

一种基于分割驱动的螺栓缺陷增强和检测方法

Yangjie Xiao, Ke Zhang, *Member, IEEE*, Jiacun Wang, *Senior, IEEE*, Xin Sheng, Yurong Guo, Meijuan Chen, Zehua Ren, Zhaoye Zheng, Zhenbing Zhao, *Member, IEEE*,

Abstract—螺栓缺陷检测对于确保输电线路的安全至关重要。然而，缺陷图像的稀缺性和数据分布的不平衡极大地限制了检测性能。为了解决这个问题，我们提出了一种基于分割驱动的螺栓缺陷编辑方法 (SBDE) 来增强数据集。首先，提出了一种螺栓属性分割模型 (Bolt-SAM)，通过 CLAHE-FFT 适配器 (CFA) 和多部分感知掩码解码器 (MAMD) 增强复杂螺栓属性的分割，生成高质量的掩码供后续编辑任务使用。其次，设计了一个掩码优化模块 (MOD) 并与图像修补模型 (LaMa) 集成，以构建螺栓缺陷属性编辑模型 (MOD-LaMa)，通过属性编辑将正常螺栓转换为缺陷螺栓。最后，提出了一种编辑恢复增强 (ERA) 策略，将编辑后的缺陷螺栓恢复并放回原始检查场景中，从而扩展缺陷检测数据集。我们构建了多个螺栓数据集并进行了广泛的实验。实验结果表明，SBDE 生成的螺栓缺陷图像明显优于最先进的图像编辑模型，并有效提高了螺栓缺陷检测性能，充分验证了所提出方法的有效性和应用潜力。项目信息及代码可在 <https://github.com/Jay-xyj/SBDE> 获取。

Index Terms—Defect attribute editing, bolt defect generation, SAM, data augmentation, bolt defect detection.

I. 介绍

BOLTS 是连接和保护输电线路的关键组件，其状况直接影响电力系统的稳定性和安全性。由于长时间暴露在恶劣的高海拔环境中，螺栓容易因自然老化、环境腐蚀和外力而出现缺失或螺母等缺陷。因此，及时准确地检测螺栓缺陷对于确保输电线路的安全运行至关重要 [1]。传统的人工检查方法效率低下，难以满足大规模输电线路的需求。随着无人机智能检查技术的快速发展，基于深度学习的缺陷检测已成为输电线路运维的主要方法 [2]–[5]。然而，当前研究和实际工程中的缺陷图像数量严重不足，限制了数据驱动检测算法的性能。因此，有效扩展缺陷数据集是一个紧迫的研究课题。

近年来，一些研究集中于通过生成模型扩展缺陷数据集 [6], [7]。这些模型拟合原始数据的分布，通过从噪声中采样直接生成缺陷图像。SDGAN [8] 提出了一种基于生成对抗网络 (GANs) 表面缺陷图像生成方法。它从工业无缺陷图像中生成高质量和多样化的缺陷图像，以扩展缺陷数

据集。然而，这些方法通常依赖于足够的缺陷图像进行训练，并且在生成过程中可能引入意外的随机噪声，导致输出偏离工业场景。这种直接生成方法缺乏对缺陷属性的精确控制，因此难以满足复杂场景中目标检测的应用需求。

值得注意的是，虽然缺陷图像较为稀少，但正常图像却足够丰富，并且与缺陷图像共享高度的特征相关性。因此，通过对正常图像进行属性编辑来生成缺陷图像是一种更高效且极具前景的解决方案。然而，实现这些方法在捕捉少样本缺陷特征时面临挑战。目前，将正常图像转换为缺陷图像已成为工业领域的研究热点。DFMGAN [9] 提出了针对少样本场景的缺陷图像生成方法。通过无缺陷图像上训练数据高效的 StyleGAN2 [10] 并结合一个缺陷感知的残差模块，它实现了高质量和多样化的缺陷图像生成。尽管这些方法已取得良好效果，但它们通常适用于结构简单或背景单一的图像。它们专注于将一类正常图像转换为相应的缺陷图像，而不是对同一幅正常图像进行精确的属性编辑 [11]–[13]。即使通过生成采样添加缺陷并重建原始图像，原始图像也不可避免地会被改变 [14]–[19]。

更重要的是，这些现有的缺陷转换方法不适用于电力领域的螺栓缺陷场景。根本原因在于螺栓图像的复杂性和质量较差，这导致了不同的缺陷转换方式。螺栓具有高度规则的几何形状和缺陷类型，要求编辑方法能够在满足工业标准的同时实现高精度属性转换 [6], [20]。此外，螺栓图像通常较小，螺栓与背景之间的像素对比度极低，大大增加了生成模型精确编辑目标区域的难度。与背景简单的表面缺陷不同，螺栓图像生成中的偏差可能直接导致结构失真，严重影响编辑质量。因此，实现螺栓的精确属性编辑需要克服许多挑战，包括目标区域的精确定位、保持背景一致性以及确保缺陷属性的精确编辑。

为了解决上述问题，我们提出了 SBDE 模型，该模型集成了 Bolt-SAM 用于螺栓属性分割，以及 MOD-LaMa 用于缺陷编辑，如图 1 所示。SBDE 通过精确分割和编辑正常螺栓的属性，将原始正常螺栓的图像转换为相应的缺陷图像。编辑后的缺陷图像通过 ERA 用于增强检测数据集，以提高缺陷检测的准确性。本文的贡献总结如下：

- 1) We propose the first ever bolt attribute segmentation model, named Bolt-SAM. We designed CFA to enhance the extraction of bolt edge features and proposed MAMD to improve the segmentation accuracy of complex bolt attributes through multi-task fine-tuning.
- 2) We design MOD and integrate it with LaMa to construct MOD-LaMa for bolt attribute removal. By combining Bolt-SAM with MOD-LaMa, we propose the first bolt defect attribute editing method, SBDE, which realizes the zero-defect-shot attribute editing and converts normal bolts into defect bolts.
- 3) We develop a defect detection data augmentation strategy, ERA, which seamlessly restores the edited bolt defect images into the original inspection scenes

Research supported by the National Natural Science Foundation of China (NSFC) under grant numbers 62076093, 61871182, 62206095, by the Fundamental Research Funds for the Central Universities under grant numbers 2022MS078, 2023JG002, 2023JC006, by the Natural Science Foundation of Hebei Province under grant number F2024502017, and by the Graduate Student Innovation Ability Training Funding Project of Hebei Provincial Department of Education under grant number CXZZBS2024167.

Y. Xiao, K. Zhang, X. Sheng, Y. Guo, M. Chen, Z. Ren, Z. Zheng and Z. Zhao are with the Department of Electronic and Communication Engineering, North China Electric Power University, Hebei, 071003, P. R. China. K. Zhang, Y. Guo and Z. Zhao are also with Hebei Key Laboratory of Power Internet of Things Technology, North China Electric Power University, Baoding 071003, Hebei, China. K. Zhang is the corresponding author (e-mail: zhangkeit@ncepu.edu.cn).

J. Wang is with the Computer Science and Software Engineering Department, Monmouth University, West Long Branch 07764, USA. (e-mail: jwang@monmouth.edu).

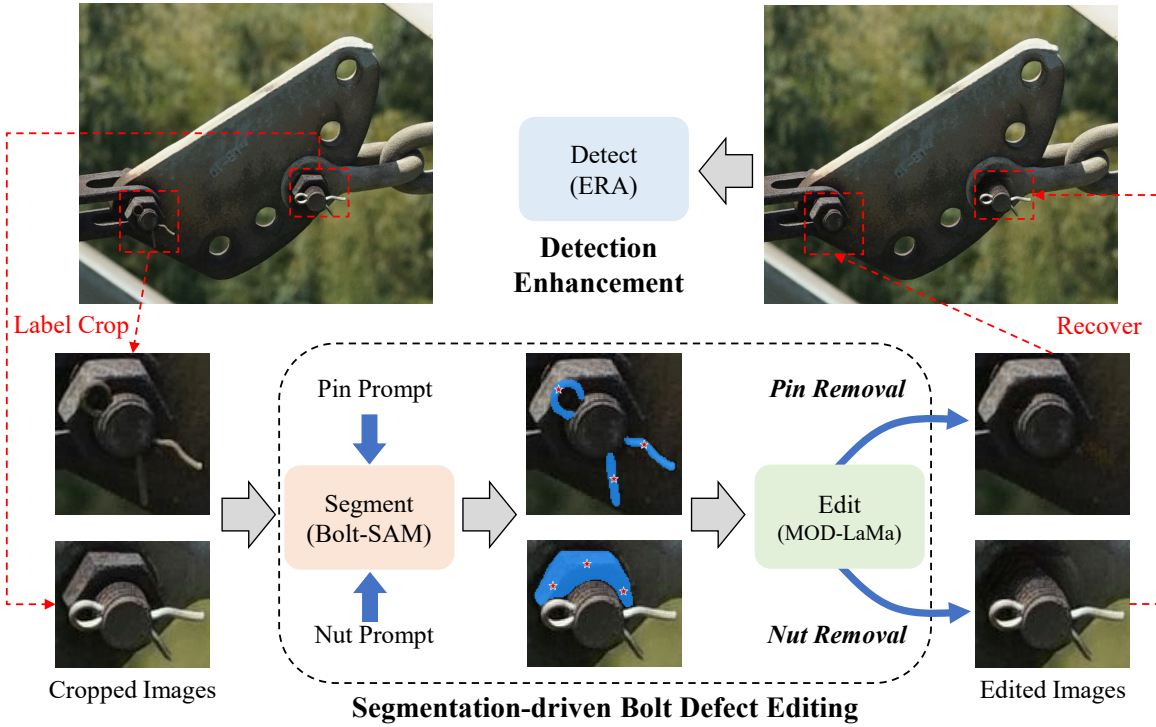


Fig. 1. 所提出方法的整体框架。SBDE 结合分割和编辑生成缺陷，以增强缺陷检测数据集。

to expand the dataset. The effectiveness of our method is validated through detection experiments.

本文的其余部分组织如下。我们首先在第二节回顾相关工作。然后，第三节详细描述了我们的方法。之后，第四节介绍了我们的实验和实验结果。最后，我们在第五节总结这项工作并讨论未来的工作。

II. 相关工作

A. 任意分割模型

Segment Anything Model (SAM) [21] 是一个强大的基础分割模型，支持多模态输入，包括点提示、框提示和掩码提示。它能够在未知领域执行零样本分割，并在各种视觉任务中表现出色。其通用性和高效性迅速使其成为分割领域研究的焦点，激发了众多研究以增强其在特定任务中的表现 [22], [23]。HQ-SAM [24] 通过融合早期和最终的 ViT 特征引入高质量的输出令牌，提高了 SAM 的分割性能，同时保持了其效率和零样本泛化能力。SAM-Adapter [25] 提出了一种简单而高效的适配器机制，将领域特定信息与视觉提示进行整合，在不需要对 SAM 网络进行微调的情况下增强了其在复杂分割任务中的表现。对于低质量图像场景，RobustSAM [26] 提出了一系列抗降质模块，以增强 SAM 在各种图像降质条件下的分割鲁棒性。

此外，一些研究进一步探索了提示对 SAM 的影响。SAMAUG [27] 提出了一种视觉点提示增强方法，通过生成增强的点提示为 SAM 提供更多用户意图信息，从而提高交互式图像分割性能。此外，SAM 的强大功能已被扩展到其他任务。Inpaint Anything (IA) [28] 提出了一种新的“点击填充”范式，将 SAM 与各种模型结合，实现目标区域的精确定位和修复。结合 SAM 及其相关优化方法，为解决螺栓缺陷编辑中属性定位和分割的挑战提供了有力支持。

B. 图像编辑

图像编辑一直是计算机视觉中的一个关键领域。作为经典的生成模型，GAN 在图像编辑任务中得到了广泛应用 [29]。StarGAN [30] 和 StarGAN v2 [31] 提出了多域翻译框架，支持多目标属性的转换。StyleGAN [32]、StyleGAN2 [10] 和 StyleGAN2-ADA (ada) [33] 通过结合风格编码和生成器架构，在图像编辑中设立了一个基准，允许生成具有语义解耦的高质量图像。此外，基于 StyleGAN 的逆向方法，如 pSp [17]、e4e [18] 和 ReStyle [19]，通过潜在空间优化实现了真实图像的精确还原和灵活编辑。

近年来，扩散模型的兴起进一步推动了生成技术的发展。DDPM [34] 首次引入了基于马尔科夫链的扩散过程，为这个领域奠定了理论基础。随后，DDIM [35] 通过加速采样优化了生成效率，并在潜在空间中实现了语义操作。稳定扩散 (SD) [36] 作为一个开源扩散模型，由于其高效的潜在空间操作，广泛用于生成和编辑任务，进一步推动了扩散模型的研究。在此基础上，提出了多种微调和反演方法，以满足不同任务的需求 [37]–[39]。DreamBooth (DB) [38] 从一小组参考图像生成高保真图像，同时保持主体特征。关于扩散反演方法，Null-text Inversion [40] 结合枢轴反演和无文本优化，以支持灵活的文本驱动的图像编辑。负提示反演 [41] 通过前向传播实现了高效编辑，而无需复杂的优化，而直接反演 [42] 采用分支优化策略，大幅加速推理同时提高编辑精度。

鉴于现有生成模型的优越性能，许多研究人员已将其应用于缺陷生成任务。SDGAN [8] 从无缺陷的图像中生成高质量和多样化的缺陷样本，以增强数据集的多样性。Defect-GAN [11] 利用层次结构实现了多样化背景和随机缺陷的高保真生成。DFMGAN [9] 结合了 StyleGAN2 和缺陷感知残差模块，在低样本情况下生成多样化的缺陷图像。StableSDG [13] 调整了 Stable Diffusion 模型的参数

和引导图像生成过程,以生成高保真度的钢表面缺陷图像。

尽管上述方法在解决工业缺陷时表现良好,但它们并不适合于螺栓缺陷编辑。与一般的表面缺陷不同,螺栓缺陷通常表现为关键属性的丧失,这需要编辑方法在保持背景一致性和保留原始结构特征的同时,精确移除目标区域。面对螺栓的专业性和复杂性,基于生成模型的属性编辑是一项巨大挑战。相比之下,基于图像修复的方法提供了一个可行的替代方案。LaMa [43] 是一个高性能的图像修复模型,通过结合快速傅里叶卷积和大掩码训练,提高了修复质量。此外,IA [28] 将 SAM 与 LaMa 结合以移除目标区域。受此启发,我们提出了 SBDE。通过精确定位并移除正常螺栓的属性,SBDE 不仅实现了有效的螺栓缺陷编辑,还为工业应用中的缺陷转换任务提供了一种新的解决方案。

C. 螺栓缺陷检测

螺栓缺陷检测是确保输电线路安全运行的关键技术之一。近年来,基于深度学习的目标检测方法已广泛应用于该领域。AVSCNet [20] 提出了自动视觉形状聚类网络,以解决输电线路中螺栓形状的复杂性。通过利用无监督聚类结合特征增强、融合和区域特征提取,它显著提高了缺失销钉的检测能力。UBDDM [1] 通过设计超小目标感知模块和多尺度特征融合策略,解决了超小目标检测问题。其两阶段检测方法能够以最低的标注成本实现精确的螺栓缺陷检测。PA-DETR [44] 集成了电力知识,将相对位置编码和多尺度特征提取模块引入 DETR [45],并结合属性和缺陷分类器来提高在视觉上具有挑战性缺陷场景中的检测准确性。

尽管现有方法在优化检测模型结构方面取得了进展,但在数据层面的探索仍然相对有限。当螺栓缺陷图像不足或数据分布不平衡时,这些方法的性能会显著下降。为了解决这个问题,我们提出了 ERA 策略来扩展缺陷检测数据集,从而有效提高不同场景下螺栓缺陷检测的性能。

III. 方法

为了解决输电线路中正常和缺陷螺栓样本间的严重不平衡问题,特别是缺陷样本的极度稀缺,我们提出了一种名为 SBDE 的分割驱动螺栓缺陷编辑方法。图 1 展示了我们提出方法的框架,其中 SBDE 包含两个核心步骤:分割和编辑。它从正常螺栓图像生成缺陷样本,并使用 ERA 策略将其恢复到检测图像中,增加数据集并提高螺栓缺陷检测模型的准确性。

在 SBDE 模型中,分割阶段的目标是从正常的螺栓图像中准确提取关键属性掩膜,为后续的缺陷编辑提供精确的分割区域。为此,我们提出了一种基于基线 RobustSAM 改进的分割模型 Bolt-SAM,其结构如图 2 所示。以螺栓的销属性为例,输入图像首先通过 CFA 模块,在图像编码器中引入自适应频域信息,从而增强图像特征。然后在掩膜解码器中,销被分为三个部分,并基于相应的点提示进行多任务微调,为每个部分生成局部掩膜。最后,一个融合网络整合每个部分的局部掩膜以生成完整的销掩膜,从而实现螺栓关键属性的准确分割。

复杂的现实环境条件,如光照变化、背景噪声和低对比度区域,会削弱模型捕捉螺栓边缘的能力,导致螺栓图像中关键属性的分割不准确。针对这个问题,我们设计了 CFA 模块,如图 2 中图像编码器左侧所示。CFA 通过简单

的 MLP 将自适应频域信息引入 SAM 的图像编码器,以增强特征表示。

具体来说,给定输入图像 $I \in \mathbb{R}^{H \times W \times C}$,CFA 首先通过对比度有限自适应直方图均衡化 (CLAHE) [46] 来增强细节和边缘特征,得到优化的图像 I_{cla} 。然后, I_{cla} 通过快速傅里叶变换 (FFT) 转换到频域,并表示为 $z = \text{FFT}(I_{cla})$ 。CFA 利用高频掩码 M_{hpf} 构建一个高通滤波器 (HPF),以提取高频边缘信息。根据 [47] 中高频掩码的设计, M_{hpf} 定义为:

$$M_{hpf}^{i,j}(\tau) = \begin{cases} 0, & \frac{4|(\frac{i}{H}-\frac{H}{2})(\frac{j}{W}-\frac{W}{2})|}{HW} \leq \tau \\ 1, & \text{otherwise} \end{cases} \quad (1)$$

其中 H 和 W 分别表示图像的高度和宽度,而 τ 控制低频区域的比例。然后使用逆快速傅里叶变换 (IFFT) 将高频信息转换回空间域,得到相应的高频分量 $I_{hfc} = \text{IFFT}(z \cdot M_{hpf})$:

最后, I_{hfc} 与 Patch Embedding 的输出相加,以获得适用于 Adapter 的输入特征 F_i 。该 Adapter 由两个线性层 (FC_{tune} 和 FC_{up}) 以及激活函数 GELU [48] 组成,用于优化特征表示和调整维度。Adapter 的最终输出是提示向量 P^i ,定义为:

$$P^i = \text{FC}_{up}(\text{GELU}(\text{FC}_{tune}^i(F_i))) \quad (2)$$

其中, FC_{tune} 是每个 Adapter 独特的可调线性层,用于生成与当前层特征相关的提示。 FC_{up} 是所有 Adapter 共享的上投影层,用于将提示向量调整为一个与 Transformer 输入特征匹配的嵌入维度。 F_i 是输入的线性映射的总和。 P^i 是附加到每个 Transformer 层上的输出提示,用于有效地指导特征学习。

1) 多部件感知掩码解码器:螺栓属性的结构和视觉复杂性增加了 SAM 分割的难度,尤其是销钉在图像中常常以分散形式出现。为了解决这个问题,我们提出了一种称为 MAMD 的分段微调架构,图 2 右侧展示了该架构。MAMD 采用多任务微调策略,将销钉分成三个部分 (pin0, pin1, pin2) 进行独立优化,并通过一个融合网络生成完整的销钉掩膜。

SAM 的框提示不适用于输电线路螺栓图像中的属性分割。由于螺栓图像的低分辨率以及像素之间缺乏显著对比,使用框提示往往会引入来自非目标区域的过多信息,干扰分割。相比之下,点提示提供了更精确的指导,使模型能够聚焦于目标区域。它对于非连续结构尤其有利。因此,我们基于点提示微调了 Bolt-SAM。通过为螺栓销的每个局部区域提供点提示,我们指导模型优化区域特征,以准确捕捉边界信息和细节特征,从而提高非连续复杂结构的处理精度。MAMD 引入了 Focal Loss 和 Dice Loss 作为损失函数。对于第 i 个子区域,局部损失函数定义为:

$$L_{local,i} = \alpha \cdot L_{focal}(M_i, GT_i) + \beta \cdot L_{Dice}(M_i, GT_i) \quad (3)$$

其中 M_i 和 GT_i 分别表示第 i 个子区域的预测掩码和真实掩码,而 α 和 β 是损失函数的权重系数。MAMD 使用一个具有两个卷积层和批量归一化层的轻量级融合网络将局部掩码特征整合为完整的销帽掩码。全局掩码的优化与局部损失函数相同。

MAMD 通过多任务微调有效解决了螺栓销非连续结构的分割问题,为后续的螺栓缺陷编辑提供了高质量的分割结果。值得注意的是,对于螺母的分割,我们只采用了一个分支进行微调。

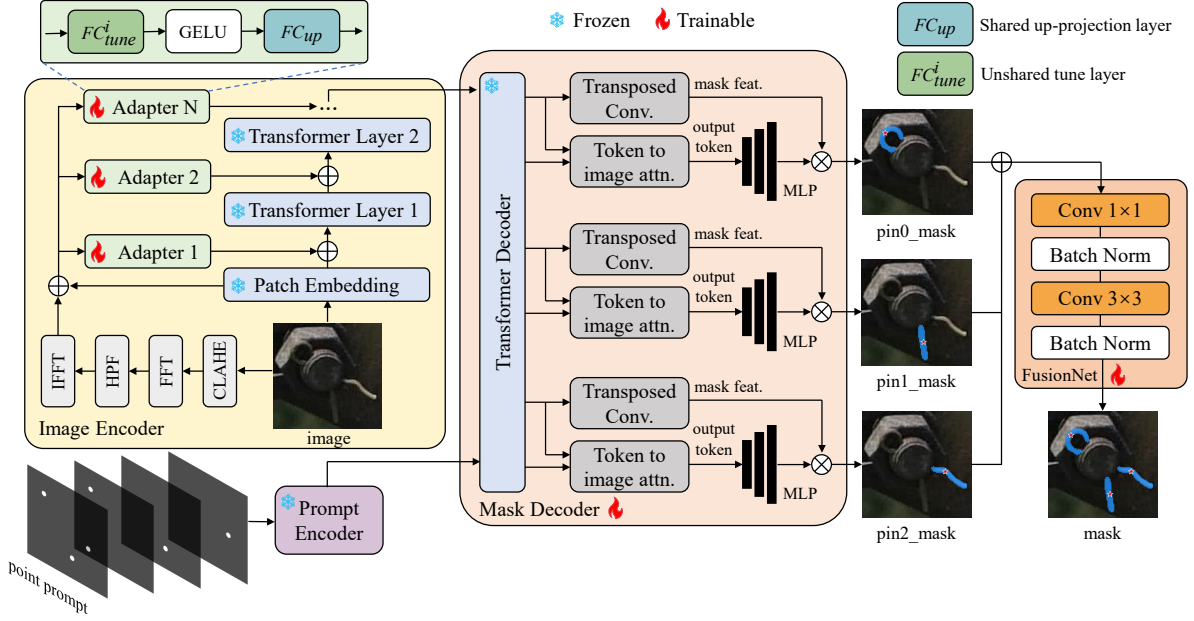


Fig. 2. 我们所提出的 Bolt-SAM 架构。我们在每个 transformer 层中加入了适配器模块，以增强特征，并采用多任务微调掩码解码器架构来提高属性分割的准确性。一个简单的 MLP 构建的融合网络被用来整合完整的属性掩码。原始图像编码器和提示编码器的参数是冻结的，而适配器模块、掩码解码器和融合网络设置为可训练的。

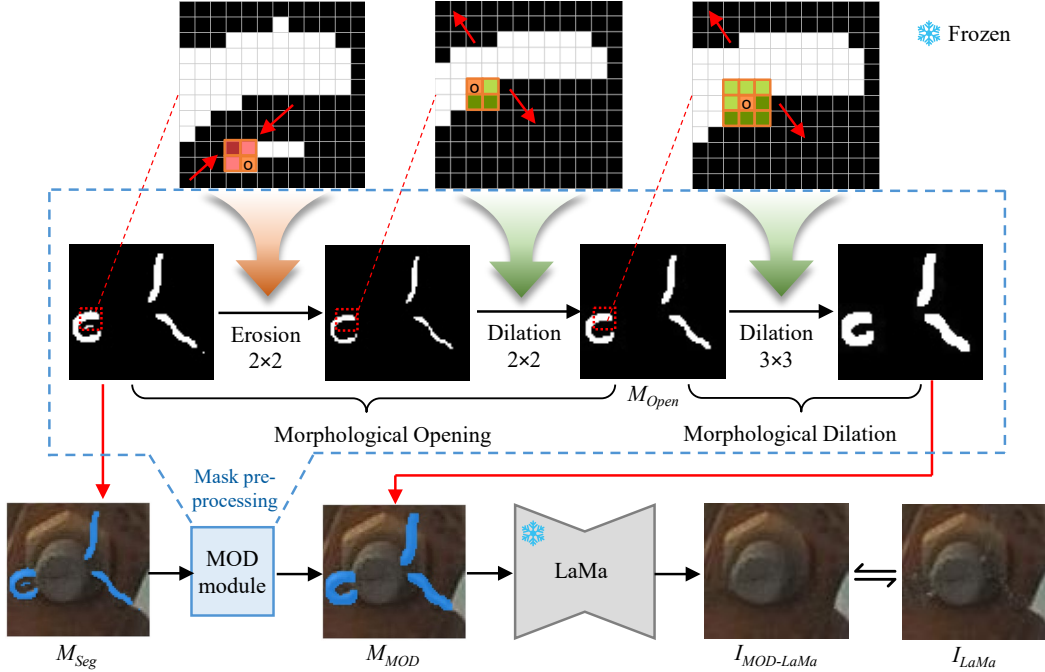


Fig. 3. MOD-LaMa 的结构。结构元素 s 的交互规则：浅色区域显示与掩码的交集，而深色区域表示没有重叠。对于侵蚀，只有当 s 的所有像素都适合在掩码内时，才保留 o （任何深红色将 o 设为 0）。对于膨胀，只要 s 的任何部分与掩码重叠，即保留 o （任何深绿色将 o 设为 1）。

A. 编辑：MOD-LaMa

螺栓特性编辑通过将正常螺栓转换为缺少销钉或螺母的缺陷螺栓来解决缺少缺陷样本的问题。受到 IA 的启发，我们采用创新的图像修复方法进行编辑。LaMa 利用其整合局部特征与全局上下文的能力，生成与背景一致的内容，使其特别适合消除螺栓缺陷。由于螺栓图像的低分辨率，Bolt-SAM 生成的掩膜通常缺乏足够的上下文覆盖，限制

了 LaMa 修复的效果。我们提出 MOD-LaMa，如图 3 所示，该方法采用形态学操作来优化掩膜。通过应用开运算和膨胀运算，该方法平滑边界并扩展目标区域，增强上下文信息。

具体来说，MOD 首先对分割掩膜执行形态学开运算（先腐蚀后膨胀），定义如下：

$$M_{open} = (M_{seg} \ominus s) \oplus s \quad (4)$$

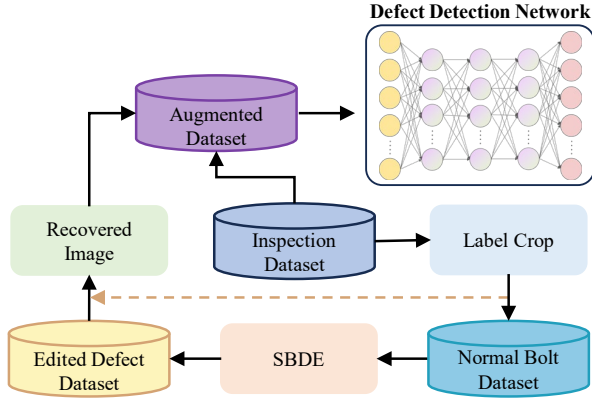


Fig. 4. ERA 的工作流程包括检查图像裁剪、螺栓编辑和缺陷恢复。它增强了用于缺陷检测网络的检查数据集。

，其中 M_{seg} 是初始分割掩膜， s 是控制腐蚀和膨胀程度的结构元素。 $\ominus$ 表示腐蚀运算，用于缩小目标区域的边界并从掩膜中去除噪声。 $\oplus$ 表示膨胀运算，通过扩展边界恢复目标区域的完整性。腐蚀和膨胀的定义如下：

$$(M_{seg} \ominus s)(x, y) = \{o \mid (s)_o \subseteq M_{seg}\} \\ = \min_{(x,j) \in s} M_{seg}(x+i, y+j) \quad (5)$$

$$(M_{open} \oplus s)(x, y) = \{o \mid (s)_o \cap M_{open} \neq \emptyset\} \\ = \min_{(x,j) \in s} M_{open}(x+i, y+j) \quad (6)$$

，其中 (x, y) 是当前结构元素的起始坐标， (i, j) 表示相对于起始点的偏移量。 s_o 是掩膜内结构元素的空间平移。然后，MOD 进一步对 M_{open} 执行膨胀运算，生成最终优化的掩膜。

$$M_{MOD} = M_{open} \oplus s \quad (7)$$

最后，LaMa 使用 MOD 优化的掩膜 M_{MOD} 与原始图像 I_{ori} 一起修复目标区域，有效实现螺栓缺陷属性编辑。

$$I_{edit} = \text{LaMa}(I_{ori}, M_{MOD}) \quad (8)$$

B. 检测：ERA

为了进一步验证 SBDE 的有效性，我们提出了 ERA 策略。ERA 基于检验图像，通过编辑缺陷样本并将其恢复到原始检验场景中来扩展螺栓缺陷检测数据集。整体过程如图 4 所示。具体来说，我们使用 LabelImg 标注检验图像中的正常螺栓 I_{ins} ，将其裁剪以提取单个正常螺栓 I_{ori} ，并使用 SBDE 编辑它们以生成缺陷螺栓 I_{edit} 。最后， I_{edit} 被恢复到它在 I_{ins} 中的原始位置，并将标签更新为缺陷，从而形成缺陷增强的检验图像 I_{aug} 。恢复操作如下：

$$I_{aug}(i, j) = \begin{cases} I_{edit}(x', y'), & \text{if } (x, y) \cap R_{box} \\ I_{ins}(x, y), & \text{otherwise} \end{cases} \quad (9)$$

其中 R_{box} 是边界框， (x', y') 指的是 I_{edit} 中的相应坐标映射。

经 ERA 增强的检测图像完好地保留了原始输电线的场景信息，只在局部螺栓区域生成缺陷。这使检测模型能够在同一背景下学习多样的缺陷特征，丰富了缺陷检测的数据分布，从而提高了检测性能。此外，ERA 在数据增强过程中避免过度依赖真实缺陷数据，为少样本（甚至零样本）缺陷的螺栓检测场景提供了一个有效的解决方案。

TABLE I
BDD 数据集的数据分布。

Class	Train	Test	Total
Inspection images	1433	337	1770
Number of instances			
Normal	3961	1018	4979
Pin losing	449	100	549
Nut losing	244	63	307
All	4654	1181	5835

TABLE II
BDG 数据集的数据分布。

Class	Number of bolt images
Normal	1393
Pin losing	517
Nut losing	300
All	2253

TABLE III
BAS 数据集的数据分布。

Attribute masks of normal bolts	Train	Test
Nut_mask	886	220
Pin_mask	886	220
Pin0_mask	886	220
Pin1_mask	886	220
Pin2_mask	886	220

IV. 实验

A. 数据集

我们通过无人机巡检采集输电线路的航空影像，并使用 LabelImg 工具标注图像中的螺栓，构建了螺栓缺陷检测数据集 (BDD)。BDD 包括三种螺栓状态：正常、销失和螺母失，这些用于场景级别缺陷检测。在此基础上，根据标注信息裁剪巡检图像以提取单个螺栓图像，进一步筛选后构建螺栓缺陷生成数据集 (BDG)。BDG 支持其他缺陷生成方法的训练，以便与我们的方法进行比较。SBDE 专注于正常螺栓的分割和编辑，仅利用来自 BDG 的正常螺栓图像，并排除任何缺陷样本。为此，我们从 BDG 筛选出正常螺栓图像，并使用 Labelme 工具为其销和螺母属性带有遮罩进行标注，以构建螺栓属性分割数据集 (BAS)。在 BAS 中，销属性被精细分为三部分：头部 (pin0)、左销 (pin1) 和右销 (pin2)。这三个数据集针对不同的任务和应用场景，其示例如图 5 所示，具体数据分布详见表 I、表 II 和表 III。

在训练过程中，Bolt-SAM 仅微调 CFA 和 MAMD 的参数，而冻结其余参数。Bolt-SAM 采用点作为提示，从真实掩码中随机采样三个点进行训练。训练配置包括学习率为 0.0001，批量大小为 2，总共 50 个轮次。实验在两块 NVIDIA 4090 显卡上进行，使用 PyTorch 2.4、CUDA 12.1 和 Python 3.10 实现。在 MOD-LaMa 中，开运算使用不对称的 2×2 结构元素，而膨胀运算则采用中心对称的 3×3 结构元素。MOD-LaMa 基于预训练的 LaMa 直接进行推理，不进行额外的微调。在 ERA 中，使用最先进的 YOLOv11 [49] 检测算法来评估 SBDE 的有效性。

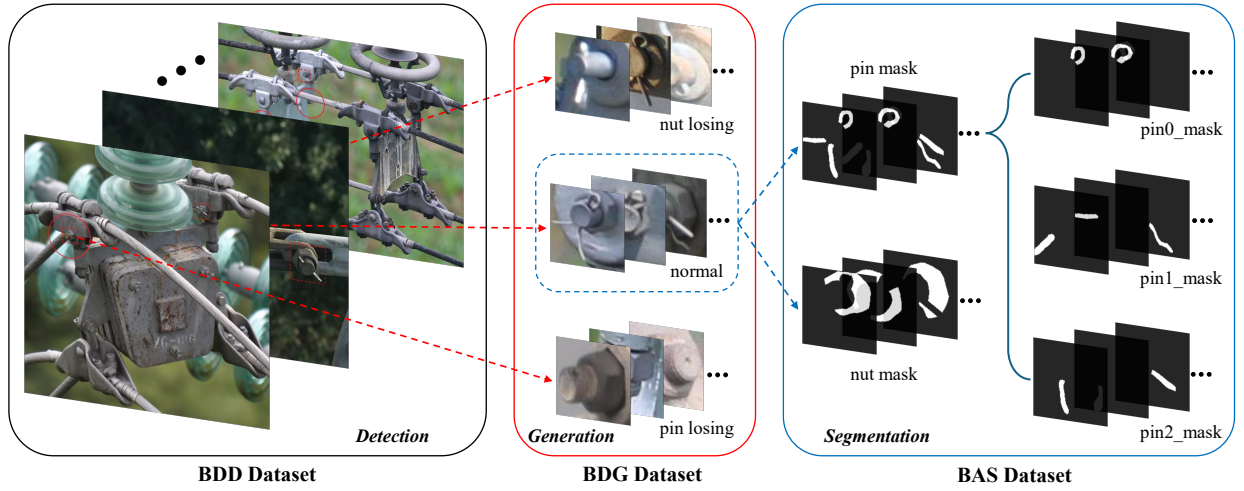


Fig. 5. 三个数据集的构建过程。它们通过标注、裁剪和过滤输电线路检修图像获得，其中 BAS 只包含正常的螺栓，并且销的属性被分成多个掩码。

为全面评估 SBDE 的性能，我们为不同任务选择了多种评估指标。对于分割任务，使用平均交并比 (mIoU)、Dice 系数 (Dice) 和像素准确率 (PA) 来评估螺栓属性掩膜的质量。对于检测任务，使用平均平均精度 (mAP)、准确率 (P) 和召回率 (R) 来验证 ERA 的螺栓缺陷检测性能。

由于 SBDE 与传统生成模型在编辑上的显著差异，在比较和消融实验中采用了不同的评估指标。对于比较实验，我们侧重于评估编辑后图像的质量及其与原始图像的整体相似性。峰值信噪比 (PSNR)、结构相似性指数测量 (SSIM) 和学习感知图像补丁相似性 (LPIPS) 被用于衡量编辑后图像与原始图像之间的差异。

然而，在消融实验中，SBDE 只编辑掩码区域，这使得传统指标不足以评价螺栓编辑的实际效果。为了解决这个问题，我们引入了两个关注属性编辑表现的指标：属性编辑准确率 (AEA) 和人类偏好得分 (HPS)。其中，AEA 使用一个预训练的螺栓缺陷分类模型来评估编辑后的螺栓图像是否能够被正确分类为目标缺陷类型，从而衡量缺陷转换的准确度。定义如下：

$$AEA = \frac{\sum_{i=1}^N \mathbb{I}(f_{cls}(I_{edit}^i) = y_{target})}{N} \quad (10)$$

其中 N 代表编辑图像的总数， I_{edit} 表示编辑过的图像， y_{target} 指目标缺陷类别， $f_{cls}(\cdot)$ 为预训练的螺栓缺陷分类模型， $\mathbb{I}(\cdot)$ 是一个指示函数，如果编辑正确则等于 1，否则为 0。HPS 是与人类视觉判断更一致的一个指标。具体规则如下：对于由 M 组消融实验配置编辑的缺陷图像，专家组对其进行排序并赋予分数 $S_{i,k} = \{1, 2, \dots, M\}$ ，其中每个分数是唯一的。分数越高，编辑性能越好。HPS 定义为：

$$HPS = \frac{\sum_{i=1}^N \sum_{k=1}^K S_{i,k}}{N \cdot K} \quad (11)$$

其中 $S_{i,k}$ 代表专家 i 给第 k 张图像的评分， N 表示专家的人数， K 代表编辑图像的总数， M 指消融实验配置的数量。

为了验证我们方法中每个模块的具体贡献，我们进行了两项消融研究，评估了 Bolt-SAM 的分割性能和 SBDE 的编辑效果。所有实验均在 BAS 数据集上进行。

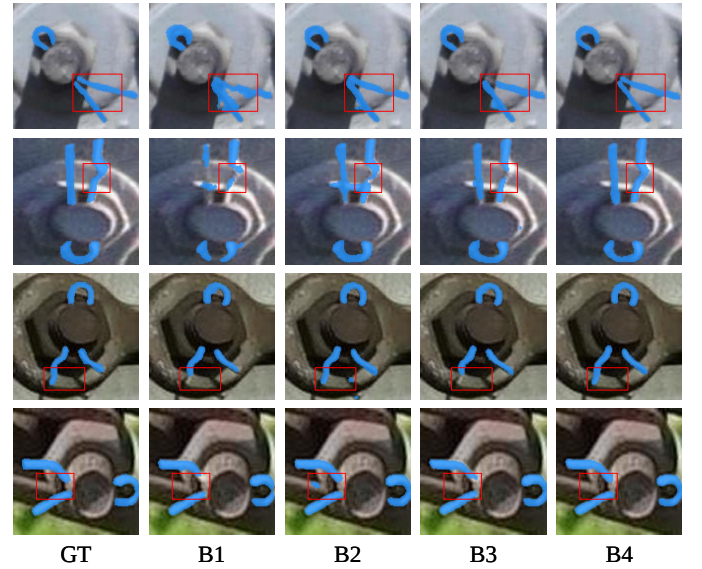


Fig. 6. 螺栓-SAM 分割基于销属性的消融结果可视化

1) 关于 Bolt-SAM 分割的消融研究：Bolt-SAM 是 SBDE 的核心算法，侧重于生成高质量的螺栓属性蒙版。为了验证 CFA 模块和 MAMD 架构的有效性，我们进行了分割消融实验，定量结果见表 IV。与基线模型 (B1) 相比，所提出的模块在所有指标上均显示出显著的改进。CFA 模块 (B2) 在螺母属性上表现出特别显著的增强，mIoU、Dice 和 PA 分别增加了 6.34、4.57 和 1.95。这是因为螺母位于螺栓的底部，容易由于光照造成的阴影效应而导致边缘模糊和不清晰，而 CFA 能有效地减轻这种影响。MAMD (B3) 是为销设计的多阶段架构，实现了更精确的销部件定位，并在 PA 指标上优于 B2。值得注意的是，PA 衡量整体像素准确性，由于螺栓图像尺寸小，PA 的改善相对有限，尤其在销属性上。总体而言，Bolt-SAM (B4) 结合了 CFA 和 MAMD，平衡了螺栓属性的独特特征，并在所有指标上取得了最佳结果。图 6 展示了改进模块在销分割上的可视化结果，其中 B4 达到了最佳分割效果。

TABLE IV
Bolt-SAM 分割的消融研究结果。最佳结果用粗体标记。

Method	CFA	MAMD	Pin			Nut		
			mIoU(%) $\uparrow$	Dice(%) $\uparrow$	PA(%) $\uparrow$	mIoU(%) $\uparrow$	Dice(%) $\uparrow$	PA(%) $\uparrow$
B1			71.16	82.80	97.48	70.44	81.90	92.76
B2	$\checkmark$		74.83	85.43	97.86	76.78	86.47	94.71
B3		$\checkmark$	73.75	84.72	97.90	-	-	-
B4	$\checkmark$	$\checkmark$	75.66	86.01	98.02	-	-	-

TABLE V
SBDE 编辑的消融研究结果。最佳结果以加粗标记。

Method	Bolt-SAM	MOD-LaMa	Pin		Nut	
			AEA(%) $\uparrow$	HPS $\uparrow$	AEA(%) $\uparrow$	HPS $\uparrow$
S1			11.83	1.17	9.92	1.08
S2	$\checkmark$		13.17	1.83	11.78	2.14
S3		$\checkmark$	78.63	3.2	69.82	2.87
S4	$\checkmark$	$\checkmark$	85.89	3.8	77.17	3.91

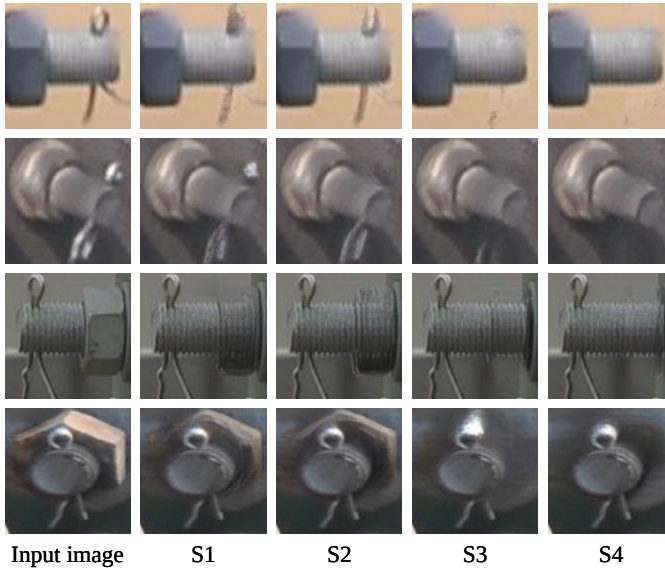


Fig. 7. SBDE 编辑的消融结果可视化

2) *SBDE* 编辑的消融研究：SBDE 通过结合 Bolt-SAM 和 MOD-LaMa 实现精确的螺栓缺陷属性编辑。为了验证这两个核心模块的贡献，我们设计了编辑任务的消融实验，采用 AEA 和 HPS 作为评价指标，如第 4.3 节所述。实验包括四种消融配置，每种配置生成 300 张编辑后的图像。这些评分由 20 位领域专家提供。定量结果如表 V 所示，图 7 展示了不同模块配置下的编辑结果的视觉比较。

基线 (S1) 使用 RobustSAM 和 LaMa 的组合，其编辑效果最差，AEA 和 HPS 均处于最低水平。引入 Bolt-SAM (S2) 后，分割精度略有提高，但总体编辑效果仍不佳，尤其是在正确编辑属性方面 (如图 7 所示，适用于 S1 和 S2)。相对而言，加入 MOD-LaMa (S3) 后，编辑性能显著提升，表明形态学优化在细化掩模区域和提高编辑质量方面起到了关键作用。AEA 的大幅提升表明属性得到了有效编辑 (如图 7 所示，适用于 S3)。SBDE (S4) 结合了 Bolt-SAM 和 MOD-LaMa 的优势，实现了所有指标的最佳性能。具

体而言，销钉和螺母的 AEA 分别达到 85.89 和 77.17，而 HPS 得分分别达到 3.8 和 3.91 的最高值，显示出卓越的编辑结果 (如图 7 所示，适用于 S4)。

总之，Bolt-SAM 增强了分割质量，MOD-LaMa 优化了编辑区域。它们的协同整合显著提高了 SBDE 在螺栓缺陷属性编辑中的表现。

为了全面验证 SBDE 在分割和编辑方面的性能优势，我们与先进的分割和编辑方法进行了详细的比较。这些对比实验分为两部分：Bolt-SAM 的分割性能比较和 SBDE 的编辑效果比较。

在分割中，我们将 Bolt-SAM 与 SAM、SAM-Adapter、HQ-SAM 和 RobustSAM 进行了比较，同时探讨了点提示数量对螺栓属性分割性能的影响。需要注意的是，点提示指的是为每个销零件设置的正提示点的数量。实验是在 BAS 数据集上进行的，并对所有比较方法进行了微调。分割性能通过三个指标来评估：mIoU、Dice 和 PA。

定量结果，如表 VI 所示，表明 Bolt-SAM 在所有指标上均显著优于其他所有方法，在使用 3-点提示时达到最佳性能 (这是本文其他实验中的默认设置)。与基线相比，Bolt-SAM 在销钉分割中分别将 mIoU、Dice 和 PA 提高了 4.50、3.21 和 0.54，并在螺母分割中提高了 6.34、4.57 和 1.95。

进一步分析点提示数量对模型性能的影响表明，增加提示数量通常能提高分割结果。这种效果在 Bolt-SAM 的引脚分割中尤为明显，相较于 1 点提示，使用 3 点提示分别使 mIoU、Dice 和 PA 提高了 2.93、3.13 和 0.41。这是因为螺栓引脚的像素太小，更多的提示可以帮助模型更好地捕捉目标区域的边界。在 SAM 的引脚分割中则相反，这反映出 SAM 的强泛化能力。

图 8 显示了不同方法的分割结果。可以观察到，Bolt-SAM 生成的掩码具有更清晰的边界，并且与真实掩码高度一致。即使在复杂背景干扰下，Bolt-SAM 也能准确捕捉螺栓属性的边缘信息，而其他方法则表现出明显的边界模糊以及目标区域的漏检和误检。

3) *SBDE* 编辑的比较实验：SBDE 的编辑原则不同于主流的编辑方法。由于目前不存在类似的方法，我们将 SBDE 与最先进的基于 GAN 和基于扩散的编辑方法进行比较，包括 StarGANv2、ReStyle+e4e、Null-text Inversion、Negative-prompt Inversion 和 Direct Inversion。所有比较方法均在 BDG 数据集上经过训练或微调，并使用 PSNR、SSIM 和 LPIPS 作为评估指标。需要注意的是，SBDE 并不依赖于任何缺陷图像。相反，这是一种基于预训练 LaMa 的零缺陷截图编辑方法，直接从正常图像生成缺陷图像。

表 VII 中的定量结果表明，SBDE 在所有指标上显著优于对比方法。与表现最好的方法相比，SBDE 将针状缺陷编辑的 PSNR 和 SSIM 分别提高了 7.48 和 5.82，同时将 LPIPS 降低了 26.98。对于螺母缺陷编辑，SBDE 将 PSNR 和 SSIM 分别提高了 4.38 和 2.88，同时将 LPIPS 降低了

TABLE VI
关于螺栓-SAM 分割的比较实验结果。最佳结果以粗体标记。

Method	Pin			Nut		
	mIoU(%) $\uparrow$	Dice(%) $\uparrow$	PA(%) $\uparrow$	mIoU(%) $\uparrow$	Dice(%) $\uparrow$	PA(%) $\uparrow$
SAM 1-point [21]	72.06	82.82	97.21	75.17	85.29	94.09
SAM 3-points [21]	71.42	82.35	97.00	75.34	85.40	94.12
SAM-Adapter 1-point [25]	62.37	75.75	96.93	41.56	53.49	87.47
SAM-Adapter 3-points [25]	67.79	80.08	97.20	51.30	63.86	88.59
HQ-SAM 1-point [24]	60.40	74.51	96.60	57.02	70.85	87.04
HQ-SAM 3-points [24]	65.28	78.33	96.68	59.10	72.71	87.05
RobustSAM 1-point [26]	69.15	81.35	97.38	69.51	81.25	92.64
RobustSAM 3-points [26]	71.16	82.80	97.48	70.44	81.90	92.76
Bolt-SAM 1-point	72.73	82.88	97.61	75.24	85.31	94.25
Bolt-SAM 3-points	75.66	86.01	98.02	76.78	86.47	94.71

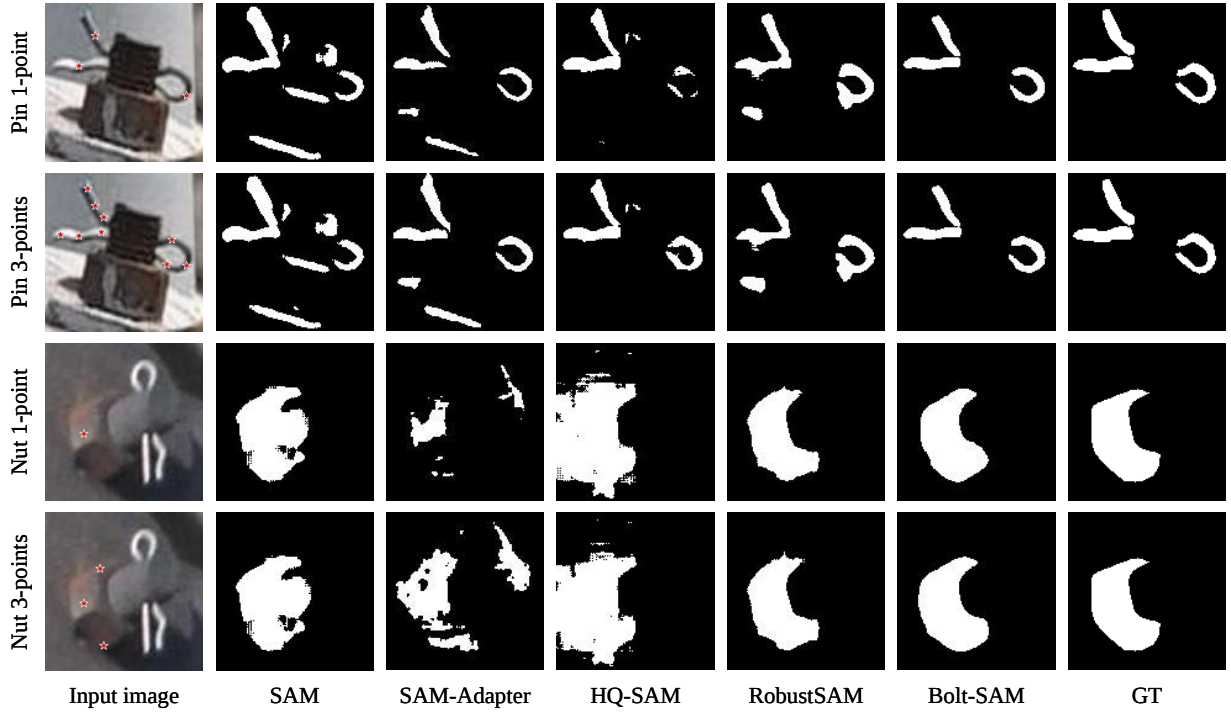


Fig. 8. Bolt-SAM 分割的对比实验可视化

21.67。这一优势来源于 SBDE 在精确编辑目标区域的同时保持图像其余部分的完整性，从而在像素级和感知质量上都优于生成方法。

图 9 显示了不同方法之间编辑结果的视觉对比。SBDE 生成的缺陷图像在视觉上与真实缺陷的感知高度一致，编辑目标区域自然且背景完好。相比之下，StarGANv2 和 ReStyle+e4e 倾向于在生成空间中搜索与相应缺陷语义对齐的潜在向量，导致与原始图像有显著偏差且收敛性差。虽然 Null-text Inversion, Negative-prompt Inversion 和 Direct Inversion 在整个图像中保持高保真度，但它们在缺陷区域的编辑效果欠佳，常常引入不必要的伪影或分散注意力的信息。这是由于在微调 Stable Diffusion 时遇到的域间隙以及训练数据差异导致的限制，阻碍了基于大型模型方法的编辑效果。

为了进一步验证 SBDE 的有效性及其编辑的缺陷图像在螺栓缺陷检测任务中的实际应用价值，我们提出了 ERA 策略以扩充 BDD 数据集，并在先进的检测模型上评估缺陷检测性能。具体而言，我们从 BDD 中筛选分辨率大于

64×64 的正常螺栓实例作为编辑对象。这些实例通过 SBDE 转化为有缺陷的螺栓，并修改其对应标签以创建带有缺陷的新检查图像，如图 1 所示。

目前，还没有专门设计用于编辑检测图像中螺栓目标的方法。现有的螺栓编辑和生成方法，当与 ERA 策略结合以恢复编辑后的螺栓至原始图像时，常常产生伪影和目标区域周围不自然的不连续性，导致性能不佳。因此，我们选择了基于复制的方法作为对照组 [50]。增强后的数据分布如表 VIII 所示。在实验中，仅对训练集进行了增强，而测试集保持不变。值得注意的是，复制后的图像与原始预编辑图像相同。我们使用 YOLOv11 作为检测模型，并使用 P、R 和 mAP 进行性能评估。所有实验均在相同的超参数设置下进行。表 IX 中的定量结果表明，SBDE 增强的数据集显著提高了螺栓缺陷检测的性能。与原始数据集相比，SBDE-Aug 在 P、R、mAP₅₀ 和 mAP_{50:95} 上分别取得了 2.2、2.3、4.4 和 1.0 的提升，有效提高了模型检测缺陷螺栓的能力。在所有指标上，SBDE-Aug 也优于

TABLE VII
SBDE 编辑的比较实验结果。最佳结果以粗体标出。

Method	Generator	Pin			Nut		
		SSIM(%) $\uparrow$	PSNR $\uparrow$	LPIPS(%) $\downarrow$	SSIM(%) $\uparrow$	PSNR $\uparrow$	LPIPS(%) $\downarrow$
StarGAN-v2 [31]	-	40.52	11.58	89.22	32.97	11.04	78.44
ReStyle+e4e [18], [19]	StyleGAN2-ada [33]	53.64	16.10	62.51	49.63	15.92	56.76
Null-text Inversion [40]	Stable Diffusion [36]	81.46	17.72	46.03	79.01	19.07	36.64
Negative Inversion [41]	+DreamBooth [38]	80.31	17.39	48.24	78.85	18.19	39.23
PnP Inversion [42]		79.13	17.19	50.55	78.23	18.12	40.07
SBDE(ours)	-	88.94	23.54	19.05	83.39	21.95	14.97

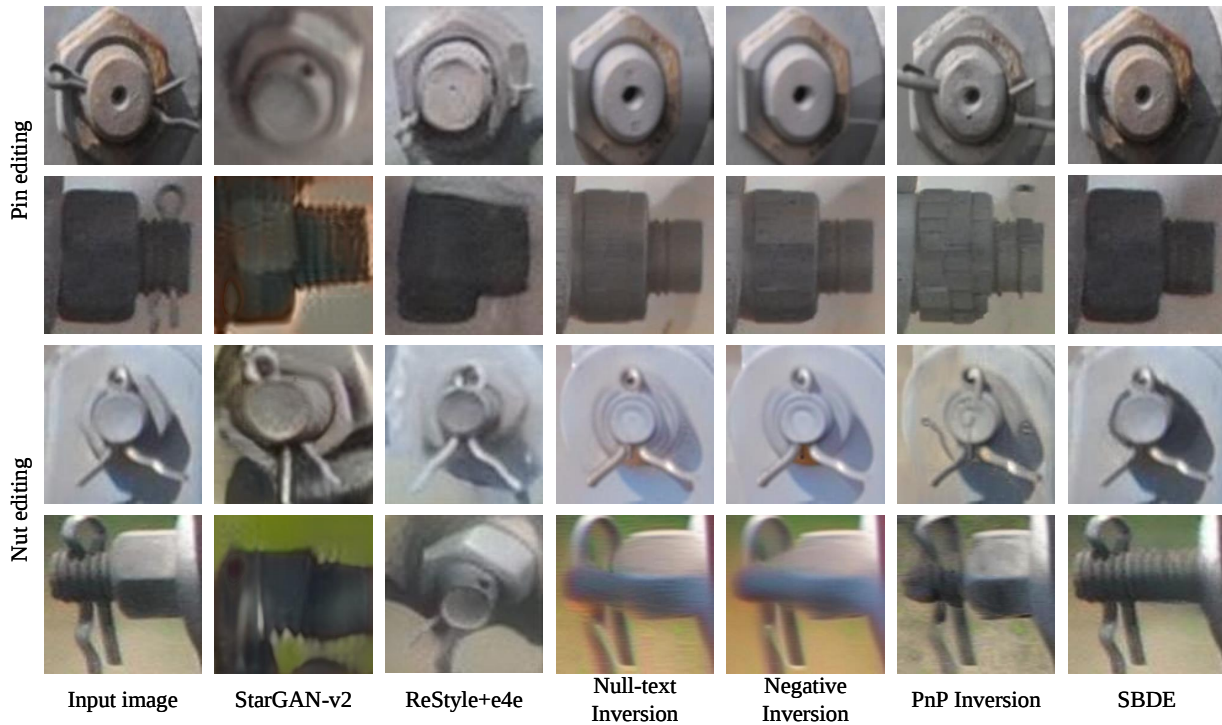


Fig. 9. SBDE 编辑对比实验的可视化

TABLE VIII
使用不同方法扩增后的训练集数据分布。

Class	Original	Copy-aug	SBDE-aug
Inspection images	1433	+759	+759
Number of instances			
Normal	3961	+2872	+1393
Pin losing	449	+192	+1179
Nut losing	244	+36	+528
All	4654	+3100	+3100

Copy-Aug, 显示出明显的优势。螺栓缺陷检测的可视化结果如图 10 所示。SBDE-Aug 取得了最佳的检测性能, 而原始检测模型和 Copy-Aug 都展示出了漏检的实例。我们的方法为有限缺陷数据下的缺陷检测提供了一种有效的解决方案, 并在实际工业应用中具有潜在的价值。

为了解决输电线路中螺栓缺陷样本稀缺和数据分布不平

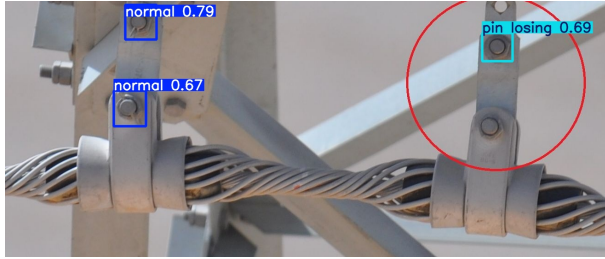
衡的问题, 本文提出了一种分割驱动的螺栓缺陷编辑方法, 称为 SBDE。其目标是将正常螺栓转化为缺陷螺栓, 以增强数据集并提高缺陷检测性能。我们提出了 Bolt-SAM, 设计了 CFA 和 MAMD 模块, 以提高复杂螺栓结构的分割精度。在编辑阶段, 我们设计了 MOD 模块以优化属性分割掩码, 并使用 MOD-LaMa 实现了精确的属性编辑, 生成相应的缺陷样本。此外, 我们还提出了 ERA 策略, 将编辑后的缺陷螺栓恢复到原始检测图像中, 从而扩充缺陷检测数据集, 并提高螺栓缺陷检测的准确性。我们提出的方法为输电线路中样本稀缺问题提供了一种创新的解决方案, 并展示了在其他领域的潜在应用性。

尽管提出的 SBDE 方法在螺栓编辑中取得了良好的效果, 但其性能仍然受到原始图像像素质量的影响。未来, 我们计划将其与生成模型结合, 利用它们在一般领域的强大生成能力以进一步提高缺陷编辑的质量。

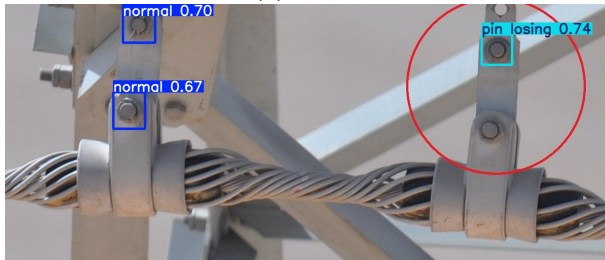
TABLE IX

不同增强方式检测结果的比较。整个数据集上的最佳结果以粗体标出。

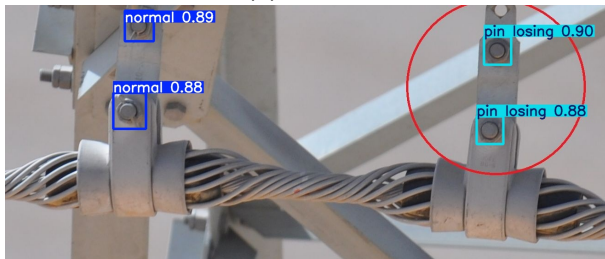
Method	Class	P(%)	R(%)	mAP ₅₀ (%)	mAP _{50:95} (%)
Original	All	86.0	76.0	83.1	46.8
	Normal	85.5	86.0	91.0	51.2
	Pin losing	76.4	70.9	77.8	39.9
	Nut losing	96.0	71.1	80.6	49.4
Copy-aug	All	84.2	77.8	84.5	46.2
	Normal	81.2	89.7	90.6	51.8
	Pin losing	74.8	70.3	79.0	40.7
	Nut losing	96.5	73.3	83.9	46.2
SBDE-aug	All	88.2	78.3	87.5	47.8
	Normal	83.1	88.7	91.2	52.4
	Pin losing	82.0	71.0	81.9	42.1
	Nut losing	99.4	75.2	89.3	48.9



(a) Original



(b) Copy-aug



(c) SBDE-aug

Fig. 10. 使用不同增强方法的检测结果可视化。深蓝色框代表正常螺栓，而浅蓝色框则表示缺少销钉的螺栓。需要注意的是，红色圆圈是我们收集的原始数据集的一部分，并未由我们处理。

REFERENCES

- [1] P. Luo, B. Wang, H. Wang, F. Ma, H. Ma, and L. Wang, "An ultrasmall bolt defect detection method for transmission line inspection," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023.
- [2] X. Tao, D. Zhang, Z. Wang, X. Liu, H. Zhang, and D. Xu, "Detection of power line insulator defects using aerial images analyzed with convolutional neural networks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 4, pp. 1486–1498, 2020.
- [3] W. Li, H. Wei, Y. Wu, J. Yang, Y. Ruan, Y. Li, and Y. Tang, "Tide: Test-time few-shot object detection," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 54, no. 11, pp. 6500–6509, 2024.
- [4] K. Zhang, R. Zhou, J. Wang, Y. Xiao, X. Guo, and C. Shi, "Transmission line component defect detection based on uav patrol images: A self-supervised hc-vit method," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 54, no. 11, pp. 6510–6521, 2024.
- [5] Y. Chen, Z.-W. Liu, G. Wen, and Y.-W. Wang, "Distributed finite-time observer for multiple line outages detection in power systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 54, no. 3, pp. 1768–1778, 2024.
- [6] K. Zhang, Y. Xiao, J. Wang, M. Du, X. Guo, R. Zhou, C. Shi, and Z. Zhao, "Dp-gan: A transmission line bolt defects generation network based on dual discriminator architecture and pseudo-enhancement strategy," *IEEE Transactions on Power Delivery*, vol. 39, no. 3, pp. 1622–1633, 2024.
- [7] Y. Ren, J. Liu, H. Zhao, and H. Zhang, "An industrial fault sample reconstruction and generation method under limited samples with missing information," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 54, no. 12, pp. 7821–7833, 2024.
- [8] S. Niu, B. Li, X. Wang, and H. Lin, "Defect image sample generation with gan for improving defect recognition," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 3, pp. 1611–1622, 2020.
- [9] Y. Duan, Y. Hong, L. Niu, and L. Zhang, "Few-shot defect image generation via defect-aware feature manipulation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 1, 2023, pp. 571–578.
- [10] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of stylegan," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 8110–8119.
- [11] G. Zhang, K. Cui, T.-Y. Hung, and S. Lu, "Defect-gan: High-fidelity defect synthesis for automated defect inspection," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2021, pp. 2524–2534.
- [12] F. Deng, J. Luo, L. Fu, Y. Huang, J. Chen, N. Li, J. Zhong, and T. L. Lam, "Dg2gan: improving defect recognition performance with generated defect image sample," *Scientific Reports*, vol. 14, no. 1, p. 14787, 2024.
- [13] Y. Tai, K. Yang, T. Peng, Z. Huang, and Z. Zhang, "Defect image sample generation with diffusion prior for steel surface defect recognition," *arXiv preprint arXiv:2405.01872*, 2024.
- [14] E. Härkönen, A. Hertzmann, J. Lehtinen, and S. Paris, "Ganspace: Discovering interpretable gan controls," *Advances in neural information processing systems*, vol. 33, pp. 9841–9850, 2020.
- [15] Y. Shen and B. Zhou, "Closed-form factorization of latent semantics in gans," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 1532–1540.
- [16] B. Wallace, A. Gokul, and N. Naik, "Edict: Exact diffusion inversion via coupled transformations," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 22 532–22 541.
- [17] E. Richardson, Y. Alaluf, O. Patashnik, Y. Nitzan, Y. Azar, S. Shapiro, and D. Cohen-Or, "Encoding in style: a stylegan encoder for image-to-image translation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 2287–2296.
- [18] O. Tov, Y. Alaluf, Y. Nitzan, O. Patashnik, and D. Cohen-Or, "Designing an encoder for stylegan image manipulation," *ACM Transactions on Graphics (TOG)*, vol. 40, no. 4, pp. 1–14, 2021.
- [19] Y. Alaluf, O. Patashnik, and D. Cohen-Or, "Restyle: A residual-based stylegan encoder via iterative refinement," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 6711–6720.
- [20] Z. Zhao, H. Qi, Y. Qi, K. Zhang, Y. Zhai, and W. Zhao, "Detection method based on automatic visual shape clustering

- for pin-missing defect in transmission lines,” *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6080–6091, 2020.
- [21] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo *et al.*, “Segment anything,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 4015–4026.
- [22] Y. Gu, Q. Wu, H. Tang, X. Mai, H. Shu, B. Li, and Y. Chen, “Lesam: Adapt segment anything model for medical lesion segmentation,” *IEEE Journal of Biomedical and Health Informatics*, 2024.
- [23] Y. Zhang, Z. Shen, and R. Jiao, “Segment anything model for medical image segmentation: Current applications and future directions,” *Computers in Biology and Medicine*, p. 108238, 2024.
- [24] L. Ke, M. Ye, M. Danelljan, Y.-W. Tai, C.-K. Tang, F. Yu *et al.*, “Segment anything in high quality,” *Advances in Neural Information Processing Systems*, vol. 36, 2024.
- [25] T. Chen, L. Zhu, C. Deng, R. Cao, Y. Wang, S. Zhang, Z. Li, L. Sun, Y. Zang, and P. Mao, “Sam-adapter: Adapting segment anything in underperformed scenes,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 3367–3375.
- [26] W.-T. Chen, Y.-J. Vong, S.-Y. Kuo, S. Ma, and J. Wang, “Robustsam: Segment anything robustly on degraded images,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 4081–4091.
- [27] H. Dai, C. Ma, Z. Yan, Z. Liu, E. Shi, Y. Li, P. Shu, X. Wei, L. Zhao, Z. Wu *et al.*, “Samaug: Point prompt augmentation for segment anything model,” *arXiv preprint arXiv:2307.01187*, 2023.
- [28] T. Yu, R. Feng, R. Feng, J. Liu, X. Jin, W. Zeng, and Z. Chen, “Inpaint anything: Segment anything meets image inpainting,” *arXiv preprint arXiv:2304.06790*, 2023.
- [29] A. Melnik, M. Miasayedzenkau, D. Makaravets, D. Pirshtuk, E. Akbulut, D. Holzmann, T. Renusch, G. Reichert, and H. Ritter, “Face generation and editing with stylegan: A survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.
- [30] Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Choo, “Stargan: Unified generative adversarial networks for multi-domain image-to-image translation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
- [31] Y. Choi, Y. Uh, J. Yoo, and J.-W. Ha, “Stargan v2: Diverse image synthesis for multiple domains,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
- [32] T. Karras, S. Laine, and T. Aila, “A style-based generator architecture for generative adversarial networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 12, pp. 4217–4228, 2021.
- [33] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, “Training generative adversarial networks with limited data,” in *Advances in Neural Information Processing Systems*, 2020.
- [34] J. Ho, A. Jain, and P. Abbeel, “Denoising diffusion probabilistic models,” *Advances in neural information processing systems*, vol. 33, pp. 6840–6851, 2020.
- [35] J. Song, C. Meng, and S. Ermon, “Denoising diffusion implicit models,” *arXiv preprint arXiv:2010.02502*, 2020.
- [36] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, “High-resolution image synthesis with latent diffusion models,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 10 684–10 695.
- [37] R. Gal, Y. Alaluf, Y. Atzmon, O. Patashnik, A. H. Bermano, G. Chechik, and D. Cohen-Or, “An image is worth one word: Personalizing text-to-image generation using textual inversion,” *arXiv preprint arXiv:2208.01618*, 2022.
- [38] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, and K. Aberman, “Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 22 500–22 510.
- [39] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, “Lora: Low-rank adaptation of large language models,” *arXiv preprint arXiv:2106.09685*, 2021.
- [40] R. Mokady, A. Hertz, K. Aberman, Y. Pritch, and D. Cohen-Or, “Null-text inversion for editing real images using guided diffusion models,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 6038–6047.
- [41] D. Miyake, A. Iohara, Y. Saito, and T. Tanaka, “Negative-prompt inversion: Fast image inversion for editing with text-guided diffusion models,” *arXiv preprint arXiv:2305.16807*, 2023.
- [42] X. Ju, A. Zeng, Y. Bian, S. Liu, and Q. Xu, “Direct inversion: Boosting diffusion-based editing with 3 lines of code,” *arXiv preprint arXiv:2310.01506*, 2023.
- [43] R. Suvorov, E. Logacheva, A. Mashikhin, A. Remizova, A. Ashukha, A. Silvestrov, N. Kong, H. Goka, K. Park, and V. Lempitsky, “Resolution-robust large mask inpainting with fourier convolutions,” in *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, 2022, pp. 2149–2159.
- [44] K. Zhang, W. Lou, J. Wang, R. Zhou, X. Guo, Y. Xiao, C. Shi, and Z. Zhao, “Pa-detr: End-to-end visually indistinguishable bolt defects detection method based on transmission line knowledge reasoning,” *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–14, 2023.
- [45] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, “End-to-end object detection with transformers,” in *European conference on computer vision*. Springer, 2020, pp. 213–229.
- [46] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, “Adaptive histogram equalization and its variations,” *Computer vision, graphics, and image processing*, vol. 39, no. 3, pp. 355–368, 1987.
- [47] W. Liu, X. Shen, C.-M. Pun, and X. Cun, “Explicit visual prompting for low-level structure segmentations,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 19 434–19 445.
- [48] D. Hendrycks and K. Gimpel, “Gaussian error linear units (gelus),” *arXiv preprint arXiv:1606.08415*, 2016.
- [49] R. Khanam and M. Hussain, “Yolov11: An overview of the key architectural enhancements,” *arXiv preprint arXiv:2410.17725*, 2024.
- [50] J. Liu, Z. Ma, Y. Qiu, X. Ni, B. Shi, and H. Liu, “Four discriminator cycle-consistent adversarial network for improving railway defective fastener inspection,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 10 636–10 645, 2021.