doi:10.12068/j.issn.1005－3026.2026.20250119

基于MSCA-Net的钢铁显微组织识别与定量表征

刘建军，吴思炜，曹光明，刘振宇
（东北大学 数字钢铁全国重点实验室，辽宁 沈阳 110819）

摘 要：传统钢铁材料金相分析高度依赖人工经验，难以对晶界及细小第二相等实现精准量化．因此，提出一种基于多尺度卷积注意力网络（multi-scale convolutional attention network，MSCA-Net）的显微组织智能识别与定量表征方法．该方法融合多尺度卷积与注意力机制，显著提升了显微组织识别精度．相较于全卷积网络（fully convolutional network，FCN），MSCA-Net在测试集上的像素准确率（pixel accuracy，PA）和频权交并比（frequency weighted intersection over union，FWIoU）分别提升2.71%和5.07%．此外，应用OpenCV对显微组织特征进行定量分析．结果表明，MSCA-Net显著降低了测量误差，实现了显微组织特征的高精度自动化表征，为材料显微组织结构分析提供了智能化新方法．

关 键 词：钢铁材料；显微组织；多尺度卷积注意力；OpenCV；定量表征
中图分类号：TP 18；TG 142.1 **文献标志码**：A **文章编号**：1005-3026(2026)06-0119-11

Recognition and Quantitative Characterization of Steel Microstructure Based on MSCA-Net

LIU Jian-jun, *WU Si-wei*, *CAO Guang-ming*, *LIU Zhen-yu*
(State Key Laboratory of Digital Steel, Northeastern University, Shenyang 110819, China. Corresponding author: WU Si-wei, E-mail: wusiwei@ral.neu.edu.cn)

Abstract: Traditional metallographic analysis of steel materials is highly dependent on manual experience, and it is difficult to accurately quantify grain boundaries and fine secondary phases. Therefore, a method for intelligent recognition and quantitative characterization of microstructure based on a multi-scale convolutional attention network (MSCA-Net) was proposed. This method integrated multi-scale convolution and attention mechanisms, which significantly improved the accuracy of microstructure recognition. Compared with those of the fully convolutional network (FCN), the pixel accuracy (PA) and frequency weighted intersection over union (FWIoU) of MSCA-Net on the test set were improved by 2.71% and 5.07%, respectively. In addition, OpenCV was applied to quantitatively analyze microstructure features. The results demonstrate that MSCA-Net significantly reduces measurement errors, achieves high-precision automated characterization of microstructure features, and provides a new intelligent method for the analysis of material microstructure.

Key words: steel material; microstructure; multi-scale convolutional attention; OpenCV; quantitative characterization

钢铁材料的力学性能依赖于其显微组织的形态与空间分布，因此，实现显微组织的精准识别与定量表征，对于调控材料组织演化与指导高质量钢铁产品的可控化生产具有重要意义[1-3]．目前，钢铁材料的显微组织表征多基于Image J，Image-Pro Plus 6.0等图像处理软件，该过程需要具备专业知识的研究者手动标定晶粒并计算特征参数[4-6]．这种依赖人工统计的方法不仅耗时耗力，且测量结果受研究者经验的影响，导致数据可重复性差．

收稿日期：2025-10-20
基金项目：智能制造系统和机器人国家科技重大专项（2025ZD1602200）．
作者简介：刘建军（1991—），男，河北唐山人，东北大学博士研究生．
通信作者：吴思炜，E-mail：wusiwei@ral.neu.edu.cn.

为解决上述问题,首先要实现对显微组织结构的精细化识别.然而,显微组织的识别是整个分析过程中最困难且最耗时的关键步骤[7-9].近年来,深度学习算法的迅速发展引起了研究者的广泛关注,其优势在于能够直接从原始数据中自动学习深层特征,从而有效避免了传统方法中烦琐的人工特征提取过程[10-12].目前,显微组织的自动识别已成为材料科学领域的研究热点.Jang等[13]提出了一种基于全卷积网络的语义分割模型,采用ResNet作为编码器来识别显微组织中的针状铁素体,该模型在测试集上的平均交并比超过85%.Banerjee等[14]提出了一种用于自动分割双相钢扫描电子显微镜图像中相的方法,有效克服了传统基于简单阈值分割方法在处理复杂组织时的局限性,显著提升了模型的准确性与鲁棒性.Azimi等[15]基于一种全卷积神经网络模型,实现了低碳钢显微组织的像素级语义分割,并引入最大投票机制以提升分割结果的准确性与一致性.虽然当前方法在显微组织的识别任务中已取得较高的识别精度,但在识别晶界和占比较少的第二相等关键显微组织结构时仍存在不足,其通常被简单归类为基体相或第二相,没有将其单独归为一类.此外,由于第二相占比较少,模型难以充分学习其特征,使识别精度偏低以及显微组织特征测量结果的准确性下降.

此外,显微组织结构的精确表征是研究材料性能的基础.目前,针对晶粒尺寸的测量主要采用截距法[16]或等效圆直径法[17]等几何测量方法进行量化分析.Gajalakshmi等[18]通过边缘检测结合截距法对晶粒数量和尺寸进行测量,并将所得结果作为支持向量回归模型的训练目标值.该模型的预测结果与人工晶粒计数相比,均方根误差最高约为5%.Banerjee等[19]通过应用边缘检测算法及形态学等图像处理技术得到封闭的晶粒轮廓,并在显微组织图像上绘制一组水平和垂直线以获取晶粒度及面积等特征.Flipon等[20]使用计数、截距和相交的方法来评估晶粒尺寸分布.在计算平均晶粒尺寸方面,与基于整体计数的方法相比,该方法具有良好的统计代表性.然而,传统的显微组织表征方法仅能对平均晶粒尺寸和相占比进行粗略估算,难以深入揭示显微组织特征对材料性能的影响规律.此外,对于单位面积晶界长度及第二相单位面积相界长度等精细结构特征,目前仍缺乏有效的数字化定量表征手段,从而限制了显微组织与性能之间关系的精确解析.

针对上述问题,本文以具有典型"铁素体-珠光体"组织的C-Mn钢为研究对象,提出一种基于多尺度卷积注意力网络的显微组织识别与定量表征方法.该方法实现了对各相的精细化识别与定量表征.同时,该方法具备良好的可扩展性,可推广应用于其他钢铁材料,如"铁素体-马氏体"双相钢和复相高强钢等.

1 实验方法

1.1 数据集

本文采用奥林巴斯(Olympus)光学金相显微镜对经4%硝酸酒精溶液腐蚀的C-Mn钢试样进行显微组织观察,在1 000倍放大下共获取21幅具有代表性的金相显微图像,其中每幅图像的分辨率为5 472像素×3 648像素.由于本文采用的特征提取网络的标准输入图像尺寸为1 024像素×1 024像素,因此需对原始图像进行裁剪,裁剪过程如图1所示.经过裁剪后,共获得315幅图像,随后,从中随机选择60幅作为测试数据,其余作为训练数据.为提升模型的识别性能,对训练集进行了数据增强处理.数据增强方法包括水平翻转和垂直翻转,以及分别旋转90°,180°和270°,最终使显微组织图像的数据量扩充至原始训练集规模的5倍.此外,在模型训练过程中从训练集中随机选择20%的数据作为验证集,用于监控训练过程,防止模型出现过拟合或欠拟合现象.

1.2 多尺度卷积注意力网络模型

1.2.1 多尺度卷积注意力模块

多尺度卷积注意力(multi-scale convolutional attention, MSCA)模块是一种专为语义分割任务设计的轻量级和高效率的注意力机制,其目的在于提升模型对复杂场景中多尺度上下文信息的感知与建模能力[21].与传统的自注意力机制相比,MSCA避免了计算复杂度随特征图尺寸平方增长的问题,更适合部署于资源受限或对实时性要求较高的工业应用场景.如图2所示,MSCA模块的核心结构采用了一种非对称的多尺度深度可分离卷积[22]结构,通过应用不同大小的卷积核,从而建立多个具有差异化感受野的分支.因此,整个模块无需引入复杂的矩阵计算或额外的参数学习机制,因而具有参数量少及计算资源消耗低等优势.尤其重要的是,MSCA模块具有良好的通用性和扩展性,可以无缝嵌入于编码器-

解码器结构中.在本文中,MSCA 被应用于编码器的多个层级,从浅层到深层逐步增强网络的特征提取能力,显著提升了模型对小目标、边缘细节以及复杂背景干扰下的目标识别性能.

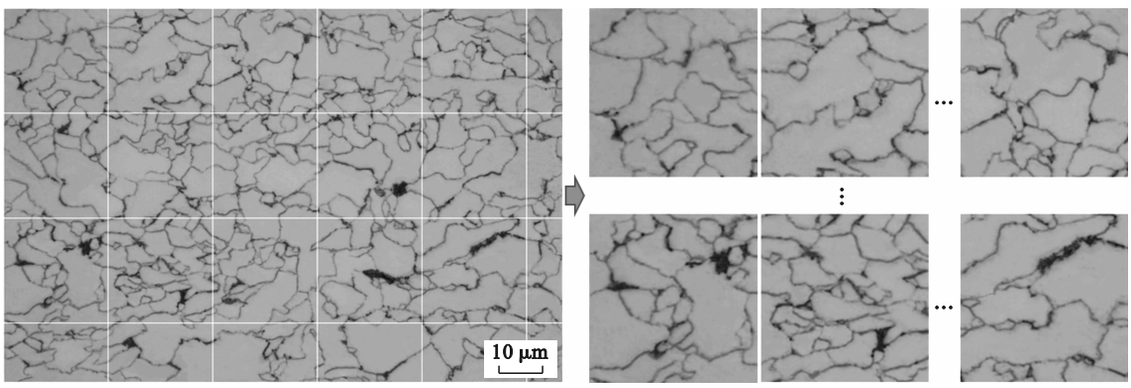
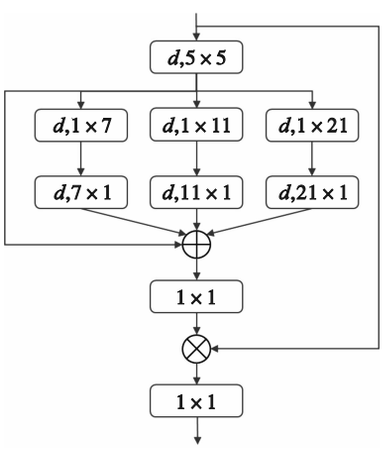


图1　显微组织图像裁剪过程

Fig. 1　Microstructure image cropping process



注:d表示深度卷积;$k_1 \times k_2$表示核尺寸;⊕表示加法;⊗表示乘法.

图2　MSCA 模块结构

Fig. 2　Structure of MSCA module

1.2.2　MSCA-Net模型结构

本文以多尺度卷积注意力模块为核心,构建了一种具有层次化特征提取能力的显微组织结构识别网络模型,从而实现钢铁显微组织图像的精准识别.图3展示了MSCA-Net模型整体结构.与 FCN 模型相比,MSCA-Net 采用典型的编码器-解码器结构,其中编码器的每个特征提取层均使用MSCA模块,通过多尺度深度可分离卷积共同获取局部细节与大范围上下文信息,从而增强了特征的语义表达能力.此外,MSCA-Net模型在编码器与解码器之间引入跳跃连接,将浅层特征与深层语义特征进行融合,以保证显微组织特征解码过程中的精确定位能力,进而提升对复杂显微组织结构的识别精度.

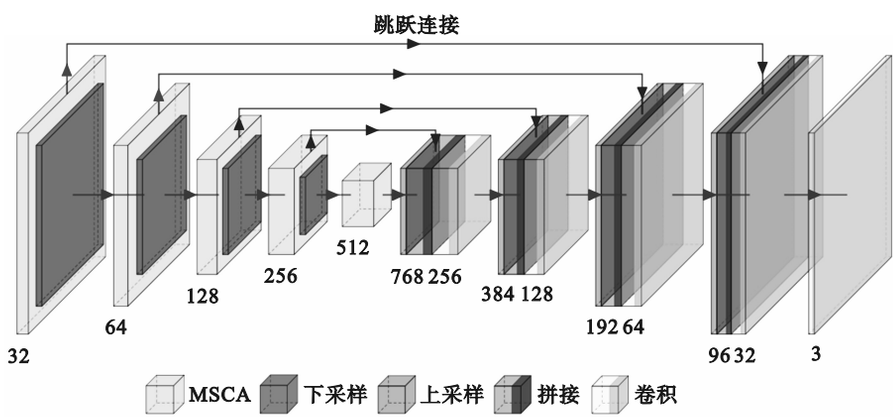


图3　MSCA-Net模型整体结构

Fig. 3　Overall structure of MSCA-Net model

此外,表1列出了MSCA-Net模型结构细节.编码器部分包含5个MSCA模块,其输出的特征图尺寸为64像素×64像素,通道数为512.随后,该特征图依次通过4个上采样模块,每个模块由1个上采样层和2个DSC组成.最终,通过一个卷积核尺寸为1×1的卷积层调整输出通道数,输出

与输入图像尺寸一致的语义分割结果.

表 1 MSCA-Net模型结构细节

Table 1 Details of MSCA-Net model structure

层类型	核尺寸	输出($C\times W\times H$)
MSCA_1	—	32×1 024×1 024
MaxPool2d_1	3×3	32×512×512
MSCA_2	—	64×512×512
MaxPool2d_2	3×3	64×256×256
MSCA_3	—	128×256×256
MaxPool2d_3	3×3	128×128×128
MSCA_4	—	256×128×128
MaxPool2d_4	3×3	256×64×64
MSCA_5	—	512×64×64
UpSample2d_1	—	512×128×128
DSC×2_1	3×3	256×128×128
UpSample2d_2	—	256×256×256
DSC×2_2	3×3	128×256×256
UpSample2d_3	—	128×512×512
DSC×2_3	3×3	64×512×512
UpSample2d_4	—	64×1 024×1 024
DSC×2_4	3×3	32×1 024×1 024
Conv2d	1×1	3×1 024×1 024

注:C为通道数;W为图像宽度(像素);H为图像高度(像素).

1.2.3 提取标尺信息

为实现将显微组织特征从像素级单位转换为具有实际物理意义的数据,首要步骤是检测标尺长度和识别数字与单位,这不仅避免了传统人工测量中因主观因素带来的误差,还提升了组织特征测量结果的客观性和可重复性.图4显示了标尺长度检测与数字及单位识别的主要过程.通过算法自动定位图像中的标尺区域,并进行精准的像素长度测量,同时识别出与之相关的数值和单位.从图4中可以看出,识别出的数值为10 μm,对应于420像素的长度.基于此,可以计算出每个像素所代表的实际长度,即约等于0.023 8 μm/像素.从而建立了像素尺度与真实物理尺度之间的线性映射关系,确保了显微组织特征测量值具有实际物理意义.

2 实验配置

2.1 模型超参数

本研究采用开源、可公开获取的Python[23]编程语言及主流深度学习与计算机视觉工具库建立模型,包括PyTorch[24]深度学习框架、Scikit-learn[25]机器学习库和OpenCV[26]图像处理库.所有模型的训练均在配备24 GB显存的NVIDIA GeForce RTX 4090 GPU上完成.训练参数的批大小(batch size)设置为2,学习率为1.2×10^{-4},优化函数为RMSprop,激活函数为ReLU.此外,在网络结构中引入批归一化(batch normalization)技术,有效缓解了梯度消失问题,进一步提高了模型的泛化性能与训练效率.

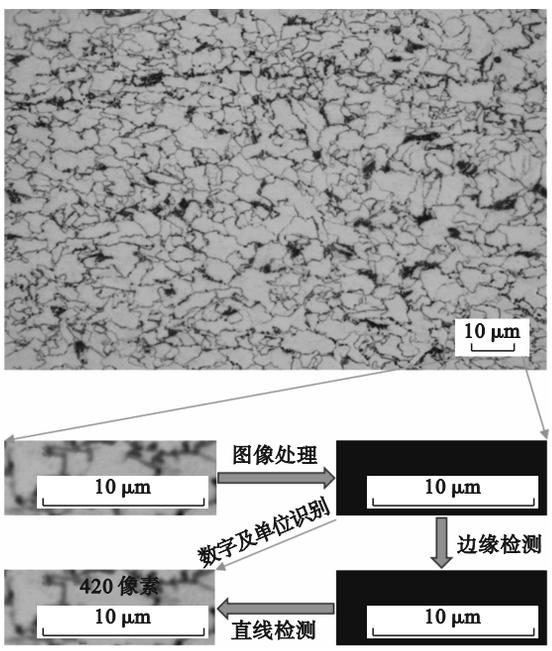


图4 标尺检测与数字及单位识别

Fig. 4 Ruler detection and recognition of numbers and units

2.2 评价指标

本文采用像素准确率和频权交并比作为模型性能的评估指标.像素准确率(PA)通过计算正确预测的像素数占总像素数的比例,直观反映模型对图像整体分类的准确性,其计算式为

$$\mathrm{PA}=\frac{\sum_{i=0}^{k}p_{ii}}{\sum_{i=0}^{k}\sum_{j=0}^{k}p_{ij}}. \tag{1}$$

其中:p_{ii}表示第i类像素被正确识别的概率;p_{ij}表示第i类像素被识别为第j类的概率;k为类别总数.

然而,在类别像素分布不均衡的情况下,像素准确率容易受占比较高的类别影响.因此,本文引入频权交并比作为补充评价指标.该指标根据各类别在数据集中出现的频率对交并比进行加权平均,使结果更贴近实际数据分布,有效降低类别不平衡带来的偏差,其计算式为

$$\mathrm{FWIoU}=\frac{1}{\sum_{i=1}^{k}\sum_{j=1}^{k}p_{ij}}\sum_{i=1}^{k}\left(\frac{\sum_{j=1}^{k}p_{ij}p_{ii}}{\sum_{j=1}^{k}p_{ij}+\sum_{j=1}^{k}p_{ji}-p_{ii}}\right). \tag{2}$$

其中,p_{ji}表示真实类别为j,但被模型预测为类别i的像素数量.

因此,本文通过结合上述两项指标,从全局分类准确性和类别平衡敏感性2个维度综合评估模型的性能.

3　结果与讨论

3.1　模型精度对比分析

为验证MSCA-Net模型在钢铁显微组织识别任务中的优越性,本文将其与FCN[27],DANet[28]和DeepLabV3+[29]模型在训练集、验证集和测试集上的像素准确率和频权交并比进行对比分析.图5显示了MSCA-Net模型与FCN,DANet和DeepLabV3+模型在训练集、验证集和测试集上的像素准确率和频权交并比结果比较.从图5中可以看出,MSCA-Net在各项指标上均优于其余3种模型,展现出较高的识别精度.具体而言,MSCA-Net模型在测试集上的像素准确率达到0.963 2,频权交并比为0.930 6,相比经典的FCN模型分别提升了2.71%和5.07%.这一提升主要得益于MSCA模块对多尺度上下文信息的有效建模以及对关键显微组织结构的注意力增强,使模型能够更准确地区分形态相似但类别不同的组织区域.此外,MSCA模块不仅在训练集和验证集上具有较高的识别精度,更在测试集上展现出优异的识别性能,这表明模型具备良好的泛化能力,能够有效适应未见过的显微组织图像样本.上述分析表明,MSCA-Net的识别性能更加优异,且稳定性强,为材料显微组织结构的高精度自动化识别提供了更加可靠的解决方案.

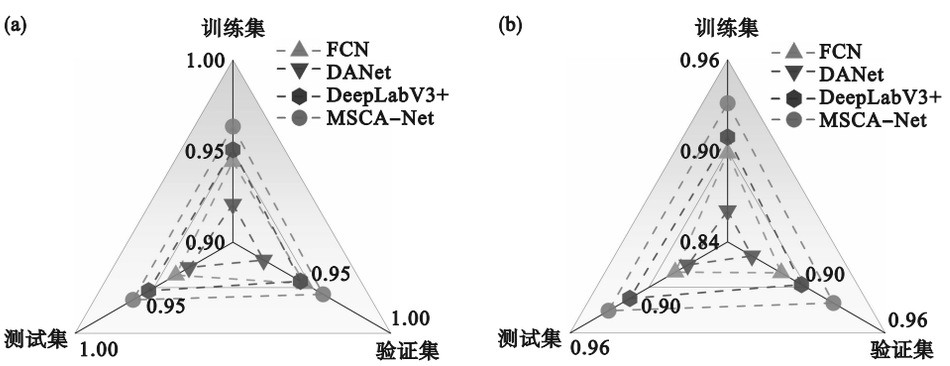


图5　FCN,DANet,DeepLabV3+和MSCA-Net模型在数据集上的结果比较

Fig. 5　Comparison of results for FCN, DANet, DeepLabV3+, and MSCA-Net models on datasets

(a)—模型像素准确率比较;(b)—模型频权交并比比较.

3.2　模型识别效果分析

图6显示了训练完成的MSCA-Net模型在显微组织图像上的识别结果.为进一步评估模型的识别准确性,图7和图8比较了FCN,DANet,DeepLabV3+与MSCA-Net模型的识别结果与真实标注(ground truth,GT)图像之间的差异,更清晰地揭示了各模型在晶界识别结果上的差异.结果表明,这几种模型对铁素体区域均具备一定的识别能力.然而,在实际金相制样过程中,由于抛光不均或腐蚀程度控制不当,部分图像中晶界呈现模糊、断裂甚至局部缺失,导致显微组织边界不完整.在这些图像质量较差的复杂区域,MSCA-Net展现出显著优于其他模型的识别准确性和稳定性.从图中可以看出,MSCA-Net模型的识别结果与GT图像高度一致,晶界识别结果连续且清晰,未出现明显断裂或错位现象.相反,其他3种模型在晶界识别中易受局部噪声和纹理干扰,尤其在相邻晶粒交汇处出现严重的识别错误,具体表现为晶界断裂,从而导致晶界未能被清晰地识别,出现晶粒合并或边界断裂等过分割或欠分割情况,严重影响了显微组织结构的几何完整性.因此,MSCA-Net模型不仅能够更精准地识别显微组织结构边界,有效抑制无关噪声的干扰,实现更精准和连续的晶界识别,还为后续显微组织的定量分析(如晶粒尺寸、单位面积晶界长度等特征量化)提供了高精度的识别结果.

图9显示了MSCA-Net模型对占比较少的第二相珠光体识别结果.为进一步分析模型在第二相细节上的识别差异,图10和图11对FCN,DANet,DeepLabV3+与MSCA-Net模型的识别结果与GT图像中的珠光体区域进行了局部放大比较.从图中可以看出,MSCA-Net模型的识别结果在组织形态完整性和边界清晰度方面均明

显优于其他模型,更加贴近实际显微组织的真实分布.在珠光体与其他相交界的复杂边界区域,FCN,DANet与DeepLabV3+模型表现出明显的识别偏差,容易出现误分割或欠分割现象.尤其在珠光体呈细小弥散分布或与铁素体边界交错的区域,其识别能力明显下降.相反,MSCA-Net模型由于其多尺度卷积注意力机制对上下文特征的有效提取,从而能够更精确地识别相界边缘,显著提升了珠光体区域的识别准确性.值得注意的是,在处理含量较低的第二相时,MSCA-Net模型展现出更强的鲁棒性,其能够有效抑制因制样过程中腐蚀不充分、图像对比度低或噪声干扰等因素带来的影响.识别结果表明,该模型在复杂显微组织结构的精细识别方面具有明显优势,为后续晶粒统计和相占比计算等定量分析提供了高精度的第二相识别结果.

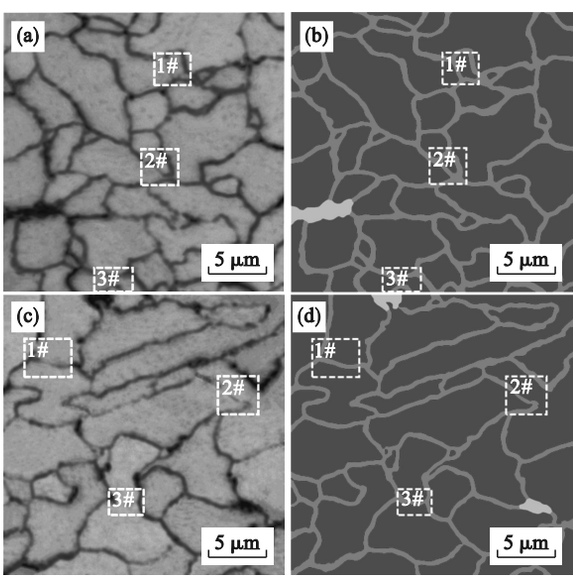


图6 MSCA-Net模型对显微组织图像数据1和2的晶界识别结果

Fig. 6 Grain boundary recognition results of microstructure image data 1 and 2 using MSCA-Net model

(a)—显微组织图像数据1;(b)—MSCA-Net模型对显微组织图像数据1的识别结果;(c)—显微组织图像数据2;(d)—MSCA-Net模型对显微组织图像数据2的识别结果.

3.3 噪声影响下的模型鲁棒性分析

为系统地评估模型对噪声的识别与抗干扰能力,本研究在测试集的显微组织图像中人为添加了噪声,并采用控制变量法固定噪声数量,以客观、定量地评价模型的抗噪性能.图12显示了对测试集中的显微组织图像添加随机斑点噪声后的示例.所添加的随机斑点噪声在显微组织图

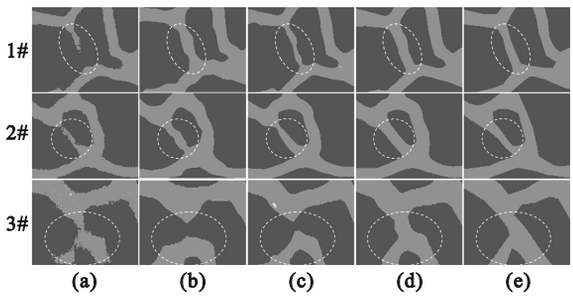


图7 图6a的晶界识别结果局部对比

Fig. 7 Local comparison of grain boundary recognition results in Fig. 6a

(a)—FCN模型;(b)—DANet模型;(c)—DeepLabV3+模型;(d)—MSCA-Net模型;(e)—GT图像.

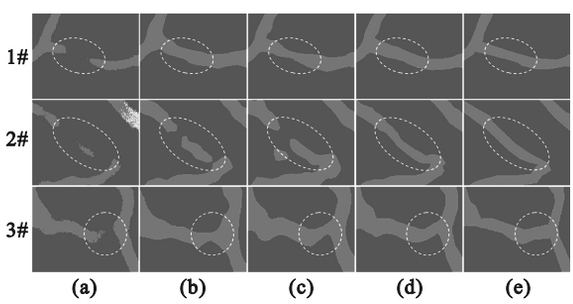


图8 图6c的晶界识别结果局部对比

Fig. 8 Local comparison of grain boundary recognition results in Fig. 6c

(a)—FCN模型;(b)—DANet模型;(c)—DeepLabV3+模型;(d)—MSCA-Net模型;(e)—GT图像.

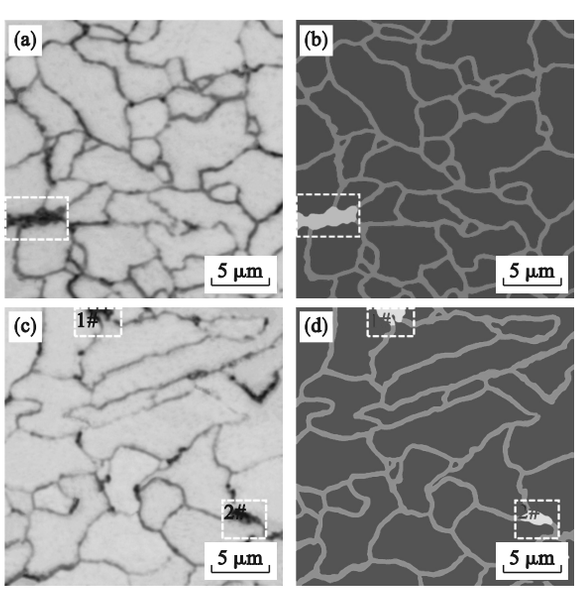


图9 MSCA-Net模型对显微组织图像数据1和2的珠光体识别结果

Fig. 9 Pearlite recognition results of microstructure image data 1 and 2 using MSCA-Net model

(a)—显微组织图像数据1;(b)—MSCA-Net模型对显微组织图像数据1的识别结果;(c)—显微组织图像数据2;(d)—MSCA-Net模型对显微组织图像数据2的识别结果.

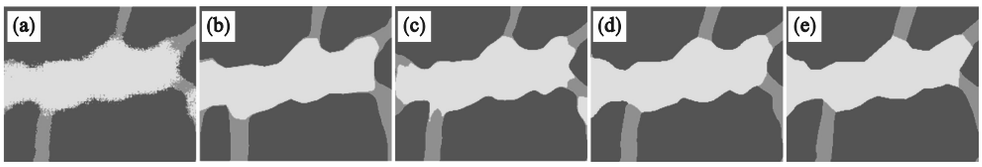


图10　图9a的珠光体识别结果局部对比

Fig. 10　Local comparison of pearlite recognition results in Fig. 9a

(a)—FCN模型;(b)—DANet模型;(c)—DeepLabV3+模型;(d)—MSCA-Net模型;(e)—GT图像.

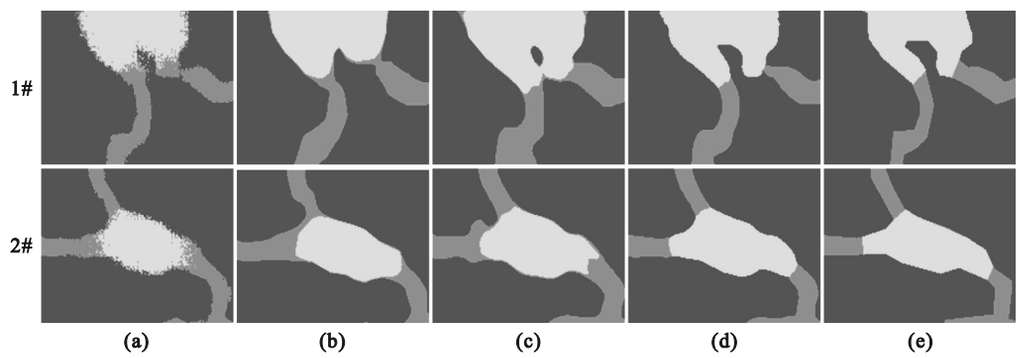


图11　图9c的珠光体识别结果局部对比

Fig. 11　Local comparison of pearlite recognition results in Fig. 9c

(a)—FCN模型;(b)—DANet模型;(c)—DeepLabV3+模型;(d)—MSCA-Net模型;(e)—GT图像.

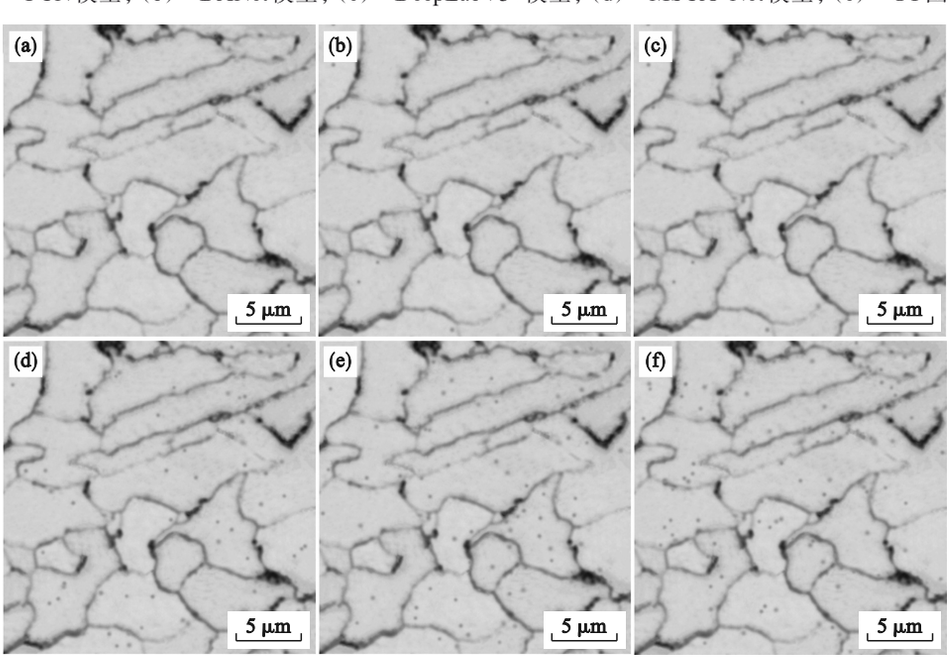


图12　添加随机噪声的显微组织图像示例

Fig. 12　Example of microstructure image with added random noise

(a)—未添加随机噪声;(b)—添加20个随机噪声;(c)—添加40个随机噪声;(d)—添加60个随机噪声;(e)—添加80个随机噪声;(f)—添加100个随机噪声.

像上呈无规律分布,其中,每个噪声包含的像素被控制在100~200像素之间,从而模拟了实际制样过程中因腐蚀不均或试剂残留等因素引起的局部干扰,该操作保证了添加噪声的可重复性.基于上述添加噪声后的测试集,本文对训练完成的FCN,DANet,DeepLabV3+与MSCA-Net模型进行了鲁棒性测试.图13显示了4种模型在不同数量噪声下的像素准确率和频权交并比变化趋势.从图中可以看出,随着噪声数量的增加,模型的性能均有所下降,说明噪声对模型的识别精度具有一定的负面影响.然而,MSCA-Net模型的像素准确率和频权交并比仍保持在0.95和0.91以上.具体而言,在噪声包含的像素为100像素时,MSCA-Net模型的像素准确率和频权交并比相较于无噪声情况分别下降了0.87%和1.48%.上述结果表明,MSCA-Net模型能够有效抑制局

部随机干扰,在复杂成像条件下仍保持较高的识别精度.此外,MSCA-Net模型在含噪声测试集上的良好识别精度主要归功于其合理的网络结构设计,该模型采用编码器-解码器结构,并且结合多尺度卷积注意力机制,通过不同尺度的深度可分离卷积获取多层级的空间上下文信息,并生成通道一致的注意力权重,实现对特征图的动态加权调制.因此,在对试样制备过程中由于腐蚀不均或成像噪声等导致的显微组织图像进行识别时,MSCA-Net展现出更强的鲁棒性和泛化能力.

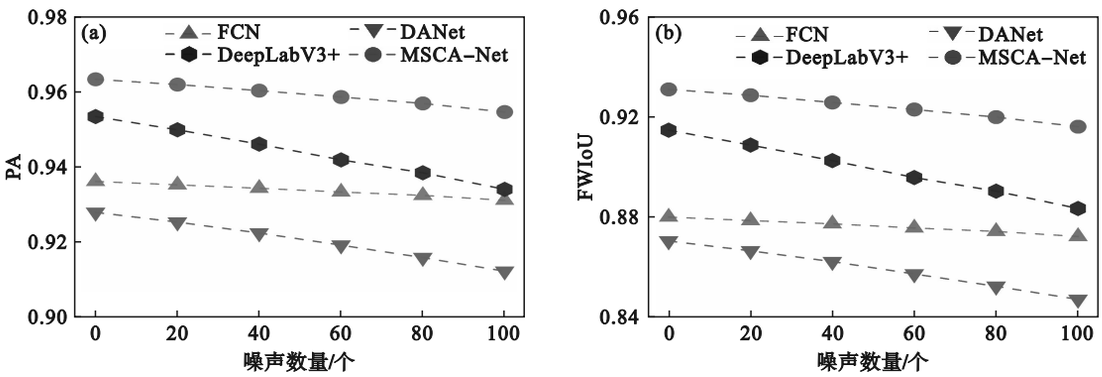


图 13 模型在不同噪声数量下的PA和FWIoU变化

Fig. 13 PA and FWIoU variations of model under different noise amounts

(a)—PA变化趋势;(b)—FWIoU变化趋势.

3.4 显微组织精细化表征

为实现从像素级到实际物理尺度的精确转换,本文使用OpenCV视觉库对经FCN,DANet,DeepLabV3+与MSCA-Net模型识别后的显微组织图像进行了定量分析.具体而言,采用轮廓检测算法自动提取4种模型输出图像中晶粒的边界轮廓,获取包括晶粒面积与周长等在内的形态参数,并将这些统计结果统一保存为CSV格式文件,为后续的深入分析与模型性能比较提供可靠的数据基础.图14展示了2个典型显微组织图像所对应的晶粒尺寸分布直方图.从统计结果可以明显看出,基于MSCA-Net模型输出的显微组织图像识别结果统计的晶粒尺寸分布曲线与GT高度吻合,峰值位置基本一致,表明其在晶粒尺寸识别方面具有较高的精度.相反,其余3种模型的分布曲线则存在显著的偏差,其峰值位置及整体分布形态均与真实情况有所偏离.上述对比结果充分说明,MSCA-Net能够有效捕捉晶粒边界特征与形态细节,在抑制图像噪声和减少误识别方面具有良好的表现,从而在晶粒分割和定量统计任务中实现了更加精准的识别效果.

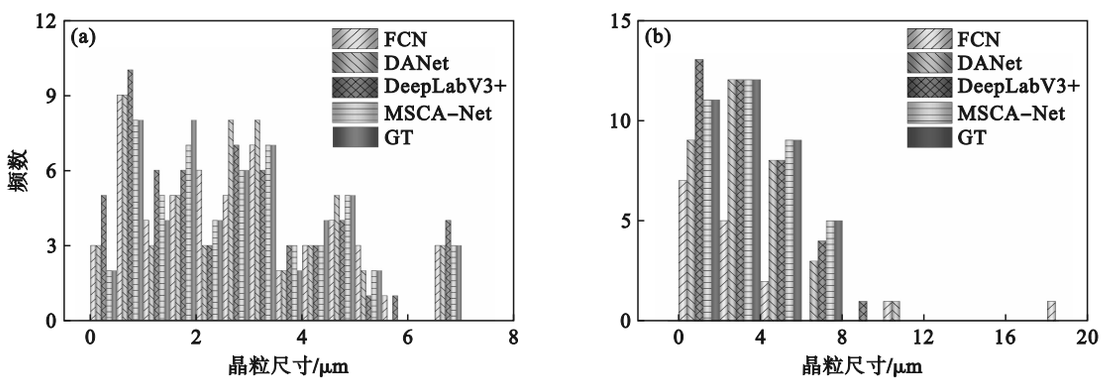


图 14 显微组织图像数据1和2的晶粒尺寸分布

Fig. 14 Grain size distribution of microstructure image data 1 and 2

(a)—显微组织图像数据1的晶粒尺寸分布;(b)—显微组织图像数据2的晶粒尺寸分布.

为实现对显微组织特征物理尺度的精细化分析,本文依据晶粒尺寸进行了分类统计:将尺寸小于2 μm的晶粒划分为小尺寸晶粒,2~4 μm的晶粒定义为中等尺寸晶粒,大于4 μm的晶粒则归类为大尺寸晶粒.图15展示了对识别结果中晶粒按尺寸区间进行统计的分布情况.可以看出,MSCA-Net模型所识别出的晶粒尺寸分布与真实数据高度一致.具体来看,在数据1

中,MSCA-Net识别得到的小、中、大尺寸晶粒占比分别为40%,36.36%和23.64%,其中小尺寸晶粒比例与真实值完全吻合,中等尺寸和大尺寸晶粒的绝对误差均为0.018 1.在数据2中,该模型的分类结果与实际分布完全一致,3种晶粒的占比分别为29.73%,32.43%和37.84%.上述分析表明,MSCA-Net在显微组织的尺寸分类与统计方面具备出色的准确性,其性能显著优于其余3种模型,展现出更强的识别能力与应用潜力.

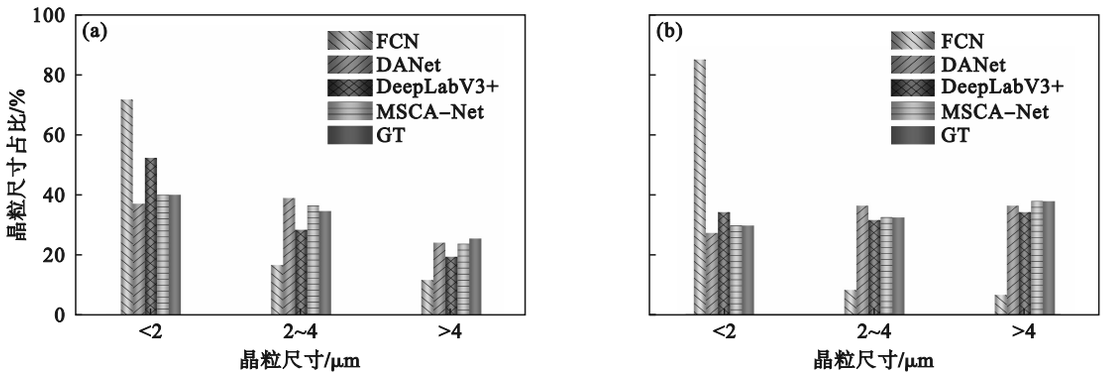


图15　晶粒尺寸分类统计结果对比

Fig. 15　Comparison of statistical results for grain size classification

(a)—显微组织图像数据1晶粒尺寸类别统计;(b)—显微组织图像数据2晶粒尺寸类别统计.

此外,为进一步评估MSCA-Net模型在显微组织识别中的准确性,本文对测试集的显微组织特征进行了精细化统计与对比分析.所统计的精细化显微组织特征包括铁素体面积分数、珠光体面积分数、晶界面积分数、铁素体相平均尺寸、珠光体相平均尺寸、单位面积相界长度和单位面积晶界长度.图16显示了基于FCN, DANet, DeepLabV3+与MSCA-Net模型识别结果所提取的特征数据与真实值之间的均方根误差(root mean square error, RMSE).

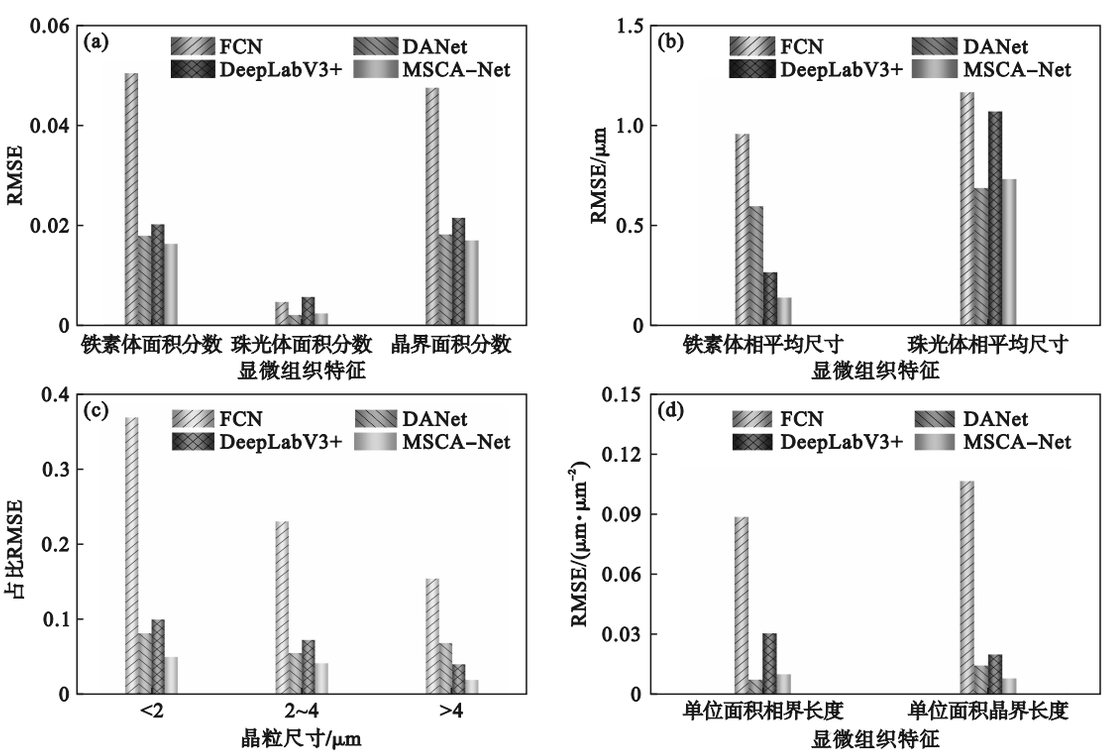


图16　基于FCN,DANet,DeepLabV3+和MSCA-Net模型的显微组织特征和晶粒尺寸的均方根误差对比

Fig. 16　RMSE comparison of microstructural features and grain sizes based on FCN, DANet, DeepLabV3+ and MSCA-Net models

(a)—铁素体、珠光体和晶界面积分数误差对比;(b)—铁素体和珠光体相平均尺寸误差对比;(c)—晶粒尺寸分类占比误差对比;(d)—单位面积相界长度和单位面积晶界长度误差对比.

从图中可以看出,基于MSCA-Net模型提取的各项特征参数的RMSE值总体上明显低于其余3种模型,这表明其识别结果更接近真实情况.相比之下,FCN,DANet和DeepLabV3+模型在复杂相界和晶界区域的识别上存在明显的偏差,导致其计算出的特征值与实际值之间误差较大.而MSCA-Net通过引入多尺度上下文注意力机制,增强了模型对局部细节与全局结构的感知能力,有效提升了特征表达的丰富性与融合能力.这不仅优化了显微组织边界的识别精度,也显著提高了各类定量特征的计算准确性.

上述结果充分验证了MSCA-Net在显微组织图像识别任务中具有更高的精确度和更强的鲁棒性.其优异的特征强化与多层次融合学习能力,使其在处理具有复杂拓扑结构的显微图像时表现出更优的泛化性能,为材料微观结构的自动量化分析提供了更为可靠的技术支持.

4 结 论

1)通过引入MSCA模块和跳跃连接,构建了MSCA-Net模型,在测试集上的像素准确率和频权交并比分别达到0.963 2和0.930 6,相比经典的FCN模型分别提升了2.71%和5.07%,实现了对复杂显微组织结构的较为准确识别.

2)MSCA-Net模型在不同噪声强度的测试集上均表现出良好的性能,其像素准确率和频权交并比始终高于0.95和0.91.这一结果表明,该模型在处理含噪显微组织图像时仍能保持较高的识别准确率,展现出良好的抗干扰能力与鲁棒性.

3)本文结合OpenCV和Tesseract-OCR对已识别的显微组织结构进行了定量表征.与基于FCN,DANet及DeepLabV3+模型识别结果的定量表征相比,基于MSCA-Net模型识别结果的定量表征更接近真实值,这进一步验证了其在显微组织精细化识别方面的优越性.

4)本文提出的MSCA-Net模型在“铁素体-珠光体”显微组织结构的精细化识别任务中表现出良好的性能.为扩展模型应用范围,后续研究可针对“铁素体-马氏体”双相钢开展.

参考文献:

[1] Yang D P, Du P J, Wu D, et al. The microstructure evolution and tensile properties of medium-Mn steel heat-treated by a two-step annealing process [J]. *Journal of Materials Science & Technology*, 2021, 75: 205-215.

[2] 李昊, 刘延辉, 王淼, 等. 显微组织对TC11合金力学性能的影响[J]. 陕西科技大学学报, 2025, 43(2): 163-169, 195.
(Li Hao, Liu Yan-hui, Wang Miao, et al. Effect of microstructure on mechanical properties of TC11 alloy [J]. *Journal of Shaanxi University of Science & Technology*, 2025, 43(2): 163-169, 195.)

[3] 田洪雷, 吴广宇, 王天, 等. 微观组织特征对服役6N01-T6铝合金材料力学性能的影响[J]. 轻金属, 2022(4): 49-54.
(Tian Hong-lei, Wu Guang-yu, Wang Tian, et al. Influence of microstructure characteristics on the mechanical properties of 6N01-T6 aluminum alloy materials in service [J]. *Light Metals*, 2022(4): 49-54.)

[4] Liu J, Aruguete D M, Jinschek J R, et al. The non-oxidative dissolution of galena nanocrystals: insights into mineral dissolution rates as a function of grain size, shape, and aggregation state [J]. *Geochimica et Cosmochimica Acta*, 2008, 72(24): 5984-5996.

[5] 朱骥飞, 张立, 徐涛, 等. 基于Image J软件的硬质合金显微组织参数化定量分析[J]. 粉末冶金材料科学与工程, 2015, 20(1): 26-31.
(Zhu Ji-fei, Zhang Li, Xu Tao, et al. Quantificational characterization of microstructural parameters of cemented carbides based on Image J software [J]. *Materials Science and Engineering of Powder Metallurgy*, 2015, 20(1): 26-31.)

[6] 黄业恩, 安建虹, 章思思, 等. Image-Pro Plus与Image J图像分析功能比较及其在生物组织结构测试中的应用[J]. 中国体视学与图像分析, 2015,20(2): 185-196.
(Huang Ye-en, An Jian-hong, Zhang Si-si, et al. Image-Pro Plus and Image J: comparison and application in image analysis of biological tissues [J]. *Chinese Journal of Stereology and Image Analysis*, 2015, 20(2): 185-196.)

[7] Chowdhury S A, Taufique M F N, Wang J, et al. Automated grain boundary (GB) segmentation and microstructural analysis in 347H stainless steel using deep learning and multimodal microscopy [J]. *Integrating Materials and Manufacturing Innovation*, 2024, 13(1): 244-256.

[8] 张永志, 辛全忠, 王永亮, 等. 基于迁移学习的钢金相组织分类与识别方法的研究[J]. 材料导报, 2021, 35(24): 24152-24157.
(Zhang Yong-zhi, Xin Quan-zhong, Wang Yong-liang, et al. Research on classification and recognition method of steel metallographic structure based on transfer learning [J]. *Materials Reports*, 2021, 35(24): 24152-24157.)

[9] Medghalchi S, Kortmann J, Lee S H, et al. Automated segmentation of large image datasets using artificial intelligence for microstructure characterisation and damage analysis[J]. *Materials & Design*, 2024, 243: 113031.

[10] Huang Z L, Wang X G, Huang L C, et al. CCNet: criss-cross attention for semantic segmentation [C]//2019 IEEE/CVF International Conference on Computer Vision (ICVV). Seoul, 2019: 603-612.

[11] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, 2017: 2261-2269.

[12] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//the European Conference on Computer Vision (ECCV). Cham: Springer, 2018: 833-851.

[13] Jang J, Van D, Jang H, et al. Residual neural network-based fully convolutional network for microstructure segmentation[J]. *Science and Technology of Welding and Joining*, 2020, 25(4): 282−289.
[14] Banerjee S, Ghosh S K, Datta S, et al. Segmentation of dual phase steel micrograph: an automated approach[J]. *Measurement*, 2013, 46(8): 2435−2440.
[15] Azimi S M, Britz D, Engstler M, et al. Advanced steel microstructural classification by deep learning methods[J]. *Scientific Reports*, 2018, 8: 2128.
[16] Li X, Cui L Y, Li J K, et al. Automation of intercept method for grain size measurement: a topological skeleton approach[J]. *Materials & Design*, 2022, 224: 111358.
[17] Bowles A, Nogita K, Dargusch M, et al. Grain size measurements in Mg−Al high pressure die castings using electron back-scattered diffraction (EBSD)[J]. *Materials Transactions*, 2004, 45(11): 3114−3119.
[18] Gajalakshmi K, Palanivel S, Nalini N J, et al. Grain size measurement in optical microstructure using support vector regression[J]. *Optik*, 2017, 138: 320−327.
[19] Banerjee S, Chakraborti P C, Saha S K. An automated methodology for grain segmentation and grain size measurement from optical micrographs[J]. *Measurement*, 2019, 140: 142−150.
[20] Flipon B, Grand V, Murgas B, et al. Grain size characterization in metallic alloys using different microscopy and post-processing techniques[J]. *Materials Characterization*, 2021, 174: 110977.
[21] Guo M H, Lu C Z, Hou Q, et al. SegNeXt: rethinking convolutional attention design for semantic segmentation[J]. *Advances in Neural Information Processing Systems*, 2022, 35: 1140−1156.
[22] Chollet F. Xception: Deep learning with depthwise separable convolutions[C]//2017 IEEE conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, 2017: 1800−1807.
[23] Python Software Foundation. Python 3.11.5[EB/OL]. (2023−08−24)[2023−09−11]. https://www. python. org/downloads/release/python−3115.
[24] Paszke A, Gross S, Massa F, et al. PyTorch: an imperative style, high-performance deep learning library[J]. *Advances in Neural Information Processing Systems*, 2019, 32: 8024−8035.
[25] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python[J]. *Journal of Machine Learning Research*, 2011, 12: 2825−2830.
[26] Bradski G. The OpenCV library[J]. *Software Tools for the Professional Programmer*, 2000, 25(11): 120−123.
[27] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, 2015: 3431−3440.
[28] Fu J, Liu J, Tian H J, et al. Dual attention network for scene segmentation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, 2019: 3146−3154.
[29] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//the European Conference on Computer Vision (ECCV). Cham: Springer, 2018: 833−851.

(上接第 118 页)

[22] 李嘉琳, 何巍华, 曲永志. PSO优化深度神经网络诊断齿轮早期点蚀故障[J]. 东北大学学报(自然科学版), 2019, 40(7): 974−979.
(Li Jia-lin, He Wei-hua, Qu Yong-zhi. Diagnosis of gear early pitting faults using PSO optimized deep neural network[J]. *Journal of Northeastern University* (*Natural Science*), 2019, 40(7): 974−979.)
[23] Mirjalili S, Mirjalili S M, Lewis A. Grey wolf optimizer[J]. *Advances in Engineering Software*, 2014, 69: 46−61.
[24] Gao Y F, Li X Y, Zhang C J, et al. An improved gray wolf optimizer for wafer probing scheduling problem[C]//2025 28th International Conference on Computer Supported Cooperative Work in Design (CSCWD). Compiegne, 2025: 819−824.
[25] 王述红, 魏崴, 韩文帅, 等. 基于CGWO算法的边坡最小安全系数全局寻优方法[J]. 东北大学学报(自然科学版), 2022, 43(7): 1033−1042.
(Wang Shu-hong, Wei Wei, Han Wen-shuai, et al. Global optimization search method for minimum safety factor of slope based on chaotic grey wolf optimization algorithm[J]. *Journal of Northeastern University* (*Natural Science*), 2022, 43(7): 1033−1042.)