

基于新型半监督分割框架的齿轮钢 残余奥氏体测量方法

邓爽^{1,2},李吉帅^{1,2},杨伟^{1,2},张佳^{1,2},顾津瑞^{1,2},李平平²,耿路峰^{1,2}

(1.中车科技创新(北京)有限公司,北京 100192;

2.中车戚墅堰机车车辆工艺研究所股份有限公司,江苏 常州 213011)

摘要:残余奥氏体作为影响齿轮钢硬度、耐磨性和尺寸稳定性的关键组织,其含量的准确量化对热处理工艺优化、产品质量控制及服役性能预测至关重要。针对传统测定残余奥氏体含量方法(如X射线衍射)依赖昂贵设备且难以实现微观组织分布的可视化定量分析的问题,提出一种先进的深度学习框架自动识别并量化齿轮钢金相图像中的残余奥氏体区域。首先,构建面向齿轮钢金相组织的语义分割图像数据集,并完成像素级精细标注;针对残余奥氏体分布细粒度高的问题,提出基于滑动窗口裁剪的图像预处理方法;针对标注数据稀缺的行业痛点,提出一种新型的半监督学习方法,即增强版混合匹配算法(MixMatch++),通过伪标签融合生成、低置信度图像筛选和混合数据增强机制,显著提升小样本条件下的分割精度。最后,设计一种加权DiceLoss损失函数,以增强网络对语义模糊像素的关注能力。实验结果表明,该框架能够为残余奥氏体的定量分析提供一种高效且可靠的解决方案,从而为齿轮钢热处理工艺优化提供高效、低成本智能检测手段。

关键词:残余奥氏体;齿轮钢;语义分割;半监督学习;微观组织分析

中图分类号:U270.4 **文献标志码:**A

参考文献引用格式:邓爽,李吉帅,杨伟,等.基于新型半监督分割框架的齿轮钢残余奥氏体测量方法[J].轨道交通材料,2026,5(3):11-18.

0 引言

齿轮钢的金相图像揭示了其微观组织特征,主要包括淬火后形成的针状或板条状马氏体、残余奥氏体(白色块状)以及可能存在的贝氏体(羽毛状或针状)。适量的残余奥氏体能提高齿轮的韧性和抗冲击性能,减少淬火裂纹倾向;但含量过高会降低表面硬度和耐磨性,导致尺寸不稳定或微裂纹产生,并降低渗碳齿轮钢的疲劳性能^[1],因此识别残余奥氏体含量具有重要意义。然而,传统测量方法(如X射线衍射与人工金相分析)存在设备成本高、检测

效率低以及难以实现空间分布表征等问题。因此,将残余奥氏体识别问题转化为基于深度学习的图像语义分割任务,成为实现其高效定量分析的重要发展方向。

自U-Net^[2]模型提出以来,基于编码器-解码器结构的语义分割方法在医学影像与遥感图像等领域取得了显著进展。随后,研究者围绕特征表达能力与多尺度建模能力,对U-Net进行了大量改进,如引入残差结构^[3-4]、注意力机制^[5]及Transformer编码器^[6-7]等,以提升复杂场景下的分割性能。然而,这类方法大多依赖大规模像素级标注数据,在实际应用中往往面临标注成本高、数据获取困难等问题。

为缓解标注数据不足的问题,半监督语义分割方法^[8-16]逐渐成为研究热点。典型方法包括基于一致性正则化的Mean Teacher(MT)^[9]模型,以及基于伪标签与数据增强的MixMatch^[12]、FixMatch^[13]等框架。这类方法通过利用大量未标注数据,引导模型

基金项目:中车重点科技项目(2026CYB495)

收稿日期:2026-01-10

第一作者简介:邓爽(1995—),男,博士,工程师,从事计算机视觉技术研究。

通信作者简介:李吉帅(1990—),男,博士,高级工程师,主要从事AI+材料、先进材料焊接及增材制造相关工作。

学习潜在数据分布,从而在小样本条件下提升分割性能。然而,现有半监督方法依赖于高置信度伪标签的生成质量。当应用于金相图像时,微观组织边界模糊、类别尺度差异大以及纹理分布复杂,导致伪标签置信度下降。

近年来,针对金属材料显微组织的语义分类、检测和分割问题,已有研究^[17-30]尝试基于全监督学习框架来解决,并在一些公开数据集(如 MetalDAM^[31]、UHCS^[32])和自建数据集上取得一定进展。后续也有研究^[33-36]引入基于 MT、MixMatch、FixMatch 等半监督学习策略以提升未标注数据的利用效率。上述方法中,CA-MT^[35]通过改进传统 MT 结构来引导模型在不同扰动输入下保持预测一致性,并引入注意力机制增强关键区域特征表达,在金相图像公开数据集上达到了目前最好的性能。杨伟等^[36]利用 FixMatch 半监督学习框架实现等温淬火球墨铸铁中残余奥氏体的识别。然而上述两种方法都假设半监督学习模块中的无标注数据在不同增强下的预测会保持稳定,且未对低置信度区域进行显式筛选或差异化处理,可能导致生成的伪标签质量较低。

尽管这些方法在多相组织分割任务中取得了一定效果,但在处理残余奥氏体这类细粒度、边界模糊且分布不均的微观组织时,仍存在以下不足:1)对高分辨率金相图像进行统一缩放输入,容易丢失关键细节信息;2)基于一致性或简单伪标签策略的方法,难以有效提升低置信度区域(尤其是组织边界)的监督质量;3)现有 MixMatch 类方法主要采用单一 MixUp^[37]增强策略,难以对复杂微观结构进行有效建模。针对上述问题,本文在现有半监督框架基础上进行改进:一方面,通过滑动窗口裁剪策略保留高分辨率图像中的细粒度结构信息;另一方面,在 MixMatch 框架基础上提出增强版混合匹配算法(MixMatch++),通过伪标签筛选与加入 CutMix^[38]的混合数据增强机制,提高低置信度区域的伪标签质量;同时,结合权重引导的 DiceLoss 损失函数^[39]设计,进一步增强模型对类别不平衡及边界区域的表征能力,从而提升残余奥氏体的定量分割精度。

1 主要研究内容

1.1 数据集构建

本文构建面向齿轮钢金相图像分析的语义分割数据集,其中所有标注图片的像素均经过金相学专家严格标注,标注信息包含 2 个语义分割类别(残

余奥氏体类别和其他类别)。采用的材料型号为 18CrNiMo7-6,热处理工艺为渗碳淬火+低温回火,并在两种显微镜下进行图像采集。图 1 显示了数据集中 4 个有标注样本的原始图片和标注后的图片(红色区域为残余奥氏体,灰色区域为其余组织),可以看出这 4 幅图像展现出的纹理和颜色信息有较大差异。最终,本数据集包含训练集样本 352 幅图片(图片分辨率为 1 944×2 580 像素),其中有标注训练样本 102 幅图片,无标注训练样本 250 幅图片,有标注测试集样本 21 幅图片。在该数据集上设计深度学习算法模型自动识别图像中的材料组成信息,可有效解决传统金相分析中人工标注效率低、主观性强等问题,实现残余奥氏体含量的智能、精准量化分析。

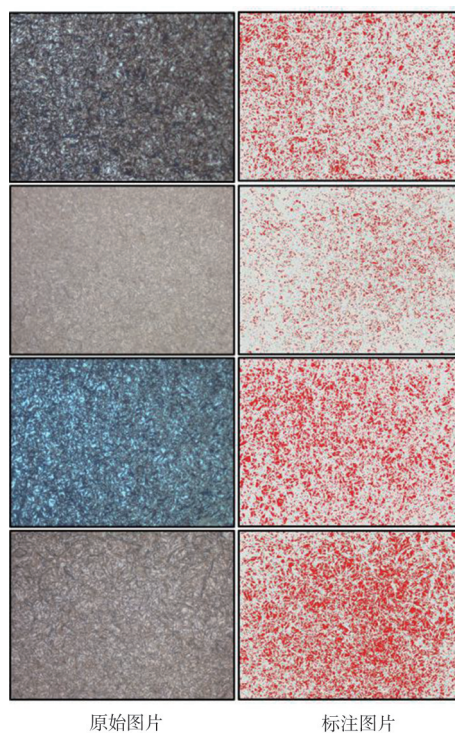


图 1 数据集中 4 个原始图片和标注图片

1.2 网络结构

本文提出的半监督语义分割模型在训练阶段的整体架构如图 2 所示。该架构并行处理有标注样本与无标注样本两条数据流。对于有标注样本,预处理流程依次包含滑动窗口裁剪与数据增强;对于无标注样本,预处理流程则在滑动窗口裁剪与数据增强之后额外引入增强版混合匹配算法 MixMatch++。数据增强操作涵盖随机翻转、随机裁剪、随机缩放、随机颜色噪声注入及随机灰度化。预处理后的图像共同输入至 Trans-UNet 骨干网络^[6]以提取深度语义特征。模型训练采用加权 DiceLoss^[39]作为分割

监督损失函数。在测试阶段,模型整体流程与训练阶段基本一致,但不再使用增强版混合匹配算法,即

对输入图像进行滑动窗口裁剪与必要的增强后直接输入 Trans-UNet 进行推理。

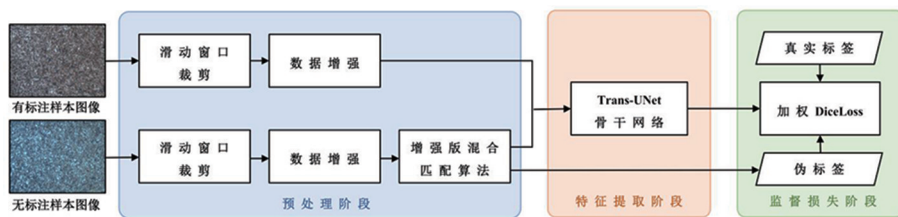


图2 训练阶段的网络架构

1.3 滑动窗口裁剪

在样本预处理阶段提出基于滑动窗口裁剪的数据增强技术,具体而言,在神经网络的训练阶段,对于一个分辨率为 $1\ 944 \times 2\ 580$ 像素的原始图片样本,在 $1\ 944$ 像素的高度上每隔 162 个像素采样一个裁剪子图起始点纵坐标,共 11 个坐标。然后在宽度为 $2\ 580$ 的像素上每隔 215 个像素采样一个裁剪子图起始点横坐标,共 11 个坐标。之后根据起始点坐标对原始图片裁剪成分辨率为 324×430 像素的子图。这 121 个裁剪子图经过数据增强操作,规范成分辨率为 320×320 像素的图片。在测试阶段,原图会基于滑动窗口均匀裁剪成 36 个不重叠的子图。原始齿轮钢金相图像样本经过滑动窗口裁剪技术

得到的子图如图3所示。

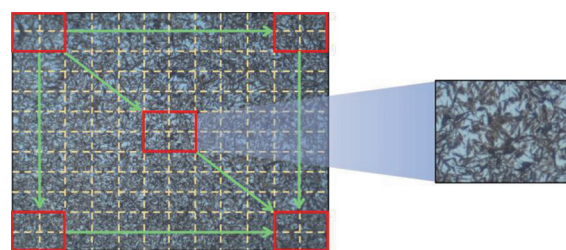


图3 图像经过滑动窗口裁剪得到的子图

1.4 增强版混合匹配算法

在样本预处理阶段中所采用的增强版混合匹配算法 (MixMatch++) 如图4所示,其包含伪标签筛选和混合数据增强两个步骤。

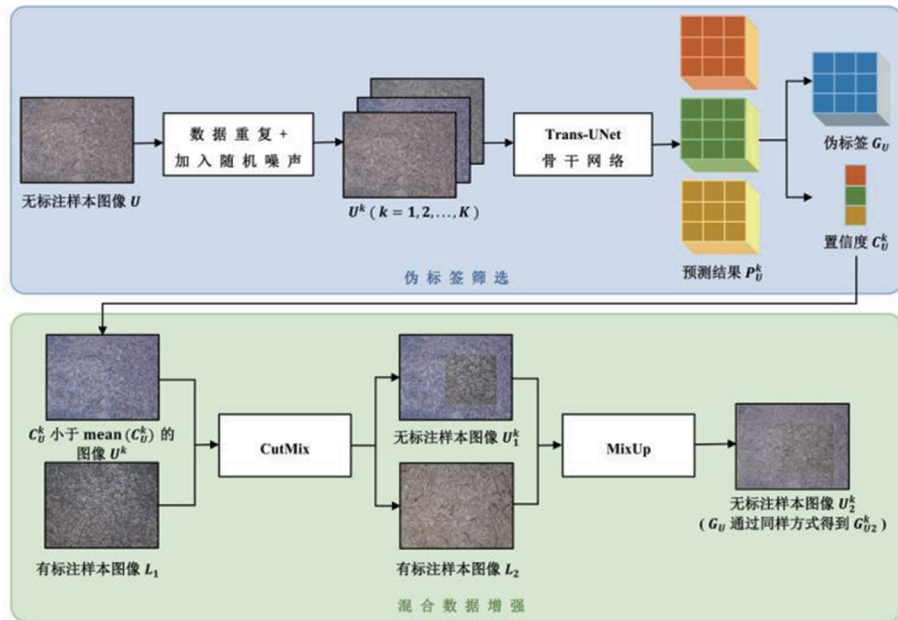


图4 MixMatch++流程

具体对于一个无标注训练图像 $U \in \mathbf{R}^{320 \times 320 \times 3}$, 先将该样本重复 K 份 ($K=3$), 然后这 K 幅图像中的每一幅图像都会加入一个均匀分布的随机噪声得到 $U^k (k=1, 2, \dots, K)$, 保证每幅图像不一样。之后每

幅图片 U^k 需要输入当前的语义分割神经网络模型得到网络的预测结果 $P_U^k \in \mathbf{R}^{320 \times 320 \times 2}$ 。根据 P_U^k 生成图像 U^k 的置信度 $C_U^k \in \mathbf{R}$ 以及软伪标签 $G_U \in \mathbf{R}^{320 \times 320 \times 2}$;

$$\begin{cases} \mathbf{E}_U^k(i,j) = 1 + \sum_{m=1}^2 (\mathbf{P}_U^k(i,j,m) \times \\ \ln(\mathbf{P}_U^k(i,j,m))) / \ln 2 \\ C_U^k = \frac{1}{320 \times 320} \sum_{i=1}^{320} \sum_{j=1}^{320} (\mathbf{E}_U^k(i,j) \times \\ \max_m(\mathbf{P}_U^k(i,j,m))) \\ \mathbf{G}_U(i,j,m) = \frac{1}{k} \sum_{k=1}^k \mathbf{P}_U^k(i,j,m) \end{cases} \quad (1)$$

式中: $\mathbf{P}_U^k(i,j,m)$ 为三维矩阵 \mathbf{P}_U^k 在 (i,j,m) 位置的值, $\mathbf{E}_U^k \in \mathbf{R}^{320 \times 320}$ 为样本 U^k 的预测信息熵。从这 K 个无标注图像 U^k (对应软伪标签为 \mathbf{G}_U) 中选择其对应置信度 C_U^k 小于 $\text{mean}(C_U^k)$ 的样本。

被挑选出来的每个图像 U^k 与任意 2 个有标注训练图像 $L_1 \in \mathbf{R}^{320 \times 320 \times 3}$ 和 $L_2 \in \mathbf{R}^{320 \times 320 \times 3}$ (对应 One-Hot 编码为 $\mathbf{G}_1 \in \mathbf{R}^{320 \times 320 \times 2}$ 和 $\mathbf{G}_2 \in \mathbf{R}^{320 \times 320 \times 2}$) 进行 CutMix^[38]+MixUp^[37] 的混合数据增强操作。首先无标注图像 U^k 和有标注图像 L_1 进行 CutMix 操作得到样本 $U_1^k \in \mathbf{R}^{320 \times 320 \times 3}$ 和其对应伪标签 $\mathbf{G}_{U_1}^k \in \mathbf{R}^{320 \times 320 \times 2}$ 。然后图像 U_1^k 和有标注图像 L_2 进行 MixUp 操作得到图像 $U_2^k \in \mathbf{R}^{320 \times 320 \times 3}$ 和其对应伪标签 $\mathbf{G}_{U_2}^k \in \mathbf{R}^{320 \times 320 \times 2}$ 。对于这 K 个无标注图像 U^k 中置信度 C_U^k 大于阈值 $\text{mean}(C_U^k)$ 的样本, 只将图像 U^k (对应伪标签为 \mathbf{G}_U) 与有标注图像 L_2 (对应真实标签为 \mathbf{G}_2) 进行 MixUp 操作得到图像 U_2^k 和其对应伪标签 $\mathbf{G}_{U_2}^k$ 。

最终将所有无标注图像 U_2^k 输入 Trans-UNet 骨干网络提取语义特征, 并和其对应的伪标签 $\mathbf{G}_{U_2}^k$ 进行分割监督。原有 MixMatch^[12] 只对所有无标注样本进行 MixUp 的简单数据增强操作, 而本论文提出的 MixMatch++ 算法在原有 MixMatch 算法基础上能针对伪标签置信度小的无标注图像样本加入 CutMix+MixUp 的融合数据增强机制, 更进一步增加了伪标签的置信度。

1.5 损失函数

本文在网络模型的训练阶段采用的监督方式为基于权重图的 DiceLoss 损失函数^[39]。具体而言, 记有标注或无标注样本经过网络输出的预测结果为 $\mathbf{P} \in \mathbf{R}^{320 \times 320 \times 3}$, 以及对应的真实标签或伪标签为 $\mathbf{G} \in \mathbf{R}^{320 \times 320 \times 2}$ 。构建一个类别权重图 $\mathbf{W}_C \in \mathbf{R}^{320 \times 320 \times 2}$ 来加强损失函数对残余奥氏体类别的注意力, \mathbf{W}_C 的构建如下所示:

$$\mathbf{W}_C(i,j,m) = 2 \times \mathbf{G}(i,j,0) + 1 \times \mathbf{G}(i,j,1) \quad (2)$$

式中: $\mathbf{W}_C(i,j,m)$ 为三维矩阵 \mathbf{W}_C 在 (i,j,m) 位置

的值; $\mathbf{G}(i,j,m)$ 为三维矩阵 \mathbf{G} 在 (i,j,m) 位置的值 ($m=0$ 表示残余奥氏体类别, $m=1$ 表示其他类别)。然后, 利用 Sobel 算子^[40] 对 \mathbf{G} 提取两个类别的边缘像素区域 S , 并根据 S 构建一个边缘权重图 $\mathbf{W}_S \in \mathbf{R}^{320 \times 320 \times 2}$:

$$\mathbf{W}_S(i,j,m) = \begin{cases} 2, & \text{if } (i,j) \in S \\ 1, & \text{if } (i,j) \notin S \end{cases} \quad (3)$$

式中: $\mathbf{W}_S(i,j,m)$ 为三维矩阵 \mathbf{W}_S 在 (i,j,m) 位置的值。最后, \mathbf{P} 与 \mathbf{G} 之间的带权 DiceLoss 损失函数 $\text{Loss}(\mathbf{P}, \mathbf{G})$ 通过如下公式得到:

$$\begin{cases} \mathbf{I} = \sum_{i=1}^{320} \sum_{j=1}^{320} (\mathbf{W}_C \odot \mathbf{W}_S \odot \mathbf{P} \odot \mathbf{G}) \\ \mathbf{Q} = \sum_{i=1}^{320} \sum_{j=1}^{320} (\mathbf{W}_C \odot \mathbf{W}_S \odot \mathbf{P}) \\ \mathbf{T} = \sum_{i=1}^{320} \sum_{j=1}^{320} (\mathbf{W}_C \odot \mathbf{W}_S \odot \mathbf{G}) \\ \text{Loss}(\mathbf{P}, \mathbf{G}) = 1 - \frac{1}{2} \sum_{m=1}^2 \frac{2 \times \mathbf{I}}{\mathbf{Q} + \mathbf{T}} \end{cases} \quad (4)$$

式中: \odot 表示矩阵的点乘操作; \mathbf{I} , \mathbf{Q} 和 \mathbf{T} 分别代表 DiceLoss 损失函数中的交集项、预测项和目标项。上述损失函数中的类别权重图 \mathbf{W}_C 和 DiceLoss 函数能减少残余奥氏体和其他类别的样本不平衡问题, 而边缘权重图 \mathbf{W}_S 的构建能加强损失函数对语义类别边缘像素的注意力。

2 实验结果

2.1 实验配置、模型复杂度和推理速度

实验配置: 采用数据批次的大小为 8, 其中有标注批次大小为 5, 无标注批次大小为 3。本论文训练最大迭代次数为 20, 梯度优化器为 Adam 优化器^[41], 基础学习率为 $1e-2$, 并在训练过程中基于余弦退火的学习率衰减策略对学习率进行逐步更新。

模型复杂度: 在推理阶段的参数量与浮点运算数 (FLOPs) 与 Trans-UNet 完全一致 (参数量为 93.23×10^6 个, FLOPs 为 50.39 GMac)。

推理速度: 在实际推理过程中, 每个子图的单次推理平均耗时为 39.92 ms (约 25.05 FPS)。对于一张完整图像, 36 个子图的推理总耗时约为 1.44 s。该处理速度能够满足齿轮钢金相分析场景中对高精度与工程实用性兼备的部署需求 (单幅图像的推理时间小于 2 s)。

2.2 消融实验

在消融实验中, 将所提出的各个子模块逐步叠加到基线模型上, 以验证各模块的有效性。基线方法仅采用“有标注样本预处理 (数据增强) + Trans-UNet +

加权 DiceLoss”对原始图像进行全监督语义分割。具体对比实验设置如下:1)仅使用基线方法,记为“基线方法”;2)在基线方法的有标注样本预处理中加入滑动窗口裁剪,记为“基线方法+”;3)进一步加入“无标注样本预处理(滑动窗口裁剪+数据增强)”,并且对无标注样本进行伪标签监督训练,以实现半监督学习,记为“基线方法++”;4)在无标注样本预处理中加入 MixMatch++算法,记为“本文方法”。上述方法均在第 1.1 节提出的数据集上进行训练和测试,评价指标包括残余奥氏体类别的交并比(RA IoU)、平均交并比(mIoU)和平均准确率(OA),最终在测试数据集上的评价指标如表 1 所示。

表 1 消融实验后的测试集分割性能指标 %

方法	RA IoU	mIoU	OA
基线方法	67.58	79.03	92.29
基线方法+	69.46	80.75	93.75
基线方法++	70.08	82.01	94.20
本文方法	70.96	82.57	94.90

从表 1 可以观察到各模块贡献的递进关系,其分别对应了数据表征增强、特征流形平滑以及决策边界优化 3 个层面的改进:1)“基线方法+”相较于“基线方法”的提升验证了滑动窗口裁剪的物理意义,一方面迫使模型学习更丰富的局部上下文特征

以应对视场受限的情况,另一方面相当于引入了显式的尺度扰动正则化,从而增强网络对细微奥氏体边界的几何不变性识别能力。2)“基线方法++”相较于“基线方法+”的提升揭示了无标注数据在半监督框架下的流形平滑效应。引入无标注样本及其伪标签进行一致性训练,实质上是通过熵最小化原则促使未标注数据在特征流形上向高置信度类别簇聚拢。3)“本文方法”相较于“基线方法++”的进一步提升则归因于 MixMatch++算法对伪标签质量的熵约束优化。在“基线方法++”中,伪标签的置信度分布往往较为尖锐且包含噪声,容易导致错误累积。MixMatch++通过混合增强与温度锐化机制,不仅构造了更具挑战性的虚拟训练样本以平滑决策边界,更关键的是其对伪标签分布进行了校准。

总体而言,“本文方法”相较于“基线方法”RA IoU 提升 3.38%,mIoU 提升 3.54%,OA 提升 2.61%,体现了从数据粒度保持到特征空间正则,再到标签质量校准的协同优化效果。两个测试集样本的原始图像和不同方法下的预测图片如图 5 所示,进一步佐证了上述机理分析:在蓝色方框标注的复杂边界区域,基线方法存在明显的过分割或欠分割粘连,而本文方法通过上述机制的联合作用,实现了对残余奥氏体形态学边界的精确解析。

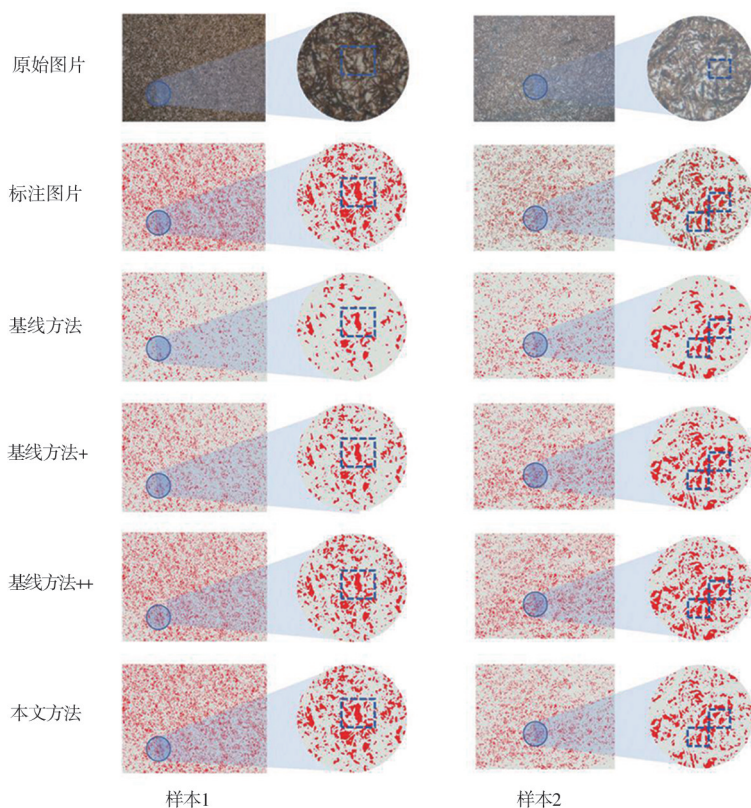


图 5 两个测试集样本的原始图像和不同方法下的预测图片

2.3 加权 DiceLoss 函数的有效性

为了评估加权 DiceLoss 函数的有效性,在训练时采用了多种损失函数进行对比实验,分别为:1) 交叉熵损失函数,记为“CELoss”;2) Focal 损失函数^[42],记为“FocalLoss”;3) 不带权重的 DiceLoss 损失函数,记为“DiceLoss”;4) 本文提出的损失函数,记为“加权 DiceLoss”。上述方法在测试时的评价指标如表 2 所示。

表 2 测试集上采用不同损失函数的分割性能指标 %

损失函数	RA IoU	mIoU	OA
CELoss	70.26	82.09	94.36
FocalLoss ^[42]	70.42	82.27	94.55
DiceLoss ^[39]	70.57	82.33	94.61
加权 DiceLoss	70.96	82.57	94.90

从表 2 的定量结果可以看出,从 CELoss 到加权 DiceLoss 的性能递进体现了损失函数在类别不平衡抑制与空间结构感知两个维度上的协同优化:1) DiceLoss 相较于 CELoss 与 FocalLoss 的提升,其机理在于 Dice 系数直接优化预测结果与真实标注之间的重叠度,能够在全局层面上保持对少数类(残余奥氏体)的召回完整性;2) 本文提出的加权 DiceLoss 在标准 DiceLoss 基础上进一步引入了类别权重图与边缘权重图,类别权重图依据标签分布频率对背景与前景类施加差异化的惩罚力度,边缘权重图则迫使网络在训练过程中向难以区分的边界像素分配更高的优化注意力。加权 DiceLoss 相较于标准 DiceLoss 在 RA IoU 上实现了 0.39% 的额外增益,且 mIoU 与 OA 同步提升。

2.4 MixMatch++ 的有效性

为了证明本论文提出的 MixMatch++ 的有效性,在两个公开金相图像语义分割数据集 MetalDAM^[31] 和 UHCS^[32] 上与其他半监督学习方法进行对比。为了保证实验结果的公平性和可比性,将所有方法在相同的实验环境和数据集设置下应用半监督语义分割任务。本论文提出的方法与其他半监督学习方法在 MetalDAM 数据集和 UHCS 数据集的测试集上的对比结果如表 3 所示,其中前 8 个方法的结果直接从文献摘录得到,MixMatch 算法和 MixMatch++ 算法的结果为本论文实验复现得到。

从表 3 中看出, MixMatch++ 在 MetalDAM 上较 MixMatch 提升 0.35%, 较当前最优方法 CA-MT 提升 0.19%, 在 UHCS 上分别提升 0.18% 与 0.33%, 其

机理在于 MixMatch++ 相对于原始 MixMatch 引入了伪标签置信度阈值筛选与混合数据增强的锐化校准。在金相图像分割任务中,无标注数据中的低置信度伪标签会向模型引入错误监督信号,导致特征空间边界模糊。 MixMatch++ 通过拒绝低置信度样本参与损失回传,实质上是对无监督训练过程施加了熵最小化约束,迫使模型仅在特征流形的高密度区域进行更新。

表 3 两个公开数据集的测试集分割性能指标 %

方法	mIoU (MetalDAM)	mIoU (UHCS)
U-Net ^[2]	62.69	51.19
MT ^[9]	63.07	52.21
Fixmatch ^[13]	63.52	53.52
Flexmatch ^[14]	63.81	53.27
BCP ^[10]	64.67	54.09
Con2Net ^[34]	64.95	54.68
CA-MT ^[35]	66.26	55.08
MixMatch	66.10	55.23
MixMatch++	66.45	55.41

为了进一步阐述 MixMatch++ 算法在训练过程中相比于 MixMatch 算法是否提升了无标注数据中伪标签的置信度,在 MetalDAM 和 UHCS 两个数据集的训练过程中的每一个迭代统计所有伪标签置信度 C_u^k 的均值,绘制成折线图,如图 6 所示。可以发现,本论文提出的 MixMatch++ 算法能在训练过程中大部分情况下都能比 MixMatch 算法有更高的伪标签置信度。

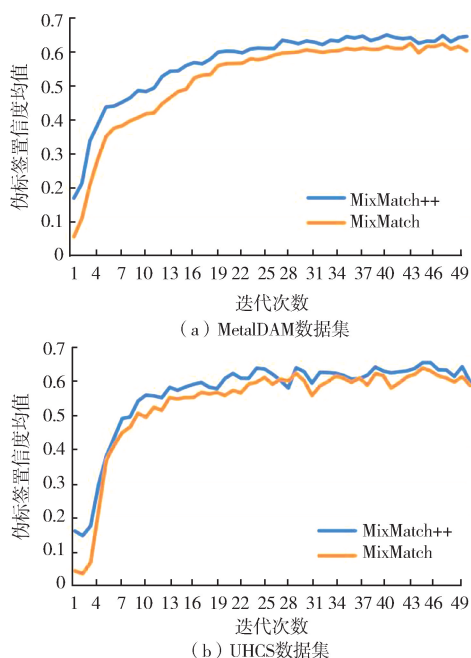


图 6 两个半监督算法在公开数据集上的伪标签置信度均值

3 结论

针对齿轮钢金相图像中残余奥氏体的测量提出一种先进的深度学习解决方案,首先,构建了精确标注的齿轮钢金相图像语义分割数据集,在齿轮钢金相图像上应用基于语义分割的神经网络;然后,提出基于滑动窗口的剪裁方法,减少图像预处理阶段残余奥氏体的信息损失;之后,利用大量无标注数据进行自我学习并提升伪标签的置信度,提出增强版混合匹配算法(MixMatch++);最后,为了增强网络对残余奥氏体区域的注意力,提出了基于权重图的DiceLoss损失函数。在构建的齿轮钢金相图像数据集上的实验结果显示了本方法在测量残余奥氏体含量的有效性。除此之外,两个公开数据集上的验证指标也证明了 MixMatch++算法能很好地利用无标注数据来提升网络性能。

参考文献:

- [1] 薛彦均,韩磊,梁江涛,等.渗碳齿轮钢中残余奥氏体转变与疲劳性能的关系[J].钢铁研究学报,2025,37(5):652-659.
- [2] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [C]//Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Cham: Springer, 2015: 234-241.
- [3] DIAKOIANNIS F I, WALDNER F, CACCETTA P, et al. ResU-Net-a: A deep learning framework for semantic segmentation of remotely sensed data[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 162: 94-114.
- [4] ALOM M Z, HASAN M, YAKOPCIC C, et al. Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation [EB/OL]. 2018; arXiv: 1802.06955. <https://arxiv.org/abs/1802.06955>.
- [5] OKTAY O, SCHLEMPER J, LE FOLGOC L, et al. Attention U-Net: Learning where to look for the pancreas [EB/OL]. 2018; arXiv: 1804.03999. <https://arxiv.org/abs/1804.03999>.
- [6] CHEN J N, LU Y Y, YU Q H, et al. TransUNet: Transformers make strong encoders for medical image segmentation [EB/OL]. 2021; arXiv: 2102.04306. <https://arxiv.org/abs/2102.04306>.
- [7] CAO H, WANG Y Y, CHEN J, et al. Swin-unet: Unet-like pure transformer for Medical image segmentation [C]//Computer Vision - ECCV 2022 Workshops. Cham: Springer, 2023: 205-218.
- [8] LAINE S, AILA T. Temporal ensembling for semi-supervised learning [J/OL]. Proceedings of the International Conference on Learning Representations (ICLR), 2017: 1-13. DOI: <https://doi.org/10.48550/arXiv.1610.02242>.
- [9] TARVAINEN A, VALPOLA H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results [C]//Neural Information Processing Systems, 2017.
- [10] BAI Y H, CHEN D W, LI Q L, et al. Bidirectional copy-paste for semi-supervised medical image segmentation [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 17-24, 2023, Vancouver, BC, Canada. IEEE, 2023: 11514-11524.
- [11] LEE D H. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks [J]. Proceedings of the International Conference on Machine Learning Workshops (ICMLW), 2013: 896-901.
- [12] BERTHELOT D, CARLINI N, GOODFELLOW I, et al. MixMatch: A holistic approach to semi-supervised learning [EB/OL]. 2019; arXiv: 1905.02249. <https://arxiv.org/abs/1905.02249>.
- [13] SOHN K, BERTHELOT D, LI C L, et al. FixMatch: Simplifying semi-supervised learning with consistency and confidence [EB/OL]. 2020; arXiv: 2001.07685. <https://arxiv.org/abs/2001.07685>.
- [14] ZHANG B, WANG Y, HOU W, 等. Flexmatch: Boosting semi-supervised learning with curriculum pseudo labeling [J/OL]. Proceedings of the Conference on Neural Information Processing Systems (NeurIPS), 2021, 34: 18408-18419. DOI: <https://doi.org/10.48550/arXiv.2110.08263>.
- [15] SOULY N, SPAMPINATO C, SHAH M. Semi supervised semantic segmentation using generative adversarial network [C]//2017 IEEE International Conference on Computer Vision (ICCV). October 22-29, 2017, Venice, Italy. IEEE, 2017: 5689-5697.
- [16] MIYATO T, MAEDA S I, KOYAMA M, et al. Virtual adversarial training: A regularization method for supervised and semi-supervised learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(8): 1979-1993.
- [17] DECOST B L, HOLM E A. A computer vision approach for automated analysis and classification of microstructural image data [J]. Computational Materials Science, 2015, 110: 126-133.
- [18] AZIMI S M, BRITZ D, ENGSTLER M, et al. Advanced steel microstructural classification by deep learning methods [J]. Scientific Reports, 2018, 8: 2128.
- [19] DECOST B L, LEI B, FRANCIS T, et al. High throughput quantitative metallography for complex microstructures using deep learning: A case study in ultrahigh carbon steel [J]. Microscopy and Microanalysis, 2019, 25(1): 21-29.
- [20] BISWAS M, PRAMANIK R, SEN S, et al. Microstructural segmentation using a union of attention guided U-Net models with different color transformed images [J]. Scientific Reports, 2023, 13: 5737.
- [21] MAJUMDAR S, SAU A, BISWAS M, et al. Metallographic image segmentation using feature pyramid based recurrent residual U-Net [J]. Computational Materials Science, 2024, 244: 113199.
- [22] MÜLLER M, STIEFEL M, BACHMANN B I, et al. Overview: Machine learning for segmentation and classification of complex steel microstructures [J]. Metals, 2024, 14(5): 553.
- [23] 刘志,赵吉轩,王司同,等.基于改进YOLO算法的DH980高强钢金相组织识别 [J/OL]. 材料与冶金学报, 1-12 [2026-05-13]. <https://link.cnki.net/urlid/21.1473.TF.20251024.1645.002>.
- [24] 刘钊,吉雨翔,李泽曦,等.面向复杂金相组织的深度学习晶界分割与晶粒参数自动化表征方法 [J/OL]. 中国机械工程, 1-12

- [2026-05-13]. <https://link.cnki.net/urlid/42.1294.TH.20260420.1933.026>.
- [25] 段献宝,何惠珍,李平平,等.基于深度卷积神经网络的钢材微观组织分类识别[J].铁道车辆,2022,60(1):43-47.
- [26] 陶玉婷,李平平,徐云涛,等.基于深度学习和数字图像处理的晶界分割与修复[J].轨道交通材料,2024,3(3):6-11.
- [27] 高金威,聂武楠,陈祯,等.基于改进型YOLOv5s模型的钢脱碳层深度测定[J].轨道交通材料,2026,5(1):7-11.
- [28] 聂武楠,高金威,陈祯,等.基于深度学习和数字图像处理的灰铸铁石墨评级方法研究[J].轨道交通材料,2026,5(1):12-15,23.
- [29] 陈祯,王志,张泽曜,等.深度学习驱动的金相图像分析技术进展与应用综述[J].轨道交通材料,2026,5(1):1-6.
- [30] 张佳,李吉帅,李富强,等.基于语义分割的铸铁基体组织含量评定研究和应用[J].轨道交通材料,2026,5(1):24-33.
- [31] 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.
- [32] DECOST B L, HECHT M D, FRANCIS T, et al. UHCSDB: Ultra-High carbon steel micrograph DataBase: Tools for exploring large heterogeneous microstructure datasets[J]. Integrating Materials and Manufacturing Innovation, 2017, 6(2): 197-205.
- [33] CHEN D L, SUN D P, FU J, et al. Semi-supervised learning framework for aluminum alloy metallographic image segmentation[J]. IEEE Access, 2021, 9: 30858-30867.
- [34] CHEN F, ZHANG Y M, GUO Y L, et al. Semi-supervised metallographic image segmentation via consistency regularization and contrastive learning[J]. IEEE Access, 2023, 11: 87398-87408.
- [35] CAO Y Y, LUO B, CHEN Y H, et al. Confidence-Aware Mean Teacher for semi-supervised metallographic image semantic segmentation[J]. Computational Materials Science, 2025, 249: 113645.
- [36] 杨伟,李吉帅,邓爽,等.多特征融合与半监督学习的等温淬火球墨铸铁残余奥氏体智能分割方法[J].轨道交通材料,2026,5(1):16-23.
- [37] ZHANG H Y, CISSE M, DAUPHIN Y N, et al. Mixup: Beyond empirical risk minimization[EB/OL]. 2017; arXiv: 1710.09412. <https://arxiv.org/abs/1710.09412>.
- [38] YUN S, HAN D, CHUN S, et al. CutMix: Regularization strategy to train strong classifiers with localizable features[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). October 27-November 2, 2019, Seoul, Korea. IEEE, 2020: 6022-6031.
- [39] LI X Y, SUN X F, MENG Y X, et al. Dice loss for data-imbalanced NLP tasks[EB/OL]. 2019; arXiv: 1911.02855. <https://arxiv.org/abs/1911.02855>.
- [40] KANOPOULOS N, VASANTHAVADA N, BAKER R L. Design of an image edge detection filter using the Sobel operator[J/OL]. IEEE Journal of Solid-state Circuits, 1988, 23: 358-367. DOI: <https://doi.org/10.1109/4.996>.
- [41] KINGMA D P, BA J. Adam: A Method for Stochastic Optimization[J/OL]. Proceedings of the International Conference on Learning Representations (ICLR), 2015: 1-15.
- [42] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J/OL]. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017: 2980-2988. DOI: <https://doi.org/10.1109/iccv.2017.324>.

Measurement of retained austenite in gear steel microstructure based on a novel semi-supervised segmentation framework

DENG Shuang^{1,2}, LI Jishuai^{1,2}, YANG Wei^{1,2}, ZHANG Jia^{1,2}, GU Jinrui^{1,2}, LI Pingping², GENG Lufeng^{1,2}

(1. CRRC Technology Innovation (Beijing) Co., Ltd., Beijing, 100192, China;

2. CRRC Qishuyan Institute Co., Ltd., Changzhou, Jiangsu 213011, China)

Abstract: Retained austenite is a key microstructure influencing the hardness, wear resistance and dimensional stability of gear steel, and its accurate quantification is crucial for optimizing heat treatment processes, controlling product quality and predicting service performance. As the traditional methods for determining the content of retained austenite (such as X-ray diffraction) rely on expensive equipment and are difficult to realize visualized quantitative analysis of the microstructure distribution, an advanced deep learning framework is proposed to achieve automatic identification and quantification of the retained austenite area in gear steel metallographic images. Firstly, a semantic segmentation image dataset for gear steel microstructure is constructed to implement pixel-level fine annotation. In response to the fine-grained distribution of residual austenite, an image preprocessing method based on sliding window cropping is then proposed. Furthermore, to address the scarcity of annotated data, a novel semi-supervised learning method is proposed, combining pseudo-label fusion, low-confidence sample filtering, and hybrid data augmentation to improve segmentation performance with limited labels. Lastly, a weighted DiceLoss function is proposed to enhance the network's attention to semantically ambiguous pixels. Test results show that this framework can provide an efficient and reliable solution for the quantitative analysis of retained austenite, thus providing an efficient and low-cost intelligent detection method for optimizing the heat treatment processes for the gear steel.

Keywords: retained austenite; gear steel; semantic segmentation; semi-supervised learning; microstructure quantification