

文章编号 1004-924X(2023)16-2406-12

联合 3D 建模与改进 CycleGAN 的故障数据集 扩增方法

李宝平, 戚恒熠*, 王满利, 魏 坡

(河南理工大学 物理与电子信息学院, 河南 焦作 454000)

摘要: 基于深度学习的设备故障检测系统性能很大程度上依赖于样本集的规模及类别多样性。由于工业生产中难以全面采集各类故障样本, 由此就有样本集扩增需求。本文提出联合 3D 模型和改进 CycleGAN 的故障数据集扩增方法。首先, 提出利用 3D 建模软件模拟生成各类故障图片, 将其作为 CycleGAN 迁移网络输入, 约束引导生成真实故障图像, 以解决样本不足及分布不均衡问题; 其次, 对 CycleGAN 网络生成器进行改进, 提出 U-ResNet 生成器, 用以解决数据集扩增过程中的边缘模糊和梯度消失问题。将该方法应用于带式输送机跑偏检测任务, 结果表明相较于其他扩增方法, 该方法训练过程中轮廓结构收敛快, 时效性好, 应用于目标检测网络准确率达到 98.1%, 较原真实数据集提升 4.5%。说明该数据集扩增方法可以满足故障数据集类别分布均衡, 图像质量高的要求。

关键词: 数据集扩增; 3D 模型; CycleGAN; U-ResNet; 带式输送机

中图分类号: TP391.4 **文献标识码:** A **doi:** 10.37188/OPE.20233116.2406

Equipment fault dataset amplification method combine 3D model with improved CycleGAN

LI Baoping, QI Hengyi*, WANG Manli, WEI Po

(College of Physics and Electronic Information, Henan Polytechnic University,
Jiaozuo 454000, China)

* Corresponding author, E-mail: mysteriq@home.hpu.edu.cn

Abstract: The performance of deep-learning-based equipment fault detection systems relies heavily on the size and class diversity of the sample set. Because it is difficult to collect all types of fault sample comprehensively in industrial production, there is a demand for sample set augmentation. A fault dataset amplification method combining 3D modeling with an improved cycle generative adversarial network (CycleGAN) is proposed. First, various equipment malfunction images generated by 3D modeling software are applied to the CycleGAN network training to guide it in generating pseudo-real images to address the problem of insufficient samples and an uneven distribution. Second, a U-ResNet generator is used in the CycleGAN network to solve the problem of edge blurring and gradient vanishing during network training. The method was applied to the task of belt conveyor deviation detection. The experimental results show that the contour structure of the method converges quickly in the training process and has good timeliness in

收稿日期: 2022-11-16; 修订日期: 2023-01-28.

基金项目: 国家自然科学基金资助项目 (No. 52074305); 河南理工大学博士基金资助项目 (No. B2021-64)

comparison with other amplification methods. The accuracy rate of the method is 98.1% when applying to the target detection network, which is 4.5% higher than that of the original real dataset. It meets the basic requirements of a balanced distribution of amplified datasets and high image quality.

Key words: dataset amplification; 3D model; CycleGAN; U-ResNet; belt conveyor

1 引言

随着人工智能技术的发展与成熟,基于深度学习的工业设备故障检测技术逐渐成为研究的热点和发展趋势,其可代替人工实时检测故障,有效节省人力资源,在部署前需大量故障样本训练模型,故障样本的采集场景多为工业设备工作现场,受限较多,难以采集全面、多样的故障样本,样本分布不均衡导致故障检测的鲁棒性下降,故亟需通过数据集扩增方法解决设备故障样本缺乏问题^[1]。

现有的工业设备故障样本扩增方法多为传统方法和深度学习方法,传统方法主要是通过图像的旋转、裁剪、镜像、尺度变换等^[2]来实现,其并未实质解决故障样本缺乏和分布不均衡问题。现在基于深度学习的数据集扩增方法主要通过生成对抗网络(Generative Adversarial Network, GAN)^[2]实现,包含噪声生成^[3-4]与风格迁移^[5-6]两种类别。

生成对抗网络由 Goodfellow 等人于 2014 年提出^[7],由多维噪声通过生成器和判别器博弈生成图像,训练过程中存在模式崩塌,梯度消失、爆炸等问题。在此基础上,Radford 在 2015 年提出由深度卷积搭建的 DCGAN 网络^[8],解决了模式崩塌的问题,使用转置卷积替代上采样,取消池化层,使网络结构更稳定,但并未解决梯度消失问题。Arjovsky 等在 2017 年提出 WGAN^[9],使用 Wasserstein 距离代替 GAN 网络中 JS 距离,有效解决了梯度消失问题,并在此基础上提出 WGAN-GP^[10],使用梯度惩罚代替权重裁剪,取得更好的训练效果。上述基于噪声的图像生成方法应用于故障样本集扩增,虽然可以提升样本的规模,但其并未解决故障样本分布平衡问题,其扩增故障集种类分布服从于原故障集分布。

Zhu 等在 2017 年提出 CycleGAN^[11],可实现非配对的两组图像无监督风格迁移,其模型由两组生成器、判别器构成,通过引入循环一致损失

函数约束源域和目标域图像一一映射。风格迁移网络应用于工程数据集扩增,主要表现为三种形式:(1)图像增强方面,如水下图像恢复^[12],遥感图像去雾^[13],高分辨率图像重建^[14];(2)不同类型图像转换,如可见光图像向红外图像的转换^[15],光学图像向声呐图像的转换^[16];(3)场景融合,如火焰图像与不同背景图像的融合^[17],故障绝缘子场景向无故障绝缘子场景的迁移^[18]。将风格迁移方法应用于故障数据集扩增,存在复杂图像轮廓结构收敛慢,不稳定的问题,另外生成样本分布取决于另一风格样本集,鲁棒性受限。

本文提出联合 3D 模型与改进 CycleGAN 的故障数据集扩增方法可以较好地解决上述问题。首先通过 3D 动态模型模拟生成各种类别故障样本,引导迁移网络生成各种类别的真实故障样本,满足样本多样化的需求;其次改进的 CycleGAN 生成器可以加速轮廓结构收敛,同时避免梯度消失问题,提高训练实效性。我们以带式输送机跑偏故障检测为研究对象,开展故障数据集扩增方法研究。通过 3D 建模软件创建带式输送机跑偏故障动态模型,将采样后的模型样本输入 CycleGAN 网络,实现 3D 故障模型图像向真实跑偏图像的迁移。实验表明该方法相比于目前最先进的基于噪声的生成网络,扩增故障样本集类别分布更加均衡,满足多样性需求。将扩增数据集应用于目标检测网络,该方法训练的模型具有更高的检测精度,也验证了将该方法应用于设备故障数据集扩增合理可行。

2 改进的 CycleGAN 网络架构

2.1 整体架构

本文以无监督图像风格迁移网络 CycleGAN 为基础进行设计,实现从 3D 模型渲染图像到真实跑偏图像的风格迁移。网络架构如图 1 所示。

CycleGAN 中包含两个映射函数,生成器 G 和生成器 F^[12,15,17,19]。其中生成器 G 负责 X 域(模

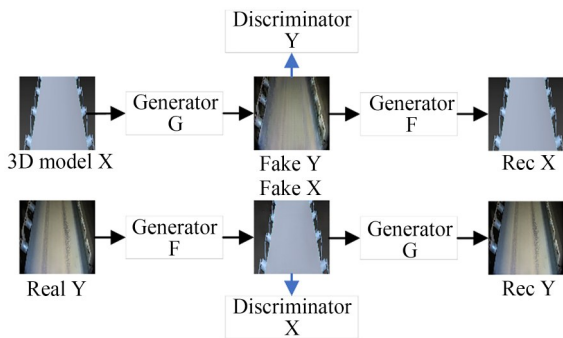


图1 CycleGAN整体架构

Fig. 1 Framework of CycleGAN

型图像)向Y域(真实图像)的图像生成,生成器F则负责Y域向X域的图像生成。如图1所示,X域的图像通过生成器G生成Fake Y,再通过生成器F重构回X域输入的原图像Rec X;Y域的图像同理。

2.2 生成器架构

CycleGAN网络在保留复杂图像轮廓方面存在收敛较慢且不稳定的缺点^[13,20],在生成图像中容易出现轮廓模糊、缺失、错位等情况。CycleGAN网络引入循环一致性损失后可在一定程度上

上约束图像的轮廓,减少源域到目标域的路径映射的可能性,提高轮廓收敛速度,由拉普拉斯金字塔^[21]可知,生成器在上采样、下采样过程中会不可逆的损失部分轮廓特征信息,导致生成图像出现轮廓模糊、缺失、错位等情况。U-net生成器使用跳连结构(Skip Connection)将下采样层与上采样层相连,下采样层提取到的特征可以直接传递到上采样层,有助于保持轮廓结构信息的完整性^[22],降低生成图像出现轮廓模糊、缺失、错位的概率。ResNet生成器直接将输入信息绕道传至输出,以保持信息的完整性,整个网络只需要学习输入、输出差别的那一部分,简化学习目标和难度,克服梯度消失问题,提高训练实效性^[23]。由此我们考虑将U-net与ResNet结合构造生成器,将有助于解决上述问题。改进的U-ResNet生成器架构如图2所示,包括编码器、解码器两部分,图3为改进生成器中卷积神经网络实现风格迁移过程的特征流程图。

与U-net生成器相比,U-ResNet把左侧编码器中的升维卷积块替换为残差块,在图像下采样后可以保留更多的特征信息;同时U-ResNet相

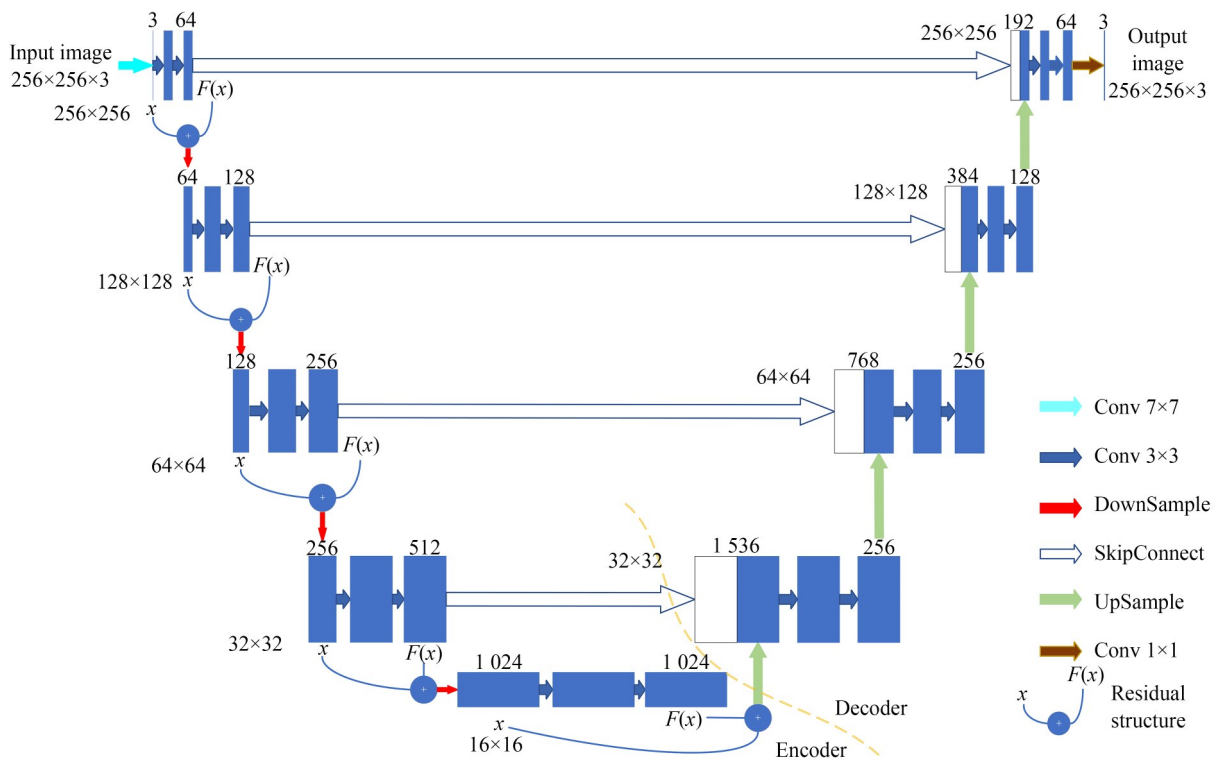


图2 改进的生成器架构

Fig. 2 Framework of improved generator

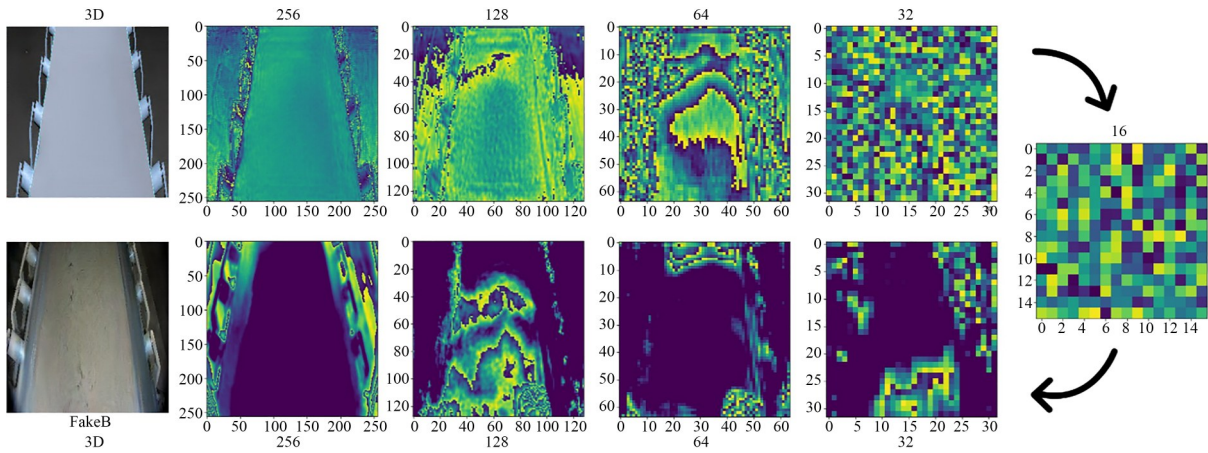


图 3 生成器卷积层特征图

Fig. 3 Feature map of improved generator

比于完全由残差块组成的生成器,每次上采样后会通过跳连结构连接对等下采样层,用以补全采样过程中丢失的轮廓特征信息。因此,U-ResNet 生成器架构相较于 U-net 和 Resnet 架构,训练过程中可以保留更多轮廓特征信息,生成图像轮廓收敛更快,具有更好的实效性。

ResNet,U-net 及 U-ResNet 生成器上下采样次数为 1 时的简化架构如图 4 所示。用符号 d 表示下采样损失系数,符号 u 表示上采样损失系数, $d, u \in [0, 1], f_n$ 表示卷积,输入轮廓特征为 x ,输出轮廓特征为 y 。

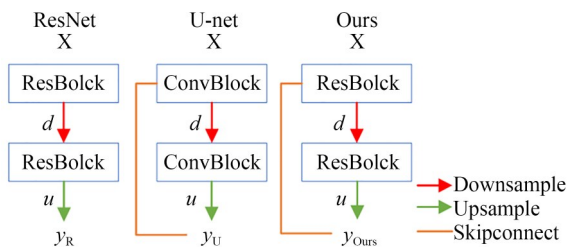


图 4 几种生成器简化架构

Fig. 4 Several generator simplification architectures

ResNet 上采样后输出轮廓特征表达式为:

$$y_R = uf_2(d(f_1(x) + x)) + ud(f_1(x) + x). \quad (1)$$

U-net 上采样后输出轮廓特征表达式为:

$$y_U = uf_2(d(f_1(x))) + f_1(x). \quad (2)$$

U-ResNet 上采样后输出轮廓特征表达式为:

$$y_{Ours} = uf_2(d(f_1(x) + x)) + ud(f_1(x) + x) + f_1(x) + x. \quad (3)$$

图 5 为各个生成器首个 epoch 训练时上采样后输出的轮廓特征图,式(3)分别与式(1)、式(2)做差即可得到跳连结构与残差结构保留的特征信息,图 5 体现出了更多的轮廓特征信息,对比式(1)~式(3)并结合图 3~图 5,U-ResNet 架构可以保留更多轮廓特征。由多层卷积神经网络特性可知随着上、下采样次数增加,U-ResNet 保留的轮廓特征远超 ResNet 和 U-net。

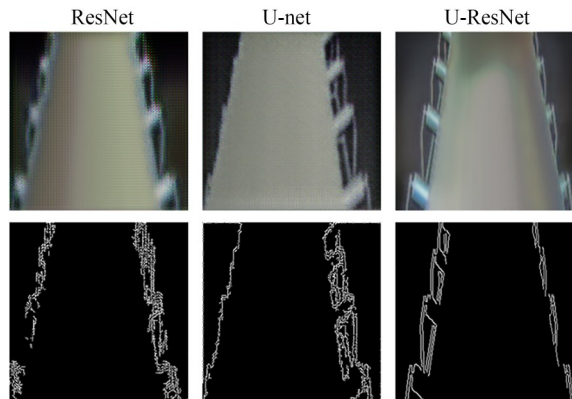


图 5 轮廓特征图

Fig. 5 Contour feature images

2.3 判别器架构

本文采用的判别器结构如图 6 所示,该判别器为马尔可夫判别器(Markovian discriminator)^[24]。与普通判别器不同,普通判别器直接将输入映射为一个实数,表示判别为真的概率;该判别器将

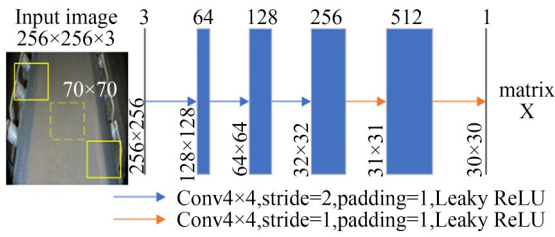


图 6 判别器架构

Fig. 6 Framework of discriminator

输入按感受域(图像块 patch)映射为一个 $N \times N$ 矩阵 X , X 代表卷积层输出的特征图, X_{ij} 表示单

个 patch 为真的概率, 矩阵 X 的均值作为判别器最终输出。在本文 CycleGAN 网络结构中, 输出 X 大小为 30×30 , 对应的感受域大小为 70×70 。

2.4 损失函数

CycleGAN 包含 3 种损失函数, 如图 7 所示, 分别是 Real Y 与 Fake Y 之间的对抗损失 (Adversarial Loss), 3D model X 与 Rec X 之间的循环一致性损失 (Cycle-consistency Loss), 以及 Real Y 与 Fake Y 之间的认同损失 (Identity Loss)。通过这 3 种损失函数约束图像实现风格迁移。

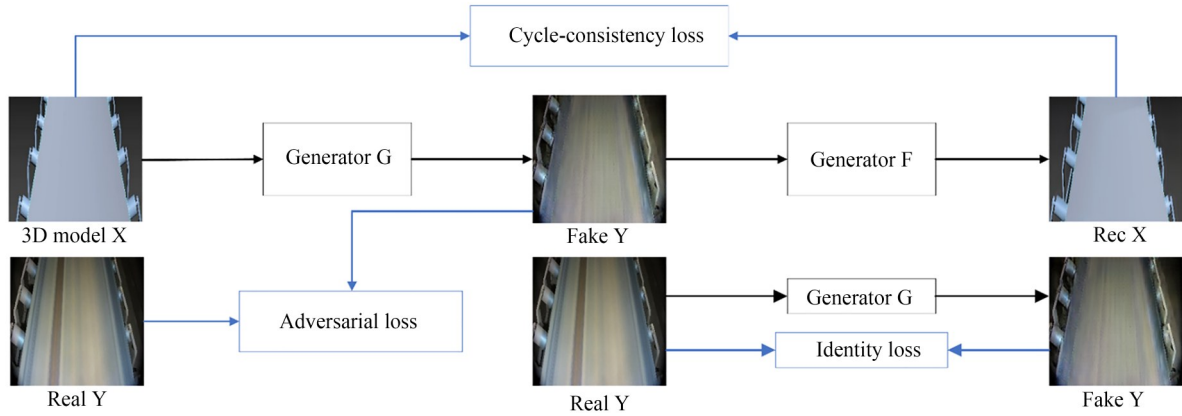


图 7 损失函数

Fig. 7 Loss function

2.4.1 对抗损失

生成器 G 依据 X 域的 3D 模型数据集和 Y 域的真实数据集生成 Fake Y, 目标是要使判别器认为 Fake Y 为真; 而判别器要与之对抗分辨出 Fake Y 为假, 如二人博弈, 迭代过程中的对抗损失可表示为:

$$L_{GAN}(G, D_Y, X, Y) = E_{y \sim P_{\text{data}}(y)} [\log D_Y(y)] + E_{x \sim P_{\text{data}}(x)} [1 - \log D_Y(G(x))], \quad (4)$$

其中: x 表示来自 3D 建模数据集的样本; y 表示来自真实跑偏数据集的样本; $G(x)$ 表示生成的样本; D_Y 是判别器, 用于将生成的图像与真实跑偏图像区分开来。生成器 G 的目的是使损失函数值最小化, 而判别器 D_Y 的目的则是最大化损失函数值, 用公式表示为:

$$\min_G \max_{D_Y} L_{GAN}(G, D_Y, X, Y). \quad (5)$$

2.4.2 循环一致性损失

将样本 x 从 X 域通过生成器 G 映射到 Y 域,

再通过重构器 F 映射回 X 域的过程称为重构, 重构的图像与样本应高度相似, 即 $F(G(x)) \approx x$ 。因此, 循环一致性损失可以定义为:

$$L_{\text{cyc}}(G, F) = E_{x \sim P_{\text{data}}(x)} [\|F(G(x)) - x\|_1]. \quad (6)$$

2.4.3 认同损失

为保证生成器 G 生成的图像是向 Y 域进行风格迁移, 将 Y 域真实跑偏图像输入到生成器 G 中, G 所生成的图像应尽可能还是它本身, 而不做其他转换, 如图 7 所示。因此 y 与 $G(y)$ 的映射损失可限制生成器 G , 表示如下:

$$L_{\text{identity}}(G, F) = E_{y \sim P_{\text{data}}(y)} [\|G(y) - y\|_1]. \quad (7)$$

2.4.4 总损失

综上, 整个 CycleGAN 的损失是由对抗损失、循环一致性损失以及认同损失综合得到^[11], 其中对抗损失计算相对熵, 循环一致损失和认同损失计算平均绝对误差。考虑机器学习基于梯

度下降和逻辑回归的思想,其最优解可以看作各层提取到的所有特征值的线性组合,因此目前损失函数大都定义为各类损失函数的线性组合。总损失函数表示如下:

$$L(G, F, D_X, D_Y) = L_{\text{GAN}}(G, D_Y, X, Y) + \alpha L_{\text{cyc}}(G, F) + \beta L_{\text{identity}}(G, F). \quad (8)$$

如果总损失不加以系数控制,损失之间直接相加可能会导致多任务的学习被某个任务所主导或学偏,当模型倾向于去拟合某个任务时,其他任务的效果往往可能受到负面影响,效果变差,因此需要在损失前添加 α, β 权重系数来平衡学习,通过实验仿真合理选择设置 α, β 。整体优化目标函数表示为:

$$G^*, F^* = \arg \min_{G, F} \max_{D_X, D_Y} L(G, F, D_X, D_Y). \quad (9)$$

3 实验设计

为检验扩增数据集的图像质量和样本均衡性,通过实验对多种数据集扩增方法进行对比分析,包括基于 3D 模型和迁移网络的多种数据集扩增方法,以及以 WGAN-GP 为代表的多种基于噪声的数据集扩增方法。为分析 CycleGAN 网络中生成器架构对轮廓结构收敛速度的影响,通过实验对 U-net, ResNet 和 U-ResNet (Ours) 三种生成器架构进行对比;为分析扩增数据集应用到检测网络中的效果,通过实验对基于 3 种生成器的 CycleGAN 扩增数据集及 WGAN-GP(目前最优噪声生成方法)扩增数据集进行对比分析。

实验平台采用 win10 系统,搭载 PyTorch 框架,处理器、显卡分别为 11th Gen Intel(R)Core (TM) i7-11800H @ 2.30 GHz, 16 G 内存, RTX 3060laptop。

3.1 实验样本集

真实跑偏样本集包含 900 张图像,包含左偏、右偏两种类型,右偏占比为 80%,样本图像是在矿井下由固定摄像机实拍的带式输送机跑偏图像。3D 建模样本集包含 1 000 张图像,包含左偏、右偏两种类型,每张图像跑偏程度不一,整体分布平衡,3D 图像是由 3Dmax2020 软件建模渲染得到,过程如下:

首先依据带式输送机物理结构建立 3D 模型

框架如图 8,其次架设物理摄像机模拟故障检测视角,制作皮带左右偏移循环动画模拟故障种类和程度如图 9,最后渲染出故障视频,经抽帧与图像处理得到 3D 建模样本集,两个样本集均由 256×256 的彩色图像组成。

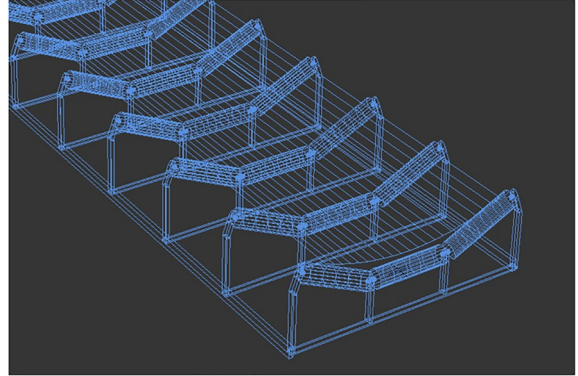


图 8 3D 建模框架

Fig. 8 Framework of 3D modeling

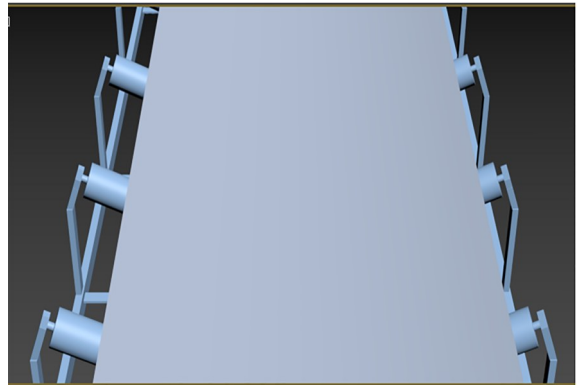


图 9 摄像机视角

Fig. 9 Camera perspective

3.2 训练配置

在 CycleGAN 网络总损失函数中,认同损失系数 β 预设为 5,其限制生成器风格迁移方向,如该生成器方向为 3D 到真实,则输入真实图像,输出也应为真实图像;循环一致性损失系数 α 预设为 10,表示循环一致损失的权重值,即原图与重构图像之间差异的权重,同时约束源域与目标域之间映射关系;对抗损失系数默认为 1。三者相辅相成,若训练中某个损失难以收敛或收敛较慢可适当增加权重值,过大的 α 和 β 会抑制映射导致损失图像多样性,对抗损失权重过大则会过分强调生成图像与真实图像完全一致,不符风格迁

移的目的。在训练中, epoch 设为 300, 在前 200 个 epoch 中, 学习率设置为 0.000 2, 在最后 100 个 epoch 中, 学习率会线性降至 0; WGAN-GP 及其他扩增方法 epoch 同样设为 300。

3.3 评价指标

本文使用 IS (Inception Score)^[25] 指标来评价扩增数据集的样本分布平衡性, 由于原真实数据集中右偏占比 80%, 且偏移程度较为单一, 分布不均衡, 因此将生成的数据集与原真实数据集取 IS 指标, IS 越小, 表示生成的数据集与原真实数据集分布相似, 故障种类的数量畸形, 不符合扩增的目的, 故 IS 指标越大越优。FID (Fréchet Inception Distance)^[25] 指标是通过计算扩增数据集与真实数据集在特征空间之间的距离来衡量的, 各扩增方法输入(噪声或 3D 图像)相同时, 较低的 FID 意味着生成图像更接近真实图像。MS_SSIM (Multi-scale Structural Similarity Index)^[26] 指标用于评价生成器轮廓结构收敛速度, PSNR (Peak Signal-to-Noise Ratio)^[27] 指标则用于评价生成图像质量。

3.3.1 IS

Inception Score 可用于评价图像的清晰度和多样性。把生成图像 x 输入 Inception V3 网络中, 将输出 1 000 维的向量 y , 向量 y 的每个维度的值对应图像属于某类的概率。

IS 评价指标表示为:

$$IS(G) = \exp\left\{E_{x \sim p_g} D_{KL}\left[P(y|X) \| P(y)\right]\right\}, \quad (10)$$

其中: $x \sim p_g$ 表示生成器生成图像, D_{KL} 表示对 $P(y|X)$ 和 $P(y)$ 求 KL 散度, 公式如下:

$$D_{KL}(P(y|X) \| P(y)) = \sum_i P(y_i|X) \log \frac{P(y_i|X)}{P(y_i)}. \quad (11)$$

IS 评价价值越大, 说明 $P(y|x)$ 与 $P(y)$ 的分布的差异越大, 生成数据集的图像质量和多样性越好。

3.3.2 FID

FID 距离计算真实样本与生成样本在特征空间之间的距离。首先利用 Inception 网络来提取特征, 然后使用高斯模型对特征空间进行建模, 再去求解两个特征之间的距离, 较低的 FID 意味着生成图像更接近真实图像。

FID 距离的计算如式 (12) 所示:

$$FID = \left\| \mu_x - \mu_y \right\|^2 + \text{Tr} \left(\sum x + \sum y \right) - 2 \left(\sum x \sum y \right)^{1/2}, \quad (12)$$

其中: x 表示真实跑偏图像, y 表示生成的图像, μ 表示均值, $\sum x$, $\sum y$ 表示协方差, Tr 表示矩阵的迹。

3.3.3 MS_SSIM 和 PSNR

MS_SSIM 和 PSNR 是广泛应用的图像评价指标。MS_SSIM 通过亮度、对比度和结构从多个尺度来衡量图像的结构相似性。PSNR 代表图像的峰值信噪比。计算公式如下:

$$l(x, y) = \frac{(2\mu_x \mu_y + c_1)}{(\mu_x^2 + \mu_y^2 + c_1)}, \quad (13)$$

$$c(x, y) = \frac{(2\sigma_{xy} + c_2)}{(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (14)$$

$$s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x \sigma_y + c_3}, \quad (15)$$

$$MS_SSIM = [l(x, y)]^{\alpha M} \cdot \prod_{j=1}^M [c(x, y)]^{\beta_j} [s(x, y)]^{\gamma_j}, \quad (16)$$

$$PSNR = 10 \cdot \log_{10} \frac{MAX_I^2}{\frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [x(i, j) - y(i, j)]^2}, \quad (17)$$

其中: x, y 表示评价对象, μ 表均值, σ 表方差, $l(x, y), c(x, y), s(x, y)$ 分别为亮度对比因子、对比度因子和结构对比因子, M 为预设的多尺度值, MAX_I 为图像最大像素值 255, m 和 n 代表图像高度和宽度。

3.4 YoloV5

为检验使用不同方法扩增的数据集应用于检测网络的效果, 本文选用 YoloV5 检测网络进行测试。实验中, 为使检测网络分类分布平衡, 将 3 种 CycleGAN 网络分别扩增 500 张左偏图像和 1 000 张混合图像; 选取 WGAN-GP 网络作为噪声生成的代表, 扩增 1 500 张混合图像 (WGAN-GP 仅能依据真实样本分布去扩增而无法针对性扩增)。扩增图像与真实图像作为整体数据集, 应用到 Yolo 目标检测网络中训练, 选择的训练模型为 YoloV5s, epoch 为 20, 训练集与验证集比例为 4: 1, 使用验证集的 MAP (Mean Average Precision) 指标评估不同生成器扩增数据集的效益。

4 实验分析

4.1 图像质量与分布

四种基于噪声的 GAN 网络在训练过程中生成的 Fake Y 如图 10 左图所示,基于 3D 模型的迁移类网络(CycleGAN、ChipGAN)训练过程中生成的 Fake Y 如图 10 右图所示。

Epoch 为 100 时,图 10 右图迁移类网络(CycleGAN, ChipGAN)存在不同情况的模糊、缺失、

错位问题,而图 10 左图图像因由多维噪声生成,并未存在缺失、错位等情况,但图像整体较为模糊、扭曲,其生成图像在皮带边界与轴相结合部分细节不如 CycleGAN 网络。在 200 个 epoch 后,3 种生成器的 CycleGAN 生成的图像轮廓均已收敛,基本稳定,但 ChipGAN 生成图像轮廓仍未完全收敛,表明迁移网络中 ChipGAN 轮廓收敛速度要低于 CycleGAN 网络。

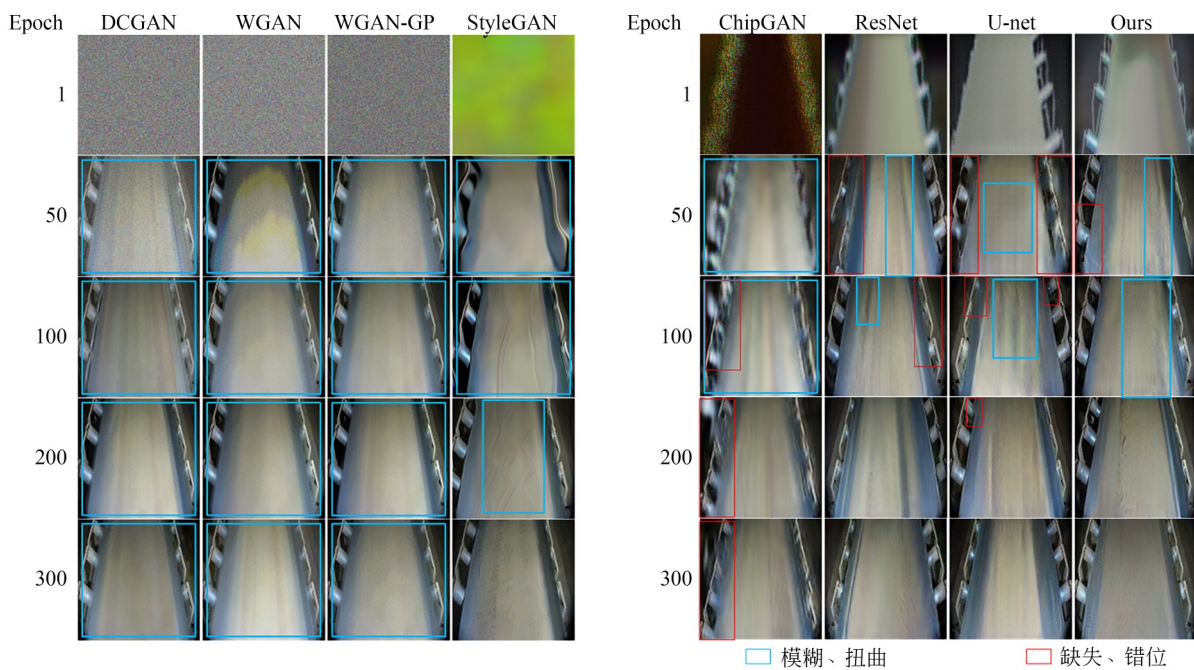


图 10 多种网络不同 epoch 下输出的 Fake Y
Fig. 10 Fake Y for multiple network outputs

下面使用 IS, FID 及 PSNR 指标来评价数据集扩增后的样本分布以及图像质量。IS 结果是通过各自扩增数据分析求得; FID 结果是通过扩增数据集与真实跑偏数据集对比分析得到; PSNR 则是 CycleGAN、ChipGAN 迁移网络在训练 300epoch 时真实图像 Real Y 与重构图像 Rec Y 对比得到,噪声生成网络因无配对图像,因此没有分析其 PSNR 指标。评价结果如表 1 所示。

IS 表示生成的数据集与原真实数据集分布相似程度,相似即右偏数量远大于左偏,不符合扩增的目的,迁移网络扩增后的数据集样本分布更加均衡,数量平均,每张图像跑偏程度不一,但分布与原真实数据集差异较大,故 IS 指标较大,

表 1 三种评价指标结果

Tab. 1 Result of three evaluation indicators

Generator/ GAN	IS ↑	FID ↓	PSNR ↑
DCGAN	1.385 7	178.805 5	/
WGAN	1.375 3	161.486 5	/
WGAN-GP	1.417 4	155.687 1	/
StyleGAN	1.397 6	164.547 3	/
ChipGAN	1.364 1	287.971 2	26.467 4
ResNet	1.484 9	232.266 9	34.595 4
U-net	1.446 9	227.299 1	33.245 7
Ours	1.519 7	192.021 2	35.096 4

即分布差异大,多样性好,更符合扩增数据集的目的。从表 2 可以看出,基于 3D 模型的迁移网络的 IS 值要显著高于基于噪声的 GAN 生成网络,在所有网络中,本文方法 IS 值最优。

FID 指标是通过计算扩增数据集与真实数据集在特征空间之间的距离来衡量的。基于噪声的 GAN 生成网络是以真实样本集为模板,因此生成图像的故障种类分布更符合真实跑偏样本集,即生成的图像右偏居多,故 FID 指标普遍优于迁移网络,但这并不符合扩增数据集使故障种类平衡分布的目的,迁移网络是由 3D 模型约束引导生产的,其故障种类分布更加均衡。在噪声生产网络中,WGAN-GP 的 FID 最优,而在所有 3D 模型迁移网络中,本文方法 FID 最优,说明在这几种迁移网络中,本文方法扩增的图像样本更接近真实样本。PSNR 指标也验证了这一点,在所有迁移网络中,本文方法 PSNR 最优。

综合 3 类指标,可知改进 U-ResNet 生成器所生成的图像相比于其他方法,生成图像质量最优,在所有迁移网络中更接近于真实图像,且满足扩增数据集平衡分布,多样化的需求。

4.2 结构收敛性

下面利用 MS_SSIM 指标分析生成器架构对图像轮廓结构收敛性影响。考虑 CycleGAN 中 Real Y 与 Fake Y 的非配对风格迁移特性,在真实跑偏样本中选取 6 张跑偏种类、程度不一的图像作为基准图像,将 CycleGAN 网络每个 epoch 生成的图像与基准图像对比得到 MS_SSIM 结果,保留最大值。把离散结果拟合为曲线,结果如图 11 所示。

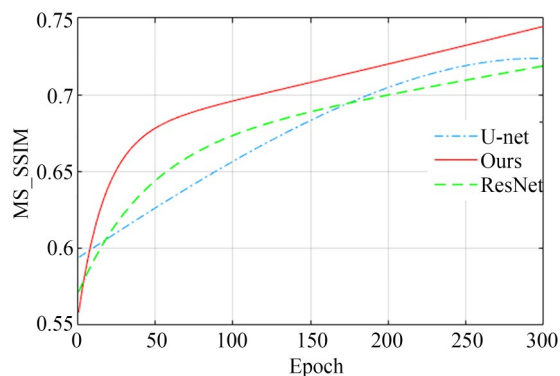


图 11 三种生成器的 MS_SSIM 指标

Fig. 11 MS_SSIM metrics for three generators

图 11 中 epoch 为 0 时, U-ResNet 的 MS_SSIM 值最低,原因是改进生成器轮廓结构的权重更高,即 epoch 为 0 时轮廓结构仍为 3D 模型,与图 10 右图第一行图像对应。随 epoch 增加,曲线上升,U-ResNet 斜率最高,即轮廓结构收敛最快,并在 120 epoch 时 MS_SSIM 已达到 0.7,而 U-net 与 ResNet 在接近 200 epoch 时 MS_SSIM 才达到 0.7,说明改进的 U-ResNet 生成器在图像轮廓结构收敛速度上更具优势。

三种生成器训练的损失收敛情况如图 12 所示,因对抗损失的特性导致整体损失是不断震荡的,整体上 U-net 与 U-ResNet 较 ResNet 生成器收敛较快。

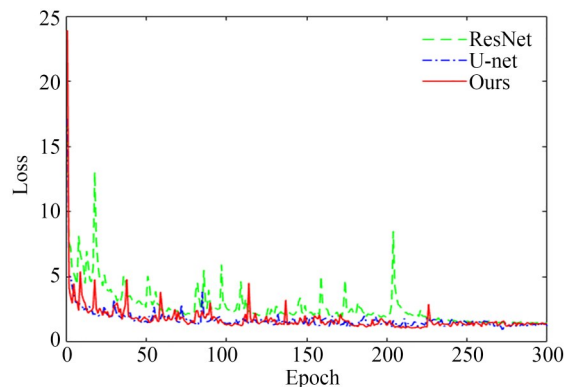


图 12 三种生成器的 loss 曲线

Fig. 12 Loss curve for three generators

4.3 检测网络

为分析不同方法得到的扩增数据集对检测网络准确率的影响,在 Yolo v5 目标检测网络中做了 5 组样本集对照实验,分别是:

- (1) 真实跑偏样本集;
- (2) 真实跑偏样本集 + WGAN-GP 扩增 1 500 张图像(WGAN-GP 扩增图像完全取决于真实跑偏数据集左右偏的分布,无法针对性扩增);
- (3) 真实跑偏样本集 + U-net 扩增的 1 500 张图像(500 左偏, 1 000 混合);
- (4) 真实跑偏样本集 + ResNet 扩增的 1500 张图像(500 左偏, 1 000 混合);
- (5) 真实跑偏样本集 + U-ResNet 扩增的 1 500 张图像(500 左偏, 1 000 混合)。

训练集与验证集比例为 4:1,其中 2-5 组中真实图像与扩增图像比例约为 2:3。基于 Yolo v5

强大的检测性能,令 epoch 为 20。选取训练后检测网络对验证集的 MAP 结果作为评价指标,结果如表 2 所示。

表 2 Yolov5 检测网络结果

Tab. 2 Results of Yolov5 detection (%)

数据集	左偏 AP	右偏 AP	MAP
真实	86.7	95.4	93.6
WGAN-GP+真实	89.8	96.1	94.8
U-net+真实	91.8	98.6	95.7
ResNet+真实	93.3	99.1	96.2
Ours+真实	96.7	99.5	98.1

第 1 组实验中,由于真实跑偏样本集中左偏样本数目远小于右偏,因此左偏检测的平均准确率小于右偏。第 2 组实验,存在同样的问题,因为 WGAN-GP 扩增样本类别分布依赖于真实样本的分布。在 3-5 组实验中,左偏样本扩增后,左偏检测准确率有较大提升,左右偏准确率更加接近,整体 MAP 也有显著提升。

由此可知,相较于基于噪声的数据集扩增方法,我们提出的基于 3D 模型约束的数据集扩增方法,可以有控制的生成各类故障样本,扩增数据集可以更趋于平滑,满足多样化需求,应用于检测网络,能显著提升网络性能。在 CycleGAN 网络中,我们提出的 U-ResNet 生成器,相较于 U-net、ResNet 生成器,MAP 值最优,进一步验证了 U-ResNet 生成器的优越性。

4.4 计算资源占用

各类生成器资源占用情况如表 3 所示。

改进方法在相同训练参数下计算资源需求增加约 20%,在 3060 6G laptop 移动端显卡上即可完成训练,可以满足工程应用。后期我们考虑可以通过削减卷积神经网络深度,或者在生成器后连接图像超分辨率增强模块以减少生成器输入输出图像大小,进一步降低计算资源占用,现输入输出图像大小为 256×256 ,考虑缩减至

表 3 生成器参数及训练显存占用

Tab. 3 Generator parameters and training memory occupancy

Model	Params size /MB	Layers of G	GPU Memory Usage/GB
ResNet	43.42	92	5.0
U-net	207.57	54	4.6
U-ResNet	41.7	75	5.8

128×128 ,节省训练资源,实现网络轻量化。

5 结 论

本文提出一种联合 3D 建模与改进 CycleGAN 的故障数据集扩增方法。输入 3D 模型故障样本引导迁移网络生成真实风格样本,达到扩增目的。同时我们对 CycleGAN 网络生成器架构进行改进,改进的 U-ResNet 生成器较 U-net、ResNet 生成器,具有更好的轮廓结构收敛性,生产图像质量更高。在带式输送机跑偏数据集扩增任务中,将本文方法与其他方法作对照实验,实验结果表明我们提出的联合 3D 建模与改进 CycleGAN 的故障数据集扩增方法要显著优于基于噪声生成的数据集扩增方法,扩增后的数据集图像质量更优,且样本分布更加均衡,满足多样性需求。应用于目标检测网络,检测准确率达到 98.1%,相较于原真实数据集提升 4.5%。

在未来,我们希望能将联合 3D 建模与改进 CycleGAN 的故障数据集扩增方法应用于更多工业场景。另外,我们也尝试在几方面对此方法进行改进:一是设计制作更精确的 3D 动态模型,采样图像可以更好地约束引导生成器,生成更接近真实的图像样本;二是我们将尝试在网络架构、生成器架构、判别器结构等方面对训练网络进行改进,提升网络性能;另外目前网络的规模仍然偏大,我们希望在网络轻量化方面做更多的研究。

参考文献:

[1] SHIN H, AHN Y, TAE S, *et al.* Enhancement of multi-class structural defect recognition using generative adversarial network [J]. *Sustainability*, 2021,

13(22): 12682.

[2] SHORTEN C, KHOSHGOFTAAR T M. A survey on image data augmentation for deep learning [J]. *Journal of Big Data*, 2019, 6(1): 1-48.

[3] CHEN L, WANG H J, MENG X H. Remote sens-

- ing image dataset expansion based on generative adversarial networks with modified shuffle attention [J]. *Sensors*, 2021, 21(14): 4867.
- [4] ZHANG H G, HU X G, MA D Z, *et al.* Insufficient data generative model for pipeline network leak detection using generative adversarial networks [J]. *IEEE Transactions on Cybernetics*, 2022, 52(7): 7107-7120.
- [5] SANDFORT V, YAN K, PICKHARDT P J, *et al.* Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks [J]. *Scientific Reports*, 2019, 9: 16884.
- [6] SU W C, YE H, CHEN S Y, *et al.* DrawingIn-Styles: portrait image generation and editing with spatially conditioned StyleGAN [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2022.
- [7] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, *et al.* Generative adversarial networks [J]. *Communications of the ACM*, 2020, 63(11): 139-144.
- [8] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks [C]. *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*, 2016.
- [9] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks [C]. *Proceedings of the 34th International Conference on Machine Learning - Volume 70. August 6-11, 2017, Sydney, NSW, Australia. New York: ACM*, 2017: 214-223.
- [10] GULRAJANI I, AHMED F, ARJOVSKY M, *et al.* Improved training of wasserstein GANs [C]. *Proceedings of the 31st International Conference on Neural Information Processing Systems. December 4 - 9, 2017, Long Beach, California, USA. New York: ACM*, 2017: 5769-5779.
- [11] ZHU J Y, PARK T, ISOLA P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks [C]. *2017 IEEE International Conference on Computer Vision (ICCV). 22-29, 2017, Venice, Italy. IEEE*, 2017: 2242-2251.
- [12] 王昊天, 刘庆省, 陈亮, 等. 改进的 CycleGAN 网络用于水下显微图像颜色校正 [J]. *光学精密工程*, 2022, 30(12)1499-1508
- WANG H T, LIU Q S, CHEN L, *et al.* Improved CycleGAN network for underwater microscopic image color correction [J]. *Opt. Precision Eng.*, 2022, 30(12)1499-1508(in Chinese)
- [13] HU A N, XIE Z, XU Y Y, *et al.* Unsupervised haze removal for high-resolution optical remote-sensing images based on improved generative adversarial networks [J]. *Remote Sensing*, 2020, 12(24): 4162.
- [14] WANG X T, YU K, WU S X, *et al.* ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks [M]. *Lecture Notes in Computer Science. Cham: Springer International Publishing*, 2019: 63-79.
- [15] 郝帅, 吴瑛琦, 马旭, 等. 基于 CycleGAN-SIFT 的可见光和红外图像匹配 [J]. *光学精密工程*, 2022, 30(5)602-614
- HAO S, WU Y Q, MA X, *et al.* Visible and infrared image matching based on CycleGAN-SIFT [J]. *Opt. Precision Eng.*, 2022, 30(5)602-614 (in Chinese)
- [16] 凡志逸, 夏伟杰, 刘雪. 基于修正 CycleGAN 的声呐图像库构建方法研究 [J]. *声学技术*, 2021, 40(6)890-894
- FAN Z M, XIA W J, LIU X. Modified CycleGAN based sonar image library construction [J]. *Technical Acoustics*, 2021, 40(6)890-894 (in Chinese)
- [17] 杨植凯, 卜乐平, 王腾, 等. 基于循环一致性对抗网络的室内火焰图像场景迁移 [J]. *光学精密工程*, 2020, 28(3)745-758
- YANG Z K, BU L P, WANG T, *et al.* Scenemi-gration of indoor flame image based on Cycle-Consistent adversarial networks [J]. *Opt. Precision Eng.*, 2020, 28(3)745-758 (in Chinese)
- [18] 崔克彬, 潘锋. 用于绝缘子故障检测的 CycleGAN 小样本库扩增方法研究 [J]. *计算机工程与科学*, 2022, 44(3)509-515
- CUI K B, PAN F. A CycleGAN small sample library amplification method for faulty insulator detection [J]. *Computer Engineering and Science*, 2022, 44(3)509-515 (in Chinese)
- [19] 杜振龙, 沈海洋, 宋国美, 等. 基于改进 CycleGAN 的图像风格迁移 [J]. *光学精密工程*, 2019, 27(8)1836-1844
- DU Z L, SHEN H Y, SONG G M, *et al.* Image

- style transfer based on improved CycleGAN [J]. *Opt. Precision Eng.*, 2019, 27(8): 1836-1844 (in Chinese)
- [20] DOU H, CHEN C, HU X, *et al.* Asymmetric CycleGAN for image-to-image translations with uneven complexities [J]. *Neurocomputing*, 2020, 415: 114-122.
- [21] PARIS S, HASINOFF S W, KAUTZ J. Local Laplacian filters [J]. *ACM Transactions on Graphics*, 2011, 30(4): 1-12.
- [22] RONNEBERGER O, FISCHER P, BROX T. *U-Net: Convolutional Networks for Biomedical Image Segmentation* [M]. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015: 234-241.
- [23] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [24] ISOLA P, ZHU J Y, ZHOU T H, *et al.* Image-to-image translation with conditional adversarial networks [C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 5967-5976.
- [25] BENNY Y, GALANTI T, BENAÏM S, *et al.* Evaluation metrics for conditional image generation [J]. *International Journal of Computer Vision*, 2021, 129(5): 1-20.
- [26] CHEN M J, BOVIK A C. Fast structural similarity index algorithm [J]. *Journal of Real-Time Image Processing*, 2011, 6(4): 281-287.
- [27] HUYNH-THU Q, GHANBARI M. Scope of validity of PSNR in image/video quality assessment [J]. *Electronics Letters*, 2008, 44(13): 800-801.

作者简介:



李宝平(1981—),男,博士,讲师,2016年于中国传媒大学获得博士学位,现河南理工大学物电学院从事教学科研工作,主要研究方向为计算机视觉及人工智能。E-mail: libaoping@hpu.edu.cn

通讯作者:



威恒熠(1999—),男,河南开封人,硕士研究生,2021年于河南理工大学获得学士学位,主要从事计算机视觉、图像处理等方面的研究。E-mail: mysteriousq@home.hpu.edu.cn