面向铝合金微观组织分割的弱监督结构约束方法

薛 正，陈大力†，李鸣春，刘士新

(东北大学 信息科学与工程学院，沈阳 110819)

摘 要: 铝合金微观组织的自动分割是材料性能评估和工业质量控制的重要环节. 现有基于深度学习的微观组织分割方法通常依赖大量像素级标注数据进行全监督训练, 然而精确标注往往需要材料专家参与, 标注成本高且效率较低. 针对该问题, 提出一种面向铝合金微观组织分析的弱监督分割方法. 该方法构建一种双分支分割网络: 分类分支在图像级标签的监督下学习全局语义信息; 结构约束分支在输入图像的监督下通过所提出的一致性约束损失进行自监督学习, 从而获得像素级约束并学习微观组织的局部细节特征. 在推断阶段, 提出结构相似性扩散算法, 对微观组织激活区域进行优化以获得更加准确的分割结果. 将所提出的方法在企业采集的铝合金微观组织数据集上进行验证, 结果表明: 该方法在 mIoU 指标上达到 76.0%, 显著优于传统弱监督分割方法, 且分割精度接近全监督模型, 同时模型参数量降低约 60%. 所提方法在工业材料微观组织自动分析中具有良好的应用前景.

关键词: 铝合金微观组织；弱监督学习；智能分割；工业人工智能；材料图像分析；深度学习

中图分类号: TP391.41；TP181 **文献标志码:** A

DOI: 10.13195/j.kzyjc.2026.0269

引用格式: 薛正, 陈大力, 李鸣春, 等. 面向铝合金微观组织分割的弱监督结构约束方法 [J]. 控制与决策, xxxx, x(x): xxxx-xxxx.

Weakly supervised structure-constrained method for aluminum alloy microstructure segmentation

XUE Zheng，*CHEN Da-li*†，*LI Ming-chun*，*LIU Shi-xin*

(College of Information Science and Engineering，Northeastern University，Shenyang 110819，China)

Abstract: Aluminum alloy microstructure segmentation is an important step in material property evaluation and industrial quality control. Existing deep learning-based microstructure segmentation methods usually rely on large amounts of pixel-level annotations for fully supervised training. However, precise annotation often requires the participation of domain experts, resulting in high labeling cost and low efficiency. To address this issue, this paper proposes a weakly supervised segmentation method for aluminum alloy microstructure analysis. The proposed method constructs a dual-branch segmentation network, in which the classification branch learns global semantic information under image-level supervision, while the structure-constrained branch performs self-supervised learning under the guidance of the input image and the proposed consistency constraint loss, thereby introducing pixel-level constraints and learning local detailed features of microstructures. In the inference stage, a structural similarity diffusion algorithm is further proposed to refine the activated microstructure regions and obtain more accurate segmentation results. The proposed method is validated on an aluminum alloy microstructure dataset collected from industry. Experimental results show that the proposed method achieves an mIoU of 76.0%, significantly outperforming conventional weakly supervised segmentation methods, while attaining performance close to that of fully supervised models. Meanwhile, the number of parameters is reduced by about 60%. These results demonstrate that the proposed method has promising application potential in automatic analysis of industrial material microstructures.

Keywords: aluminum alloy microstructure； weakly supervised learning； intelligent segmentation； industrial artificial intelligence；material image analysis；deep learning

0 引 言

铝合金因其密度低、强度高以及良好的耐腐蚀

收稿日期: 2026-03-30；**录用日期:** 2026-09-20.

基金项目: 国家自然科学基金项目 (61773104)；中央高校基本科研业务费专项资金项目 (N2404011, N2404008).

责任编委: 贲晛烨.

†通信作者. E-mail: chendali@ise.neu.edu.cn

本文附带电子附录文件, 可登录本刊官网该文“资源附件”区自行下载阅览.

性能, 被广泛应用于航空航天、轨道交通和汽车制造等重要工业领域. 材料的宏观性能在很大程度上取决于其微观组织结构, 因此, 对铝合金微观组织进行准确分析是材料性能评估和生产质量控制中的关键环节[1]. 在传统金相分析过程中, 研究人员通常通过光学显微镜获取材料的金相图像, 并依赖专家经验对不同组织相进行分割和统计. 这种人工分析方式不仅效率较低, 而且容易受到主观因素影响, 难以满足现代工业生产对高效率和高一致性的要求. 因此, 研究高效、自动化的微观组织图像分析方法具有重要意义[2].

近年来, 随着深度学习技术的发展, 基于卷积神经网络的图像分析方法在材料微观组织分割领域得到了广泛应用[3,4]. 其中, 语义分割方法能够对图像中的不同组织相进行像素级划分, 从而实现对材料微观结构的精细分析, 在材料科学和工业检测领域具有重要应用价值[5-10]. 然而, 大多数深度学习分割方法依赖大量像素级标注数据进行全监督训练. 对于金相图像而言, 像素级标注通常需要材料领域专家逐像素进行标记, 标注过程复杂且耗时, 获取大规模高质量标注数据具有较高成本, 这在一定程度上限制了深度学习方法在材料微观组织分析中的应用[11].

为降低标注成本, 弱监督语义分割方法逐渐成为近年来研究的热点[12,13]. 该类方法通常仅利用图像级标签等弱标注信息进行训练, 通过类激活图 (Class Activation Map, CAM)[14] 等技术实现对目标区域的粗定位, 并进一步生成伪标签以训练分割网络. 相比全监督方法, 弱监督方法能够显著降低数据标注成本, 因此在图像的密集预测任务中得到了广泛研究. 然而, 现有弱监督分割方法大多采用两阶段学习框架, 即首先训练分类网络生成类激活图, 然后利用生成的伪标签训练分割网络. 这类方法通常存在训练流程复杂、误差传播以及模型参数规模较大等问题. 此外, 由于缺乏有效的像素级约束, 类激活图往往只关注目标的判别性区域, 难以完整覆盖目标结构, 从而影响最终分割结果的精度.

针对上述问题, 本文围绕挖掘微观组织图像的结构约束, 提出一种面向铝合金微观组织分析的弱监督分割方法. 该方法通过构建双分支网络结构, 使模型在训练过程中能够同时利用图像级标注信息和图像本身所蕴含的结构特征. 其中, 一条分支侧重于学习图像整体语义信息, 以保证模型具备稳定的类别判别能力; 另一条分支则利用图像内部的结构关系对特征学习过程进行约束, 从而引导网络逐步关注微观组织的细节区域. 在模型推断阶段, 通过引入结构相似性扩散算法对初始激活结果进行进一步调整, 使得预测区域能够更加完整地覆盖目标组织. 与传统弱监督分割方法相比, 所提出的方法无需依赖复杂流程即可获得较高精度的分割结果, 同时模型规模更加轻量, 因而在工业材料微观组织自动分析任务中具有良好的应用前景.

本文的主要贡献如下:

1) 提出一种面向铝合金微观组织分析的弱监督分割方法, 在仅利用图像级标签进行训练的情况下实现对微观组织的像素级分割;

2) 构建一种双分支分割网络结构, 通过分类分支学习全局语义信息, 并利用结构约束分支在输入图像的监督下学习微观组织的局部细节特征;

3) 提出一种基于结构相似性的扩散算法, 用于优化微观组织激活区域, 从而进一步提升分割精度;

4) 在企业采集的铝合金微观组织数据集上进行了实验验证, 结果表明所提出方法在 mIoU 指标上达到 76.0%, 在保持较小模型规模的同时获得了接近全监督方法的分割性能.

1 相关工作

1.1 弱监督分割方法

弱监督语义分割 (Weakly Supervised Semantic Segmentation, WSSS) 旨在利用不完整或弱形式标注信息实现像素级图像分割. 与传统全监督分割方法依赖大量像素级标注不同, 弱监督分割通常只需要图像级标签、点标注或边界框等较弱的监督信息, 从而显著降低数据标注成本. 早期研究主要采用多实例学习 (Multiple Instance Learning, MIL)[15] 和自步学习 (Self-Paced Learning)[16] 等策略, 通过在图像级标签约束下逐步挖掘潜在的像素级目标区域. 这类方法能够在一定程度上实现弱监督分割, 但由于缺乏明确的空间约束, 模型往往难以准确定位完整目标区域, 分割精度仍然有限.

近年来, 基于类激活图的技术路线成为弱监督分割研究的主流框架. 该类方法通常首先训练分类网络生成类激活图以定位目标区域, 然后通过区域扩展或语义传播策略生成伪标签, 最后利用伪标签训练分割网络. 然而, 由于分类网络更关注判别性区域, 生成的类激活图往往呈现出明显的稀疏性. 针对这一问题, 研究者提出了多种改进方法. 例如, IRN 方法[17] 通过学习像素间的亲和关系来实现区域传播, 从而扩展激活区域并生成更加完整的伪标签; SEAM 方法[18] 引入自监督一致性约束, 使模型在不同图像变换下保持激活区域的一致性, 从而提升

CAM 质量; WeCLIP 方法[19] 利用 CLIP 跨模态表示能力, 通过视觉与文本语义信息的结合提高弱监督分割性能; MoRe 方法[20] 将 Transformer 结构引入弱监督分割, 通过全局注意力机制增强特征建模能力. 这些方法在一定程度上缓解了 CAM 稀疏性问题, 但多数方法仍然依赖复杂的多阶段训练流程或额外网络模块, 导致模型结构复杂、训练成本较高.

与上述方法相比, 本文方法在训练流程和模型结构上更加简洁, 主要体现在以下两个方面: (1) 相比 IRN、SEAM 等两阶段弱监督方法, 本文方法无需单独训练伪标签生成网络或第二阶段分割网络, 而是通过统一网络直接实现像素级预测, 属于端到端的单阶段弱监督分割方法; (2) 相比 MoRe、WeCLIP 等依赖大型模型或复杂流程的方法, 本文方法仅在训练阶段引入结构约束损失, 并在推断阶段增加结构相似性扩散算法, 即可实现分割性能提升, 模型参数规模较小、训练流程简单, 具有较好的灵活性和工程可实现性.

1.2　微观组织图像分割方法

微观组织图像分割是材料图像分析中的重要研究方向, 其目标是从金相图像中分割不同组织结构, 以支持材料性能评估和工艺优化. 随着深度学习的发展, 越来越多研究采用卷积神经网络进行微观组织图像分割. 例如, Li 等[21] 提出了一种结合多任务学习与生成对抗网络的方法, 同时实现晶界检测与第二相分割, 通过联合优化边界信息与区域信息提升了模型对复杂微观组织的分割能力. 此外, 部分研究还通过引入注意力机制和多尺度特征融合策略提升模型对复杂组织结构的表达能力. 例如, Sun 等[22] 提出通道多尺度自适应注意力网络, 通过融合多尺度卷积与注意力机制增强模型对复杂纹理结构的建模能力, 从而提升微观组织分割性能.

针对微观组织图像中标注样本不足以及分割结果可靠性不足的问题, 部分研究进一步从不确定性建模和弱监督学习角度开展探索. 例如, Li 等[23] 提出一种基于 Dirichlet 分布的不确定性估计方法, 用于铝合金金相图像分割与测量, 在获得分割结果的同时实现对预测置信度的刻画, 从而提高模型在复杂微观组织场景下的稳定性. 除此之外, Han 等[24] 提出一种基于伪标签的弱监督分割方法, 通过少量标注样本与伪标签迭代优化实现纹理图像分割. 该方法在降低标注成本方面具有一定优势, 但其监督形式仍依赖像素级涂鸦标注, 需要人工提供空间位置先验信息, 因此标注成本仍然高于仅使用图像级标签的弱监督方法.

与上述方法相比, 本文方法在监督方式和模型结构方面具有明显优势, 主要体现在两个方面: (1) 现有微观组织分割方法大多依赖像素级标注、部分像素级标注或涂鸦级标注, 而本文方法仅需要图像级标签即可完成模型训练, 进一步降低了数据标注成本; (2) 许多现有方法依赖复杂的编码器—解码器结构或额外的不确定性建模模块以提高分割性能, 而本文方法采用结构约束的弱监督分割网络, 模型结构更加轻量, 训练过程相对简单, 对计算资源要求较低, 更适用于工业材料图像分析场景.

2　方　法

2.1　方法整体框架

本节论述所提出的面向铝合金微观组织分析的弱监督结构约束分割方法. 该方法仅利用图像级标签进行训练, 通过联合建模微观组织的低层结构一致性与高层结构关联性, 实现对组织区域的像素级分割. 整体方法主要包括两个核心组成部分: 双分支结构约束分割网络和结构相似性扩散算法.

双分支结构约束分割网络由共享主干和两个任务分支构成. 其中, 分类分支用于学习微观组织图像的全局类别信息, 以保证网络具备稳定的图像级判别能力; 结构约束激活分支则生成类激活图, 并结合结构一致性约束损失, 利用输入图像自身的低层结构信息对激活响应进行约束, 从而提升组织区域的定位完整性. 该网络在训练阶段完成图像级监督学习, 并在推断阶段输出初始类激活图.

在此基础上, 结构相似性扩散算法进一步利用网络中间层特征所蕴含的高层结构关系, 对初始类激活图进行扩散优化. 该模块通过保留深层特征空间中的正相关结构连接, 将高置信度响应传播到具有相似结构属性的区域, 从而补全弱监督条件下容易缺失的组织响应区域. 方法整体流程如图 1 所示, 下面分别介绍各组成部分.

2.2　双分支结构约束分割网络

本节主要介绍所提出的双分支结构约束分割网络. 该网络由共享特征提取模块、分类分支和结构约束激活分支组成. 前者用于提取图像表征, 后两个分支分别承担图像级分类与区域定位任务.

设输入微观组织图像为

$$I \in \mathbb{R}^{H\times W\times 3}. \tag{1}$$

H和W分别表示图像的高度和宽度. 输入图像首先经过主干网络提取深层特征:

$$F = f_\theta(I). \tag{2}$$

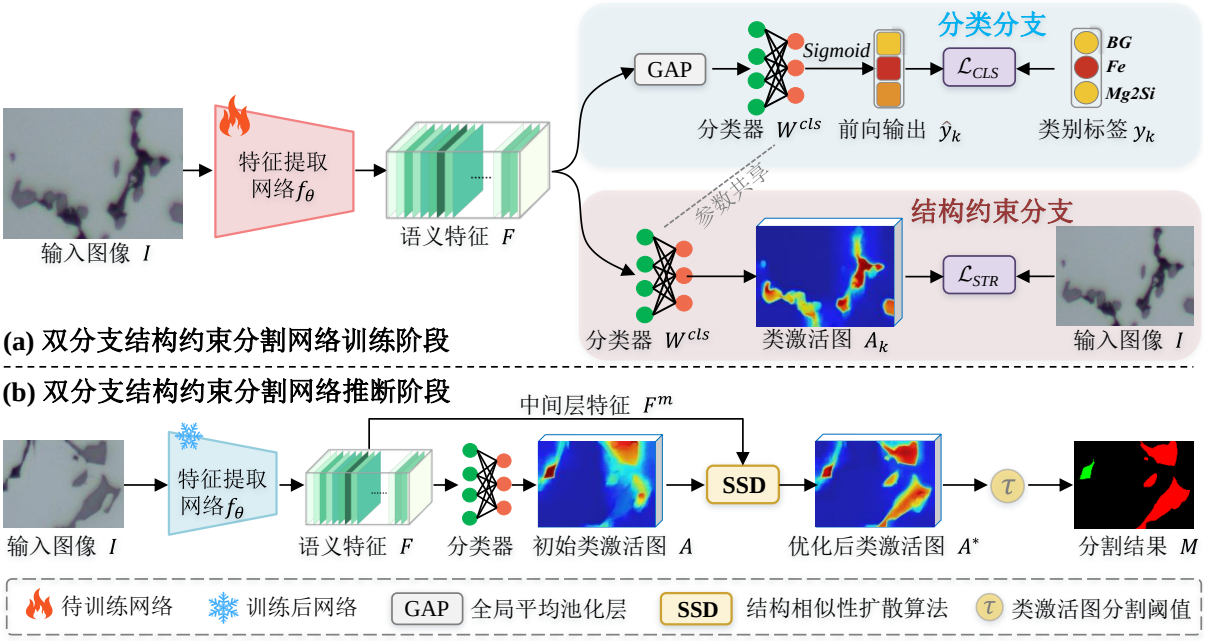

图 1 弱监督结构约束分割方法训练和推断图

其中, $f_\theta(\cdot)$表示参数为θ的特征提取网络, 输出特征图满足

$$F \in \mathbb{R}^{H' \times W' \times C}. \tag{3}$$

C表示特征通道数, H'和W'表示下采样后的空间尺寸. 深层特征F随后被分别送入两个分支用于学习全局语义和目标结构.

2.2.1 分类分支

分类分支用于在图像级标签监督下学习微观组织图像的全局判别信息. 首先, 对特征图F进行全局平均池化, 得到各通道的全局描述量

$$z_c = \frac{1}{H'W'} \sum_{x=1}^{H'} \sum_{y=1}^{W'} F_c(x,y), \ c = 1, 2, \ldots, C. \tag{4}$$

进一步将各通道的全局描述量组成向量, 记为

$$\mathbf{z} = [z_1, z_2, \ldots, z_C]^\top \in \mathbb{R}^C. \tag{5}$$

随后, 通过线性分类器获得图像级预测结果

$$\hat{\mathbf{y}} = \sigma(W^{\text{cls}}\mathbf{z} + \mathbf{b}^{\text{cls}}). \tag{6}$$

其中, $W^{\text{cls}} \in \mathbb{R}^{K \times C}$和$\mathbf{b}^{\text{cls}} \in \mathbb{R}^K$分别表示分类器的权重矩阵与偏置向量, K为类别数, $\sigma(\cdot)$表示 Sigmoid 函数. 分类分支采用多标签分类损失进行优化:

$$L_{\text{cls}} = -\sum_{k=1}^{K} [y_k \log(\hat{y}_k) + (1 - y_k) \log(1 - \hat{y}_k)]. \tag{7}$$

其中, y_k和$\hat{y}_k$分别表示第k类的真实图像级标签与预测概率.

2.2.2 结构约束激活分支

为了在弱监督条件下获得微观组织区域的空间定位信息, 本文利用分类器权重直接从共享特征图中生成类激活图. 记W^{cls}的第k行对应的权重向量为

$$\mathbf{w}_k = [w_{k,1}, w_{k,2}, \ldots, w_{k,C}]^\top \in \mathbb{R}^C. \tag{8}$$

则类别k对应的类激活图A_k可由分类权重与特征图加权求和表示为

$$A_k(x,y) = \sum_{c=1}^{C} w_{k,c} F_c(x,y). \tag{9}$$

其中, $w_{k,c}$表示类别k在第c个特征通道上的分类权重, $F_c(x,y)$表示特征图F在第c个通道、位置(x,y)处的响应值. 相应地, 有

$$A_k \in \mathbb{R}^{H' \times W'}. \tag{10}$$

为了便于后续计算, 对类激活图进行归一化处理:

$$\tilde{A}_k(x,y) = \frac{A_k(x,y) - \min(A_k)}{\max(A_k) - \min(A_k) + \varepsilon}. \tag{11}$$

其中, ε为防止分母为零的极小常数.

2.2.3 结构一致性约束损失

由于分类监督更倾向于突出目标的局部判别性区域, 直接生成的类激活图通常难以完整覆盖微观组织结构. 为此, 本文在训练阶段引入结构一致性约束损失, 利用输入图像的低层结构信息对类激活图进行约束学习.

考虑到类激活图的空间分辨率为$H' \times W'$, 首先将输入图像下采样至相同尺度, 得到

$$I^{\downarrow} \in \mathbb{R}^{H' \times W' \times 3}. \tag{12}$$

设位置p与q处的像素向量分别为$I_p^{\downarrow}$与$I_q^{\downarrow}$, 则两位置之间的结构相似度定义为

$$S_{pq}=\exp\left(-\frac{\|I_p^{\downarrow}-I_q^{\downarrow}\|_2^2}{\sigma^2}\right). \tag{13}$$

其中, $S_{pq}\in[0,1]$表示位置p与q之间的结构相似度, σ为控制相似度衰减速度的参数. 将当前目标类别对应的归一化激活图$\tilde{A}_k$展开为向量形式, 记为

$$\mathbf{a}=\mathrm{vec}(\tilde{A}_k)\in\mathbb{R}^M. \tag{14}$$

其中, $M=H'W'$表示特征图中的像素总数, a_p表示位置p处的激活响应.

理想情况下, 若两个位置在图像中具有较高的结构相似度, 则它们在类激活图中的响应也应保持接近. 基于此, 结构一致性约束损失定义为

$$L_{\mathrm{str}}=\frac{1}{M}\sum_{p}\sum_{q\in\mathcal{N}(p)}S_{pq}(a_p-a_q)^2. \tag{15}$$

其中, $\mathcal{N}(p)$表示位置p在特征图空间中的局部邻域, N表示邻域大小. 该损失通过约束结构相似像素之间的响应一致性, 抑制类激活图的局部稀疏现象, 从而促进组织区域的完整定位.

2.2.4　分割网络的训练和推断

在训练阶段, 双分支结构约束分割网络通过图像级分类损失与结构一致性约束损失进行联合优化, 总体目标函数写为

$$L=L_{\mathrm{cls}}+\lambda L_{\mathrm{str}}. \tag{16}$$

其中, λ为平衡两项损失的权重系数.

在推断阶段, 首先利用结构约束激活分支生成初始类激活图, 并将其作为后续分割结果生成的基础. 若不引入进一步的后处理, 初始类激活图经过阈值化即可得到弱监督分割结果; 进一步地, 本文引入结构相似性扩散算法对其进行优化, 以获得更完整的组织区域响应.

2.3　结构相似性扩散算法

尽管结构一致性约束损失能够在训练阶段提升类激活图的结构完整性, 但弱监督模型生成的激活图仍然主要聚焦于目标的局部判别性区域, 难以完整覆盖微观组织结构. 为进一步优化类激活图的空间分布, 本文在推断阶段引入结构相似性扩散算法. 与训练阶段基于输入图像低层结构信息构建显式一致性约束不同, 该模块利用网络中间层特征所蕴含的高层结构关系, 对初始类激活图进行扩散优化. 详细流程如图 2 所示.

为构建深层结构关系, 本文从主干网络的中间层提取特征表示, 并将其作为结构相似性建模的基础. 具体而言, 本文选取骨干网络的第四层输出特征作为中间层特征F^{m}. 该特征与初始类激活图空间分辨率一致, 因此无需额外尺度对齐. 相较于浅层特征,

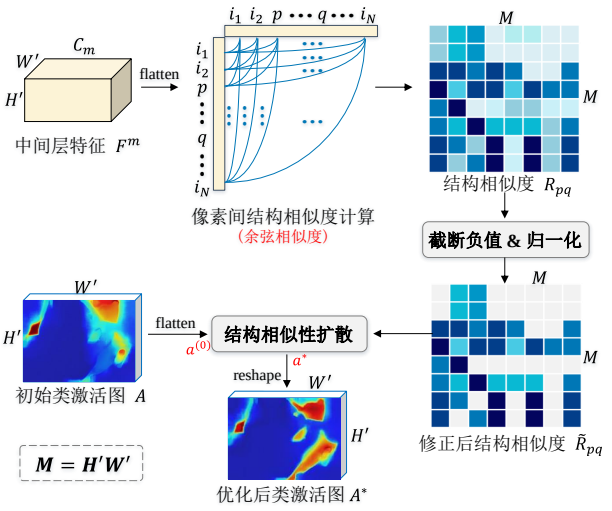

图 2　结构相似性扩散算法流程图

中间层特征能够表征更稳定的组织区域关联关系; 相较于过深层特征, 其仍保留较充分的空间结构细节, 因此更适合用于引导类激活图的结构传播. 记该中间层特征图表示为

$$F^{\mathrm{m}}\in\mathbb{R}^{H'\times W'\times C_m}. \tag{17}$$

其中, C_m表示中间层特征的通道数. 设位置p处对应的特征向量为$\mathbf{f}_p\in\mathbb{R}^{C_m}$, 则位置$p$与$q$之间的深层结构相似度定义为

$$R_{pq}=\frac{\mathbf{f}_p^{\top}\mathbf{f}_q}{\|\mathbf{f}_p\|_2\,\|\mathbf{f}_q\|_2}. \tag{18}$$

其中, $R_{pq}\in[-1,1]$表示位置p与q在深层特征空间中的余弦相似度.

为避免低相关或无位置对应关系的特征连接对扩散过程产生干扰, 本文首先对深层结构相似度进行非负截断, 仅保留具有正相关性的结构连接关系:

$$\hat{R}_{pq}=\mathrm{ReLU}(R_{pq}). \tag{19}$$

其中, $\hat{R}_{pq}\geqslant 0$表示位置p与q之间的非负结构关联强度. 进一步地, 为保证扩散过程中的数值稳定性与传播权重分布的合理性, 对非负结构连接权重进行归一化处理, 得到转移权重

$$\tilde{R}_{pq}=\frac{\hat{R}_{pq}}{\sum_{j=1}^{M}\hat{R}_{pj}+\varepsilon}. \tag{20}$$

其中, ε为防止分母为零的极小常数, $M=H'W'$表示特征图中的像素总数. 由此, $\tilde{R}_{pq}$表示位置p从位置q接收信息时的归一化传播权重.

设由结构约束激活分支生成的初始类激活图为A_k, 为便于在深层结构相似性图上进行迭代传播与更新, 将其展开为向量形式

$$\mathbf{a}^{(0)}=\mathrm{vec}(A_k). \tag{21}$$

则基于深层结构相似性的扩散过程可表示为

$$a_p^{(t+1)} = \sum_{q=1}^{M} \tilde{R}_{pq} a_q^{(t)},\ t = 0, 1, \ldots, T-1. \quad (22)$$

其中, $a_p^{(t)}$表示第t次迭代后位置p处的激活响应, T表示扩散迭代次数. 经过T次扩散后, 可以得到优化后的激活响应

$$\mathbf{a}^* = \mathbf{a}^{(T)}. \quad (23)$$

进一步将优化后的向量表示恢复为空间分布形式, 从而得到经过结构相似性扩散后的类激活图

$$A^* = \mathrm{reshape}(\mathbf{a}^*). \quad (24)$$

最后, 通过阈值分割获得组织区域预测结果

$$M(x, y) = \begin{cases} 1, & A^*(x, y) > \tau \\ 0, & \text{otherwise} \end{cases} \quad (25)$$

其中, τ为阈值参数.

通过结构相似性扩散算法, 模型能够充分利用网络中间层特征所蕴含的高层结构关系, 对初始类激活图进行进一步优化, 从而在无需额外监督的情况下获得更加完整的组织分割结果.

3 实 验

3.1 数据集

本文采用铝合金金相图像数据集验证所提方法. 该数据集共包含 833 张由光学显微镜采集并经裁剪处理得到的微观组织图像, 分辨率统一为256×192像素. 数据集按 8:2 进行固定随机划分, 训练集和测试集分别包含 666 张和 167 张图像. 为保证实验结果的可复现性, 本文将随机种子固定设置为 0, 所有对比实验、消融实验和超参数分析均采用相同的数据划分. 在材料专家指导下, 数据集同时提供图像级标注与像素级标注. 图像中主要包含 3 类组织: 浅灰色 Al 基体、黑色骨架状 Mg_2Si 相以及深灰色块状 $FeMnAl_6$相. 其中, Al 基体视为背景, Mg_2Si 相和 $FeMnAl_6$相视为前景类别. 图像级标注用于弱监督训练, 像素级标注仅用于性能评估和公平对比.

3.2 实验细节

3.2.1 网络设置与训练策略

所提双分支结构约束分割网络采用 ResNet50 作为主干, 并使用 ImageNet 预训练参数初始化, 输出步长设为 16. 训练时采用 SGD 优化器, 初始学习率为 0.1, 动量为 0.9, 权重衰减为4×10^{-5}, 训练 30 个 epoch, 批大小设为 16. 数据增强包括随机尺度缩放、随机翻转和归一化, 其中输入尺度随机设为 320–640 像素, 再裁剪为480 × 480. 测试阶段采用多尺度推断生成类激活图, 并结合 CRF 后处理优化边界.

3.2.2 评价指标与超参数设置

本文采用 mIoU 作为主要评价指标, 同时使用 mDice、mRecall、mSpec.和 mPrec.进行辅助分析. 关键超参数包括结构一致性损失权重λ、扩散次数T、分割阈值τ和局部邻域大小N. 综合实验结果, 最终设定$\lambda = 0.3$、$T = 14$、$\tau = 0.5$、$N = 5$. 结构相似度参数σ固定为 1, 不再作为独立超参数进行调节, 以避免引入额外的参数搜索过程.

3.3 性能实验

3.3.1 定量比较

本文以 DeepLabV2 在全监督设置下的结果作为性能上界参考, 并选取 5 种弱监督方法 CAM、IRN、SEAM、WeCLIP 及 MoRe 进行比较, 结果见表 1. DeepLabV2 取得最佳结果, mIoU 和 mDice 分别为 76.6% 和 85.9%. 在弱监督方法中, CAM、IRN 和 SEAM 的 mIoU 分别为 52.7%、59.2% 和 53.3%, 说明仅依赖分类定位或有限传播难以充分恢复复杂微观组织. WeCLIP 在本数据集上表现较弱, mIoU 仅为 45.1%; MoRe 达到 68.2%, 表现出较强竞争力. 相比之下, 本文方法取得 76.0% 的 mIoU 和 85.4% 的 mDice, 在所有弱监督方法中表现最佳, 并已接近全监督 DeepLabV2. 与此同时, 本文方法在 mPrec.上达到 85.8%, 显著优于其他弱监督方法, 说明其能够在保证区域完整性的同时有效抑制背景误检.

表 1 不同方法在测试集上的定量比较

方法	mDice	mIoU	mRecall	mSpec.	mPrec.
DeepLabV2	85.9	76.6	81.3	91.6	91.5
CAM[14]	64.7	52.7	85.1	92.0	56.5
IRN[17]	71.3	59.2	84.7	92.1	63.9
SEAM[18]	65.4	53.3	85.3	92.1	57.2
WeCLIP[19]	54.8	45.1	74.1	96.0	63.7
MoRe[20]	79.3	68.2	**93.6**	**98.2**	71.0
Ours	**85.4**	**76.0**	85.1	93.5	**85.8**

3.3.2 定性比较

图 3 给出了不同方法在测试集上的典型可视化结果, 其中绿色和红色分别表示两类前景组织. CAM 只能定位局部判别区域, IRN 和 SEAM 虽能扩展目标范围, 但在细长骨架结构和复杂边界处仍存在明显漏检. WeCLIP 容易产生较大范围误激活, MoRe 在区域完整性上有所改善, 但边界仍不够准确. 相比之下, 本文方法能够更完整地恢复 Mg_2Si 骨架状结构和 $FeMnAl_6$相块状结构, 并有效减少背景误分割, 结果与真实标注更为接近, 说明所提结构一致性约束和结构相似性扩散能够提升局部形态建模与空间结构恢复能力.

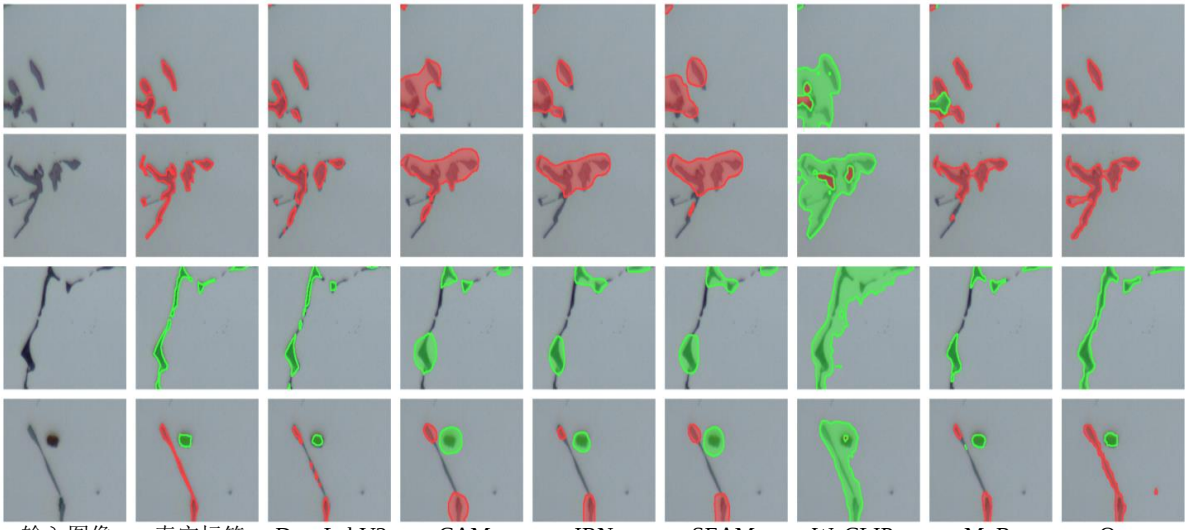

图 3　不同方法在测试集上的定性比较

3.4　消融实验

3.4.1　部件消融

为验证各组成模块的作用, 本文比较了基线模型L_{cls}、加入结构一致性损失的$L_{cls}+L_{str}$、加入结构相似性扩散的$L_{cls}+SSD$以及完整模型$L_{cls}+L_{str}+SSD$, 结果见表 2 和图 4. 由表 2 可知, 基线模型的 mIoU 仅为 52.7%. 引入L_{str}后, mIoU 提升至 69.5%, 表明结构约束能够有效增强区域定位能力; 仅加入 SSD 时, mIoU 达到 68.3%, 说明推断阶段的结构传播同样有效. 联合两者后, 模型取得 76.0% 的 mIoU 和 85.4% 的 mDice, 性能最佳. 图 4 的可视化结果也表明, 两模块分别改善区域完整性与边界质量, 具有良好的互补性.

表 2　不同模块组合的消融实验结果

方法	mDice	mIoU	mRecall	mSpec.	mPrec.
L_{cls}	64.7	52.7	85.1	92.0	56.5
$L_{cls}+L_{str}$	80.3	69.5	90.7	96.0	73.4
$L_{cls}+SSD$	79.5	68.3	85.3	92.8	75.0
$L_{cls}+L_{str}+SSD$	85.4	76.0	85.1	93.5	85.8

需要注意的是, 相较于仅引入L_{str}的模型, 完整模型的 mRecall 略有下降, 但 mPrec.、mIoU 和 mDice 均进一步提升. 这表明L_{str}主要增强组织区域的覆盖能力, 而 SSD 进一步利用中间层结构相似关系优化激活区域分布, 减少部分边界附近的误激活, 使模型在区域完整性与预测准确性之间取得了更好的平衡.

3.4.2　超参数分析

图 5 分析了λ、T、τ和N对模型性能的影响. 其中, λ用于平衡分类损失与结构一致性约束损失, T表示结构相似性扩散的迭代次数, τ用于控制最终前景区域的判定阈值, N表示结构一致性建模中的局部邻域大小. 结果表明: 当$\lambda=0.3$时性能最佳, 过小难以发挥结构约束作用, 过大则会抑制类别判别; 扩散次数T在 14 附近达到最佳; 阈值$\tau=0.5$时前景保留与背景抑制较为平衡; 局部邻域大小$N=5$时最有利于建模微观组织相似性. 基于此, 本文最终采用$\lambda=0.3$、$T=14$、$\tau=0.5$和$N=5$. 综合来看, 本文方法对超参数变化总体较为稳定. 除λ过大时会因结构约束过强而导致性能下降外, T、τ和N在最优值附近均表现出较小波动, 说明所提出方法不存在明显的超参数敏感问题, 具有较好的鲁棒性与可调性.

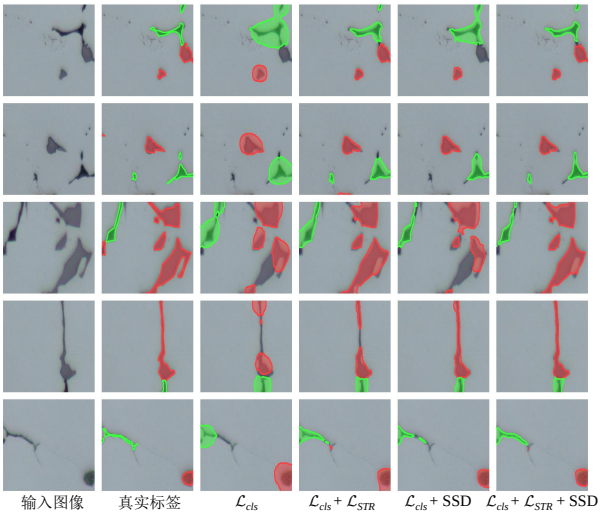

图 4　不同模块组合的定性消融结果

3.5　不同组织类型的分割性能分析

为进一步分析模型在不同类型微观组织上的分割表现, 本文分别统计了 Mg_2Si 相和 $FeMnAl_6$相的分割结果, 见表 3 和图 6. Mg_2Si 相通常呈细长骨架状, 结构复杂, 分割难度相对更高; $FeMnAl_6$相则多

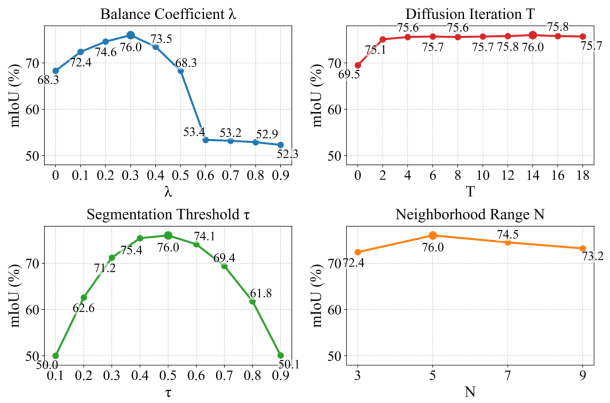

图 5 超参数分析结果

呈块状分布, 整体边界相对清晰. 通过分类别统计, 可以进一步考察本文方法在不同形态组织上的适应能力.

表 3 不同方法在 $FeMnAl_6$ 相和 Mg_2Si 相上的分割性能比较

方法	$FeMnAl_6$					Mg_2Si				
	Dice	IoU	Recall	Spec.	Prec.	Dice	IoU	Recall	Spec.	Prec.
DeepLabV2	81.9	69.3	74.7	99.8	90.5	76.5	62.0	69.5	99.9	85.1
CAM[14]	50.5	33.8	76.8	96.2	37.6	47.0	30.7	84.2	98.0	32.6
IRN[17]	60.8	43.7	75.5	97.8	50.9	55.2	38.1	82.0	98.7	41.6
SEAM[18]	53.1	36.1	76.9	96.6	40.5	46.3	30.1	84.5	97.9	31.8
WeCLIP[19]	53.9	36.9	40.0	**99.7**	82.3	15.9	8.6	92.5	88.6	8.7
MoRe[20]	73.9	58.6	**90.0**	98.4	62.7	65.4	48.6	**93.3**	98.9	50.4
Ours	**82.2**	**69.8**	81.6	99.5	**82.8**	**74.9**	**59.8**	74.3	**99.7**	**75.5**

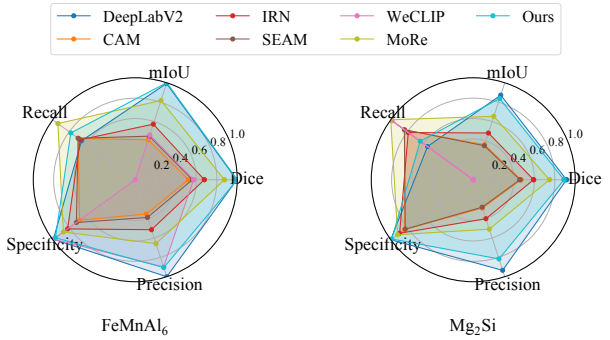

图 6 不同方法在 Mg_2Si 相和 $FeMnAl_6$ 相上的分割性能雷达图

从表 3 可以看出, 本文方法在 $FeMnAl_6$ 相上的 IoU 达到 69.8%, 在 Mg_2Si 相上的 IoU 达到 59.8%, 均优于现有弱监督对比方法. 相较于弱监督方法中表现较好的 MoRe, 本文方法在 $FeMnAl_6$ 相和 Mg_2Si 相上的 IoU 均提升了 11.2%, 说明整体性能提升并非主要由某一类组织单独贡献, 而是来源于两类前景组织的共同改善. 进一步结合 Dice 指标来看, 本文方法在 $FeMnAl_6$ 相和 Mg_2Si 相上的 Dice 分别达到 82.2% 和 74.9%, 同样高于其他弱监督方法, 表明所提方法既能够提升块状 $FeMnAl_6$ 相的分割准确性, 也能够改善细长骨架状 Mg_2Si 相的区域恢复能力. 图 6 中的雷达图进一步表明, 本文方法在两类组织上均具有较优的综合性能, 尤其在 IoU、Spec.和 Prec.方面优势较为明显. 上述结果说明, 结构一致性约束和结构相似性扩散不仅能够改善局部区域定位, 还能增强模型对不同形态组织的适应能力, 从而提升铝合金微观组织弱监督分割的整体性能.

3.6 模型效率分析

表 4 比较了不同方法的参数规模、推理速度和分割性能. 其中, 推理速度测试在 NVIDIA RTX 3090 GPU 上完成, FPS 统计采用单卡、单张图像输入的测试方式. 从参数规模来看, 本文完整模型的参数量为 23.5M, 而全监督 DeepLabV2 的参数量为 61.6M, 本文方法的参数量仅为其 38.1%, 降低约 60%. 与其他弱监督方法 WeCLIP 和 MoRe 相比, 本文方法也具有更轻量的模型规模. 完整模型的推理速度为 8.93 fps, 虽略低于不含 SSD 时的 9.33 fps, 但分割性能提升更为显著; 引入 SSD 后, mIoU 由 69.5% 提升至 76.0%, 说明该模块仅以较小的推理代价带来了明显收益. 总体来看, 本文方法在较低参数量和计算开销下取得了接近全监督方法的性能, 兼顾了精度与效率, 更适用于资源受限条件下的工业材料图像分析任务.

表 4 不同方法的模型效率分析

方法	主要组件	#Para(M)	Infer(FPS)	mIoU(%)
DeepLabV2	L_{ce}	61.6	8.11	76.6
WeCLIP	L_{cls}	155.6	1.98	45.1
MoRe	L_{cls}	94.7	16.06	68.2
Ours	L_{cls}	23.5	9.33	52.7
	$L_{cls}+L_{str}$	23.5	9.33	69.5
	$L_{cls}+SSD$	23.5	8.93	68.3
	$L_{cls}+L_{str}+SSD$	23.5	8.93	76.0

4 结 论

针对铝合金微观组织分割中像素级标注成本高、弱监督结果区域不完整和边界不准确的问题, 本文提出了一种弱监督结构约束分割方法. 该方法通过构建双分支结构约束分割网络, 在图像级监督下联合学习类别语义信息与局部结构信息, 并在推断阶段引入结构相似性扩散算法, 对初始激活图进行优化, 从而提升组织区域定位的完整性与边界质量. 实验结果表明, 本文方法在铝合金微观组织数据集上取得了优于现有弱监督方法的分割效果, 并接近全监督方法的性能. 消融实验和不同组织类型的分割性能分析进一步验证了结构约束学习与结构相似

性扩散的有效性及互补性. 总体而言, 本文工作为低标注成本条件下的铝合金微观组织自动分割提供了一种有效思路, 对弱监督分割方法在工业材料图像分析中的应用具有一定参考价值.

需要指出的是, 本文方法主要在当前铝合金微观组织数据集上进行了验证, 对于更复杂组织类型、多类别微观结构以及跨设备、跨批次采集数据下的泛化能力, 仍需进一步研究. 后续将结合更多来源的工业材料图像, 对所提方法的鲁棒性和推广能力进行进一步验证.

参考文献 (References)

[1] Luengo J, Moreno R, Sevillano I, et al. A tutorial on the segmentation of metallographic images: Taxonomy, new MetalDAM dataset, deep learning-based ensemble model, experimental analysis and challenges[J]. Information Fusion, 2022, 78: 232-253.

[2] Dutta T, Das D, Banerjee S, et al. An automated morphological classification of ferrite-martensite dual-phase microstructures[J]. Measurement, 2019, 137: 595-603.

[3] Zhang K G, Chen W, Liu W K, et al. A framework for supervised and unsupervised segmentation and classification of materials microstructure images[J]. Acta Materialia, 2025, 301: 121588.

[4] Zhang X C, Zhang B, Guo G H, et al. In situ 3-D surface reconstruction for metal deformation evolution via SEM[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 5043212.

[5] 王鹏, 赵春晖, 周华良, 等. 基于多时相巡检图像的变电设备抗干扰缺陷检测[J]. 控制与决策, 2024, 39(3): 885-892.
(Wang P, Zhao C H, Zhou H L, et al. Anti-interference defect detection of substation equipment based on multi-temporal inspection images[J]. Control and Decision, 2024, 39(3): 885-892.)

[6] 李玉, 崔书琳, 赵泉华. 基于优化 RDD 分区的 Spark 并行 *K*-means 大尺度遥感图像分割[J]. 控制与决策, 2024, 39(5): 1612-1619.
(Li Y, Cui S L, Zhao Q H. Spark parallel *K*-means large scale remote sensing image segmentation based on optimized RDD partition[J]. Control and Decision, 2024, 39(5): 1612-1619.)

[7] 王维, 王显鹏, 宋相满. 基于自适应多目标进化 CNN 的图像分割方法[J]. 控制与决策, 2024, 39(4): 1185-1193.
(Wang W, Wang X P, Song X M. An image segmentation method based on adaptive multi-objective evolutionary CNN[J]. Control and Decision, 2024, 39(4): 1185-1193.)

[8] 张泽辉, 吴富龙, 李帆雅, 等. 面向复杂工业场景的人体姿态快速估计方法[J]. 控制与决策, 2026, 41(9): 2631-2640.
(Zhang Z H, Wu F L, Li F Y, et al. Fast human pose estimation method for complex industrial environments[J]. Control and Decision, 2026, 41(9): 2631-2640.)

[9] 韩佳伶, 陈芋渝. 协同自进化的粒子群优化算法及其在图像分割的应用[J]. 控制与决策, 2026, 41(6): 1743-1752.
(Han J L, Chen Y Y. Collaborative self-evolutionary particle swarm optimization algorithm and its application in image segmentation[J]. Control and Decision, 2026, 41(6): 1743-1752.)

[10] 褚菲, 王建文, 马小平. 复杂工业过程安全运行控制方法研究综述: 现状、挑战与展望[J]. 控制与决策, 2026, 41(6): 1489-1508.
(Chu F, Wang J W, Ma X P. A review of safety operation control methods for complex industrial processes: Current status, challenges and future prospects[J]. Control and Decision, 2026, 41(6): 1489-1508.)

[11] Yeom J, Stan T, Hong S, et al. Segmentation of experimental datasets via convolutional neural networks trained on phase field simulations[J]. Acta Materialia, 2021, 214: 116990.

[12] Jiang P T, Han L H, Hou Q B, et al. Online attention accumulation for weakly supervised semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(10): 7062-7077.

[13] Na J, Kim S J, Kim H, et al. A unified microstructure segmentation approach *via* human-in-the-loop machine learning[J]. Acta Materialia, 2023, 255: 119086.

[14] Zhou B L, Khosla A, Lapedriza A, et al. Learning deep features for discriminative localization[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016: 2921-2929.

[15] Vezhnevets A, Buhmann J M. Towards weakly supervised semantic segmentation by means of multiple instance and multitask learning[C]. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, 2010: 3249-3256.

[16] Wei Y C, Liang X D, Chen Y P, et al. STC: A simple to complex framework for weakly-supervised semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(11): 2314-2320.

[17] Ahn J, Cho S, Kwak S. Weakly supervised learning of instance segmentation with inter-pixel relations[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, 2020: 2204-2213.

[18] Wang Y D, Zhang J, Kan M N, et al. Self-supervised equivariant attention mechanism for weakly supervised semantic segmentation[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, 2020: 12272-12281.

[19] Zhang B F, Yu S Y, Wei Y C, et al. Frozen CLIP: A strong backbone for weakly supervised semantic segmentation[C]. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, 2024:

3796-3806.

[20] Yang Z W, Meng Y C, Fu K X, et al. MoRe: Class patch attention needs regularization for weakly supervised semantic segmentation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2025, 39(9): 9400-9408.

[21] Li M C, Chen D L, Liu S X, et al. Grain boundary detection and second phase segmentation based on multi-task learning and generative adversarial network[J]. Measurement, 2020, 162: 107857.

[22] Sun Y L, Huang X Y. CMAA: Channel-wise multi-scale adaptive attention network for metallographic image semantic segmentation[J]. Expert Systems with Applications, 2025, 276: 126925.

[23] Li M C, Liu Y, Chen D L, et al. Uncertainty estimation method based on Dirichlet distribution for microstructure segmentation and measurement in aluminum alloy metallographic images[J]. Measurement, 2025, 246: 116744.

[24] Han Y X, Li R Q, Wang B, et al. A pseudo-labeling based weakly supervised segmentation method for few-shot texture images[J]. Expert Systems with Applications, 2024, 238: 122110.

作者简介

薛正 (1998−), 男, 博士生, 主要研究方向为图像语义分割、弱监督学习、工业人工智能等, E-mail: xuez@mails.neu.edu.cn;

陈大力 (1981−), 男, 教授, 通信作者, 主要研究方向为机器学习、图像处理与分析、工业人工智能等, E-mail: chendali@ise.neu.edu.cn;

李鸣春 (1993−), 男, 博士后, 主要研究方向为图像处理与分析、工业人工智能、医学图像处理等, E-mail: limingchun@syu.edu.cn;

刘士新 (1968−), 男, 教授, 主要研究方向为智能优化决策、人工智能、机器学习、生产计划与调度等, E-mail: sxliu@mail.neu.edu.cn.