

特色专题

人工智能在金属材料组织图像识别与定量分析中的应用

沈春光^{1,2}, 孙硕^{1,2}, 徐伟^{3*}, 郑士建^{1,2*}

摘要 基于人工智能(artificial intelligence, AI)技术的微观组织识别及定量化兼具高精度和高效率的优势,有力推动了高通量组织分析技术的发展。聚焦 AI 辅助金属材料组织图像分析这一新兴领域,以微观组织由定性分析逐步向精细定量分析的发展为脉络,系统综述了传统机器学习分类算法、深度学习分类算法、目标检测算法、语义分割算法在金属材料微观组织分类、识别以及定量化方面的研究进展,尤其重点论述了广泛采用的语义分割算法的研究现状;同时,针对 AI 算法在材料微观组织图像分析领域面临的组织复杂度高、标注样本匮乏等瓶颈问题,介绍了数据增强、模型架构改进等方面的创新策略及其应用效果。最后,总结和展望了基于 AI 的微观组织图像分析方法目前存在的不足以及未来的发展方向。

关键词 人工智能;金属材料;微观组织;计算机视觉;图像识别;定量分析

在金属材料研发中,微观组织一直扮演着如同基因图谱般的核心角色,如同一座桥梁既连接着前端成分工艺,又控制着后端服役性能^[1-2]。为精准指导高性能合金设计及物理机制深化,实现微观组织的精确识别及其定量化尤为重要。早期的显微组织识别与定量化主要依赖于光学显微镜(optical microscope, OM)、扫描电镜(scanning electron microscope, SEM)等设备的组织图像,并结合图像处理软件对其中定量组织信息进行统计分析^[3]。由于具有操作简单、成本低等特点,该方法被广泛应用于金属材料微观组织

分析,但受限于主观人工经验,其仅能可靠地应用于形貌简单的微观组织,难以准确分析复杂的多相组织;此外,由于不同操作者之间存在主观经验差异,使得定量结果往往具有较低的可重复性与可比较性,尤其对于复杂组织,这种人工经验偏差更为显著,严重制约了组织定量化的可靠性。为了实现微观组织的高精度分析,以电子背散射衍射(electron backscatter diffraction, EBSD)为代表的基于晶体学的表征技术逐渐兴起并得到了广泛应用^[4-6]。尽管该类技术在微观组织分析中具有明显的精度优势,但实验效率低、实验成本高昂,难以应用于大规模图像数据的高通量分析。

随着大数据与 AI 时代的到来^[7],基于数据驱动的材料研发范式得到了业内学者的广泛关注与认可^[8-11],也为金属材料微观组织的高通量精准分析提供了全新的技术路径。如图 1 所示^[12-15],近 10 余年来,基于

1. 河北工业大学高性能轧辊材料与复合成形全国重点实验室, 天津 300401

2. 河北工业大学天津市材料层状复合与界面控制技术重点实验室, 天津 300401

3. 东北大学数字钢铁全国重点实验室, 沈阳 110819

收稿日期: 2025-09-29; 修回日期: 2025-11-10

基金项目: 国家自然科学基金项目(52404400 & U23A20543); 河北省创新能力提升计划项目(244A7626D)

作者简介: 沈春光, 副教授, 研究方向为基于人工智能的钢铁材料计算设计, 电子信箱: cgshen@hebut.edu.cn; 徐伟(通信作者), 教授, 研究方向为金属材料基因工程, 电子信箱: xuwei@ral.neu.edu.cn; 郑士建(共同通信作者), 教授, 研究方向为金属材料, 电子信箱: sjzheng@hebut.edu.cn

引用格式: 沈春光, 孙硕, 徐伟, 等. 人工智能在金属材料组织图像识别与定量分析中的应用[J]. 科技导报, 2025, 43(24): 44-60; doi:10.3981/j.issn.1000-7857.2025.09.00122

数据驱动的显微组织分析已由传统机器学习(machine learning, ML)算法逐渐过渡到深度学习(deep learning, DL)算法,相应地,建模过程中组织图像特征提取也由浅层手动特征转变为基于卷积操作的深层次自动特征,进而提升了显微组织分析模型的识别精度。同时,随着模型算法的迭代更新,显微组织分析也由最初的定性研究逐渐完善为精细的定量统计。目前,数据驱动的显微组织分析技术已在钢铁材料、钛合金、铝合金等多种典型金属材料中获得了广泛应用^[16-17],且还催生出一批嵌入 AI 技术的显微组织分析平台或软件^[18],极大加速了显微组织的高精度、高通量分析。然而,数据驱动技术在材料显微组织分析领域的应用也同样充满挑战。不同于交通、医疗等应用场景,材料领域的显微组织通常具有更高的

复杂度,这对数据集构建与模型架构选择等方面提出了更高要求;此外,显微组织图像中需要识别的物相信息量通常较大,数据标签的创建需要耗费较大人工成本,致使标注数据较为匮乏,同时显微组织图像中各类物相比比例通常差距显著,导致样本类别不平衡,上述因素均严重制约了 AI 模型的精度与鲁棒性。针对上述挑战,业内学者在数据集构建、模型网络结构设计等方面提出了相应的应对策略,较大程度提升了 AI 技术在材料组织图像分析领域的应用效果。本文通过对近年来相关领域的公开文献进行总结,对早期非 AI 技术以及近些年发展的 ML 技术、DL 技术在材料显微组织分类、识别与定量分析方面的研究进展进行系统综述,并对该领域的未来发展提出展望和建议。

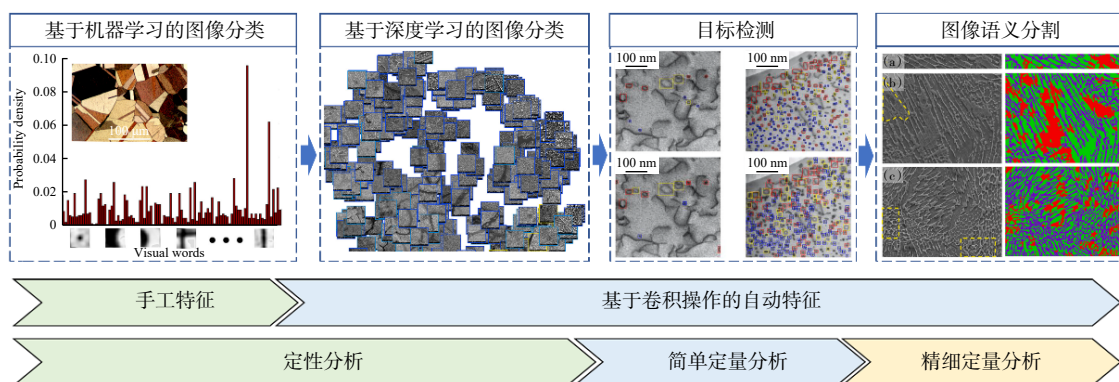


图1 金属材料显微组织分析中常用的 AI 技术及其特点

1 基于非 ML 算法的显微组织图像分析研究

在 AI 技术兴起之前,传统的显微组织图像数据分析方法多依赖人工经验操作或晶体学表征技术。人工经验操作方面,学者根据材料领域知识,通过人工主观经验对组织图像中各类微观结构进行定性分类^[19-20],随后通过手动测量或使用图像处理软件对微观结构的形貌、尺寸等信息进行定量分析^[21-22]。该方法操作简便,在组织图像分析中应用极其广泛。然而,受限于人类视觉的有限分辨能力,该人工经验主导的分析方法通常仅可准确应用于较为简单的显微组织,例如双相钢^[23],对更为复杂显微组织的分析效果往往并不理想。晶体学表征技术的快速发展有效解决了金属材料中复杂显微组织的识别与定量分析,例如常用的 EBSD^[24]与透射电子显微镜(transmission

electron microscope, TEM)中选区电子衍射(selected area electron diffraction, SAED)^[25],可实现微纳尺度下各类物相、析出相颗粒等微观结构的精准识别及其定量分析^[26-27]。但上述表征技术通常实验效率较低、实验成本较高,并不适用于组织图像数据的大规模、高通量分析,进而制约了所得定量结果的统计意义。除此之外,边缘检测等传统计算机视觉方法也已经被广泛应用于显微组织图像分析^[28],但其检测精度高度依赖人工阈值和图像质量,同时对复杂显微组织的应用效果并不理想。综上所述,传统组织分析方法难以兼顾高精度与高效率,亟需结合新兴的 AI 技术探索全新的组织图像分析范式,实现其精准高效分析,进而推进新材料研发。

2 基于 ML 算法的微观组织图像分类研究

自 Paul Raccuglia 等^[29]的《Nature》封面文章发表之后,基于统计学习理论的 ML 方法在材料科学领域得到了巨大关注并获得了广泛应用。以监督学习为代表的 ML 可以通过挖掘训练数据获得一个参数化的函数近似,从而建立起输入与输出之间的定量映射关系,为微观组织图像的分类预测研究提供了全新的技术路径。常用的 ML 分类算法包括支持向量机(support vector machine, SVM)^[30]、人工神经网络(artificial neural network, ANN)^[31]、随机森林(random forest, RF)^[32]等,表 1 中给出了各分类算法的简介及其优缺点。由于 ML 算法无法直接处理图像数据,因此该研究阶段主要采用人工预处理的手动提取的微观组织特征,通过 ML 分类算法建立其与组织类别之间的映射关系。图 2 中给出该阶段的常用建模思路^[33-34]。首先,通过样品制备与组织表征收集图像数据,随后对组织数据进行图像处理并手动提取形貌、纹理等特征参数。在获得特征参数之后,需要通过归一化等方

法对数据进行预处理,消除不同特征之间的数值维度差异,并通过特征工程方法筛选高关联度特征,用以构建最终的分类预测模型。Gola 等^[33, 35]应用上述建模方法对低碳钢微观组织进行了研究,首先通过二值化等图像处理方法提取 OM、SEM 图像中的组织形貌参数,随后应用 SVM 分类算法建立起组织形貌参数与组织类别(马氏体、珠光体、贝氏体)之间的映射关系,进而实现了组织图像分类,其分类精度可达到 85% 以上。该工作通过将参数化微观组织信息与 ML 模型相结合,实现了可靠的组织分类,在数值型数据层面上建立组织特征参数与相类别之间的关系,形成了基于早期 ML 算法的组织图像数据分析范式。与之相似, Müller 等^[34]基于 Haralick 特征、局部二值模式等纹理参数,采用 SVM 分类算法预测了不同冷速下获得的不同贝氏体组织; DeCost 等^[13]通过视觉词袋(bag of visual features)方法自动提取了组织图像中微观结构特征,并应用 SVM 模型建立起组织图像与其材料类别之间的关系,实现了自动分类和相似图像检索。

表 1 常用 ML 算法简介

算法	核心思想	优点	缺点
SVM	寻找最大间隔超平面,核技巧处理非线性问题	善于处理高维、非线性数据	可解释性差,对缺失数据较为敏感
ANN	模拟神经网络,通过多层连接学习复杂的非线性映射	能拟合极其复杂的非线性关系	可解释性差,小样本数据下易于过拟合
RF	Bagging集成,构建多棵决策树,通过投票得出结果	可处理高维特征,具有一定可解释性	在噪声较大的数据上易于过拟合

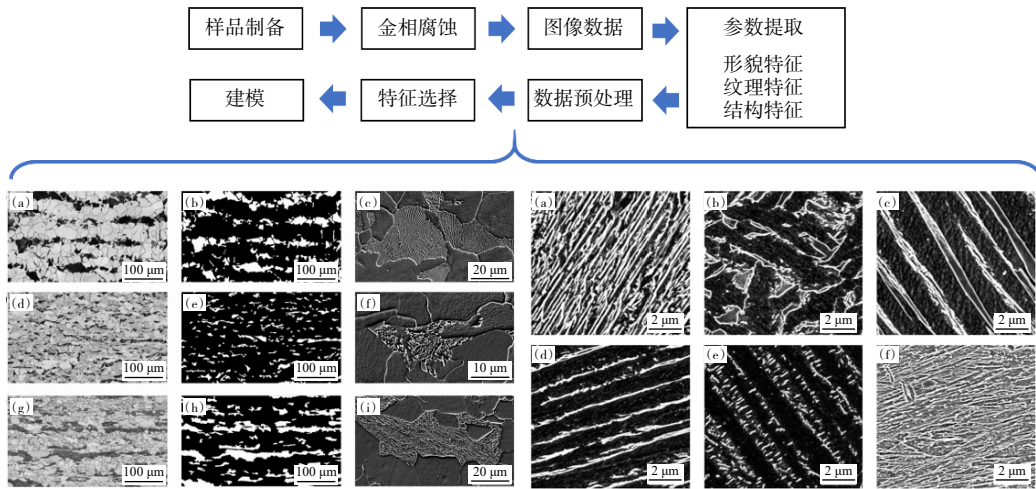


图 2 基于 ML 算法的微观组织分类模型构建及应用

通过合理地提取微观组织关键特征, ML 分类算法有能力搭建起组织特征及类别之间的关系, 实现相类别的可靠分类。此外, 该类模型还具有较好的物理机制可解释性, 可根据所建模型逆向分析出不同类别组织中最为核心的微观结构特征。同时, 传统 ML 算法对计算资源要求较低, 即便是普通的办公计算机也可实现模型训练与应用。此外, 数据是 ML 建模的基础, 由于该阶段的 ML 算法结构相对简单, 所含超参数的数量也较少, 因此模型训练过程中对数据量的要求较低, 并不会出现下文所述 DL 算法所面临的小样本问题。然而, 该类模型的精度与合理性高度依赖于手工提取的微观组织特征, 而组织特征的提取又高度依赖于操作者经验、图像质量等因素, 因此导致该类人工主观经验主导的图像分类方法难以可靠地应用于复杂的微观组织, 此外, 由于人工干预引起的数据质量差异还会导致该类模型的鲁棒性和泛化能力较差。同时, 受限传统分类算法有限的特征维度处理能力, 使得特征工程阶段会不可避免舍弃部分特征, 导致难以完整描述组织形貌、纹理特点, 进而制约了模型性能与应用效果。综上所述, 基于 ML 的组织分类模型的物理可解释性强, 数据依赖性较低, 但其模型精度与合理性严重依赖于基于主观经验的特征工程, 使其难以有效地处理复杂微观组织, 仍需发展更为客观、智能的无损失图像特征提取技术, 用以提升复杂组织图像数据的分类精度。

3 基于 DL 算法的微观组织图像分类研究

随着 AI 技术的快速发展, 以卷积神经网络(convolutional neural networks, CNN)^[36-37]为代表的 DL 算法在金属材料微观组织分析中获得了广泛应用。相较于传统 ML 算法的手动特征工程, DL 算法可通过卷积、池化等操作, 自动学习从低级(边缘、颜色)到中级(纹理、形貌)再到高级(晶粒形态等)的多维度图像特征表示, 实现图像核心特征的无损失提取, 进而提升微观组织分类精度及其对复杂微观组织的处理能力。目前, DL 组织图像分类模型的主流建模思路如图 3 所示, 主要分为 2 条技术路径: 其一, 利用 CNN 等 DL 网络自动提取图像核心特征, 随后利用局部特征聚合描述符(vector of local aggregated descriptors,

VLAD)^[38]等特征表示方法将高维特征转化为特征向量, 最后再利用传统 ML 分类方法建立起特征向量与组织类别之间的关系; 其二, 利用现有的 CNN 架构, 在卷积层后面直接连入全连接层, 利用全连接层建立起 CNN 特征与组织类别之间的关系。对于技术路径一, DeCost 等做了较为系统的工作, 他们结合 CNN 模型与 VLAD 等纹理特征提取算法对高碳钢组织图像进行了特征提取, 并利用 RF 分类器实现了微观组织图像的准确分类^[39]; 此外, 他们在高碳钢组织分类工作中还证实, 基于 CNN 特征的分类模型比传统尺度不变特征变换(scale-invariant feature transform, SIFT)特征的分类模型具有更高的精度^[12]。对于技术路径二, Mulewicz 等^[40]基于卷积层+全连接层的深度卷积神经网络(deep convolution neural networks, DCNN)模型结构, 实现了不同钢种中 7 类微观组织的高精度分类, 其分类精度可达到 99.8%; 与之相似, Larmuseau 等^[41]利用 ResNet50 模型对各类钢铁显微组织图像进行了可靠分类, 并且证实 DL 模型可获得比人工经验更高的分类精度, 尤其是对于复杂的马氏体组织, DL 模型的精度优势更为明显。

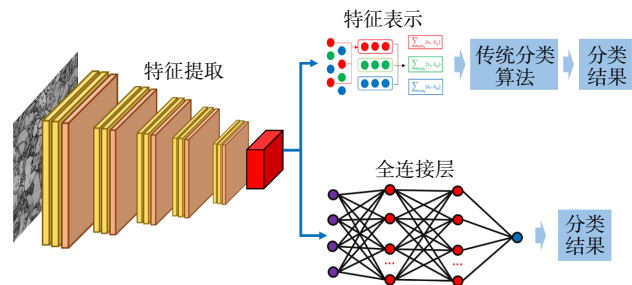


图 3 DL 组织图像分类模型的主要建模策略

CNN 等 DL 模型的应用标志着组织图像数据分析正式迈入了智能时代, 摒弃了微观组织分析对人工经验的依赖性。不像其他监督学习任务, 图像分类数据集中样本收集简单, 并且给图像分配类别标签的人工操作成本也相对较低, 因此, 基于 DL 的组织图像分类研究中小样本问题相对不突出, 通常可以获得较为充足的数据来训练高精度的 CNN。此外, 由于 DL 算法中采用多层卷积操作, 使其可以深度感知目标图像中各物相的形貌、纹理等局部特征, 构建出由低级到高级的层次化特征表示, 从而高效捕获图像的核心模式与空间结构, 进而有效克服了前文所述 ML 算法对复杂图像处理能力不足的困境, 大幅提升了 AI 技术

对复杂组织图像的处理能力,从而获得了优于人工经验的图像分类精度^[41],是当前进行组织图像分类研究的首选技术之一。尽管在当前特定的小范围组织分类研究中,标记数据匮乏的问题并不显著,但若开发针对多钢种、多材料体系的复杂组织图像分类模型时,仍需警惕样本量不足问题。此外,目前分类模型功能过于单一,仅可对组织图像进行简单定性分类,无法对图像中的定量组织信息开展深入分析,例如提取关键的相含量、尺寸等定量信息,严重制约了微观组织智能分析对理解组织性能关系乃至成分工艺设计的指导意义。因此,亟需在现有 DL 模型基础之上,通过采用计算机视觉领域的其他类别算法,实现微观组织图像的定性及定量分析。

4 基于目标检测算法的微观组织图像分析研究

前面所述的组织图像分类研究仅关注于简单的定性分析,而真正的材料研发过程则更关注量化的微观组织信息,例如含量、尺寸等。传统的组织定量分析多依赖人工经验主导的图像分析软件,尽管该方法在较为简单的组织图像中可以获得可靠应用,但其较低的人工操作效率仍会大幅制约图像数据的处理分析数量,进而降低了所得定量结果的统计意义。

此外,相似于图像分类领域所面临的困境,人工经验指导的图像定量分析仍难以应用于复杂组织。

为了在组织定性分析的基础上实现初步定量分析,DL 技术仍是实现该目标的有效途径。目标检测^[42]是 DL 领域的核心任务之一,可通过标记边界框实现目标物体的精准定位以及类别识别,在医疗、自动驾驶等诸多领域应用广泛^[43-45]。目标检测算法通过卷积操作可实现复杂组织图像特征的无损失提取,随后通过对目标区域的分类与回归操作,进而获取类别、位置以及定量信息,为微观组织定量信息的高通量计算提供技术路径。具体算法方面,YOLO(you only look once)系列是目标检测领域最为常用的算法之一^[46],表 2 是不同版本的 YOLO 算法的简介。如图 4 所示^[15,47],Shen 等^[47]应用 YOLOv3 模型对 FeCrAl 合金中辐照缺陷的位置与尺寸演化进行检测,获得了与人工标注基本一致的检测及定量精度;在他们的另一个工作中,采用区域卷积神经网络(region-based convolutional neural networks, R-CNN)模型对 TEM 图像中缺陷进行了更为细致的分类检测及其定量分析^[15];与之相似,Anderson 等^[48]利用 R-CNN 模型准确检测了 Inconel X-750 合金 TEM 图像中的辐照氦气泡,并定量统计了气泡的尺寸分布信息;Perera 等^[49]利用 YOLOv5s 算法检测了奥氏体和 Ti-6Al-4V 合金中气孔与第二相颗粒的尺寸和分布位置。

表 2 不同版本 YOLO 算法的简介

版本	核心特征	优势	局限
YOLOv1	开创性地将目标检测重构为单阶段回归问题,实现端到端优化	全局上下文推理有效降低背景误检率	空间定位精度粗糙,密集物体检测效果差
YOLOv3	采用三尺度特征金字塔结构,实现更有效的多尺度目标检测	多尺度检测能力强,兼具优异的速度和精度	锚框依赖性强,特征融合能力有限
YOLOv4	系统整合BoF和BoS优化技巧,构建高性能训练框架	推理零成本提升精度,模型应用广	超参数相互作用,训练过程较复杂
YOLOv7	提出可训练BoF概念	兼具优异的模型计算速度和精度,参数利用率高	模型结构复杂性高

目前,主流的目标检测算法多以监督学习为主,模型训练之前需要事先对组织图像中目标物体进行人工标记,形成数据集。然而,不同于交通、医疗等领域的目标检测任务,材料组织图像中通常含有极高数量密度的目标相,并且尺寸细小,这大大增加了标签图的创建成本,致使当前材料领域的目标检测往往面

临着标注样本量不足的问题。为了缓解上述瓶颈问题,近些年的研究也提出了数据增强、模型架构改进等策略。例如,Lynch 等^[50]针对 TEM 图像的缺陷检测任务,提出了一种 SIGMA-ML 数据合成方法,基于合成数据训练的 YOLOv7 模型在精确率、召回率等指标上与人工标注数据所训练的模型相当,证实了合成

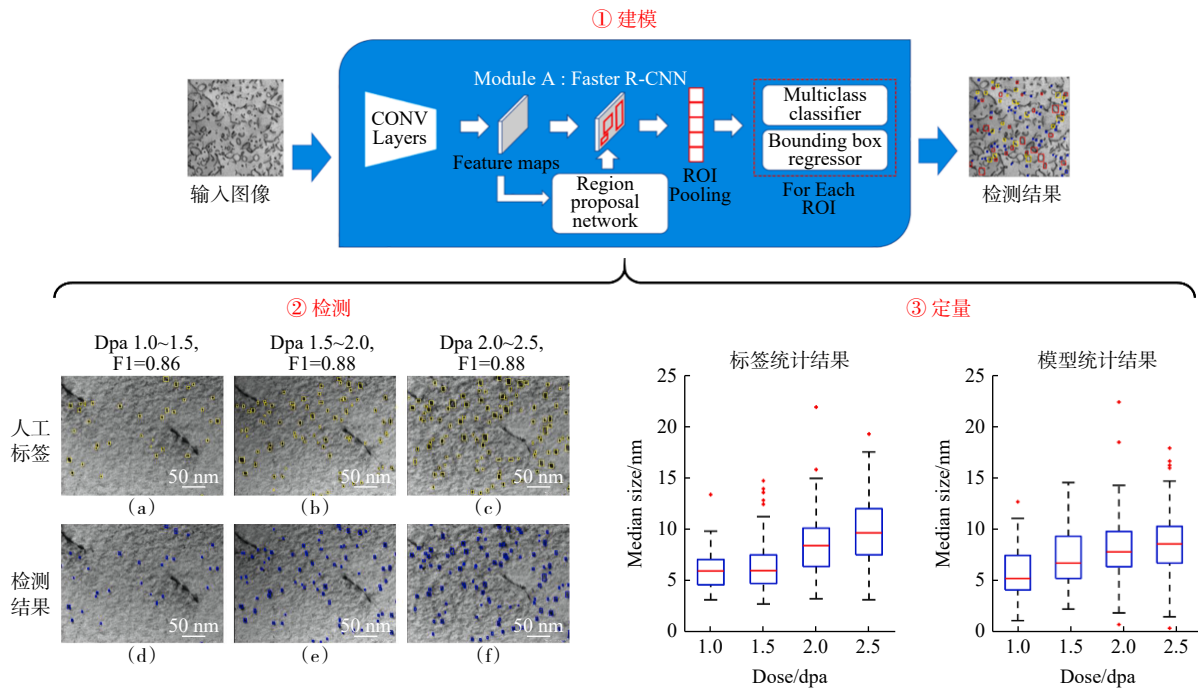


图4 基于目标检测算法的微观组织分析与量化

数据可大幅减少目标检测算法对人工标注的依赖。此外, Wu 等^[51]提出了一种迁移学习策略,该模型直接采用在大数据集上预训练的 MSCOCO 权重,使模型可学习到图像的基础特征,避免模型从头训练,降低对标注数据的依赖性,随后再将预训练模型在训练数据上进行参数微调,从而使训练模型可以对 TEM 图像中位错线与位错环进行准确定量分析。尽管该工作将迁移学习思想应用于小样本下的 YOLO 语义分割模型,但其对基于 YOLO 网络架构的目标检测任务仍具有较高的指导意义。目标检测算法也同样难以可靠应用于复杂微观组织图像,造成该问题的根源是数据集构建过程中的主观人工标注。例如, Li 等^[52]在研究中邀请了 5 位领域学者对 TEM 图像中的位错环进行标注,但不同学者所统计的位错环尺寸和数量均存在差异,这种由于人的主观因素所带来的数据集质量的不确定性,会严重影响所训练模型的客观性与精度,进而制约了其对复杂数据的分析能力。针对该问题, Jacobs^[53]提出建立具有共识性的图像数据标签,旨在通过较大规模学者的集体投票来产生更可靠、更为共识的数据标签,从而克服个人标注的主观性,建立起高可靠性模型。

通过上述列举的文献案例可知,目标检测算法可在对物相进行定性分类的基础上,进一步统计分析物

相的数量、尺寸等定量信息,为后续组织性能关系构建和材料设计提供基础。然而,受制于目标检测算法应用边界框标定物相的特点,使得该算法仅可应用于气孔、缺陷、颗粒等离散分布的物相,难以应用于连续交错分布、形态复杂的物相(例如复相钢中各类相);此外,定量组织信息提取方面也仅能基于边界框获得粗略的尺寸数据,无法对物相的精细轮廓进行检测,难以获取更为精细的含量、形貌信息。上述不足限制了目标检测算法对微观组织分析的应用范围及其对定量组织信息的挖掘深度。因此,在目标检测算法基础上,还需引入对物相边界轮廓分析能力更强的计算机视觉算法,以实现微观组织信息的全面、精细定量描述。

5 基于语义分割算法的微观组织图像分析研究

为了实现对材料中各类相、析出颗粒、缺陷等微观结构的统一分析,语义分割算法成为处理该问题的不二选择,是当前微观组织识别与定量分析领域应用最为广泛的机器视觉技术,也是本文的论述重点。语义分割算法同样基于卷积操作自动提取组织图像的核心特征,采用“端到端”网络结构,训练模型可输出

与输入图像相同尺寸的预测图像^[54],该图中不同物相被标记不同颜色,通过对输出图像进行细致的像素分析,便可得到系统的组织定量信息。目前,图像分割算法在多类金属材料的多样微观结构识别与量化工作中得到了普遍应用,大幅加速了组织图像数据的处理效率,但在进一步深入应用过程中也不断暴露出材料领域组织图像复杂、样本量有限等瓶颈问题。下面通过综述近些年组织图像语义分割领域发表的文献,系统介绍语义分割算法在微观组织识别方面的应用、存在问题及其应对策略。

语义分割算法是计算机视觉领域的重要分支之一,其核心目的在于对输入图像中各个像素点进行分

类,实现目标对象形态与边界的像素级精细预测,最终输出与输入图像尺寸一致的预测图像^[55]。图5给出了基于语义分割算法的显微组织识别与量化操作流程^[56-59]。对于语义分割算法,目前绝大多数研究均采用监督学习算法,建模过程中需要预先通过人工标注,建立包含“图像-标签”的数据集,随后通过模型训练,学习输入图像与标签图像之间的映射关系,典型的算法包括全卷积神经网络(fully convolutional neural networks, FCNN)^[60]、U-Net^[61]、SegNet^[62]等,表3中给出了各类语义分割算法的简介。2017年Azimi等^[56]基于FCNN模型架构建立了面向钢铁材料的组织识别模型,该模型通过卷积与池化操作自行提取SEM

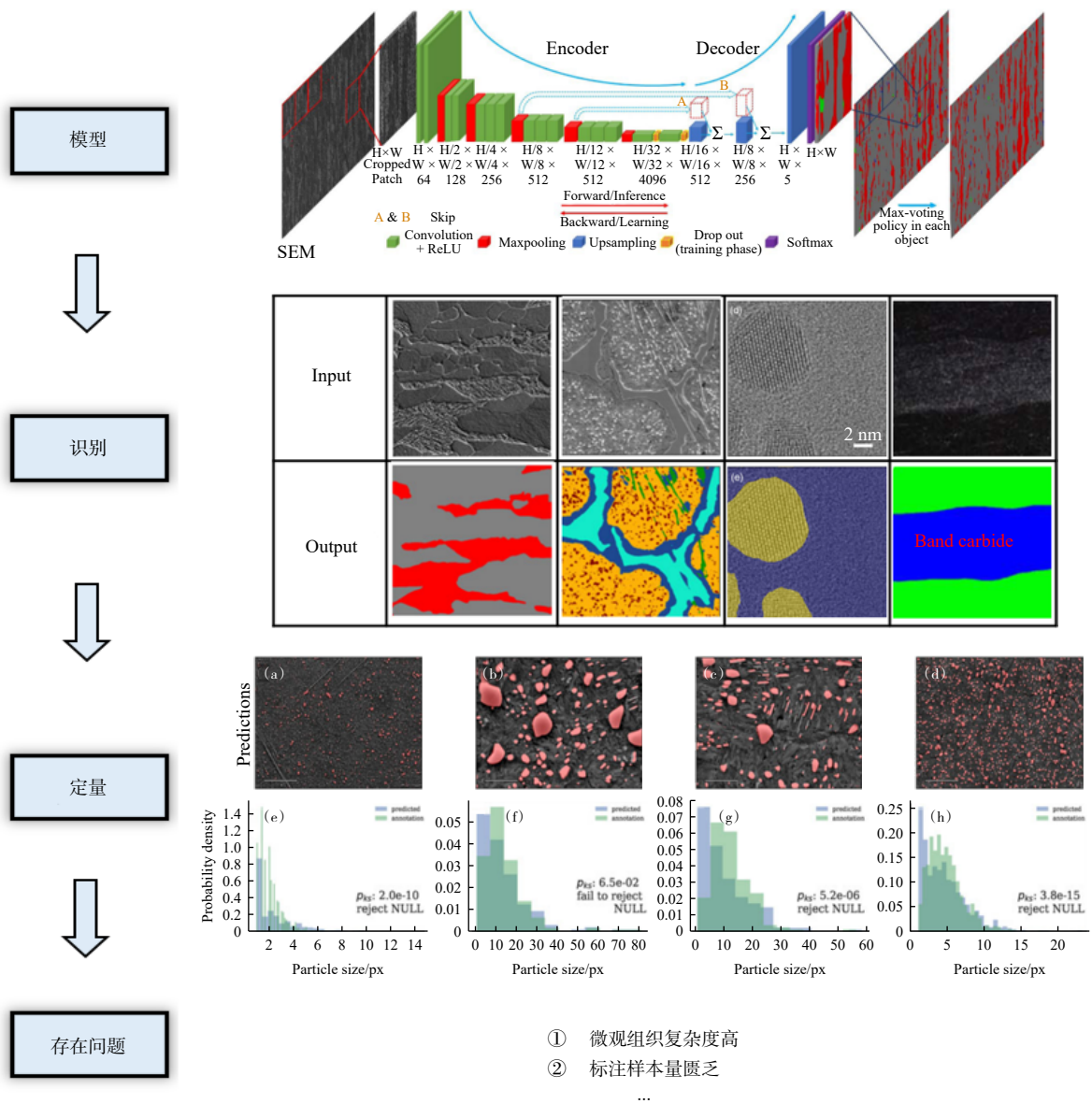


图5 基于语义分割算法的显微组织识别与量化流程

表 3 常用语义分割算法的简介

算法	核心思想	优势	局限
FCNN	将CNN中的全连接层替换为卷积层	实现端到端的语义分割,支持任意尺寸输入	下采样过程中细节丢失严重,导致特征图分辨率低,边界模糊
U-Net	采用编码器-解码器结构和跳跃连接	跳跃连接提升边界精细分割能力,小样本数据建模能力强	跳跃连接增加内存消耗,对大型数据集的处理能力不佳
SegNet	解码器采用非线性上采样,保留更多边界信息	目标边界识别能力强	解码器部分相对简单,特征重建能力可能不足
DeepLabv3+	通过强大的编码器-解码器结构融合多尺度上下文信息,并精准恢复物体边界	多尺度处理能力强,边界分割精度高	模型参数量大,计算复杂度相对较高,对硬件要求高

图像中的高维特征,随后通过上采样操作将高维特征图恢复到输入图像尺寸,最后采用最大投票策略获得输出图像中每个像素的预测标签,通过该方法实现了对 SEM 图像中马氏体、回火马氏体、贝氏体以及珠光体的准确识别。与之相似,DeCost 等^[57]应用 PixelNet 模型准确识别了超高碳钢中各类组成相,Zhang 等^[63]基于 CNN 模型架构识别了不同类型钛合金的微观组织。除了常见的相识别之外,该类语义分割算法还被应用于晶界识别^[64]、重结晶组织分析^[65]、碳化物识别^[59]、增材制造样品中冶金缺陷识别^[66]、TEM 图像中的物相识别^[58, 67]等多个场景。基于模型输出的预测结果图,通过对各类目标的像素信息进行统计计算,便可获得丰富的显微组织定量信息。例如,DeCost 等^[57]基于超高碳钢组织的识别结果,计算了各类相的晶粒尺寸;Zhao 等^[66]利用 TransUNet 模型分割了增材制造马氏体钢金相图中的冶金缺陷,并计算了各类缺陷的尺寸、形貌、分布等定量信息。

相较于目标检测算法,图像语义分割算法可获得像素级的精确位置信息和轮廓形貌信息,具有更强的边缘与细节处理能力,因此可获得丰富的微观组织信息。然而,由于语义分割算法可获得像素级的识别精度,使其相较于目标检测算法需要消耗更高的计算资源;目前主流的语义分割算法均属于监督学习,这意味着需要对每张训练图像进行像素级标注,极高的标注成本引起了严重的小样本问题;此外,复杂微观组织中晶界、相界模糊不清、难以辨别,对语义分割算法的像素级精细识别提出了巨大挑战,使得当前该方法对复杂微观组织的识别效果并不理想。针对组织复杂度高和标注样本匮乏 2 大问题,学者在后续研究中提出了不同应对策略,下文对其进行详细介绍。

5.1 语义分割算法在复杂组织图像分析方面的应用

以钢铁为代表的大多数金属材料均由复相组织构成,不同相之间形貌差别并不明显,组织复杂度高,通过常规的 OM、SEM 等表征技术无法清晰准确地区分组织中各个相,致使难以建立起高质量数据集,制约了语义分割算法在复杂微观组织分析的应用效果。

目前,学界对于复杂显微组织识别的研究多关注于如何结合高精度表征手段创建准确可靠的标签图像。针对这个问题,Shen 等^[68]在前期工作中提出了 EBSD 指导下的 DL 建模方法,并将其成功应用于具有复杂显微组织的实际工程钢种。如图 6 所示^[68-69],该方法的核心思路在于通过 SEM-EBSD 原位表征实验,将 EBSD 实验产生的相图用作原位 SEM 图像的标签,实现了对复杂微观组织的像素级精准标定,进而建立起高质量数据集。该方法在奥氏体不锈钢(发生部分马氏体相变)和淬火 & 配分(quenching and partitioning, Q&P)钢的组织识别中得到了应用,并获得了与 EBSD 实验基本一致的组织量化结果。该方法创新性地采用 EBSD 分析精准标定复杂钢铁组织中各类相,攻克了复杂组织智能识别中精准标签获取难的瓶颈。基于获得的高质量数据集,便可训练得到适用于复杂微观组织的图像识别模型,为高复杂度微观组织的智能识别提供了普适性的建模框架。同样基于 EBSD 信息的辅助,Durmaz 等综合利用相图、KAM(kernel average misorientation)图和 GOS (grain orientation spread)图提供的位错密度、晶体取向、晶界衬度等信息,精确标定了低碳钢中板条贝氏体,并通过训练 U-Net 模型实现了对复相组织中复杂板条贝氏体的准确识别^[70];Frank 等通过将 EBSD 与 OM、SEM 表征信息进行联用,实现了对各类马氏体、贝氏体等复杂复相组织的准确识别^[34, 71-73]。标注

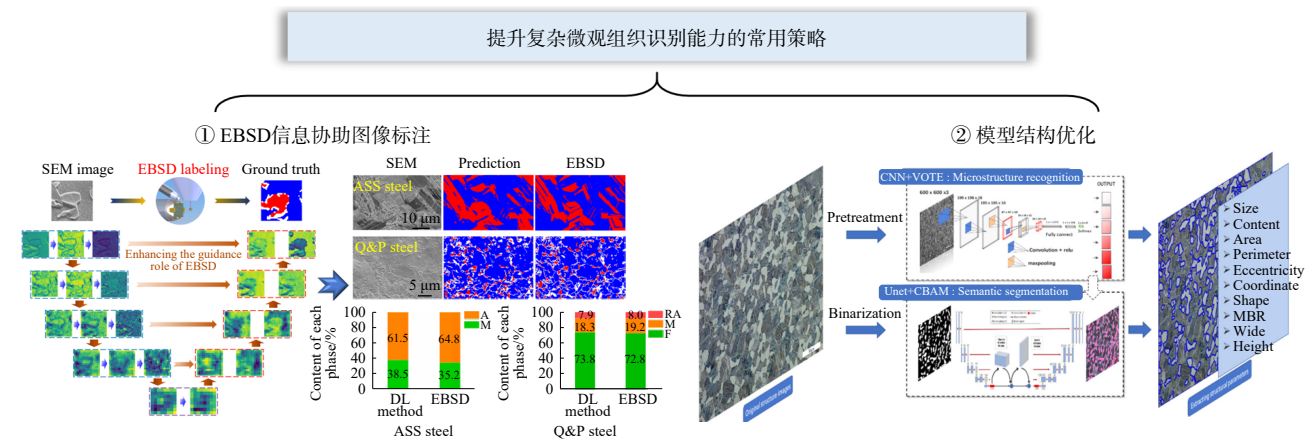


图6 提升复杂显微组织识别能力的常用策略

质量直接决定了语义分割模型的上限,通过运用材料领域的表征测试技术精细分类训练图像中复杂显微组织,可为模型训练提供更为可靠的监督信号和类别边界信息,有利于算法学习到图像像素特征与组织类别之间的精准映射关系,进而提升了语义分割模型对复杂显微组织的识别能力。然而,该类方法需要依赖较为复杂的原位表征实验,同时标签数据的获取成本也较高,使其更适用于构建针对特定钢种体系的小数据集,难以扩展应用至建立复杂、多材料体系的大规模数据集。

除了提升标签图像精度,改进语义分割模型计算框架也是提升其对复杂组织识别能力的有效途径(图6)^[68-69]。目前,结合图像分类模型与语义分割模型是常用手段之一。Zhao等^[69]通过结合图像分类模型与分割模型准确识别与定量分析了钛合金显微组织。该计算框架中首先采用CNN+VOTE模型对钛合金组织图像进行分类,随后基于分类结果,采用自适应阈值二值化算法对图像进行处理,得到仅保留目标组织的二值化图像,降低其他背景相对语义分割建模的影响,最后采用U-Net算法对目标结构进行分割,并且在U-Net算法中还嵌入了卷积块注意力模块(convolutional block attention module, CBAM)注意力机制,通过对通道和空间的双维度注意力聚焦,精准定位目标组织特征,提升分割精度。Ackermann等^[74]也通过2阶段DL框架可靠识别了贝氏体钢中复杂的马氏体-奥氏体混合组织(马奥岛, M-A island),该计算框架中首先应用分类模型将马奥岛根据形貌特征进行分类,随后针对不同形貌的马奥岛建立各自的语

义分割模型。上述方法的创新思路为,通过预先建立的分类模型对复杂组织识别任务进行分解,从而降低复杂显微组织的识别难度。除了联用不同算法,对已有算法的网络结构进行改进设计也可以提升模型对复杂组织的识别效果。例如,Zhou等^[75]针对高温合金中晶界析出相干扰晶界识别的问题,对传统U-Net网络结构进行优化,设计了单编码器+双解码器的网络结构,2个解码器分别用于分割晶界和析出相,同时结合图像后处理技术,实现了对高温合金中复杂晶界的高精度识别。Li等^[76]针对铝合金中Si颗粒与晶界的识别问题,通过改进RCF(richer convolutional features)网络结构提出了多任务学习模型,其中低尺度特征用于分割简单的Si颗粒,高尺度特征则用于分割更为复杂的晶界,同时结合生成对抗网络(generative adversarial network, GAN)与图像后处理技术,实现对上述复杂组织的准确分割。

综上所述,针对复杂显微组织识别难题,历过多年发展,材料领域学者借助专业知识,在高质量数据集构建、模型结构设计等方面均取得了创新进展,提升了语义分割算法对工程金属材料显微组织的识别精度。相较于模型结构设计,材料领域学者更熟悉和适用高质量数据集构建策略,可以充分发挥材料领域知识来提升复杂组织数据的标注精度,但受限于实验效率和成本,该类方法多适用于特定材料体系下的小数据集构建。而对于大规模、多材料体系的复杂组织识别任务,模型架构设计则是更优选择。

5.2 语义分割算法在小样本组织图像分析方面的应用

语义分割算法在材料领域应用还面临着严峻的

小样本问题, 该问题主要来源于极其耗时的人工标注。尽管通过组织表征可以快速收集大量的组织图像数据, 但由于图像中通常含有较高比例的相界和晶界, 使得单张图像的标注工作量极大。标注数据的匮乏将严重影响模型训练过程中特征学习的多样性以

及评价指标的稳定性, 最终降低了模型的精度和泛化能力。

如图 7 所示^[14, 77-81], 针对小样本问题, 运用计算模拟手段合成大量仿真的图像数据、扩充数据集, 是最为直接且有效的手段。目前, 最为基础且有效的数据

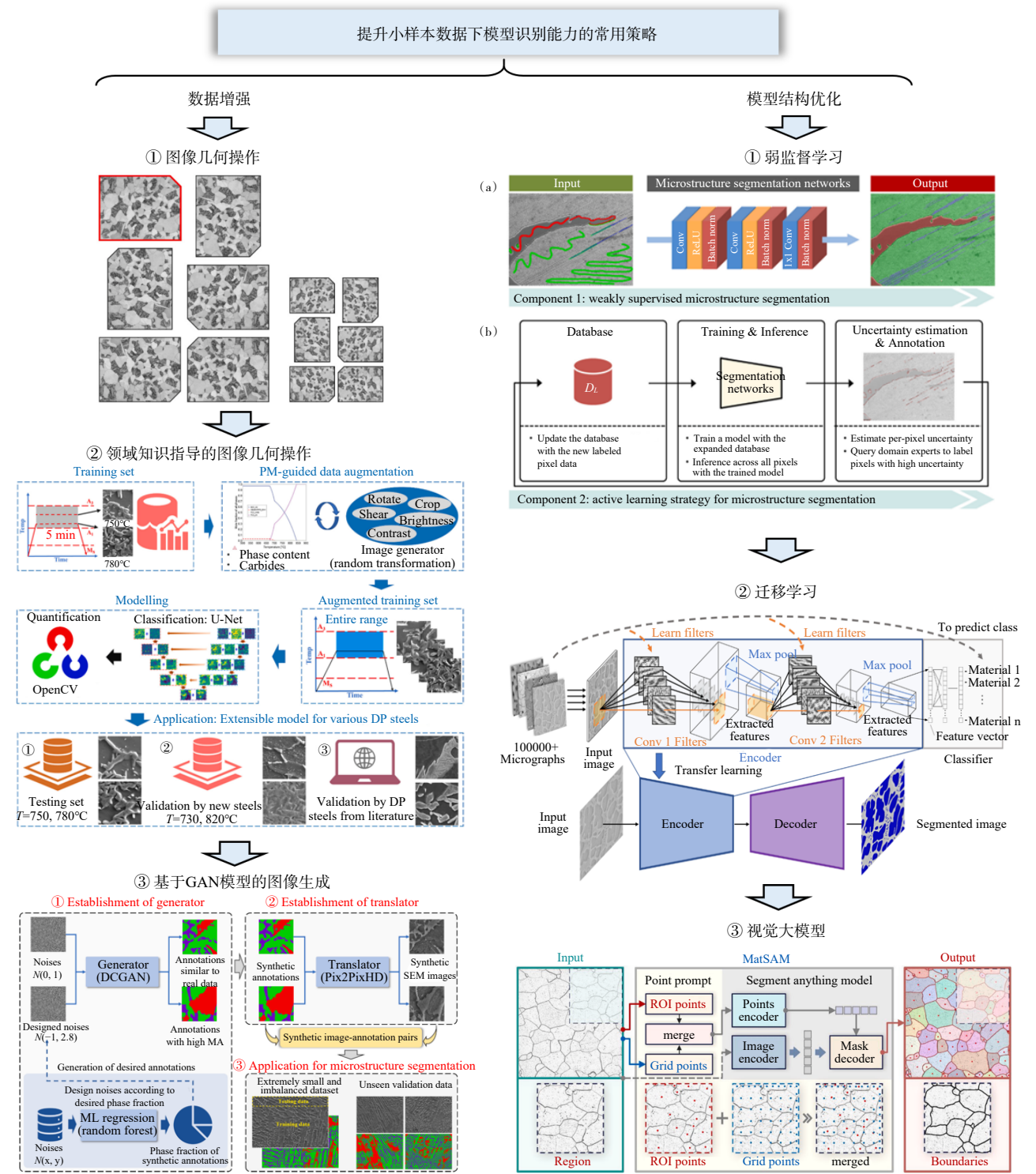


图 7 提升小样本数据下模型微组织识别能力的常用策略

增强方法为图像几何操作,该方法通过将原始训练图像及其标签进行同步操作(水平翻转、镜像翻转等),进而生成“新图像数据”,快速扩充数据集。例如,在 Ostormujof 等^[81]的工作中,将训练图像数据沿着不同坐标轴进行 90°~270°的随机翻转,形成新样本,大幅扩充了数据集;Shen 等^[68]也采用了图像翻转、随机裁剪的方式进行了训练数据增强。该方法由于操作简单已经得到了极为普遍的应用,几乎成为了语义分割数据集构建的默认操作^[56,58]。在常规图像几何操作的基础上,将其与材料领域的物理冶金原理相结合,还可以实现不同条件显微组织的模拟生成。Shen 等^[78]提出了热力学指导下的组织图像数据增强方法。首先,通过热力学计算确定目标工艺参数下双相钢的关键显微组织信息,例如相含量分布特征、碳化物析出情况等,随后通过随机图像处理(翻转、随机裁剪等)寻找与计算组织信息相匹配的图像数据,进而获得了目标工艺下的显微组织图像。利用该方法,基于 750℃、780℃ 退火温度下组织数据,成功模拟生成了 730℃ 和 820℃ 下的组织图像,有效扩充了数据集,所训练的模型实现了对该双相钢整个退火温度区间内马氏体和铁素体的精准识别。尽管基于图像几何操作的数据增强方法可以快速扩充训练样本数据,但所谓的“生成图像”均来自原有有限的训练样本,本质上并未增加新的数据信息,仍会制约模型的鲁棒性和扩展应用能力。为生成全新的图像数据,Sandfeld 等^[82]提出了一种用于模拟 TEM 图像中位错结构的参数化图像生成模型。首先,该方法选择无位错的基体区域作为生成图像的背景,随后利用样条状曲线模拟不同形貌的位错结构,最后将人工模拟位错绘制于背景图像上,生成对应的标签图像,进而获得全新的 TEM 图像及其标签。Sandfeld 等^[83]还利用该模型生成了包含裂纹与晶界的组织图像数据,大幅地丰富了数据集并提升了模型训练效果。然而,该类方法仅适用于生成位错线、晶界等简单的线型微观组织,难以用于生成形貌复杂的多相组织,普适性较为有限。

为实现复杂组织的可靠模拟,近年来,以 GAN 为代表的生成式 AI 算法成为一种有力工具^[84-86]。Han 等^[87]应用 HP-VAE-GAN 模型生成了高仿真的超高碳钢显微组织,且通过结合真实数据与模拟数据提升了组织图像分类模型的精度,验证了生成组织图像数

据的可靠性。Cao 等^[88]利用 GAN 模型生成了不同工艺参数下的增材制造钛合金组织,生成图像的定量组织信息与真实组织图像基本一致。上述工作充分证实了 GAN 模型在复杂显微组织生成方面的独特优势。然而,现有的 GAN 模型只能生成组织图像数据,并无法生成组织图像所对应的标签图像,仍然无法解决语义分割任务中的小样本问题。为此,Shen 等^[14]提出了一种可以同时生成组织图像及其标签图的两步 GAN 模型架构。该架构由生成模型和翻译模型组成,其中生成模型采用深度卷积生成对抗网络(deep convolution generative adversarial network, DCGAN)算法,可通过输入随机噪声生成彩色标签图像,翻译模型则采用条件生成式 Pix2Pix 算法,可通过输入标签图像生成对应的显微组织图像,最终联用 2 个 GAN 模型,便可实现组织图像及其标签图像的同步生成。该方法被成功应用于小样本数据下的增材制造钢铁显微组织识别任务,并且证实了生成数据对语义分割模型精度的提升作用。与之相似,Ma 等^[89]将蒙特卡罗模拟与 Pix2Pix 图像风格迁移 GAN 模型相结合,也实现了纯铁多晶组织及其标签的生成,并通过结合真实数据与生成数据也进一步提升了晶界识别的精度。综上所述,基于图像几何变换操作的图像增强方法本质上并无法引入新信息,对模型泛化能力的提升效果有限,但是优点在于操作简单,实践门槛较低。相较之下,基于 GAN 的图像增强方法可生成全新数据,更大地丰富原始数据集样本信息,提升模型对域外数据的鲁棒性,并且将其与材料领域计算模拟方法相结合还可生成不同条件下“图像-标签”对,是未来极有潜力的研究方向,但实践门槛相对较高。

除了数据增强外,算法开发或模型架构改进也是提升小样本数据下模型能力的有效途径,如图 7 所示^[14,77-81]。基于涂鸦标注的弱监督模型可大幅降低语义分割模型对数据量的需求^[90-91],Na 等^[80]提出了一种结合弱监督语义分割算法与主动学习策略的高普适性显微组织分割模型,其中,弱监督算法仅以领域专家的涂鸦标注为监督信息进行训练,大幅降低标注工作量;主动学习过程中基于边际采样评估像素不确定性,进而引导领域专家对关键区域进行标注,随后更新数据库并重新训练弱监督模型,通过循环往复操作,使用最小的标注成本获得了最优的模型性能。与

之相似, Gupta 等^[92]应用基于涂鸦标注的弱监督模型实现了对不同工艺条件下碳钢组织的准确识别; 迁移学习是材料领域面临小样本问题时的常用建模策略, 已在性能预测与合金设计方面获得应用^[93-94], 鉴于此, 学者也将该策略应用于小样本组织识别与定量化。Stuckner 等^[77]应用迁移学习策略构建了适用于镍基合金等多类材料的微观组织识别模型。该模型首先在含有十余万张图像数据的 MicroNet 数据集上预训练编码器, 随后将其迁移于 U-Net 等语义分割模型, 仅使用少量训练数据微调模型参数, 便可获得高精度模型。Alrfou 等^[95]同样利用迁移学习策略, 结合 CNN 和 Swin Transformer 预训练权重, 实现了组织图像的特征提取及其精准分割。最近, 语义分割技术正在快速向视觉大模型的方向转变, 例如 SAM(segment anything model)等基础分割模型的兴起^[96], 极大缓解了语义分割模型对大量精细标注数据的严重依赖。Ma 等^[97]通过结合 SAM 大模型与领域知识, 在无需额外模型训练的条件下, 建立起了具有优异泛化能力的合金组织分割模型, 并获得了与传统监督模型极为接近的识别精度。Li 等^[79]基于 SAM 视觉大模型开发出 MatSAM 模型, 同样无需额外训练与标注, 便可处理多晶、多相、缺陷等多种微观结构的识别任务, 并适用于 OM、SEM 等多种成像方式, 证实了视觉大模型技术的有效性。综上所述, 弱监督学习、迁移学习以及视觉大模型均可提升 AI 技术对小样本组织图像数据的识别精度。弱监督学习虽然可以通过涂鸦标注等方法降低标注成本, 但该类模型的性能上限通常较为有限。迁移学习凭借预训练模型可快速适应新任务, 但模型性能对源域和目标域的数据分布差异较为敏感, 预训练模型选择不当易导致模型性能不佳。相较之下, 视觉大模型凭借极大规模数据的预训练使其获得了强大的泛化能力, 显著提升了小样本下组织识别的精度, 是语义分割领域未来的重点发展方向之一, 但较于前 2 种方法, 其计算资源需求和部署门槛更高。

6 总结与展望

近年来, 基于数据驱动的 AI 技术在金属材料微观组织识别与定量化研究方面取得了令人瞩目的进

展。随着图像分类、目标检测、语义分割等 AI 算法的不断应用, 促使组织图像数据的智能分析完成了由浅层定性分类向深度定量统计的转变。然而, 金属材料领域中标注数据的稀缺性以及组织图像的高复杂性长期制约 AI 技术的应用效果。针对小样本问题, 学者在样本生成、模型架构改进等方面做出了多种创新性工作, 且随着 AI 算法的快速发展, 相关研究工作的重心也逐渐由前期的数据增强向新兴的迁移学习、视觉大模型转变, 尤其是最近发展的视觉大模型, 颠覆了传统监督学习算法的建模范式, 极大程度降低了建模对数据的依赖性。针对组织图像复杂度高的问题, 学者从高质量数据集构建、模型改进等方向着手提出应对策略, 并且从应用效果来讲, 高质量数据集提供的精准标签更易于训练高鲁棒性 AI 模型, 同时更有利于发挥材料领域学者的专业优势, 是解决该类问题的首选策略。

从图像智能分析技术发展而言, 材料图像数据智能分析在未来发展中应紧跟 AI 领域前沿技术, 将最近的算法、理念应用于本领域, 实现事半功倍的效果。目前, SAM 等语义分割大模型的应用极大缓解了标注数据稀缺的难题, 有望替代传统监督学习算法。相对于大模型, 零样本学习与少样本学习算法也是克服材料领域小样本问题的有效技术, 然而现有研究针对零/少样本学习在材料组织分析领域的探索尚显不足^[98]。此外, 在图像识别算法结构中融入物理冶金学知识, 也是未来极具前景的发展方向, 例如本文中列举的 EBSD 指导下的 DL 建模方法以及热力学指导的图像数据增强就是一种初步尝试。除了上述结合, 材料领域知识还可以嵌入或指导损失函数设计、模型架构改进等方面, 形成物理机制约束。例如, 损失函数方面, 设计过程中可考虑加入物理机制惩罚项, 通过引入形态学与拓扑学约束, 提升 AI 模型对晶粒边界等特定位置识别的精准度与连续性; 模型架构方面, 可以考虑在网络结构中引入先进的注意力机制, 通过添加通道注意力或空间注意力模块, 使得模型更为关注目标区域, 进而提升模型性能。将领域知识与 AI 模型相结合可有力提升 AI 模型的性能和合理性, 发展出材料科学领域特有的智能组织识别与定量化体系。

从指导新材料研发的角度而言, 材料图像数据智

能分析在未来发展中应进一步加强与现有材料集成计算设计体系的结合。目前,基于AI的微观组织分析模型多被视作游离于集成计算设计框架以外的一套单独计算体系,对新材料研发或理论机制深化的贡献多局限于方法论层面,尚未广泛应用于指导合金设计。鉴于此,微观组织智能分析在后续发展可关注于如何将该计算工具合理地嵌入材料研发体系,切实指导新材料开发。例如,目前基于生成式AI模型的合金设计工作多涉及生成微观组织图像数据^[99-100],可考虑将图像识别与量化模型嵌入该框架,用以提取合成图像数据中的定量组织信息,为建立物理冶金信息指导的ML预测与设计模型提供基础。通过将图像智能分析模型与现有的集成计算设计框架进行结合,实现金属材料领域多模态数据的深度挖掘,进而更好地指导原型合金设计。此外,在后续原型合金向工程构件转化过程中,可考虑应用图像识别技术对不同尺寸(原型合金、中试级、工业级)样品中微观组织进行精细分析,定量不同尺寸下组织信息差异,并将其与工艺参数信息、力学性能进行关联,揭示微观组织的尺寸效应规律并加速工程转化效率。

参考文献 (References)

- [1] 董瀚,王毛球,翁宇庆.高性能钢的M³组织调控理论与技术[J]. *钢铁*, 2010, 45(7): 1-7.
- [2] 董瀚,廉心桐,胡春东,等.钢的高性能化理论与技术进展[J]. *金属学报*, 2020, 56(4): 558-582.
- [3] Samuels L E. Light Microscopy of carbon steels[M]. Materials Park, Ohio: ASM International, 1999.
- [4] Humphreys F J. Characterisation of fine-scale microstructures by electron backscatter diffraction (EBSD)[J]. *Scripta Materialia*, 2004, 51(8): 771-776.
- [5] Long F, Kerr D, Domizzi G, et al. Microstructure characterization of a hydride blister in Zircaloy-4 by EBSD and TEM[J]. *Acta Materialia*, 2017, 129: 450-461.
- [6] Winkelmann A, Cios G, Tokarski T, et al. EBSD orientation analysis based on experimental Kikuchi reference patterns[J]. *Acta Materialia*, 2020, 188: 376-385.
- [7] 吴文峻,廖星创,赵金琨. DeepSeek技术创新与通用人工智能发展趋势[J]. *科技导报*, 2025, 43(6): 14-20.
- [8] 宿彦京,付华栋,白洋,等.中国材料基因工程研究进展[J]. *金属学报*, 2020, 56(10): 1313-1323.
- [9] 王晨充,徐伟.综述:合金设计中物理模型与人工智能的集成与发展[J]. *金属学报*, 2025, 61(4): 541-560.
- [10] 王国栋.生成式AI(AIGC)助力钢铁行业转型升级[J]. *科技导报*, 2025, 43(13): 1-2.
- [11] 谢建新,宿彦京,薛德祯,等.机器学习在材料研发中的应用[J]. *金属学报*, 2021, 57(11): 1343-1361.
- [12] DeCost B L, Francis T, Holm E A. Exploring the microstructure manifold: Image texture representations applied to ultra-high carbon steel microstructures[J]. *Acta Materialia*, 2017, 133: 30-40.
- [13] 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.
- [14] Shen C G, Zhao J X, Huang M H, et al. Generation of micrograph-annotation pairs for steel microstructure recognition using the hybrid deep generative model in the case of an extremely small and imbalanced dataset[J]. *Materials Characterization*, 2024, 217: 114407.
- [15] Shen M R, Li G Z, Wu D X, et al. Multi defect detection and analysis of electron microscopy images with deep learning[J]. *Computational Materials Science*, 2021, 199: 110576.
- [16] Che L, He Z P, Zheng K Y, et al. Deep learning in alloy material microstructures: Application and prospects[J]. *Materials Today Communications*, 2023, 37: 107531.
- [17] 沈园竣,马飞跃,谭廖麟.人工智能在金相分析中的发展现状与展望[C]//第十二届中国金属学会青年学术年会暨首届“碳中和”冶金青年科学家沙龙论文集(二).赣州:中国金属学会, 2024: 334.
- [18] 尹万健,刘森,龚新高.人工智能驱动下的材料科学范式变革:方法、平台与应用前沿[J]. *科学通报*, 2025, 70(24): 4012-4014.
- [19] Hwang J H, Trang T T T, Lee O, et al. Improvement of strength-ductility balance of B2-strengthened lightweight steel[J]. *Acta Materialia*, 2020, 191: 1-12.
- [20] Song H, Yoo J, Kim S H, et al. Novel ultra-high-strength Cu-containing medium-Mn duplex lightweight steels[J]. *Acta Materialia*, 2017, 135: 215-225.
- [21] Zhi H H, Li J S, Li W M, et al. Simultaneously enhancing strength-ductility synergy and strain hardenability via Si-alloying in medium-Al FeMnAlC lightweight steels[J]. *Acta Materialia*, 2023, 245: 118611.
- [22] Dong G Q, Wang J L, Wang L Y, et al. Transitions in the thermal stability of pre-strained austenite-competing effects between defect density and slip band spacing[J]. *Scripta Materialia*, 2023, 223: 115077.
- [23] Liu L, He J Z, Li L J, et al. Making ultrahigh-strength dual-phase steels tough: Experiment and simulation[J]. *Jour-*

- nal of Materials Science & Technology, 2025, 226: 302–316.
- [24] Venables J A, Harland C J. Electron back-scattering patterns: A new technique for obtaining crystallographic information in the scanning electron microscope[J]. *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics*, 1973, 27(5): 1193–1200.
- [25] Asadi A M, Jafari E M. Transmission electron microscopy as best technique for characterization in nanotechnology[J]. *Synthesis and Reactivity in Inorganic, Metal–Organic, and Nano–Metal Chemistry*, 2015, 45(3): 323–326.
- [26] Yang M J, King D J M, Povstugar I, et al. Precipitation behavior in G-phase strengthened ferritic stainless steels[J]. *Acta Materialia*, 2021, 205: 116542.
- [27] Zambrano O A, Ton K, Islam A. Towards the understanding of hydrogen embrittlement in pipeline steels using EBSD[J]. *International Journal of Hydrogen Energy*, 2025, 174: 151470.
- [28] Ding L J, Goshtasby A. On the canny edge detector[J]. *Pattern Recognition*, 2001, 34(3): 721–725.
- [29] Raccuglia P, Elbert K C, Adler P D F, et al. Machine-learning-assisted materials discovery using failed experiments[J]. *Nature*, 2016, 533(7601): 73–76.
- [30] Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, et al. A comprehensive survey on support vector machine classification: Applications, challenges and trends[J]. *Neurocomputing*, 2020, 408: 189–215.
- [31] Graupe D. Principles of artificial neural networks[M]. London: World Scientific Publishing Co. Pte. Ltd., 2007.
- [32] Breiman L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5–32.
- [33] Gola J, Britz D, Staudt T, et al. Advanced microstructure classification by data mining methods[J]. *Computational Materials Science*, 2018, 148: 324–335.
- [34] Müller M, Britz D, Ulrich L, et al. Classification of bainitic structures using textural parameters and machine learning techniques[J]. *Metals*, 2020, 10(5): 630.
- [35] Gola J, Webel J, Britz D, et al. Objective microstructure classification by support vector machine (SVM) using a combination of morphological parameters and textural features for low carbon steels[J]. *Computational Materials Science*, 2019, 160: 186–196.
- [36] Santry D J. Demystifying deep learning: An introduction to the mathematics of neural networks[M]. USA: Wiley-IEEE Press, 2024.
- [37] 山世光, 阚美娜, 刘昕, 等. 深度学习: 多层神经网络的复兴与变革[J]. *科技导报*, 2016, 34(14): 60–70.
- [38] Zhang J, Cao Y Y, Wu Q. Vector of locally and adaptively aggregated descriptors for image feature representation[J]. *Pattern Recognition*, 2021, 116: 107952.
- [39] Ling J L, Hutchinson M, Antono E, et al. Building data-driven models with microstructural images: Generalization and interpretability[J]. *Materials Discovery*, 2017, 10: 19–28.
- [40] Mulewicz B, Korpala G, Kusiak J, et al. Autonomous interpretation of the microstructure of steels and special alloys[J]. *Materials Science Forum*, 2019, 949: 24–31.
- [41] Larmuseau M, Sluydts M, Theuwissen K, et al. Race against the Machine: Can deep learning recognize microstructures as well as the trained human eye?[J]. *Scripta Materialia*, 2021, 193: 33–37.
- [42] Amjoud A B, Amrouch M. Object detection using deep learning, CNNs and vision transformers: A review[J]. *IEEE Access*, 2023, 11: 35479–35516.
- [43] Li G, Wang Y G, He B, et al. Low-light multimodal object detection: A survey[J]. *Computer Science Review*, 2025, 58: 100804.
- [44] Cai J Q, Wu Y Q. Three-dimensional object detection for autonomous driving via deep learning: A review[J]. *Engineering Applications of Artificial Intelligence*, 2025, 161: 112238.
- [45] Albuquerque C, Henriques R, Castelli M. Deep learning-based object detection algorithms in medical imaging: Systematic review[J]. *Heliyon*, 2025, 11(1): e41137.
- [46] Murat A A, Kiran M S. A comprehensive review on YOLO versions for object detection[J]. *Engineering Science and Technology, an International Journal*, 2025, 70: 102161.
- [47] Shen M R, Li G Z, Wu D X, et al. A deep learning based automatic defect analysis framework for *in situ* TEM ion irradiations[J]. *Computational Materials Science*, 2021, 197: 110560.
- [48] Anderson C M, Klein J, Rajakumar H, et al. Automated detection of helium bubbles in irradiated X-750[J]. *Ultramicroscopy*, 2020, 217: 113068.
- [49] Perera R, Guzzetti D, Agrawal V. Optimized and autonomous machine learning framework for characterizing pores, particles, grains and grain boundaries in microstructural images[J]. *Computational Materials Science*, 2021, 196: 110524.
- [50] Lynch M J, Jacobs R, Bruno G A, et al. Accelerating domain-aware electron microscopy analysis using deep learning models with synthetic data and image-wide confidence scoring[J]. *NPJ Computational Materials*, 2025, 11: 261.
- [51] Wu M, Sharapov J, Anderson M, et al. Quantifying disloca-

- tion-type defects in post irradiation examination *via* transfer learning[J]. *Scientific Reports*, 2025, 15: 15889.
- [52] Li W, Field K G, Morgan D. Automated defect analysis in electron microscopic images[J]. *NPJ Computational Materials*, 2018, 4: 36.
- [53] Jacobs R. Deep learning object detection in materials science: Current state and future directions[J]. *Computational Materials Science*, 2022, 211: 111527.
- [54] Garcia-Garcia A, Orts-Escolano S, Oprea S, et al. A survey on deep learning techniques for image and video semantic segmentation[J]. *Applied Soft Computing*, 2018, 70: 41–65.
- [55] Yu H S, Yang Z G, Tan L, et al. Methods and datasets on semantic segmentation: A review[J]. *Neurocomputing*, 2018, 304: 82–103.
- [56] Azimi S M, Britz D, Engstler M, et al. Advanced steel microstructural classification by deep learning methods[J]. *Scientific Reports*, 2018, 8: 2128.
- [57] 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.
- [58] Groschner C K, Choi C, Scott M C. Machine learning pipeline for segmentation and defect identification from high-resolution transmission electron microscopy data[J]. *Microscopy and Microanalysis*, 2021, 27(3): 549–556.
- [59] Yu S P, Zhang S S, Shu C Q, et al. Automatic assessment of band carbide in bearing steels based on machine learning[J]. *Engineering Research Express*, 2025, 7(1): 015261.
- [60] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2015: 3431–3440.
- [61] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation[C]//Proceedings of Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Berlin, Germany: Springer International Publishing, 2015: 234–241.
- [62] Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481–2495.
- [63] Zhang J H, Yu J T, Wei X R, et al. A soft scanning electron microscopy for efficient segmentation of alloy microstructures based on a new self-supervised pre-training deep learning network[J]. *Materials Characterization*, 2024, 218: 114532.
- [64] Liang R Z, Tang L, Wang N, et al. A deep learning fusion model for *in situ* grain size analysis of nickel-based superalloys[J]. *Materials Characterization*, 2025, 229: 115518.
- [65] 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.
- [66] Zhao J X, Shen C G, Huang M H, et al. Deep learning accelerated micrograph-based porosity defect quantification in additively manufactured steels for uncovering a generic process-defect-properties relation[J]. *Materials Characterization*, 2025, 225: 115094.
- [67] Horwath J P, Zakharov D N, Mégret R, et al. Understanding important features of deep learning models for segmentation of high-resolution transmission electron microscopy images[J]. *NPJ Computational Materials*, 2020, 6: 108.
- [68] Shen C G, Wang C C, Huang M H, et al. A generic high-throughput microstructure classification and quantification method for regular SEM images of complex steel microstructures combining EBSD labeling and deep learning[J]. *Journal of Materials Science & Technology*, 2021, 93: 191–204.
- [69] Zhao P L, Wang Y W, Jiang B Y, et al. A new method for classifying and segmenting material microstructure based on machine learning[J]. *Materials & Design*, 2023, 227: 111775.
- [70] Durmaz A R, Müller M, Lei B, et al. A deep learning approach for complex microstructure inference[J]. *Nature Communications*, 2021, 12: 6272.
- [71] 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.
- [72] Bachmann B I, Müller M, Stiefel M, et al. Efficient phase segmentation of light-optical microscopy images of highly complex microstructures using a correlative approach in combination with deep learning techniques[J]. *Metals*, 2024, 14(9): 1051.
- [73] Bachmann B I, Müller M, Britz D, et al. Reproducible quantification of the microstructure of complex quenched and quenched and tempered steels using modern methods of machine learning[J]. *Metals*, 2023, 13(8): 1395.
- [74] Ackermann M, Iren D, Wesselmecking S, et al. Automated segmentation of martensite-austenite islands in bainitic steel[J]. *Materials Characterization*, 2022, 191: 112091.
- [75] 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.
- [76] 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.
- [77] Stuckner J, Harder B, Smith T M. Microstructure segmentation with deep learning encoders pre-trained on a large microscopy dataset[J]. *NPJ Computational Materials*, 2022, 8: 200.
- [78] Shen C G, Wei X L, Wang C C, et al. A deep learning method for extensible microstructural quantification of DP steel enhanced by physical metallurgy-guided data augmentation[J]. *Materials Characterization*, 2021, 180: 111392.
- [79] Li C T, Han X, Yao C, et al. A novel training-free approach to efficiently extracting material microstructures *via* visual large model[J]. *Acta Materialia*, 2025, 290: 120962.
- [80] 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.
- [81] Ostormujof T M, Purohit R P R, Breumier S, et al. Deep Learning for automated phase segmentation in EBSD maps. A case study in Dual Phase steel microstructures[J]. *Materials Characterization*, 2022, 184: 111638.
- [82] Govind K, Oliveros D, Dlouhy A, et al. Deep learning of crystalline defects from TEM images: A solution for the problem of 'never enough training data'[J]. *Machine Learning: Science and Technology*, 2024, 5(1): 015006.
- [83] Trampert P, Rubinstein D, Boughorbel F, et al. Deep neural networks for analysis of microscopy images: Synthetic data generation and adaptive sampling[J]. *Crystals*, 2021, 11(3): 258.
- [84] Aggarwal A, Mittal M, Battineni G. Generative adversarial network: An overview of theory and applications[J]. *International Journal of Information Management Data Insights*, 2021, 1(1): 100004.
- [85] Shamsolmoali P, Zareapoor M, Granger E, et al. Image synthesis with adversarial networks: A comprehensive survey and case studies[J]. *Information Fusion*, 2021, 72: 126–146.
- [86] Öcal A, Özbakır L. Supervised deep convolutional generative adversarial networks[J]. *Neurocomputing*, 2021, 449: 389–398.
- [87] Han Y X, Liu Y H, Chen Q C. Data augmentation in material images using the improved HP-VAE-GAN[J]. *Computational Materials Science*, 2023, 226: 112250.
- [88] Cao Z H, Liu Q, Liu Q C, et al. A machine learning method to quantitatively predict alpha phase morphology in additively manufactured Ti-6Al-4V[J]. *NPJ Computational Materials*, 2023, 9: 195.
- [89] 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: 125.
- [90] Pan Z Y, Sun H C, Jiang P, et al. CC4S: Encouraging certainty and consistency in scribble-supervised semantic segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, 46(12): 8918–8935.
- [91] Lin D, Dai J F, Jia J Y, et al. ScribbleSup: Scribble-supervised convolutional networks for semantic segmentation [C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway, NJ: IEEE, 2016: 3159–3167.
- [92] Gupta A K, Ganguly S. A weakly supervised fully convolutional network and posterior conditional random field for segmentation of phases in multiclass steel microstructure[J]. *Materials Today Communications*, 2025, 48: 113309.
- [93] Wei X L, van der Zwaag S, Jia Z X, et al. On the use of transfer modeling to design new steels with excellent rotating bending fatigue resistance even in the case of very small calibration datasets[J]. *Acta Materialia*, 2022, 235: 118103.
- [94] Zhu D X, Wu H H, Hou F G, et al. A transfer learning strategy for tensile strength prediction in austenitic stainless steel across temperatures[J]. *Scripta Materialia*, 2024, 251: 116210.
- [95] Alrfou K, Zhao T, Kordijazi A. Transfer Learning for Microstructure Segmentation with CS-UNet: A Hybrid Algorithm with Transformer and CNN Encoders, 2023, arXiv: 2308.13917.
- [96] Kirillov A, Mintun E, Ravi N, et al. Segment anything[C]//*Proceedings of IEEE/CVF International Conference on Computer Vision (ICCV)*. Piscataway, NJ: IEEE, 2024: 3992–4003.
- [97] Ma X D, Zhang Y Q, Wang C C, et al. Alloy microstructure segmentation through SAM and domain knowledge without extra training[J]. *Scripta Materialia*, 2025, 260: 116581.
- [98] Tung P Y, Harrison R J. Efficient microstructure segmentation in three-dimensional imaging: Combining few-shot learning with the segment anything model[J]. *Next Materials*, 2025, 8: 100663.
- [99] Attari V, Khatamsaz D, Allaire D, et al. Towards inverse microstructure-centered materials design using generative phase-field modeling and deep variational autoencoders[J]. *Acta Materialia*, 2023, 259: 119204.
- [100] Kusampudi N, Diehl M. Inverse design of dual-phase steel microstructures using generative machine learning model and Bayesian optimization[J]. *International Journal of Plasticity*, 2023, 171: 103776.

Application of artificial intelligence in microstructure image recognition and quantification of metallic materials

SHEN Chunguang^{1,2}, SUN Shuo^{1,2}, XU Wei^{3*}, ZHENG Shijian^{1,2*}

1. State Key Laboratory of High-Performance Roll Materials and Composite Forming, Hebei University of Technology, Tianjin 300401, China
2. Tianjin Key Laboratory of Material Layered Composite and Interface Control Technology, Hebei University of Technology, Tianjin 300401, China
3. State Key Laboratory of Digital Steel, Northeastern University, Shenyang 110819, China

Abstract Artificial intelligence (AI) technology is profoundly transforming the research paradigms in the field of materials science, driving the analysis methods for material microstructures to shift from traditional human-experience-dominated approaches to data-driven intelligent recognition. AI-based microstructure recognition and quantification, characterized by high accuracy and efficiency, have significantly advanced the development of high-throughput microstructure analysis techniques. This review focuses on the emerging field of AI-assisted microstructure analysis of metallic materials. Following the development from qualitative analysis toward refined quantitative analysis of microstructures, it systematically summarizes the research progress in traditional machine learning algorithms, deep learning-based classification, object detection, and semantic segmentation algorithms for the classification, recognition, and quantification of metallic material microstructures. Particular emphasis is placed on the current state of widely adopted semantic segmentation algorithms. Meanwhile, addressing the challenges faced by semantic segmentation in this domain, such as high microstructural complexity and limited annotated samples, the innovative strategies proposed by researchers in data augmentation and model architecture improvements, along with their enhanced performance, are discussed. Finally, the existing limitations and future directions of AI-based microstructure analysis methods are summarized and outlooked.

Keywords artificial intelligence; metallic materials; microstructure; computer vision; image recognition; quantification ●



(责任编辑 魏来)