(p(y|x)\|p(y))\right) \tag{5}$$

式中：$IS(img)$为一组生成图像集合的IS分数；N为所有图像的数量；D_{KL}为KL散度；$p(y|x)$为每张图像的条件分布；$p(y)$为对应的边缘分布。

由于货车装载状态标准图像智能识别模型是基于YOLOv5目标检测模型实现，则将生成的图像作为增量的数据集，以目标检测模型的评价指标F1-score直接衡量生成图像的有效性，$F1$计算公式如公式(6)所示。

$$F1=\frac{2\times(precision\times recall)}{precision+recall} \tag{6}$$

式中：$F1$为精确率和召回率的调和平均数，能综合、全面地评估误报和漏报情况；$recall$为召回率，为正确预测的正样本数量占所有正样本的数量比例；$precision$为精确率，表示模型预测的正样本数量占模型所有预测数量的比例。

3.3　结果分析

将基于启发式－生成对抗网络方法，与人工生成方法、基于图像增强方法、基于图像迁移方法共4种方法生成的图像与原图比较，不同图像生成方法效果比较如图6所示。图像增强和图像迁移方法，仅对图像做像素级变换，以及对图像全局特性做调整，无法生成新的正样本。相比于手动合成特殊项点目标与背景图像融合效果不佳，本算法生成的图像较为自然、逼真。对所生成的图像质量和实际应用于标准图像智能识别

表3　特殊货物原始图像统计数量

Tab.3　Statistical results of original images of special cargo

特殊货物类型	轮式、履带式货物	预应力梁	设备类货物（不包含卷钢的球形、柱形货物）
顶部图像数量/张	120	6	12

表4　特殊项点原始图像统计数量

Tab.4　Statistical Count of Original Images for Special Inspection Points

异物类型	石块	草垫	泡沫胶	稻草	桶	长条板	盖子
图像数量/张	135	246	281	231	27	105	32
异物类型	扒乘人员	垃圾袋	空袋子	纸盒	瓶子	棍状物	破损
图像数量/张	314	400	321	334	508	364	87
异物类型	闸瓦	堆状残货	纸片	槽帮残货	铁丝/绳子	饭盒	木块
图像数量/张	51	415	369	352	258	193	62



图6　不同图像生成方法效果比较

Fig.6　Comparison of different image generation methods

项点目标检测模型的提升效果两方面进行比较，将不同方法生成的数据抽取相同数量的图像，与原始图像合并、划分为训练和验证数据集，输入到同样的基于YOLOv5 m的目标检测模型中，模型使用的参数和预训练模型保持一致。将生成100张图像的时间取平均，计算生成单张图像所需的时间(包含推理、生成、用于目标检测的标注数据的时间)。

特殊货物各种方法的结果比较如表5所示。

特殊项点各种方法的结果比较如表6所示。

由以上结论可以看出，基于本研究提出的图像生成方法，所生成的特殊货物和特殊项点的图像质量最好、样式最多样，且相比于其他方法，本研究方法生成的图像作为目标检测的新增数据集，明显提升了智能识别效果，推理、合成一张图像的时间也较短，表明该算法架构能够在短时间生成较为有效、多样的图像数据。其中，对于原始样本数量极少的特殊货物数据集，本方法生成的图像能显著提升模型识别效果，在实际生产环境中，使用重新选取新丰镇一个月的数据，特殊货物的识别精确率由52%提升至63%；对于已有一定样本数量的特殊项点图像，本方法扩充的样本对模型识别的数值提升较小，但是由于该方法扩增了特殊项点物体的位移、尺寸、旋转角度等样本特征，在实际生产环境中，由于特殊项点的样本较多，重新选取新丰镇一天的数据，特殊项点的精确率由84%提升至93%，误报和漏报情况有所减少。

本研究的图像生成方法的时间成本，主要在于前期的掩码标注和生成算法的训练学习。借助R-ISAT工具的帮助，标注一张图像前景目标的分割掩码时间能压缩至平均7 s。YOLOv5学习掩码特征、训练轮数约180轮数据，需要20 h；GAN网络部分训练大约280轮，需要大约34 h，排除前期系统研发时间，共计54 h。在推理和生成阶段，平均每张图像推理出前景目标约45 s。相

表5　特殊货物各种方法的结果比较

Tab.5　Comparison of results for each method on special cargo

方法评价指标	原始数据	本研究方法	人工生成	图像增强方法	图像迁移方法
IS	6.1	6.8	6.5	6.1	6.3
*F*1	0.56	0.81	0.79	0.74	0.69
单张推理时间/s	—	45	252	15	40

表6　特殊项点各种方法的结果比较

Tab.6　Comparison of results for each method on special inspection items

方法评价指标	原始数据	本研究方法	人工生成	图像增强方法	图像迁移方法
IS	6.7	7.8	7.7	6.6	6.8
*F*1	0.92	0.98	0.96	0.95	0.92
单张推理时间/s	—	46	381	17	41

比而言，传统合成方法不需要模型学习的时间，图像迁移的模型学习时间大约为37 h。因此，虽然研究提出的图像生成算法存在较多开发程序和训练模型的时间，但相比于人工合成方案，生成图像环节节省了82%以上的时间；相比于其他生成图像的算法速度更快，对智能识别模型有更多的效果提升，说明所提出的方法能有效解决由于人力不足以及技术人员疲劳因素导致的生成图像不一致及正样本图像少的问题，能高效生成高质量、多样性和有工程应用价值的图像。

4　结束语

铁路货运正加快向信息化、智能化、智慧化转型升级，对发挥好技防保安全作用提出了更高要求。铁路货车装载状态图像数量与质量，直接决定了图像智能识别准确率。通过铁路货车装载状态图像生成技术应用，能够解决因正类样本不足造成识别模型泛化能力差、识别准确率低的问题，提升铁路货车装载状态图像智能识别能力和识别效果。后续将进一步深化跨站图像比对识别、多源数据融合识别等技术难题，为推动铁路货检智能化提供技术支撑。

参考文献：

[1] 石　磊. 铁路货车装载状态标准图像智能识别技术体系设计研究[J]. 铁道货运，2022，40(11)：34−39，48.
SHI Lei. Design of Intelligent Recognition Technology System for Standard Images of Wagon Loading Status[J]. Railway Freight Transport，2022，40(11)：34−39，48.

[2] 石　磊. 关于铁路货检作业优化对策的探讨[J]. 铁道货运，2018，36(10)：33−36.
SHI Lei. A Study on the Strategies to Optimize Railway Cargo Inspection Operation[J]. Railway Freight Transport，2018，36(10)：33−36.

[3] 马崇昇，唐　娉，赵理君，等. 深度学习图像数据增广方法研究综述[J]. 中国图象图形学报，2021，26(3)：487−502.
MA Dongao，TANG Ping，ZHAO Lijun，et al. Review of Data Augmentation for Image in Deep Learning[J]. Journal of Image and Graphics，2021，26(3)：487−502.

[4] 刘朝辉，杨　杰，陈智超. 基于深度学习的轨道表面异物识别方法[J]. 中国铁道科学，2023，44(3)：23−33.
LIU Zhaohui，YANG Jie，CHEN Zhichao. Foreign Object Recognition Method for Track Surface Based on Deep Learning[J]. China Railway Science，2023，44(3)：23−33.

[5] ZHENG S，LU C J，NARASIMHAN S G. TPSeNCE：Towards Artifact−Free Realistic Rain Generation for Deraining and Object Detection in Rain[C]//2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV).Waikoloa：IEEE，2024：5382−5391.

[6] 肖子轩，刘启钢，叶　飞，等. 基于Dropout−PGGAN的铁路货检图像数据增强方法研究[J]. 铁道运输与经济，2024，46(6)：108−117.
XIAO Zixuan，LIU Qigang，YE Fei，et al. Data Enhancement Method for Railway Freight Inspection Images Based on Dropout−PGGAN[J]. Railway Transport and Economy，2024，46(6)：108−117.

[7] 刘泽润，尹宇飞，薛文灏，等. 基于扩散模型的条件引导图像生成综述[J]. 浙江大学学报(理学版)，2023，50(6)：651−667.
LIU Zerun，YIN Yufei，XUE Wenhao，et al. A Review of Conditional Image Generation Based on Diffusion Models[J]. Journal of Zhejiang University (Science Edition)，2023，50(6)：651−667.

[8] KIRILLOV A，MINTUN E，RAVI N，et al.Segment Anything[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris：IEEE，2023：3992−4003.

[9] 马冬梅，郭智浩，罗晓芸. 改进YOLOv5s−Seg的高效实时实例分割模型[J]. 计算机工程与应用，2024，60(16)：258−268.
MA Dongmei，GUO Zhihao，LUO Xiaoyun. Improved Efficient Real−Time Instance Segmentation Model Based on YOLOv5s−Seg[J]. Computer Engineering and Applications，2024，60(16)：258−268.

[10] REMEZ T，HUANG J，BROWN M. Learning to Segment Via Cut-and-paste. In Proceedings of the European Conference on Computer Vision (ECCV)[EB/OL].[2024−10−08]. https://openaccess.thecvf.com/content_ECCV_2018/html/Tal_Remez_Learning_to_Segment_ECCV_2018_paper.html.

[11] PÉREZ P，GANGNET M，BLAKE A. Poisson Image Editing[J]. ACM Transactions on Graphics，2003，22(3)：313−318.

[12] SZEGEDY C，VANHOUCKE V，IOFFE S，et al. Rethinking the Inception Architecture for Computer Vision[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas：IEEE，2016：2818−2826.

（责任编辑：金　颖）

Scene-Constrained Image Sample Generation Technology for Railway Wagon Loading Status

GE Yue[1], WANG Zhijing[2], KE Xiangxi[3], SUN Wenqiao[1], WANG Xiaowei[4]

(1. Transportation & Economics Research Institute, China Academy of Railway Sciences Corporation Limited, Beijing 100081, China; 2. Beijing China Railway Sciences Passenger and Freight Transportation Technology Co., Ltd., China Academy of Railway Sciences Corporation Limited, Beijing 100081, China; 3. Liuzhou Railway Station, China Railway Nanning Group Co., Ltd., Liuzhou 545001, Guangxi, China; 4. Passenger Transport Department, China Railway Kunming Group Co., Ltd., Kunming 650011, Yunnan, China)

Abstract: Railway wagon loading status images are critical carriers of information reflecting the operational status of freight wagons, and the intelligence of these images directly impacts both train operation and cargo safety. Due to the impacts of cargo types, wagon types, and methods for loading safety, certain scenarios suffer from a scarcity of wagon loading status images, hindering the construction of models for object detection at inspection items in these contexts. Considering the characteristics of degraded railway wagon loading status images, this study proposed a multi-technology fusion model for generating railway wagon loading status images to satisfy the requirements of railway freight inspection, integrating several methods with a generative adversarial network (GAN) and instance segmentation. Specifically, the model incorporated generative and adversarial-based constraints, an object extraction method of foreground items, and a heuristic method for object placement at inspection items. This approach enriched the railway wagon loading status image database for special scenarios and provided technical support for constructing reliable and intelligent detection models.

Keywords: Railway Freight; Cargo Inspection; Wagon Loading Status; Image Generation; Object Detection; Generative Adversarial Network