

DOI:10.19651/j.cnki.emt.2416729

基于感兴趣区域约束的漏缆卡扣重构缺陷检测*

任帅 李毅 肖武龙 王逸涵 李柏林

(西南交通大学机械工程学院 成都 610031)

摘要: 针对自监督漏缆卡扣缺陷算法中因背景重构误差造成的误检测,本文提出一种基于感兴趣区域约束的双阶段漏落卡扣缺陷重构方法。针对目标检测算法定位到的分类卡扣区域图像,首先嵌入分割网络用于分离卡扣及漏缆区域。随后,其对应掩码被作为线索引导堆叠对抗生成网络对卡扣区域进行重构,保证背景一致性的同时,精细化重构卡扣缺陷区域。此外,通过嵌入深度残差块和改进损失函数迫使生成网络更专注于感兴趣区域的重构。最后,训练完成的网络被用于卡扣图像的重构,并根据重构前后图像的相似度分数判定是否存在缺陷。在漏缆卡扣数据集上,定量结果表明本文算法对多类型卡扣缺陷识别的准确率 92.3%,召回率为 93.4%,高于其他自监督卡扣重构方法。可视化结果表明所提卡扣重构方法背景重构误差较少。

关键词: 漏缆卡扣;缺陷检测;感兴趣区域引导;堆叠对抗生成网络;注意力机制

中图分类号: TP391.41; U285.6; TN913 **文献标识码:** A **国家标准学科分类代码:** 460.5030

Detecting reconstruction defects in leaky cable snaps based on region of interest constraints

Ren Shuai Li Yi Xiao Wulong Wang Yihan Li Bailin

(School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: Aiming to address the false detections caused by background reconstruction errors in the self-supervised leaky cable snap defect detection algorithm, this paper proposes a two-stage leaky drop snap defect reconstruction method based on region of interest constraints. For the classified snap region images localized by the target detection algorithm, a segmentation network is first employed to differentiate between snap and cable leakage regions. Subsequently, the corresponding masks are utilized to guide a stacked adversarial generative network to reconstruct the snap regions, ensuring high-quality reconstruction of the defect areas while maintaining background consistency. Additionally, the generative network is optimized to place greater emphasis on reconstructing the regions of interest by integrating deep residual blocks and refining the loss function. Ultimately, the trained network is deployed for the reconstruction of snap images, determining the presence of defects based on the similarity scores of the images before and after reconstruction. Quantitative results on the leaky cable snap dataset demonstrate that the proposed algorithm achieves a defect recognition accuracy of 92.3% and a recall rate of 93.4%, surpassing the performance of other self-supervised snap reconstruction methods. Visualization results further indicate a reduction in background reconstruction errors in the proposed method.

Keywords: leaky cable snap; defect detection; area of interest guidance; stacked adversarial generative networks; attention mechanism

0 引言

近年来,随着中国铁路事业的迅猛发展,高速铁路安全运行的重要性日益凸显。高速铁路移动通信网络系统在确保中国高速铁路无线通信畅通的过程中扮演着关键角色,

对高速铁路的安全运行至关重要^[1-2]。在高速铁路轨道系统中,通讯线路的覆盖主要依赖于通信泄露同轴电缆^[3]。漏缆通常由卡扣固定在隧道壁,受到高速列车运行时的压力和热波影响,卡扣会出现损坏或松动,对列车的安全运行造成隐患^[4-5]。因此,漏缆卡扣缺陷检测是铁路自动化巡检

收稿日期:2024-08-24

* 基金项目:四川省科技研发计划重点项目(2021YFN0020)资助

中的重要内容之一。

目前,针对漏缆卡扣的缺陷检测主要采用全监督方式的特征设计和深度学习的方法^[6-10]。尽管上述方法在缺陷检测中取得了令人满意的结果,但全监督算法的训练容易受到标注样本数量以及缺陷样本与正常样本之间均衡性的限制,存在过拟合的风险。此外,标注过程繁琐复杂,时间成本高且标注结果存在不够客观、易受个体主观影响等问题。更重要的是,全监督方法难以识别训练样本中未出现的缺陷类型。但实际场景中,漏缆卡扣存在种类多样、缺陷样式参差且数量稀少等特点。针对全监督缺陷检测方法的不足,李毅等^[11]提出了一种基于对抗生成网络的无监督漏缆卡扣缺陷检测算法。该方法利用对抗生成网络学习大量正常样本并将其用于重构缺陷样本,并通过对比重构前后样本来对缺陷进行判别。但该重构方法也会因图像背景重构误差而导致一定数量的误检测。

在无监督缺陷检测方法中,降低背景重构误差是降低其误检测风险的重要手段。如 Zavrtanik 等^[12]利用分块、随机去除的方式处理训练图像,迫使网络以领域信息对缺陷区域进行重构,从而减少背景区域重建误差。Xia 等^[13]提出利用额外注意力网络判断重构前后图像的差异图进行缺陷判别,抑制背景误差对缺陷判别的干扰。Kim 等^[14]则采用循环生成对抗网络(cycle generative adversarial network, CycleGAN)重构缺陷图像,其通过堆叠额外的对抗生成网络来提升图像的重构性能。针对结构,高港等^[15]则提出以嵌入坐标注意力的方式来约束网络对背景区域的重构噪声。尽管上述方法能在一定程度上改善了缺陷重构的性能,但嵌入注意力机制这种被动策略仍然无法有效地限制背景重构噪声的生成。

针对上述问题,本文受 Jabbar 等^[16]中堆叠生成对抗网络(stacked generative adversarial networks, StackGAN)的启发,提出基于感兴趣区域约束的漏缆卡扣缺陷重构方法,并将其用于卡扣的缺陷检测。该方法在 StackGAN 的基础上,通过预测感兴趣大致区域并将对应掩码作为线索引导缺陷卡扣重构,抑制背景重构噪声的生成。在 StackGAN 的基础上,本文进一步改进重构误差函数,使其更适用于卡扣等弱纹理图像的重构,并迫使网络方向在过程中更关注感兴趣区域,而非背景区域。不同于前文所提改进方法,本文以线索指引的方式进行目标区域重构,对比可视化结果表明背景重构误差更少。

1 漏缆卡扣缺陷检测方法

本研究提出的基于感兴趣区域约束的漏缆卡扣缺陷重构方法主要用于减少重构过程中的背景误差,降低误检率。如图 1 所示,该算法主要由 3 个部分组成:漏缆卡扣的定位、漏缆卡扣的重构以及漏缆卡扣的缺陷判断。

1)卡扣定位:在原始的漏缆图像中,卡扣所占区域较小,图像背景复杂且存在较大干扰物。为了提升卡扣重构

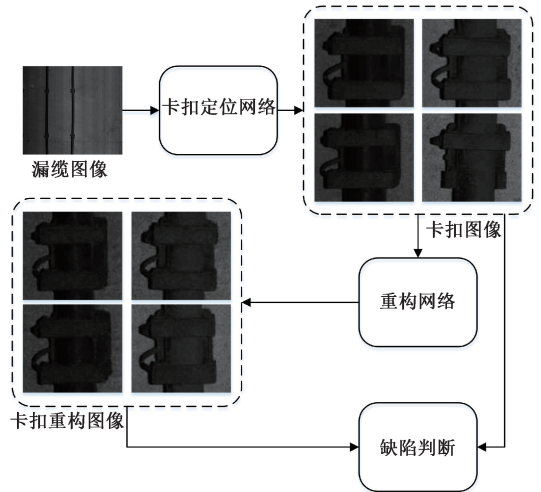


图 1 漏缆卡扣自监督缺陷检测流程

Fig. 1 Self-supervised defect detection process for leaky cable snaps

效果,提升缺陷检测精度,本文首先利用目标检测算法对卡扣区域进行定位并截取。

2)卡扣重构:针对定位到的卡扣区域,本文利用重构网络对其进行重构并利用重构前后的图像差异进行缺陷判别。

3)缺陷判断:计算漏缆图像中所有卡扣重构前后的差异数值,并利用差异阈值判断该卡扣是否存在缺陷。

1.1 漏缆卡扣定位

本文迁移现有较为优秀的目标检测模型,通过对精度、召回率以及计算开销等指标进行全面的评估来筛选最合适的卡扣定位算法。本文在卡扣数据集上测试了基础目标检测算法 Faster R-CNN、SSD、YOLOv5 以及现行优秀的目标检测方法 YOLOv7、YOLOv8 和 RT-DETR。对比实验结果,本文选择 RT-DETR^[17]算法进行漏缆卡扣定位。同时,鉴于部署对计算硬件性能的要求,本文采用 MobileNetV4^[18]作为定位方法的主干网络,以减少定位算法的参数数量和计算开销。

1.2 漏缆卡扣重构模型

背景重构误差是自监督缺陷检测算法中误检测的主要原因,为了提升卡扣缺陷重构效果,降低误检率,本文提出一种双阶段漏缆卡扣缺陷重构方法。该方法的主要结构如图 2 所示。

对于定位截取的漏缆卡扣图像 I_k , 预设一个掩码分隔网络对其卡扣及漏缆区域进行提取,得到初始的掩码预测图 I_{mask}^p 。对形态学处理后的掩码预测图 \hat{I}_{mask}^p , 将其与卡扣图像 I_k 进行合并,作为第一个对抗生成网络的输入。其中掩码预测图 \hat{I}_{mask}^p 主要作为重构索引,用于减少背景区域的重构误差。在第一个对抗生成网络中,生成器 G_1 和判别器 D_1 采用对抗训练的方式,以对应类的无缺陷卡扣图像 I_{gt} 作为监督,使生成器 G_1 将缺陷卡扣图像 I_k 尽可能相似地

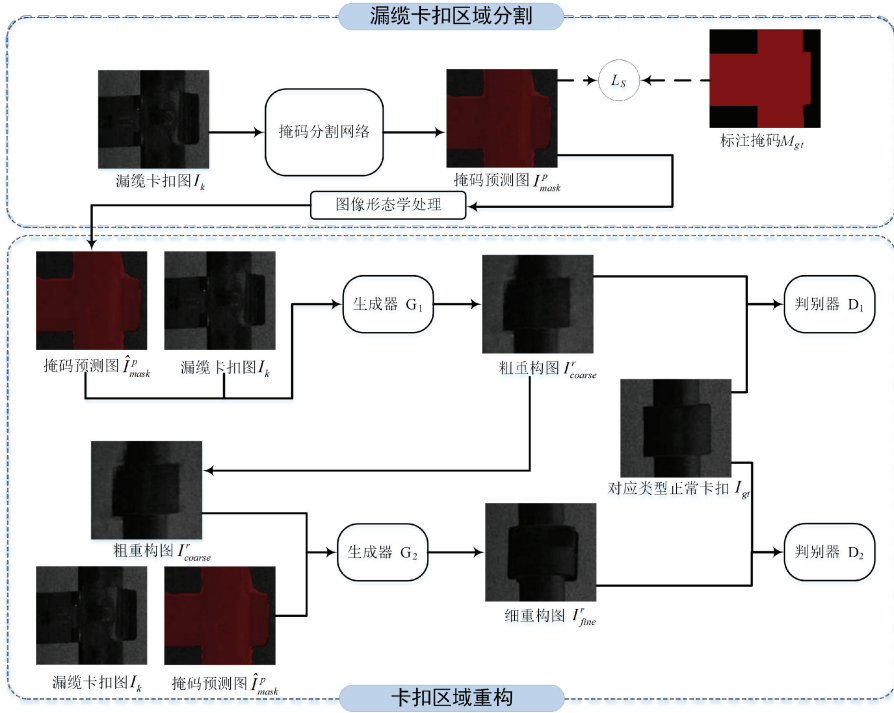


图 2 漏缆卡扣重构模型

Fig. 2 Leakage cable snap reconfiguration model

重构为无缺陷图像。驱动第一个对抗生成网络的损失函数主要由重构损失 \mathcal{L}_{re}^1 和对抗损失 \mathcal{L}_{ad}^1 两部分组成,定义如下:

$$\mathcal{L}^1 = \mathcal{L}_{re}^1 + \mathcal{L}_{ad}^1 \quad (1)$$

$$\mathcal{L}_{re}^1 = \mathcal{L}_{l1}^1 + (1 - \mathcal{L}_{SSIM}) \quad (2)$$

其中, \mathcal{L}_{SSIM} 表示结构相似度损失, \mathcal{L}_{l1}^1 表示卡扣区域内 L_1 损失:

$$\mathcal{L}_{l1}^1 = \alpha \hat{I}_{mask}^p \cdot \cdot + \beta (\sim \hat{I}_{mask}^p) \cdot \cdot \quad (3)$$

其中, \hat{I}_{mask}^p 表示输入的二值化掩码, $\cdot \cdot$ 表示输入图像的 L_1 损失,输入为正常卡扣图 I_{gt} 和粗重构图 I_{coarse}^r 。 α 和 β 分别表示权重系数。

$$\mathcal{L}_{ad}^1 = \mathbb{E}[\log(D_1(\cdot))] + [\log(1 - D_1(\cdot))] \quad (4)$$

其中, D_1 表示这一级对抗生成网络中的判别器。 (\cdot) 和 $(\hat{\cdot})$ 分别表示判别器输入的正常卡扣图 I_{gt} 和粗重构图 I_{coarse}^r 。

对于第一级对抗生成网络中生成器输出的粗重构图 I_{coarse}^r , 本文进一步将其与卡扣图像 I_k 和掩码图 \hat{I}_{mask}^p 进行拼接,作为下一级对抗生成网络训练的输入。通过再一次引入无缺陷卡扣图像 I_{gt} 作为监督,对粗预测进行重构,以对抗训练的方式细化重构结果,从而减少上一级重构过程中的背景误差。这一级网络的训练损失 \mathcal{L}^2 由重构损失、对抗损失和潜在空间损失三部分组成。

$$\mathcal{L}^2 = \mathcal{L}_{re}^2 + \mathcal{L}_{ad}^2 + \mathcal{L}_{lat}^2 \quad (5)$$

其中,重构误差 \mathcal{L}_{re}^2 依旧由卡扣区域的 L_1 损失和结构相似度损失组成。

$$\mathcal{L}_{re}^2 = \mathcal{L}_{l1}^2 + (1 - \mathcal{L}_{SSIM}) \quad (6)$$

对抗损失 \mathcal{L}_{ad}^2 表示为:

$$\mathcal{L}_{ad}^2 = \mathbb{E}[\log(D_2(\cdot))] + [\log(1 - D_2(\cdot))] \quad (7)$$

其中, (\cdot) 和 $(\hat{\cdot})$ 分别表示判别器输入的正常卡扣图 I_{gt} 和精重构图 I_{fine}^r 。

图像潜在空间损失 \mathcal{L}_{lat}^2 可表示为:

$$\mathcal{L}_{lat}^2 = \|\delta_{I_{gt}^r}, \delta_{I_{fine}^r}\|_2 \quad (8)$$

其中, $\|\cdot\|_2$ 表示 L_2 正则, $\delta_{I_{gt}^r}$ 、 $\delta_{I_{fine}^r}$ 分别表示 I_{gt} 和 I_{fine}^r 在判别器 G_2 的潜在空间中的向量表示。

因此联合两个对抗生成网络训练所采用的损失函数 \mathcal{L} 为对各级对抗生成网络损失的加权求和:

$$\mathcal{L} = \lambda_1 \mathcal{L}^1 + \lambda_2 \mathcal{L}^2 \quad (9)$$

1.3 网络结构及组成

掩码分割网络:掩码分割网络用于分割卡扣图像的卡扣及漏缆区域,输出二值化掩码。分割网络主要由编码器和解码器组成。编码器用于提取输入图像 I_k 的潜在空间的特征表示。解码器则将潜在空间表示映射为与输入图像相同尺寸的二值化掩码。掩码分割网络的详细结构如表 1 所示。

少量的二值化分割掩码存在一些空腔和噪声。为了去除这些噪声,本文采用形态学处理开闭运算。对于初始分割结果 I_{pre_mask} ,形态学处理后的结果 \hat{I}_{pre_mask} 由下过程得到:

表 1 掩码分割网络潜在特征空间结构

Table 1 Potential feature space structure for mask segmentation networks

网络层	类型	输出维度
输入		128×128×3
Conv1	Conv3×3+BN+Relu	128×128×64
Conv2	Conv3×3+BN+Relu	64×64×128
Conv3	Conv3×3+BN+Relu	32×32×256
Conv4	Conv3×3+BN+Tanh	16×16×512
DeConv1	ConvTrans3×3+BN+Relu	32×32×256
DeConv2	ConvTrans3×3+BN+Relu	64×64×128
DeConv3	ConvTrans3×3+BN+Relu	128×128×64
输出	Conv1×1+Tanh	128×128×3

$$\hat{I}_{pre_mask} = (I_{pre_mask} \ominus Z) \oplus Z \quad (10)$$

$$\hat{I}_{pre_mask} = (I_{pre_mask} \oplus Z) \ominus Z \quad (11)$$

其中, \oplus 为膨胀操作, \ominus 为腐蚀操作。

重构网络: 本文所提卡扣重构网络的主要由两个对抗生成网络组成。前一个对抗生成网络利用无缺陷卡扣做监督, 生成缺陷卡扣图像的粗重构 I'_{coarse} 。后一对抗生成网络再次利用无缺陷卡扣做监督, 将粗预测与缺陷卡扣图合并作为输入。通过对比粗重构前后图像, 精细化调整背景区域的重构效果, 减少背景重构误差。每一个对抗生成网络都有生成器和判别器组成, 其中生成器网络的详细结构如表 2 所示, 判别器的详细网络结构如表 3 所示。

表 2 生成器网络结构

Table 2 Generator network structure

网络层	类型	输出维度
输入		128×128×4
Conv1	Conv3×3+BN+Relu	128×128×64
Conv2	Conv3×3+BN+Relu	64×64×128
Conv3	Conv3×3+BN+Relu	32×32×256
Conv4	Conv3×3+BN+Relu	16×16×512
Conv5	Conv3×3+BN+Relu	8×8×1 024
Conv6	RE×3	8×8×1 024
DeConv1	ConvTrans3×3+BN+Relu	16×16×512
DeConv2	ConvTrans3×3+BN+Relu	32×32×256
DeConv3	ConvTrans3×3+BN+Relu	64×64×128
DeConv4	ConvTrans3×3+BN+Relu	128×128×64
输出	Conv1×1+Tanh	128×128×3

生成器由编码器和解码器组成, 采用深度卷积神经网络架构。编码器主要由卷积层、归一化层和激活函数层, 将尺度 128×128 的输入图像抽象表征为 8×8 潜在空间向

量。相反, 解码器由反卷积层、归一化层和激活函数层组成, 用于将图像的潜在空间表示复映射为重构输出。在映射过程中, 跳跃连接还被用于在上采样过程中传递多尺度特征信息, 弥补特征下采样过程中缺失的特征信息。此外, 本文在在编码器和解码器的中间嵌入 3 个深度瓶颈残差块 (residualblock and ELA, RE)。RE 模块使用残差块结构和高效局部注意力 (efficient local attention, ELA) 对潜在空间向量进行精炼, 逐步扩大单个神经元的感受野, 并给予重要的通道特征和纹理区域更多关注。

残差结构^[19]已被证明能够较好地处理深度卷积神经网络训练中梯度消失和梯度爆炸等问题。而在漏缆卡扣图像重构中, 注意力可以帮助生成器更好地聚焦于感兴趣区域, 借助于卡扣区域掩码的指引, 从而产生更理想的重构结果。因此本文引入 ELA^[20]注意力机制并将其嵌入至残差结构中。深度瓶颈残差块的详细结构如图 3 所示。

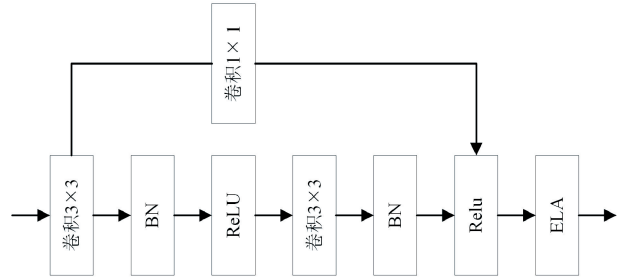


图 3 深度瓶颈层残差块

Fig. 3 Residualblock and ELA

ELA 注意力机制的详细结构如图 4 所示。对于输入特征, ELA 首先在平行的两个特征空间中沿不同方向对其进行池化。获得感兴趣区域在潜在表达空间中的坐标编码。

$$z_c^h(h) = \frac{1}{H} \sum_{0 \leq i \leq W} x_c(h, i) \quad (12)$$

$$z_c^w(w) = \frac{1}{W} \sum_{0 \leq j \leq H} x_c(j, w) \quad (13)$$

式中: W 为该层级特征图的宽度, H 为该层级特征图的高度。 $x_c(h, i)$ 表示输入特征沿 H 方向的输入, $x_c(j, w)$ 表示输入特征沿 W 方向的输入。

接下来, 采用一维卷积编码不同特征空间的坐标响应, 并通过 GropNorm 归一化操作对位置信息进行处理, 获得位置注意力在水平和垂直方向上的表示。

$$y^h = \sigma(G_n(F_h(z_h))) \quad (14)$$

$$y^w = \sigma(G_n(F_w(z_w))) \quad (15)$$

式中: σ 表示 Sigmoid 非线性激活函数。 F_h 和 F_w 分别表示不同特征空间的一维卷积, 其卷积核尺寸被设为 7×7 。

最终, 对于给定输入 x_c , ELA 注意力的输出 Y 表示为:

$$Y = x_c \times y^h \times y^w \quad (16)$$

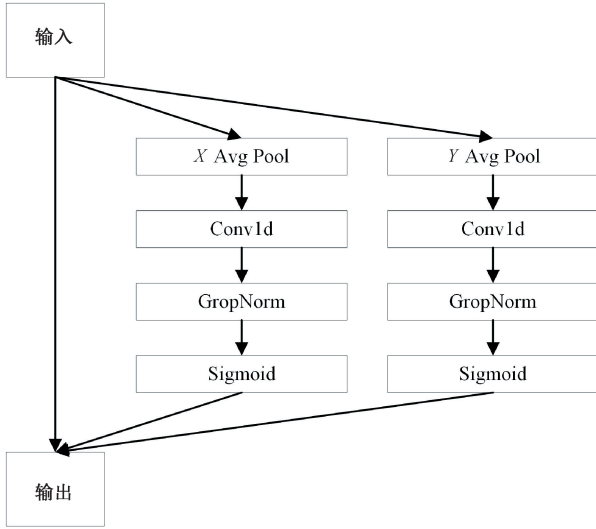


图 4 ELA 注意力机制

Fig. 4 ELA attention mechanisms

重构网络的判别器主要用于将预测图像和无缺陷图像再次映射为潜在空间表示,以用于真假判别。其详细的网络结构如表 3 所示。

表 3 判别器网络结构

Table 3 Discriminator network structure

网络层	类型	输出维度
输入		128×128×3
Conv1	Conv4×4+BN+LeakeyRelu	64×64×64
Conv2	Conv4×4+BN+LeakeyRelu	32×32×128
Conv3	Conv4×4+BN+LeakeyRelu	16×16×256
Conv4	Conv4×4+BN+LeakeyRelu	8×8×512
输出	Conv4×4+BN+Tanh	8×8×1

1.4 缺陷判别

如图 1 所示,本文迁移训练好的卡扣定位网络与重构网络用于检测未标注漏缆图像中缺陷卡扣。由于受到采集条件的限制,采集到的漏缆图像中漏缆卡扣区域较小,导致定位分类后的卡扣图像边缘模糊、噪声增加和背景区分度较低。采用直方图距离,峰值信噪比,差异哈希算法和残差等方法,难以有效地依据重构前后的图像相似度来区分正常与缺陷卡扣图像。为增强对不同类型缺陷卡扣的检测效果,本文采用了余弦相似度(cosin similarity, COSIN)和结构相似性(structure similarity index measure, SSIM)的卡扣缺陷检测方法,余弦相似度的计算方式如下:

$$\cos(\theta) = \frac{\sum_{i=1}^n (\mathbf{x}_i \times \mathbf{y}_i)}{\sqrt{\sum_{i=1}^n (\mathbf{x}_i)^2} \times \sqrt{\sum_{i=1}^n (\mathbf{y}_i)^2}} \quad (17)$$

式中: x_i 和 y_i 分别表示输入向量的分量。

利用余弦相似度,本文分别计算输入的卡扣图像与重构后图像的图像余弦相似度 \cos^l ,同时计算二者潜在向量空间表示的余弦相似度 \cos^g 。

$$\cos_{total} = \alpha \cos^l + (1 - \alpha) \cos^g \quad (18)$$

式中: α 为权重系数, $\alpha \in (0, 1)$ 。

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (19)$$

其中, μ_x 和 σ_x^2 分别表示真实图像的均值和方差, μ_y 和 σ_y^2 分别表示修复图像的均值和方差, σ_{xy} 为修复图像与真实图像的协方差, c_1 和 c_2 为常数。

本文同时依据 \cos_{total} 和 $SSIM(x, y)$ 的值,来判断漏缆卡扣得缺陷。

$$K = \begin{cases} \text{无缺陷,} & \cos_{total} > 0.965 \&\& SSIM(x, y) > 0.96 \\ \text{有缺陷,} & \text{其他} \end{cases} \quad (20)$$

2 实 验

2.1 实验数据集

本研究采集的图像来自中国铁道科学院,图像均由综合巡检车在实际线路上运行时所采集。如图 5 和 6 所示,该数据集包含了 4 种不同类型的漏缆卡扣,图像尺寸为 2 048×2 011,共 5 052 张漏缆图像,包含缺陷卡扣的漏缆图像共 445 张,均为正常卡扣的漏缆图像 4 607 张。

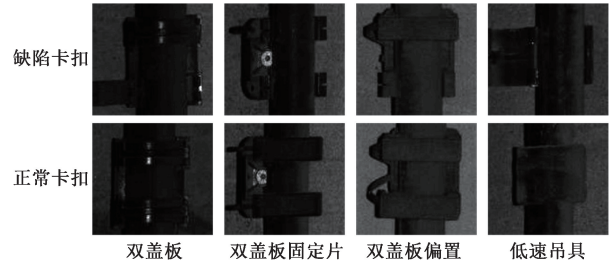


图 5 漏缆卡扣图

Fig. 5 Leakage cable snap diagram

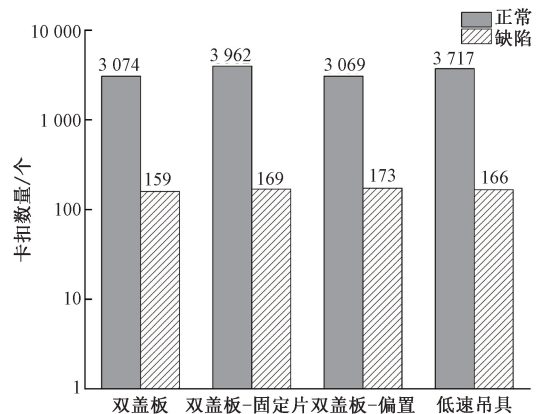


图 6 漏缆卡扣数据

Fig. 6 Leakage cable snap data

为了评估定位算法的回归性能,本文随机挑选并标注了 5 052 张漏缆图像中的 3 536(70%)张作为训练集,剩下的 1 516 张漏缆图像作为验证集,并额外挑选未标注的图像 1 500 张作为测试集。为了评估后续重构和缺陷检测的结果,本文将卡扣定位后截取尺寸设定为 128×128 。

为了训练卡扣重构算法,本文将 8 693 卡扣图像作为训练集和 2 900 张作为验证集,并随机挑选其中的正常卡扣作为监督。同时,本文挑选 2 500 张卡扣图像评估算法的缺陷检测性能,其中正常卡扣图像 1 833 张和缺陷卡扣图像 667 张。

2.2 实验平台及参数设置

本文所有相关实验均在设备 Dell Precision 5820 Tower 上进行,其 CPU 型号为 Intel i9-10900X,内存为 64 G,GPU 型号为 GeForce RTX3090,显存为 24 G。实验采用的操作系统均为 Windows10,编程语言为 Python3.8,训练及调试环境为 PyTorch 1.13.1,cuda11.7,编程平台为 Pycharm,定位、重构以及相关对比算法均在相同环境下运行。所有网络的训练均采用 SGD 优化器,初始学习率设置为 0.001,权重衰减系数为 0.0005,动量衰减为 0.9,训练周期为 200 轮,批量大小为 4。参考文献[16],联合损失函数 \mathcal{L} 各级损失的权重 λ_1 和 λ_2 被分别设置为 5 和 1。

2.3 评价指标

在卡扣定位和分类实验中,本文采用精确度(precision,P)、召回率(recall,R)、平均精度(mean average precision,mAP)、模型参数量(parameters)、计算量(giga floating point operations per second,GFLOPS)、帧率(frames per second,FPS)等指标对模型进行评价。在卡扣

重构实验中,本文分别采用绝对平均误差(mean absolute error,MAE),结构相似度 SSIM 和余弦相似度 COSIN 等指标评估重构样本与无缺陷样本的相似度。由于缺少与缺陷卡扣相对应无缺陷卡扣图像,因此本文选择将重构卡扣与对应类型的 400 张无缺陷样本计算相似度并取其平均值。在上述相似度指标当中,SSIM 和 COSIN 值越接近 1 表示图像对间相似度越高,越接近 0 则表明相似度越小。相反,MAE 指标越低则表明相似度越高。

在卡扣的缺陷判别中,本文分别采用精确度、召回率以及曲线下面积(area under the curve,AUC)值来评估不同算法的缺陷检测性能。AUC 被定义为受试者工作特性(receiver operating characteristic curve,ROC)曲线下的面积。AUC 越接近 1 表明模型的缺陷检测效果越好。

2.4 卡扣定位实验

为了挑选合适的目标算法,本文在卡扣数据集上训练并测试了基础目标算法 Faster-RCNN、SSD、YOLOv5-M 和现行主流的目标检测算法 YOLOv7、YOLOv8-M 以及 RT-DETR。详细实验结果如表 4 所示,可以看出:

1)在所有目标定位算法中,RT-DETR 取得了最高的精度、召回率以及平均精度,它们分别为 97.6%、94.8%、98.8%。同时,该算法的检测速度为 110 fps。

2)在平均精度上,Faster-RCNN 为 91.3%和 SSD 为 82.3%显著低于 RT-DETR 算法,YOLOv5-M、YOLOv7、YOLOv8-M 则略低于 RT-DETR 算法。同时,这些算法的检测速度均低于 RT-DETR。

3)将 RT-DETR 主干换为 MobileNetV4 后,该算法在平均精度 98.6%和召回率 93.5%略低于 RT-DETR 的情况下,参数量和计算量分别降低了 42.5%,30.2%。

表 4 部分主流目标检测算法对比

Table 4 Comparison of some mainstream target detection algorithms

模型	精确度/ %	召回率/ %	平均精度/ %	参数量/ M	计算量/ G	帧率/ fps
Faster-RCNN	90.8	92.7	91.3	41.3	173.6	63
SSD	84.2	75.6	82.3	36.1	98.6	54
YOLOv5-M	94.5	92.4	96.8	24.9	62.1	85
YOLOv7	95.2	93.6	97.1	36.6	103.5	92
YOLOv8-M	96.4	92.8	97.6	25.9	78.9	97
RT-DETR	97.6	94.8	98.8	20.2	58.6	110
改进 RT-DETR	97.3	93.5	98.6	11.6	40.9	128

在实际检测时,卡扣检测需要满足高精度和高检测帧率等要求。综合巡检车的巡检速度约为 35 m/s,而每张漏缆图像沿铁路方向的视场宽度约为 2 m。此外,巡检车还需要同步检测钢轨、铁路信号盒、铁路扣件等其他设备缺陷。基于上述卡扣定位结果的对比,本文挑选改进 RT-DETR 作为卡扣定位算法。

2.5 卡扣重构对比实验

为了评估本文所提算法的卡扣重构性能,在卡扣数据集上,本文将其与卡扣缺陷检测文献[11],CycleGAN、跳跃生成对抗网络异常检测(skip generative adversarial network for anomaly detection, Skip-GANomaly)和 StackGAN 等无监督重构方法进行对比。各方法重构结果

客观评价指标如表 5 所示,可视化对比结果如图 7 所示。相关实验可以看出:

1)相较于其他算法,本文重构算法在卡扣数据集上取得了最高 SSIM 和 COSIN 值为 0.971 和 0.975,最低的 MAE 值为 1.094。

2)相较于文献[11]中的卡扣重构方法,本文所提方法

的重构性能在 SSIM 和 COSIN 指标上分别提高了 2.9% 和 1.1%,在 MAE 指标上显著地降低了 2.218。

3)从图 7 展示的重构可视化结果可以看出,针对不同类型的卡扣,本文所提算法取得了较好的重构效果。更重要的是,本文算法计算得到的卡扣重构图像中,背景区域的变化更小。



图 7 缺陷卡扣重构对比实验图

Fig. 7 Comparative experimental diagram of defective snap reconfiguration

表 5 漏缆卡扣重构对比结果

Table 5 Leakage cable snap reconfiguration comparison results

方法	SSIM	COSIN	MAE
CycleGAN	0.872	0.921	7.123
Skip-GANomaly	0.917	0.939	5.847
文献[11]	0.942	0.964	3.312
StackGAN	0.939	0.944	4.621
本文算法	0.971	0.975	1.094

为了评估本文所提算法对卡扣缺陷检测的性能影响,本文定量对比了一些重构方法和分类方法在卡扣数据集上的缺陷检测性能,其结果如表 6 所示。其中,分类算法的结果由全监督标注数据集训练所得。

表 6 的实验结果表明:

1)本文所提重构方法的缺陷检测性能优于异常检测算法 Skip-GANomaly,文献[11]和 Stack-GAN 重构算法,

表 6 漏缆卡扣缺陷检测对比

Table 6 Comparison of leaky cable snap defect detection

模型	精确度	召回率	AUC
YOLOv5-M	0.921	0.929	0.972
YOLOv8-M	0.928	0.931	0.974
Skip-GANomaly	0.881	89.100	0.931
文献[11]	0.911	0.919	0.961
Stack-GAN	0.891	0.897	0.931
本文算法	0.923	0.934	0.972

与分类算法 YOLOv5-M 和 YOLOv8-M 相当。

2)本文所提重构方法卡扣缺陷检测的精确度和召回率分别为 0.923 和 0.934,AUC 值为 0.972。

3)本文所提方法相较于文献[11]中所提的自监督卡扣缺陷检测方法,精度提升了 1.2%,召回率提升了 1.5%。相较于改进前的 Stack-GAN 方法,本文方法在精度

(+3.2%)、召回率(+3.7%)和 AUC(+4.1%)值上均取得了较大提升。

本文对文献[11]和 StackGAN 在正常卡扣与缺陷卡扣判断能力方面进行了比较。如图 8~10 所示,展示了测试集中 2 500 张卡扣图像的相似度分数分布情况。

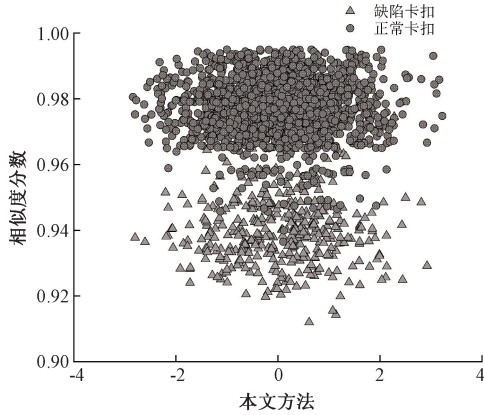


图 8 本文方法的相似度分数分布

Fig. 8 Distribution of similarity scores for the methods in this paper

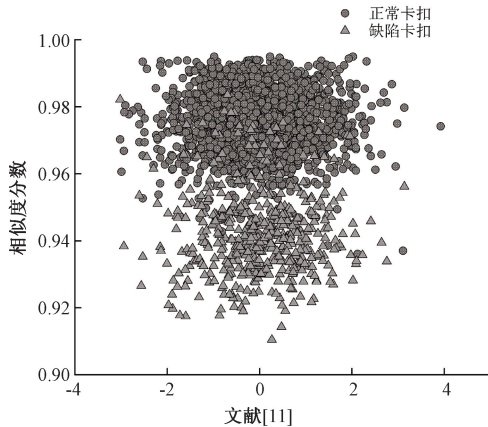


图 9 文献[11]的相似度分数分布

Fig. 9 Distribution of similarity scores in literature [11]

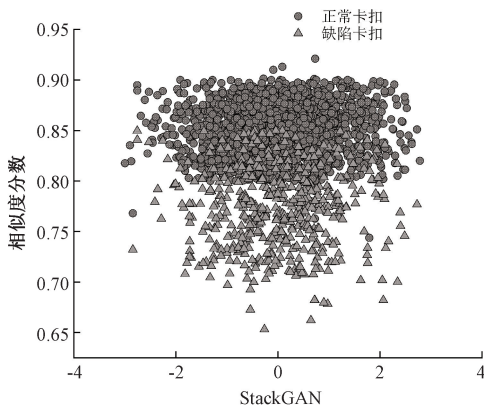


图 10 StackGAN 的的相似度分数分布

Fig. 10 Distribution of similarity scores for StackGAN

根据图 8~10 的分析结果可以看出,本文提出的方法在区分正常卡扣与缺陷卡扣方面,较文献[11]和 StackGAN 表现出更优越的性能。这表明本文的方法在处理此类分类任务时,具有更高的准确性和有效性。

2.6 消融实验

在本文所提基于重点区域索引的双阶段缺陷卡扣重构方法中,分割掩码对重构区域有指引作用,同时生成网络嵌入的深度瓶颈残差块 RE 也能够促使网络关注重点区域。为了分别确定这些改进对卡扣重构和缺陷识别的具体影响,本文进行了相关消融实验,具体实验结果分别如表 7 和 8 所示。

表 7 模块消融实验对比表

Table 7 Comparison table of modular ablation experiments

RE	掩码	SSIM	COSIN	MAE	AUC
×	×	0.939	0.944	4.621	0.931
✓	×	0.945	0.949	3.145	0.945
×	✓	0.955	0.958	2.154	0.959
✓	✓	0.971	0.975	1.094	0.972

表 8 深度瓶颈层数量对重构实验的影响

Table 8 Impact of the number of RE on reconfiguration experiments

RE 数量	SSIM	COSIN	MAE
1	0.940	0.944	4.621
2	0.941	0.944	4.287
3	0.945	0.949	3.145
4	0.946	0.951	3.026

从表 7 消融实验结果可以看出:

1)从加入 RE 后,SSIM、COSIN 和 AUC 分别提升了 0.6%、0.5%和 1.4%,同时 MAE 降低 1.476。这证明 RE 模块在生成器中的应用有助于降低图像采集过程中的噪声,帮助生成器更好地理解图像的全局结构和局部细节,从而产生更准确的重构结果,减少图像生成后的伪影和色彩差异。

2)加入掩码图后模型在 SSIM、COSIN 和 AUC 分别提升了 1.6%、1.4%和 2.8%和同时 MAE 降低 2.467。掩码图的使用使网络能够集中关注漏缆卡扣区域,减少背景干扰,从而提高图像重构的逼真程度。

3)在同时加入 RE 和掩码图后,本文模型的 MAE 减少了 3.527,而 SSIM、COSIN 和 AUC 分别提升了 3.2%、3.1%和 4.1%。逐步引入各项改进措施后,本文模型表现出明显的性能提升趋势,表明改进后的模型能有效提升漏缆卡扣的重构质量。

表 8 中,本文还研究了不同数量的瓶颈残差块对于重构实验结果的影响,从中可以看出:

1)随着残差块数量的增加,模型的效效果逐渐增加。当 RE 数量为 4 时,MAE 分别为 3.026,SSIM 和 COSIN 分别为 0.946 和 0.951。

2)在对比 RE 数量为 3 和 4 时,SSIM 和 COSIN 的提升仅分别为 0.1% 和 0.2%,而 MAE 的降低幅度也仅为 0.119。这表明,当 RE 数量超过 3 时,模型性能的提升非常有限。此外,考虑到模型的计算复杂度和训练时间,为了高效且快速地进行漏缆卡扣重构,本文的方法选择了使用 3 个深度瓶颈层的残差块。

3 结 论

针对现行自监督卡扣缺陷检测方法由于重构背景误差而导致检测召回率低的问题,本文提出一种基于重点区域索引的双阶段漏缆卡扣重构缺陷检测算法。与基于分块重构 GAN 方法不同,本文为重构过程提供卡扣及漏缆区域掩码作为索引,指导网络对重点区域而非全局进行重构。此外,坐标注意力机制被嵌入到生成器中进一步强化网络对重点区域进行关注。更重要的是,双阶段重构过程被采用,通过对一阶重构前后图像并将其用于二次重构来修正背景噪声和卡扣细节,进一步保证重构前后图像背景区域在纹理和语义上的一致性。定量实验结果表明,相较于基于分块重构 GAN 方法的卡扣重构方法,本文所提的卡扣重构方法在结构相似度和余弦相似度指标上分别提高了 2.9% 和 1.1%,由此导致的卡扣缺陷检测精度、召回率和 AUC 分别提高了 1.2%、1.5% 和 1.1%。不同类型卡扣的可视化结果表明,所提的重构方法背景误差较小。对比其他自监督和监督缺陷检测方法,所提算法具有竞争力。

参考文献

[1] 屈毅,张卫军,杨锐. 铁路通信网密码技术应用研究[J]. 中国铁路,2023(10):91-96.
QU Y, ZHANG W J, YANG R. Research on application of railway communication network cryptography[J]. China Railway,2023(10):91-96.

[2] 张辉,宋雅男,王耀南,等. 钢轨缺陷无损检测与评估技术综述[J]. 仪器仪表学报,2019,40(2):11-25.
ZHANG H, SONG Y N, WANG Y N, et al. Review of rail defect non-destructive testing and evaluation [J]. Chinese Journal of Scientific Instrument, 2019,40(2): 11-25.

[3] LIU P H, FENG J R, GE W T, et al. A system-level performance evaluation for a 5G system under a leaky coaxial cable MIMO channel for high-speed trains in the railway tunnel[J]. Electronics,2022,11(8):1185.

[4] 何玉娟,宋鑫,李斌. 基于 RFID 定位的铁路隧道漏缆视频巡检方案[J]. 中国铁路,2018(11):104-108.
HE Y J, SONG X, LI B. RFID-based video inspection

plan for leaky coaxial cables in railway tunnels[J]. China Railway, 2018(11):104-108.

[5] 杜晓燕,常凯,许鹏飞. 高速铁路路基电缆槽对路基结构的影响分析及优化建议[J]. 铁道建筑,2019,59(3): 69-72.
DU X Y, CHANG K, XU P F. Analysis about influence of cable trough on subgrade structure of high speed railway and optimization suggestion[J]. Railway Engineering,2019,59(3):69-72.

[6] 张云佐,杨攀亮,李汶轩. 基于改进 SSD 算法的铁路隧道漏缆卡扣检测[J]. 激光与光电子学进展,2021, 58(22):391-398.
ZHANG Y Z, YANG P L, LI W X. Detection of cable leakage fixture in railway tunnel based on improved SSD algorithm[J]. Laser & Optoelectronics Progress, 2021, 58(22): 391-398.

[7] 王逸涵. 通信漏缆卡扣缺陷检测算法研究[D]. 成都:西南交通大学,2021.
WANG Y H. The research of railway communication cable fixtures detection algorithm [D]. Chengdu: Southwest Jiaotong University,2021.

[8] 张云佐,宋洲臣,杨攀亮. 基于 MD-LBP 关联方向特征的铁路隧道漏缆卡扣检测算法[J]. 北京交通大学学报,2021,45(5):101-107.
ZHANG Y Z, SONG ZH CH, YANG P L. Algorithm for detecting leaky cable buckle in railway tunnels based on MD-LBP and correlated directional features[J]. Journal of Beijing Jiaotong University, 2021,45(5):101-107.

[9] 彭一潇. 高速铁路隧道漏缆卡扣缺陷识别方法研究[D]. 北京:北京交通大学,2021.
PENG Y X. Research on defect identification method of leaky cable fixture in high-speed railway tunnel[D]. Beijing: Beijing Jiaotong University,2021.

[10] 马海月. 基于深度学习的隧道漏缆卡扣故障检测系统研究与实现[D]. 石家庄:河北科技大学,2020.
MA H Y. Research and realization of tunnel leaky cable fixture fault detection system based on deep learning[D]. Shijiazhuang: Hebei University of Science and Technology,2020.

[11] 李毅,许知建,王逸涵,等. 基于分块重构 GAN 的铁路隧道漏缆卡扣缺陷检测[J]. 铁道标准设计,2024, 68(2):176-184.
LI Y, XU ZH J, WANG Y H, et al. Defect detection of leakage cable fastener in railway tunnel based on block reconstruction GAN [J]. Railway Standard Design, 2024,68(2):176-184.

[12] ZAVRTANIK V, KRISTAN M, SKOČAJ D.

- Reconstruction by inpainting for visual anomaly detection [J]. Pattern Recognition, 2021, 112: 107706.
- [13] XIA D Q, YU ZH L, DENG X Y. A real-time unsupervised two-stage framework for fabric defect detection [C]. 2021 3rd International Academic Exchange Conference on Science and Technology Innovation(IAECST), 2021: 535-538.
- [14] KIM M, JO H, RA M, et al. Weakly-supervised defect segmentation on periodic textures using CycleGAN[J]. IEEE Access, 2020, 8: 176202-176216.
- [15] 高港,魏利胜,朱圣博. 正负样本差异特征双径向融合
的自监督缺陷检测方法[J]. 电子测量与仪器学报,
2024,38(5):201-209.
GAO G, WEI L SH, ZHU SH B. Self-supervised
defect detection based on biradial fusion of differential
features between positive and negative samples [J].
Journal of Electronic Measurement and
Instrumentation, 2024, 38(5):201-209.
- [16] JABBAR A, LI X, ASSAM M, et al. AFD-
StackGAN: Automatic mask generation network for
face de-occlusion using StackGAN[J]. Sensors, 2022,
22(5): 1747.
- [17] ZHAO Y, LYU W Y, XU SH L, et al. Detrs beat
yolos on real-time object detection [C]. IEEE/CVF
Conference on Computer Vision and Pattern
Recognition, 2024: 16965-16974.
- [18] QIN D F, LEICHNER C, DELAKIS M, et al.
MobileNetV4-universal models for the mobile
ecosystem [J]. ArXiv preprint arXiv: 2404.
10518, 2024.
- [19] HE K M, ZHANG X Y, REN SH Q, et al. Deep
residual learning for image recognition [C]. IEEE
Conference on Computer Vision and Pattern
Recognition, 2016: 770-778.
- [20] XU W, WAN Y. ELA: Efficient local attention for
deep convolutional neural networks [J]. ArXiv
preprint arXiv:2403.01123, 2024.

作者简介

任帅, 硕士研究生, 主要研究方向为机器视觉, 图像处理。

E-mail: rs00605@163.com

李柏林(通信作者), 博士, 教授, 主要研究方向为机器视觉、图像处理。

E-mail: blli62@swjtu.edu.cn