

DOI:10.19651/j.cnki.emt.2519357

基于 Pix2PixGAN 的钨晶粒晶界提取方法

文金瑞¹ 吴晓红¹ 滕奇志¹ 何海波²

(1. 四川大学电子信息学院 成都 610065; 2. 成都西图科技有限公司 成都 610065)

摘要: 针对钨钨杆扫描电子显微镜图像分割中面临的晶粒粘连区分困难及晶界遮挡等挑战,提出了一种基于 Pix2PixGAN 框架的改进方法,用于实现钨晶粒晶界的准确提取。首先,使用边缘引导注意力 EGA 模块替换标准跳跃连接,结合拉普拉斯特征图提取,以增强多尺度晶界特征表达能力;其次,采用高效上采样块进行特征上采样,有效减轻了棋盘伪影并促进了不同层级特征的融合;将原始 L2 损失函数代替为加权二元交叉熵损失与加权交并比损失的组合损失,聚焦边缘像素优化;最后,引入梯度惩罚,增强生成器稳定性与多样性。实验结果显示,改进模型的 F1 值 (F1-score) 为 72.47%,召回率 (Recall) 为 77.21%,准确率 (Precision) 为 68.32%,相较于基础 Pix2PixGAN 模型分别提高了 13.02%、6.49%、16.87%,且在 F1 值与精确率上高于 RCF、RINDNet、UCTransNet 和 MEGANet 模型,验证了该模型在晶界提取的有效性。

关键词: 钨钨杆;扫描电子显微镜;晶界提取;边缘引导注意力;高效上采样

中图分类号: TG132.3;TP391.4;TP18;TN04 **文献标识码:** A **国家标准学科分类代码:** 520.6

Tungsten grain boundary extraction method based on Pix2PixGAN

Wen Jinrui¹ Wu Xiaohong¹ Teng Qizhi¹ He Haibo²

(1. School of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China;

2. Chengdu Xitu Technology Co., Ltd., Chengdu 610065, China)

Abstract: To address the challenges encountered in segmenting scanning electron microscope images of lanthanum tungsten rods—such as difficulties in distinguishing adhered grains and occlusion of grain boundaries—an improved method based on the Pix2PixGAN framework is proposed to achieve high-precision extraction of tungsten grain boundaries. First, the standard skip connection is replaced with an edge guided attention module, integrated with Laplacian feature map extraction, to enhance the multi-scale representation of grain boundary features. Second, an efficient upsampling convolution block is introduced for feature upsampling, effectively mitigating checkerboard artifacts and facilitating the fusion of multi-level features. The original L2 loss function is substituted with a combined loss function comprising weighted binary cross-entropy loss and weighted intersection-over-union loss, emphasizing the optimization of edge pixels. Finally, gradient penalty is incorporated to improve the stability and diversity of the generator. Experimental results demonstrate that the improved model achieves an F1-score of 72.47%, a recall rate of 77.21%, and a precision of 68.32%, representing improvements of 13.02%, 6.49%, and 16.87%, respectively, over the baseline Pix2PixGAN model. Furthermore, the proposed method surpasses RCF, RINDNet, UCTransNet, and MEGANet in terms of F1-score and precision, confirming its effectiveness in grain boundary extraction.

Keywords: tungsten-lanthanum rods; scanning electron microscope; grain boundary extraction; edge-guided attention; efficient upsampling

0 引言

钨钨杆 (W-0.7%La₂O₃) 是一种典型的稀土掺杂钨基合金,广泛应用于高温结构材料领域。为能更好的将钨钨杆应用到实际工程中,需深入分析其合金性能。由于合金

性能与合金微观结构存在强关联性,构建性能-结构的定量映射关系需精准表征微观特征。钨作为钨钨杆的主要组分,其晶粒尺寸、形貌及空间分布对合金性能有重要影响。目前,金相图像作为常用表征手段,可直接观测钨钨杆中钨晶粒特征,进而为性能-结构关联建立提供依据。

利用金相图像分析钨晶粒特征前,需要在图像中将钨晶粒目标准确提取出来。虽然现有的通用的图像分割技术和合金图像晶粒分割技术已较为成熟,但在应对钨钨杆扫描电子显微镜(scanning electron microscope, SEM)图像时,现有通用方法存在局限性,需要针对性改进。合金晶粒分割主要有两种思路:一是以晶粒整体区域为识别目标进行提取^[1],二是以晶界为切入点进行边界检测^[2]。早期的阈值分割法(如 Otsu 算法)和边缘提取(如 Canny)计算简单,但其性能极度依赖图像灰度分布,难以应对复杂晶粒图像,难以解决钨钨杆 SEM 图像中的核心挑战,泛化能力有限。这一局限性促使研究者寻求更鲁棒的自动化分割方法,以适应复杂金相图像的分析需求。

近年来,深度学习技术通过自动提取图像的隐含特征,获取高级语义信息,有效地解决了自动特征学习的难题^[3],在图像分类^[4],图像分割^[5]以及图像边缘检测^[6]等计算机视觉任务中逐渐成为主流。其中,卷积神经网络(convolutional neural networks, CNN)凭借其强大的局部特征提取能力和层次化特征表示,成为处理合金图像分割^[7-9]、合金晶界检测^[10-12]等任务的重要架构。Liu 等^[13]提出了 RCF 模型,借多层卷积递归计算与跨层连接融合多尺度特征,提升边缘检测精度;Pu 等^[14]提出了 RINDNet 模型,通过检测反射率、照明、法线和深度 4 种边缘建立联系、共享特征、强化边缘提取能力;Wang 等^[15]提出的 UCTransNet 引入 Channel Transformer 模块代替跳跃连接,捕获更复杂的通道依赖以增强语义信息。这些模型在合金图像语义分割任务中都有着不错的潜力。

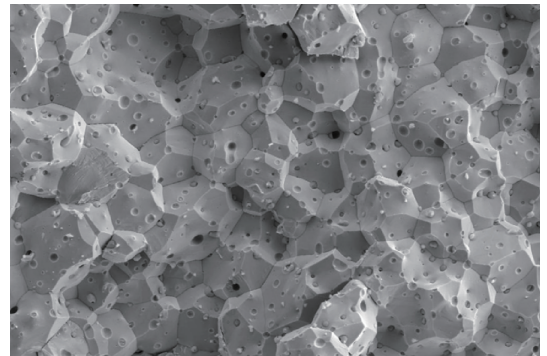
尽管这些方法提升了特征建模和边界提取能力,传统 CNN 仍面临两大局限:依赖大规模成对标注数据和生成模式单一。为此,条件对抗生成网络(conditional generative adversarial network, cGAN)在对抗生成网络(generative adversarial network, GAN)架构中引入条件输入,通过对抗训练从而提升生成质量^[16],并降低对大规模成对数据的依赖^[17]。例如,Isola 等^[18]基于 cGAN 提出 Pix2PixGAN,结合 U-Net 结构实现图像到目标的高质量映射;生成器以条件输入捕获局部与全局信息,判别器采用 PatchGAN 增强细节判别,该框架已被应用于晶界生成辅助晶粒分析^[19];Li 等^[20]提出的 SemiRCF-GAN 结合 RCF 以半监督模式补全被硅晶粒遮掩的铝晶粒晶界。值得注意的是,CNN 边缘检测算法的设计思想仍具重要参考价值。Zhou 等^[21]提出的 MPUNet 通过多分支任务解决沉淀物覆盖导致的晶界过分割,辅以修剪策略优化闭合边界。上述研究为复杂结构下的晶粒晶界提取提供了重要启发,但在应对晶界遮掩、晶棱干扰和细粒结构密集等问题时,仍存在精度与鲁棒性不足的问题。尤其在钨钨杆扫描电子显微镜 SEM 图像中,现有方法尚难以实现晶界的连续、准确提取。

为提升对钨钨杆 SEM 图像中的钨晶粒晶界的识别能力,本文参考了 Pix2PixGAN 架构对捕获局部与全局信息

的优势,通过晶界提取的方式实现钨晶粒的目标提取,提出了一种针对钨钨杆 SEM 图像的 Pix2Pix-EG 模型。改进生成器并引入边缘引导注意力(edge guided attention, EGA)^[22]模块代替跳跃连接,实现多尺度特征的深度融合。同时采用高效上采样块(efficient up-convolution block, EUCB)^[23]缓解棋盘伪影问题,并增强不同层特征的关联性。最后,在自建钨钨杆 SEM 图像测试集上开展实验,结果表明所提出的方法对钨钨杆 SEM 图像边缘提取的有效性,具有良好的应用前景。

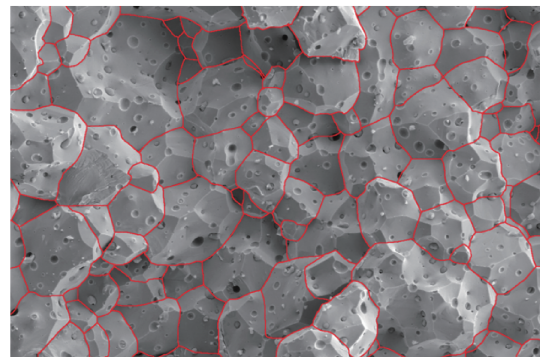
1 钨钨杆 SEM 图像分割难点

为了进一步说明钨钨杆 SEM 图像的分割难点,图 1(a)展示了所采集样本的代表性图像。图中钨晶粒的棱边与真实晶界在灰度表现上区分度较低,此外,晶粒表面及其晶界还分布有大量孔洞和细小晶粒,这些结构增加了提取完整晶粒晶界的难度。图 1(b)展示的是专家鉴定后勾勒出的钨晶粒准确晶界定位,这为后续的标签制作与评估测定给出了重要依据。



(a) 钨钨杆 SEM 图像

(a) Lanthanated tungsten rod SEM images



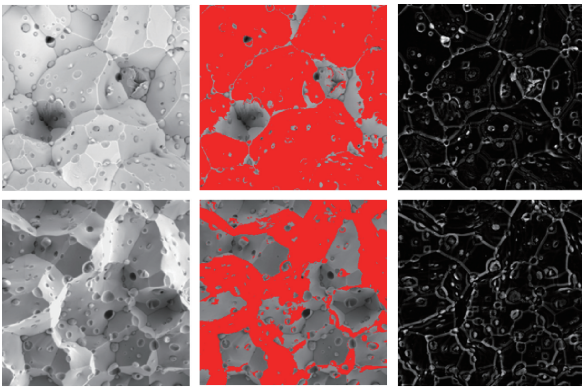
(b) 人工标注晶界图像

(b) Manually annotated grain boundary images

图 1 钨钨杆 SEM 图像与晶界定位图像

Fig. 1 SEM images of lanthanum tungsten rods and grain boundary positioning images

图 2 展示了钨钨杆 SEM 图像在传统图像处理方法下的分割效果。



(a) 原始图像 (b) 阈值分割 (c) 边缘提取
(a) Original image (b) Threshold segmentation (c) Edge extraction

图 2 传统图像处理方法提取钨晶粒

Fig. 2 Conventional image processing methods extract tungsten grains

在图 2 中, 阈值分割方法难以适应图像中灰度分布的局部变化, 常导致晶粒区域分割不完整或过度分割, 特别是在灰度过渡平缓区域识别效果较差; 边缘提取则因晶粒棱边与真实晶界在图像中灰度表现相近, 区分度不高, 导致边界响应混乱, 出现断裂和误检。整体来看, 传统方法难以有效应对图像中的结构复杂性, 限制了晶粒的准确提取。鉴于传统方法在晶界提取中的局限性, 钨钨杆 SEM 图像分割任务亟需一种自动化且鲁棒性强的分割方法。

2 基于 Pix2PixGAN 钨晶粒晶界提取算法

2.1 整体网络模型

近年来, Pix2PixGAN 作为图像翻译的主流, 有着不少的改进模型以及策略。Pix2Pix-HD^[16] 通过级联两个生成

器, 采用多尺度判别器的同时引入特征匹配损失; Pix2Pix-Diffusion^[24] 将 Pix2Pix 与扩散模型结合, 提升生成质量, 将风格控制在条件范围以内。Pix2Pix-Res^[25] 用 ResNet 替换 UNet, 增强特征提取能力。Pix2Pix-ResTrans^[26] 在前者基础上, 引入 Transformer 层建立长程依赖, 提升对图像全局上下文的理解。Pix2Pix-GP^[27] 则引入 WGAN 中的梯度惩罚 (gradient penalty, GP)^[28], 稳定训练过程, 有效缓解模式崩溃和判别器过强问题。现有改进大多着眼于提升高分辨率图像生成的稳定性和性能, 但是缺少针对合金图像晶界提取的优化。因此需要对 Pix2PixGAN 进行一系列的定制化改进。

Pix2PixGAN 模型在处理钨钨杆 SEM 图像时存在以下局限性: 细小晶粒遮掩导致晶界断裂、灰度平滑区域的晶界与晶棱特征区分困难、反卷积上采样易引发棋盘伪影, 以及损失函数对边缘像素关注不足。这些问题严重制约了钨晶粒晶界的连续性提取精度与模型鲁棒性, 亟需从网络架构与特征建模机制进行针对性优化。

为此, 本文提出了针对性优化方案, 整体网络结构如图 3 所示。优化后的生成器保留 UNet 架构并采用 Res2Net50^[29] 作为主干网络。边缘引导注意力 EGA 模块替代传统跳跃连接, 通过构建编码器低级特征与解码器高级特征的语义桥接, 跨尺度增强晶界信息保留能力, 有效化解小晶粒遮导致的晶界断裂问题。采用高效上采样块 EUCB 消除反卷积产生的棋盘伪影, 同步提升跨层特征融合与边缘细节还原能力。在损失函数设计上, 以加权二元交叉熵 (weighted binary cross entropy, WBCE) 和加权交并比 (weighted IoU, WIoU) 替换 L2 损失强化稀疏晶界学习, 辅以 L1 损失约束边界几何精度。判别器沿用 PatchGAN 结构增强局部真实性判别能力, 结合 GP^[28] 机制缓解生成模式单一与模式崩溃问题。实验表明, 在自建的钨钨杆 SEM 图像数据

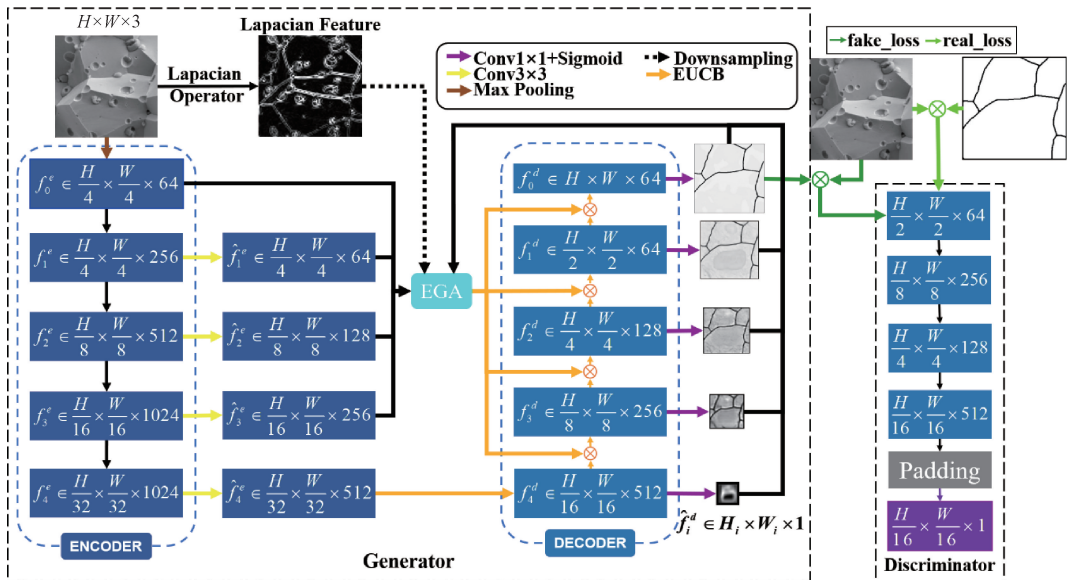


图 3 整体网络结构

Fig. 3 Overall network structure

集上提升细小晶粒遮掩修复能力、抑制晶界误分割现象,并增强灰度平滑区的晶界-晶棱区分度。

2.2 EGA 模块

图 4 所示的是 EGA 模块结构,该模块替代了 UNet 框架第 1~4 层的跳跃连接,通过多源特征融合强化边缘信息传递,并为上采样过程提供注意力引导。该模块同步接收三路输入:由残差块组下采样并修正通道所得的编码视

觉特征 \hat{f}_i^c 、上游路径上采样后经 1×1 卷积与 Sigmoid 处理生成的解码预测特征 \hat{f}_{i+1}^d 、以及以及从拉普拉斯金字塔首层高频特征 f_i^l 下采样适配的分辨率特征 f_i^l 。为抑制噪声干扰,原始图像预先进行高斯滤波平滑处理;鉴于多次高斯滤波导致边缘退化,仅选用金字塔首层特征 f_i^l ,提供给 EGA 模块。

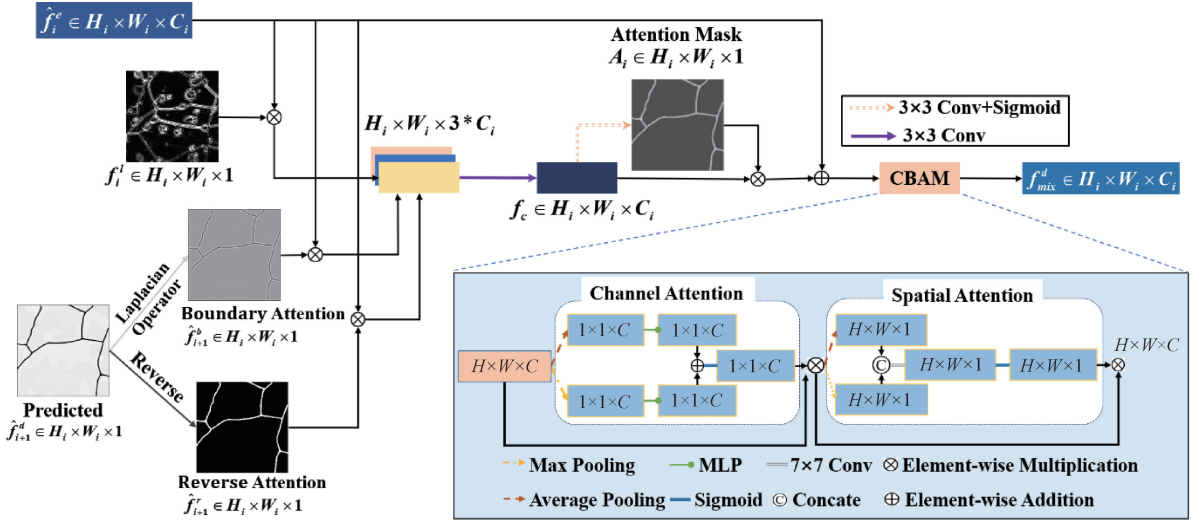


图 4 EGA 模块结构

Fig. 4 EGA module structure

在第 i 层处理中,首先计算解码特征 \hat{f}_{i+1}^d 的反向注意力 \hat{f}_{i+1}^r 和边缘注意力 \hat{f}_{i+1}^b ,其中 $\hat{f}_{i+1}^r = 1 - \hat{f}_{i+1}^d$, \hat{f}_{i+1}^b 是 \hat{f}_{i+1}^d 通过高斯-拉普拉斯二阶微分算子提取得到。随后将编码特征 \hat{f}_i^c 分别与 \hat{f}_{i+1}^r 、 \hat{f}_{i+1}^b 及 f_i^l 执行元素级乘法操作,拼接结果经卷积层融合为统一特征 f_i^c 。

$$f_i^c = \text{Conv}([\hat{f}_{i+1}^r \otimes \hat{f}_i^c, \hat{f}_{i+1}^b \otimes \hat{f}_i^c, (f_i^l \otimes \hat{f}_i^c)]) \quad (1)$$

为过滤噪声与冗余细节,进一步由 f_i^c 经卷积和 Sigmoid 激活生成注意力图 A_i ,引导模型的注意力集中在重要区域,抑制背景噪声和冗余信息。然后通过加权残差连接 \hat{f}_i^c 得到注意力特征 f_i^a 。

$$A_i = \text{sigmoid}(\text{Conv}(f_i^c)) \quad (2)$$

$$f_i^a = \hat{f}_i^c + (f_i^c \otimes A_i) \quad (3)$$

最后,将 f_i^a 输入空间-通道双注意力模块(convolutional block attention module, CBAM)优化,进一步提升模型对重要特征的关注能力。最终输出混合特征 f_{mix}^d 。

2.3 EUCB 上采样

在语义分割任务中,上采样模块需在恢复空间分辨率的同时保持特征语义的一致性。传统卷积上采样容易引入棋盘格伪影,而基于插值的方法的上采样不能够很好地结合不同层的特征信息。针对上述问题,EUCB 模块提出三重优化设计:1)以深度可分离卷积替代标准卷积,在降

低计算量的同时显著抑制伪影生成;2)串联批归一化与激活层,增强训练稳定性以保障语义一致性;3)引入轻量化通道调控机制,动态提升跨层特征的选择性融合能力。具体实现如图 5 所示。

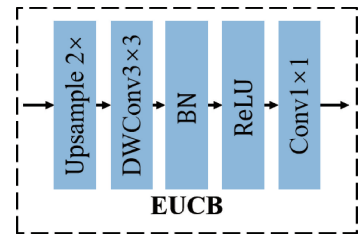


图 5 EUCB 结构

Fig. 5 EUCB structure

模块接收上游解码特征 f_{i+1}^d 和 EGA 输出的混合特征 f_{mix}^d ,经通道维度拼接后,依次通过双线性上采样、深度可分离卷积、批归一化与激活层。其中深度可分离卷积利用每个卷积核仅处理一个通道的数据,从而降低了计算参数量。为克服深度卷积的通道间信息隔离,加入通道混洗操作促进跨通道交互,最终由逐点卷积整合特征,输出高分辨率解码特征 f_i^d 。

2.4 损失函数

Pix2PixGAN 采用的是 L1+L2 损失组合的混合损失来训练生成器。L1 损失和 L2 损失的计算方式如下:

$$L1 = \sum_{i,j} |M_{i,j} - P_{i,j}| \quad (4)$$

$$L2 = \sum_{i,j} (M_{i,j} - P_{i,j})^2 \quad (5)$$

其中, $M_{i,j}$ 表示标签中坐标为 (i,j) 的像素值, $P_{i,j}$ 表示预测中坐标为 (i,j) 的像素值。

L2 损失因边缘像素稀疏导致损失被非边缘主导, 强制平滑效果使边缘模糊, 且难以建模边缘的连续性及空间关系, 在边缘提取任务中效果有限。为此, 本文采用混合损失策略, 以 WBCE 和 WIoU 取代传统 L2 损失。WBCE 通过显著增加边缘像素误判的惩罚权重, 有效缓解样本不平衡问题, 聚焦于关键边缘的准确提取; WIoU 通过位置敏感惩罚机制, 重点严惩真实边缘附近的误判像素, 从而提升定位精度, 抑制伪影, 保障边缘结构的完整性。加权二元交叉熵和加权交并比公式如下:

$$BCE(P_{i,j}, M_{i,j}) = -M_{i,j} \times \ln(\sigma(P_{i,j})) - (1 - M_{i,j}) \times \ln(1 - \sigma(P_{i,j})) \quad (6)$$

$$WBCE = \frac{\sum_{i,j} \omega_{i,j} \times BCE(P_{i,j}, M_{i,j})}{\sum_{i,j} \omega_{i,j}} \quad (7)$$

$$WIoU = 1 - \frac{\sum_{i,j} \omega_{i,j} \times (P_{i,j} \times M_{i,j}) + 1}{\sum_{i,j} \omega_{i,j} \times (P_{i,j} + M_{i,j} - P_{i,j} \times M_{i,j}) + 1} \quad (8)$$

其中式(6)中: σ 表示 Sigmoid 函数, 式(7)、(8)中: 权重 $\omega_{i,j}$ 是通过 31×31 大小的平均池化对输入标签计算获取。公式如下:

$$\omega_{i,j} = 1 + 5 * |AvgPool_{31 \times 31}(M)_{i,j} - M_{i,j}| \quad (9)$$

最终组合成混合损失 S_{loss_i} 。公式如下:

$$S_{Loss_i} = \frac{1}{N} \sum_{n=1}^N (WBCE + WIoU) \quad (10)$$

图 6 是生成器输出的不同尺度下的特征概率图。由图可知, 下采样进行到最后一层后, 其边缘特征丢失严重, 因此选定使用前 4 个尺度的特征概率图 \hat{f}_i^d 作为损失计算的参数并与对应尺寸的真实标签 gt_i 进行损失计算。计算出的 S_{loss_i} 进行求和计算, 输出混合损失 S_{loss} :

$$S_{loss} = S_{loss_0} + S_{loss_1} + S_{loss_2} + S_{loss_3} \quad (11)$$

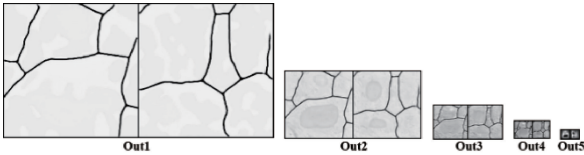


图 6 不同尺度特征概率图

Fig. 6 Different dimension feature probability graph

最后结合 L1 损失, 组成最终生成器损失 G_{loss} 。

$$G_{loss} = S_{loss} + \lambda_{L_1} \times L1_{loss}, \lambda_{L_1} = 40 \quad (12)$$

2.5 梯度惩罚 GP

训练不稳定是 GAN 常见的一个问题。当判别器能力

过强时, 会将生成样本大概率判定为假, 导致生成器因梯度信号极弱而难以更新参数, 即梯度消失; 同时, 生成器为“骗过”判别器, 可能只生成少数易被接受的样本, 出现生成结果单一化的模式崩溃问题。

为解决这些问题, 本文为判别器引入了梯度惩罚 GP, 通过惩罚梯度范数偏离 1 的情况, 防止判别器对输入变化过于敏感即梯度范数过大或不敏感即梯度范数过小, 从而保持稳定的梯度信号提供给生成器, 达到一个有效的对抗生成过程。GP 核心公式如式(13)~(14)所示。

$$GP = \lambda_{GP} \times (\nabla_x D(\hat{x})_2 - 1)^2 \quad (13)$$

$$\hat{x} = \epsilon M + (1 - \epsilon) G(P) \quad (14)$$

其中, M 为标签图像, P 为预测图像, ϵ 是一个在 $[0, 1]$ 的随机数, $\nabla_x D(\hat{x})$ 是判别器对插值样本 \hat{x} 的梯度。判别器损失 D_{loss} 如式(15)所示。

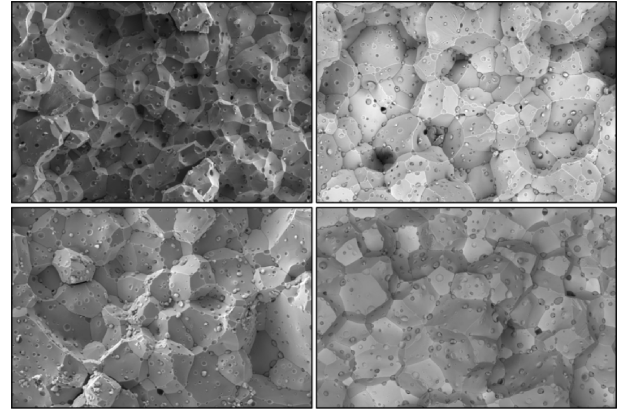
$$D_{loss} = -E(D(M)) + E(D(P)) + \lambda_{GP} \times E[(\nabla_x D(\hat{x})_2 - 1)^2] \quad (15)$$

其中, λ_{GP} 设置为 100, E 表示求均值。

3 实验结果及分析

3.1 数据集制作

为了保证模型的泛化性, 镧钨杆 SEM 图像选用了背散射电子(back scattered electron, BSE)与二次电子(secondary electrons, SE)两种不同信号拍摄的 SEM 图像。图 7 为两种信号拍摄的镧钨杆 SEM 图像。



(a) SE信号

(b) BSE信号

(a) SE signal

(b) BSE signal

图 7 不同信号拍摄的镧钨杆 SEM 图像

Fig. 7 Different signal of lanthanum tungsten rod SEM images

完整视域下的镧钨杆 SEM 图像大小为 $1024 \text{ pixel} \times 690 \text{ pixel}$ (去除表头信息)。为降低训练时间与计算成本, 图像统一裁剪为 $256 \text{ pixel} \times 256 \text{ pixel}$ 大小。同时, 为最大限度保留裁剪区域的边缘晶粒信息, 采用步长 128 pixel、沿水平和垂直方向的重叠滑动窗口裁剪策略。人工标注的标签图也采用相同的大小和策略。最后按照 Pix2PixGAN 模型输入要求, 将原图与标签图进行拼接, 制作成

512 pixel×256 pixel 大小的训练数据集,标签示例如图 8 所示。

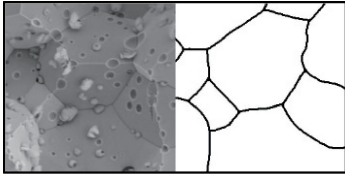


图 8 训练标签

Fig. 8 Training labels

具体数据集数量划分如表 1 所示。

表 1 数据集分布

Table 1 Data set distribution

数据集	A	B
train set	2 400 张	600 张
test set	600 张	150 张

3.2 评估指标

为了客观地评价模型性能,本文选择 F1 值,准确率(Precision)以及召回率(Recall)作为边缘提取效果的评估指标,计算公式如表 2 所示。其中,TP 表示边缘像素正确预测的数量;FP 表示边缘像素错误预测的数量;TN 表示背景像素的正确预测数量;FN 表示背景像素的错误预测数量。

表 2 指标评价

Table 2 Evaluation metrics

指标	公式
F1 值	$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$
Precision	$\frac{\text{TP}}{\text{TP} + \text{FP}}$
Recall	$\frac{\text{TP}}{\text{TP} + \text{FN}}$

3.3 实验环境及参数设置

本文基于 Pytorch1.10.1、Cuda11.1 的开发环境完成网络的搭建以及模型的训练。硬件环境以及相关训练参数设置如表 3 所示。

表 3 训练相关参数

Table 3 Experimental parameter setting

参数	数值
CPU	Inter(R)Core(TM)i5-4590
GPU	GeForce RTX 3090
Epochs	200
Batch-size	2
Learning-rate	0.000 2
Img-size	256×256

3.4 损失函数的选择实验

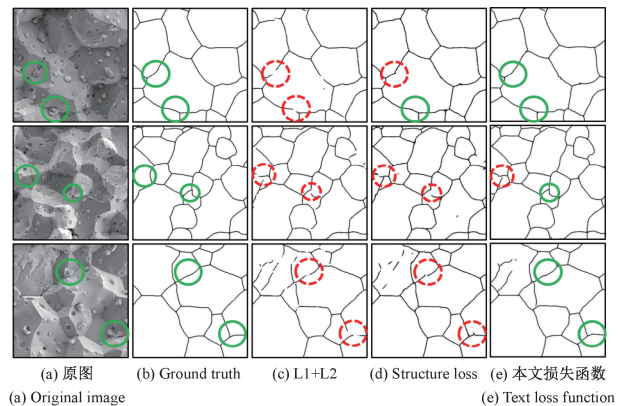
为了验证所选损失函数的合理性与有效性,本文对比了 L1+L2, Structure loss 两种损失函数组合,并在验证集上进行了指标检验。从表 4 中不同损失函数组合的性能指标对比结果可以看出,本文在各项指标上均达到了最高。

表 4 不同损失函数对比

Table 4 Comparison of different loss functions %

损失函数	F1	Recall	Precision
L1+L2	67.76	69.81	65.83
Structure loss	69.08	72.08	66.32
L1+Structure loss	72.47	77.21	68.32

图 9 是不同损失函数的晶界提取情况。可以看出:在遮掩晶粒上的识别,L1+L2 组合损失函数基本上没有识别出完整晶界;在晶粒叠加情况下的晶界提取上, Structure loss 也存在无法完整提取的情况,本文损失函数的具体情况,在上一部分的不同网络对比实验中详细描述过。整体比下来,本文采用的组合损失函数是有效的。



(a) Original image (b) Ground truth (c) L1+L2 (d) Structure loss (e) Text loss function

图 9 不同损失函数的提取结果

Fig. 9 Extraction results of different loss functions

3.5 算法消融实验及分析

本文共对原始 Pix2Pix 进行了 3 处改进,为了验证 Res2Net 和两个模块各自的有效性,对各个改进点及组合策略逐一进行消融实验。实验结果如表 5 所示。

表 5 消融实验

Table 5 Ablation experiment

模型	EGA	EUCB	F1	Recall	Precision
Baseline	×	×	52.91	64.94	44.65
+Res2Net	×	×	59.56	70.72	51.45
+EGA	✓	×	65.08	70.95	60.11
+EUCB	×	✓	64.97	73.02	58.53
本文模型	✓	✓	75.85	76.22	75.50

从表中可以看出, Res2Net 强化生成器性能, 有效地提高了模型召回率。在加入 EGA 模块后召回率提升并不明显, 但准确率提升幅度较大, 提升了 8.66%, 说明模型对边缘像素的定位更为精确, 符合 EGA 模块强化边缘信息学习的任务。在引入 EUCB 模块后, 模型的召回率有所提升, 解决了传统上采样问题, 保证了 EGA 模块的输入质量, 但是在准确率上提升不明显。最后本文模型的性能上, 召回率和准确率相比未改进前提升明显。

3.6 相关改进算法对比及分析

本文主要比较了 Pix2Pix-Res, Pix2Pix-GP 以及 Pix2Pix-ResTrans 3 种改进模型。具体性能对比如表 6 所示。

从表 6 中可以看出, 各种改进方案都在原有模型的基础上有着较好的提升, 说明针对本文任务, 这些改进思路都是可取的。其中 Pix2Pix-ResTrans 相较于原模型在召回率上提升 9.56%, 准确率上提升了 10.67%。本文模型相比于 Pix2Pix-ResTrans 在召回率上提升了 1.72%, 准确率上提升 10.18%, 说明本文模型针对钨钨杆 SEM 图像具有针对性, 能准确定位晶界位置。

图 10 是各个改进模型的对比图像, 图中正确提取的部分采用实线圆圈, 错误提取的部分采用虚线圆圈。在图

表 6 不同 Pix2Pix 改进模型对比

Table 6 Comparison of different Pix2Pix improved models

模型	F1/%	Recall/ %	Precision/ %	推理 时间/s
Pix2PixGAN	52.91	64.94	44.65	5.573 3
Pix2Pix-Res	59.56	70.72	51.45	7.102 1
Pix2Pix-GP	56.28	67.99	48.02	7.690 7
Pix2Pix-ResTrans	69.60	74.50	65.32	8.582 6
本文模型	75.85	76.22	75.50	10.008 5

像 A 中, 本文模型和 Pix2Pix-ResTrans 均能有效补全遮掩, 其他改进模型在图 A 中存在错误提取或无法连接的情况。Pix2Pix 虽然将遮掩处连接, 但是其晶界是包含了细小晶粒晶界的复合晶界, 并非有效晶界。在图像 B 中, 本文模型与 Pix2Pix-Res 能够提取处中间晶粒叠加导致的晶界模糊问题, 而其他改进模型在此处存在些许断裂。但是针对图像 B 中左侧圈画位置的晶界识别上, 所有模型都存在误提取的情况。在图 C 中, 本文模型依旧能对遮掩位置的晶界进行修复, 但 Pix2Pix-ResTrans 却没能完整提取出晶界, 同时所有模型对研磨划痕仍然缺少一定的判断能力, 错误的当成晶界提取出来。

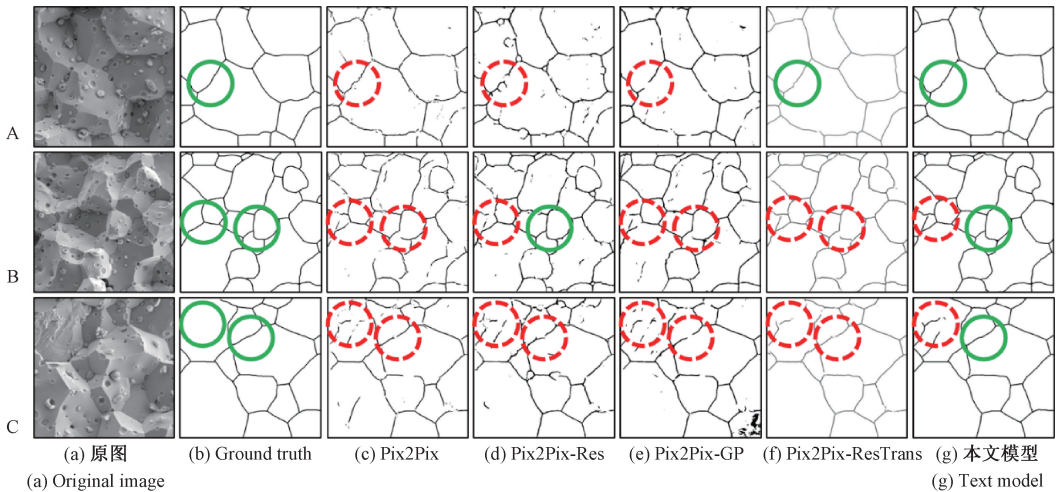


图 10 不同 Pix2Pix 改进策略的对比

Fig. 10 Comparison of different Pix2Pix improvement models

3.7 算法对比实验及分析

在钨晶粒晶界提取任务中, 本文方法对比了 5 种深度学习方法。使用测试集检验各个模型的性能, 测试集是由 750 张不同信号的钨钨杆 SEM 裁剪图像组成, 如表 1 中 test data 所示。在测试集上进行各项实验, 其中用粗体标注了每一列最好结果。分割性能评价指标如表 7 所示。

本文模型在 F1 值和准确率上取得了最优性能, 召回率上排在第 3。RCF 与 UCTransNet 采用了边缘区域扩展策略以增强边界检测的鲁棒性, 并减少了真实边界的漏检, 因此均在召回率上表现优异。但该策略也可能引入较

多伪边缘, 导致准确率下降, 从而影响像素级边界定位的精度, 难以满足对边界定位精度要求较高的精细分割任务。为了更全面地反映模型在边缘提取任务中的实际表现, 需综合考虑 F1 值、准确率与召回率的表现。与 RCF 对比, 除召回率外, 其余指标提升 28.32%、39.52%; 与 Pix2PixGAN 对比, 各项指标提升 13.02%、16.87%、6.49%; 与 RINDNet 对比, 各项指标提升 6.51%、4.58%、7.9%, 说明针对钨钨杆 SEM 图像, 本文模型所使用的 EGA 模块能够有效的将边缘信息提取出来, 并依靠 GAN 的对抗生成, 来辅助生成完整晶界, 而 RINDNet 提取的多

表 7 不同方法的边缘提取性能

Table 7 Edge extraction performance of different methods

模型	F1/%	Recall/%	Precision/%	推理时间/s
RCF(2017)	44.15	95.16	28.80	7.565 4
Pix2PixGAN(2017)	59.45	70.72	51.45	7.102 1
RINDNet(2021)	65.96	72.63	60.42	12.740 8
UCTransNet(2022)	42.23	94.17	27.23	20.428 8
MEGANet(2024)	56.51	70.37	47.30	11.112 8
本文模型	72.47	77.21	68.32	10.008 5

种边缘类型会错误地引导模型将部分棱边识别出来;与 UCTransNet 对比,除召回率外,其余指标提升 30.24%,41.09%;与 MEGANet 对比,各项指标提升 15.96%,21.02%,6.84%。最终实验数据对比表明:EGA 模块,EUCCB 以及复合损失函数训练生成器能够有效地增强网络对钨晶粒晶界的特征,更好表达晶界信息,更有效地区别出棱边-晶界和补充缺失的遮掩晶界。由于本文模型在 Pix2PixGAN 上进行了模块优化,因此在推理速度上略慢于 Pix2PixGAN,整体推理速度排在所有模型第 3,且优于同样采用 EGA 模块的 MEGANet。

图 11 是不同模型同本文模型的结果对比,图中正确提取的部分采用实线圆圈,错误提取的部分采用虚线圆圈。在图 A 和 B 中圈画位置主要是细小晶粒遮掩晶界问

题。本文模型与 UCTransNet 均能够得到完整、不断开的晶界,而其余网络均存在断裂晶界与小晶粒晶界提取的情况。但由于 UCTransNet 采用泛化预测,其生成的晶界仍然有着遮掩晶粒的形貌特征,这样得到的晶界不精准。本文模型能够有效解决小晶粒遮掩的情况,从而在保证晶界完整的同时,并于真实晶界近似吻合。在图 C、D 和 E 中圈画位置主要为晶棱与晶界区分度低以及晶粒叠加问题。本文模型在图 C 中存在误分割情况,但在另外图 D 与 E 中能够较完整的提取出正确的晶界;其他网络在图 C 中存在优异表现,但在图 D 与 E 中都难以提取出圈画位置的有效晶界,甚至存在“细碎噪声”。本文模型针对晶棱与晶界的差异化识别仍存在欠缺,但整体的识别准确度优于其他模型。

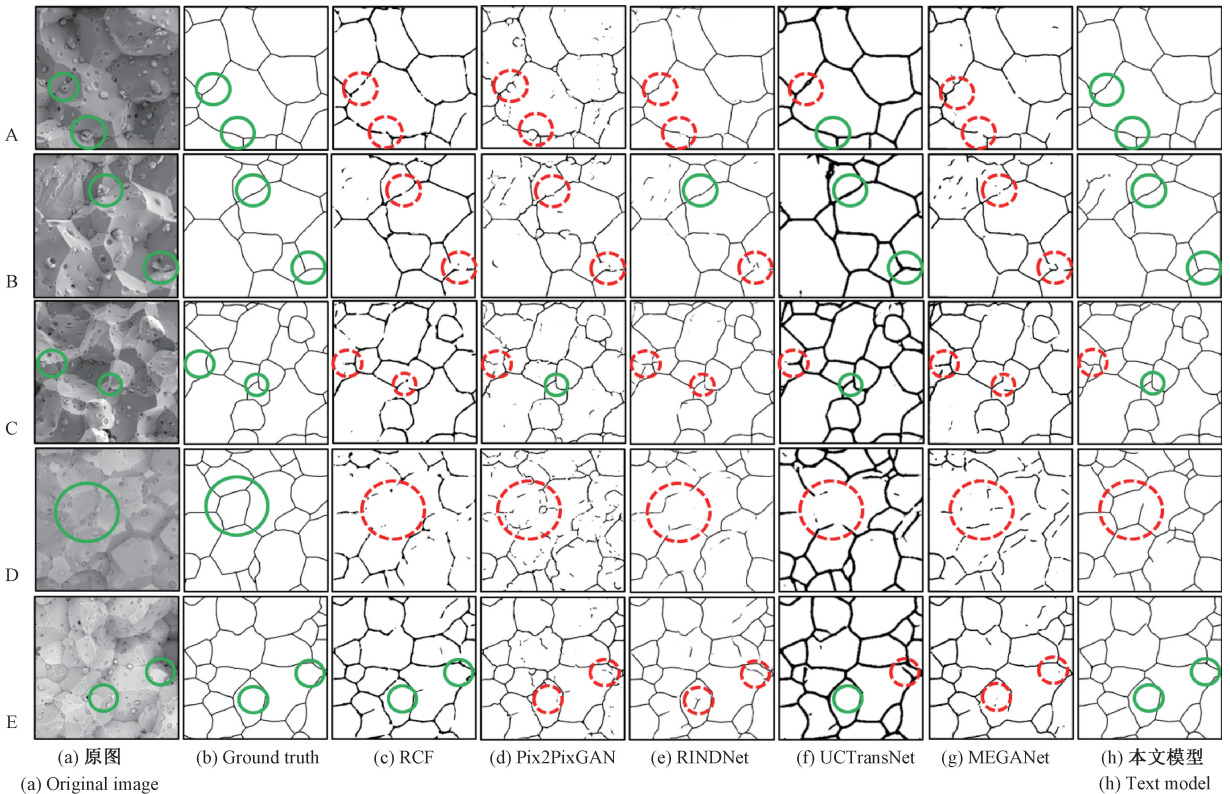


图 11 本文模型与其他模型的钨晶粒提取结果对比

Fig. 11 Comparison of tungsten grain extraction results between the proposed model and other models

图 12 展示的是完整视域下镧钨杆 SEM 图像的钨晶粒提取结果,具体由相应的裁剪图像拼接得到。图中可以看出,RCF 在多处被遮掩的晶界提取结果断裂,且对晶粒与晶界对比度低的图像提取困难;Pix2PixGAN 也存在相似问题,将遮掩的小晶粒晶界错误的提取出;RINDNet 能够较完整的提取出标签位置的晶界,但存在较多欠分割;UCTransNet 提取的晶界较为完整,且没有过多晶粒

导致的伪边界噪声,但在晶界定位精度不足,在晶粒与晶界区分度低的图像中仍存在欠分割;MEGANet 在除晶界定位精度以外,与 UCTransNet 存在的问题类似;本文模型在多张图像上的晶界提取结果上表现优异,且晶界预测结果与标签高度重合。尽管偶有晶粒误判为晶界的情况,但整体表现优于其他对比模型,能够验证本文模型的实用性。

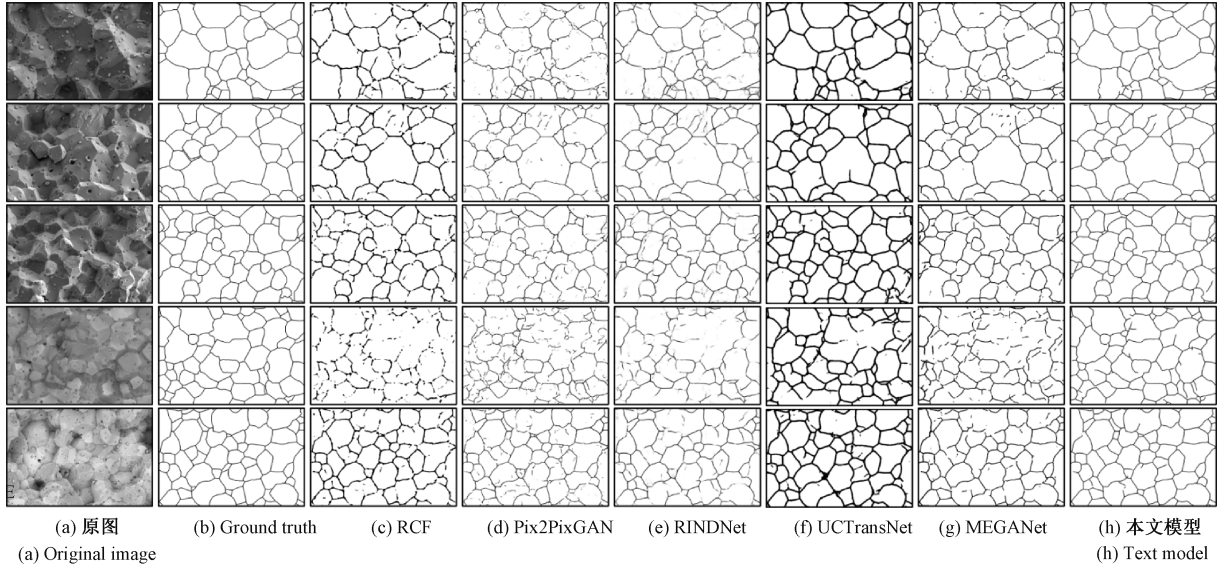


图 12 完整视域下的镧钨杆 SEM 图像预测结果对比

Fig. 12 Comparison of prediction results of lanthanum tungsten rod SEM images in complete field of view

4 结 论

为了高效精准地提取镧钨杆 SEM 图像中的钨晶粒晶界,提出了一种基于 Pix2PixGAN 的钨晶粒晶界提取模型 Pix2Pix-EG。算法模型保留 UNet 架构,使用 Res2Net 作为主干网络增强特征提取能力;EGA 模块代替传统跳跃连接,优化多尺度信息融合以保留晶界细节;采用 EUCB 模块抑制上采样伪影并强化特征关联性;选用 WBCE 与 WIoU 混合损失函数替代 L2 损失,有效解决边缘像素稀疏性问题;引入梯度惩罚策略提升了模型训练的稳定性与生成多样性。在自建数据集上的实验充分表明,所提模型在晶界提取精度与连续性均优于对比模型。后续工作将继续研究针对遮掩任务的对抗生成网络模型,扩充及改善自建镧钨杆 SEM 图像数据集,最后将算法更好的应用于实际工程中。

参考文献

- [1] LIN F X, FANG H X, LIU H, et al. Automatic detection of grains in partially recrystallized microstructures using deep learning [J]. Materials Characterization, 2025, 219: 114576-114591.
- [2] XU X Y, YU Z F, CHEN W Y, et al. Automated analysis of grain morphology in TEM images using

convolutional neural network with CHAC algorithm[J]. Journal of Nuclear Materials, 2024, 588: 154813.

- [3] LIU G Q, YAO S, ZHOU Y N, et al. Boundary-guided global-Local feature fusion network for polyp segmentation[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 1-12.
- [4] 张兴伟,刘韵婷,陈浩,等.基于 U-Net 判别器的轮胎图像缺陷检测方法研究[J].电子测量技术,2024, 47(16):139-146.
ZHANG X W, LIU Y T, CHEN H, et al. Research on tire image defect detection method based on U-Net discriminator[J]. Electric Measurement Technology, 2024, 47(16): 139-146.
- [5] 曾中华,曹东.基于改进 Unet 的矿石图像分割[J].电子测量技术,2023,46(21):176-182.
ZENG ZH H, CAO D. Ore image segmentation based on improved Unet [J]. Electric Measurement Technology, 2023, 46(21): 176-182.
- [6] 杨彬,赵倩,赵琰.融合边缘检测的遥感图像超分辨率重建算法[J].电子测量技术,2023,46(10):136-143.
YANG B, ZHAO Q, ZHAO Y. Super-resolution reconstruction of remote sensing image based on edge

- detection [J]. *Electric Measurement Technology*, 2023, 46(10): 136-143.
- [7] LIU S, AMIN-AHMADI B, LIU R, et al. Automated phase segmentation and quantification of high-resolution TEM image for alloy design [J]. *Materials Characterization*, 2023, 199: 112779.
- [8] HARIKRISHNA K, DAVIDSON M J, REDDY G D. New method for microstructure segmentation and automatic grain size determination using computer vision technology during the hot deformation of an Al-Zn-Mg powder metallurgy alloy [J]. *Journal of Materials Engineering and Performance*, 2023, 34(1): 1-11.
- [9] RAO H D, LIU D, JIN F, et al. A machine learning approach for segmentation and characterization of microtextured regions in a near- α Titanium alloy [J]. *Crystals*, 2023, 13(10): 1422.
- [10] PATRICK M J, ECKSTEIN J K, LOPEZ J R, et al. Automated grain boundary detection for bright-field transmission electron microscopy images via U-Net [J]. *Microscopy and Microanalysis*, 2023, 29 (6): 1968-1979.
- [11] WANG Y H, HE Q, XIE ZH. Grain boundary extraction method based on pixel relationship [J]. *Measurement*, 2022, 202: 111796.
- [12] WARREN P, RAJU N, PRASAD A, et al. Grain and grain boundary segmentation using machine learning with real and generated datasets [J]. *Computational Materials Science*, 2024, 233: 112739.
- [13] LIU Y, CHENG M M, HU X W, et al. Richer convolutional features for edge detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(8): 1939-1946.
- [14] PU M Y, HUANG Y P, GUAN Q J, et al. Rindnet: Edge detection for discontinuity in reflectance, illumination, normal and depth [C]. *IEEE/CVF International Conference on Computer Vision*, 2021: 6879-6888.
- [15] WANG H, CAO P, WANG J Q, et al. Utransnet: Rethinking the skip connections in u-net from a channel-wise perspective with transformer [C]. *AAAI Conference on Artificial Intelligence*, 2022, 36 (3): 2441-2449.
- [16] WANG T C, LIU M Y, ZHU J Y, et al. High-resolution image synthesis and semantic manipulation with conditional gans [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 8798-8807.
- [17] MA B Y, WEI X Y, LIU C N, et al. Data augmentation in microscopic images for material data mining [J]. *NPJ Computational Materials*, 2020, 6(1): 125.
- [18] ISOLA P, ZHU J Y, ZHOU T H, et al. Image-to-image translation with conditional adversarial networks [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 1125-1134.
- [19] ANANTATAMUKALA A, KRISHNA K V M, DAHOTRE N B. Generative adversarial networks assisted machine learning based automated quantification of grain size from scanning electron microscope back scatter images [J]. *Materials Characterization*, 2023, 206: 113396.
- [20] LI M C, CHEN D L, LIU S X, et al. Semisupervised boundary detection for aluminum grains combined with transfer learning and region growing [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 34(9): 6158-6172.
- [21] ZHOU P, ZHANG X Y, SHEN X J, et al. Multi-phase material microscopic image segmentation for microstructure analysis of superalloys via modified U-Net and rectify strategies [J]. *Computational Materials Science*, 2024, 242: 113063.
- [22] BUI N T, HOANG D H, NGUYEN Q T, et al. Meganet: Multi-scale edge-guided attention network for weak boundary polyp segmentation [C]. *IEEE/CVF Winter Conference on Applications of Computer Vision*, 2024: 7985-7994.
- [23] RAHMAN M M, MUNIR M, MARCULESCU R. EMCAD: Efficient multi-scale convolutional attention decoding for medical image segmentation [C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024: 11769-11779.
- [24] ZHANG L, RAO AN Y, AGRAWALA M. Adding conditional control to text-to-image diffusion models [C]. *IEEE/CVF International Conference on Computer Vision*, 2023: 3836-3847.
- [25] JADHAV P, SAWAL M, ZAGADE A, et al. Pix2pix generative adversarial network with resnet for document image denoising [C]. *2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2022: 1489-1494.
- [26] ZHU Z Y, LIN P F, ZHONG L L, et al. Based on TransRes-Pix2Pix network to generate the OBL image during SMILE surgery [J]. *Frontiers in Cell and Developmental Biology*, 2025, 13: 1598475.
- [27] TIREL L, ALI A M, HASHIM H A. Novel hybrid integrated Pix2Pix and WGAN model with gradient penalty for binary images denoising [J]. *Systems and*

- Soft Computing, 2024, 6: 200122.
- [28] CUI S B, JIANG Y. Effective Lipschitz constraint enforcement for Wasserstein GAN training[C]. 2017 2nd IEEE International Conference on Computational Intelligence and Applications(ICCIA), 2017: 74-78.
- [29] GAO S H, CHENG M M, ZHAO K, et al. Res2net: A new multi-scale backbone architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(2): 652-662.

作者简介

文金瑞, 硕士研究生, 主要研究方向为图像处理、计算机

视觉等。

E-mail: 1571304298@qq.com

吴晓红, 博士, 副教授, 主要研究方向为图像处理与模式识别、计算机视觉、无线通信等。

E-mail: wxh@scu.edu.cn

滕奇志(通信作者), 博士, 教授, 主要研究方向为数字图像处理与模式识别、三维图像重建及分析、计算机应用等。

E-mail: qxteng@scu.edu.cn

何海波, 本科, 高级工程师, 主要研究方向为图像处理、计算机应用。

E-mail: 351813868@qq.com