

$\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 材料 X 射线微观结构 图像的分割算法*

李欣泽^{1,2} 蔡宇佳³ 于 强²

1(中国科学院国家空间科学中心 北京 100190)

2(中国科学院大学计算机科学与技术学院 北京 100049)

3(桂林信息科技学院电子工程学院 桂林 541199)

摘 要 在微重力环境下进行材料科学实验可以消除重力对材料实验过程的干扰, 揭示材料生长过程的本征规律, 获得具有更高性能的材料. 中国空间站高温材料科学实验柜配备了 X 射线透射成像模块, 能够实现在微重力条件下对材料凝固过程的实时成像和观察. 受空间站实验条件等因素的限制, X 射线透射成像模块拍摄到的材料 X 射线图像模糊不清, 图像中的微观结构难以直接观察. 本文提出了一种专门针对 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 材料凝固过程中生成的微观结构图像而设计的分割算法 GFF-UNet++, 并从图像分割性能及材料科学两个方面对 GFF-UNet++ 算法进行全面评估. 实验结果表明, 相比 UNet, UNet++, DC-UNet, UNet3+, Pretrained-Microscopy-Models 等图像分割算法, GFF-UNet++ 在图像分割任务中的各项图像分割指标上均有明显提升, 能够更准确分割出 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 材料在生长过程中形成的微观结构. 这为材料的微观结构分割研究提供了新的思路和方法, 具有重要应用价值.

关键词 空间高温材料实验, 深度学习, X 射线图像分割, 材料微观结构

中图分类号 V45

Segmentation Algorithm of X-ray Microstructure Image of $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ Material

LI Xinze^{1,2} CAI Yujia³ YU Qiang²

1(National Space Science Center, Chinese Academy of Sciences, Beijing 100190)

2(School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049)

3(Department of Electronic Engineering, Guilin Institute of Information Technology, Guilin 541199)

Abstract Conducting materials science experiments in a microgravity environment mitigates the influence of gravity, enabling the study of intrinsic material growth mechanisms and the fabrication of ma-

* 广西重点研发计划项目 (AB22035048, AB23026105), 桂林市重大专项 (20220103-1), 广西科技基地和人才专项项目 (AD24010012), 和中国载人航天项目 (YYMT1201-EXP02) 共同资助

2024-12-13 收到原稿, 2025-05-13 收到修定稿

©The Author(s) 2025. This is an open access article under the CC-BY 4.0 License
(<https://creativecommons.org/licenses/by/4.0/>)

terials with enhanced properties. The high-temperature materials research rack aboard the Chinese Space Station is equipped with an X-ray transmission imaging module, facilitating real-time imaging and observation of material solidification processes under microgravity. However, due to the constraints of the space station's experimental conditions, the X-ray images acquired by this module often exhibit blurriness, making direct observation of microstructures challenging. To address this issue, the GFF-UNet++ image segmentation algorithm is proposed, specifically tailored for analyzing the microstructures formed during the solidification of $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ material. The algorithm's effectiveness is rigorously evaluated in terms of both image segmentation performance and its relevance to materials science applications. The experimental results demonstrate that in the image segmentation task, GFF-UNet++ outperforms established algorithms such as UNet, UNet++, DC-UNet, UNet3+ and Pretrained-Microscopy-Models, achieving notable improvements across various image segmentation metrics. Furthermore, the microstructures formed during the growth of the $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ can be segmented more accurately. This provides new ideas and methods for the study of the microstructure segmentation of materials, and has important application value.

Key words Space high-temperature material experiment, Deep learning, X-ray image segmentation, Microstructure of materials

0 引言

微重力材料科学的主要目标是在微重力环境下,通过选择合适的生长条件制备更优质的材料.当前,空间材料科学对于高温下晶体生长界面形态演化方面的研究相对不足.越来越多的研究证明,材料凝固过程中,微观结构的生长与晶体生长界面形态演化密切相关^[1].捕捉材料凝固过程中的微观结构生长行为是一种精确获取晶体生长界面形态演化过程的方式.

中国空间站上的高温材料科学实验柜配备了X射线透射成像模块,这是国际上首次在空间站中使用X射线透射成像设备进行材料科学实验.借助X射线透射成像模块,能够对晶体生长过程中形成的微观结构进行深入细致的观察和分析,从而掌握材料的形成机理,并揭示其形成机制.但是受空间站体积、功率、重量、防护要求等因素的限制,X射线透射成像模块直接生成的原位观察图像存在图像模糊退化、噪声污染严重、对比度低等问题,不能直接满足原位观察的需要.因此,开展高温材料X射线原位观察图像处理算法研究,对于进行空间站X射线原位观察实验和探究材料生长动力学至关重要.

当前,已有不少研究将深度学习方法应用于材料科学^[2-12]. Hirabayashi 等^[11]将基于神经网络的语义分割应用于三维成像获得的多晶陶瓷电子图像. Stuckner 等^[8]提出的 MicroUnet 模型通过将特征提

取到特征向量中的学习滤波器,并结合迁移学习,将滤波器应用于 UNet 分割模型,实现了更准确的微观结构分割. Yang 等^[13]提出了基于 UNet++改进版网络,加入注意力门和残差模块,用于提升分割性能,在仅使用少量训练图像的情况下,能够高精度区分单层石墨烯 (SLG) 与石墨. Liu 等^[14]提出一种新的基于 Transformer 的图像分割模型 Swin-UNet++, 用于精确识别 2.25Cr1Mo0.25V 钢断裂面上具有复杂形貌微观特征,帮助理解氢脆 (Hydrogen Embrittlement, HE) 机理,并评价合金抗 HE 性能. Shu 等^[15]提出一种利用超像素和纹理特征聚类的无监督图像分割方法,用于识别高强度钢材料的复杂微观组织.

图像处理算法的进步使得在材料微观结构上进行精确研究成为可能.然而由于条状微观结构生长过程的复杂性,利用图像分割算法实现其精确的像素级分割仍然是一项极具挑战性的任务.这种分割的准确性直接影响到材料科学的深入研究,尤其是在描述微观结构形态变化时,精确的像素级分割十分关键^[8,11,12].

本文选用 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 作为研究对象,使用高温材料科学实验柜中的 X 射线透射成像模块,获取 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 凝固过程中条状微观结构的 X 射线图像,并提出一种针对 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 材料凝固过程中产生的条状微观结构的图像分割模型,为材料科学领域的图像分析提供更加有效的工具.

1 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 材料熔化和凝固实验

中国空间站上的高温材料科学实验柜 (High-Temperature Materials Research Rack, HTMR) 支持开展金属合金、无机功能材料等多种材料的熔体生长和凝固科学实验. X 射线透射成像模块用于材料实验过程的实时成像如图 1 所示. 其中 X 射线透射成像模块包括能够发射 X 射线的 X 射线单元、样品盒、加热丝与带有相同间距沟槽的耐高温陶瓷片及 X 射线探测器. 耐高温陶瓷片由上陶瓷片和下陶瓷片构成, 下陶瓷片上放置实验材料. X 射线源发射 X 射线

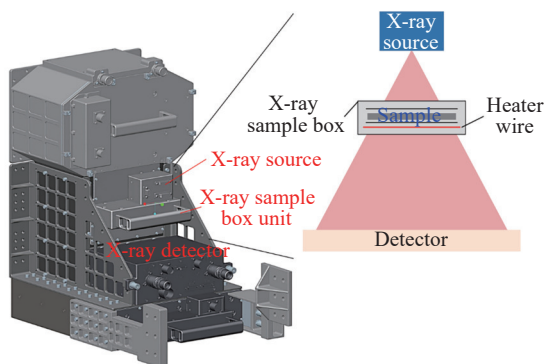


图 1 X 射线透射成像模块结构与工作原理

Fig. 1 Structure and schematic of X-ray transmission imaging module

束, 与样品盒中的物质发生相互作用. 最终透过样品的 X 射线信号被探测器接收, 并转化为图像数据.

由于空间站条件的限制, X 射线透射成像模块拍摄到的实验过程 X 射线图像模糊不清, 难以观察到实验过程中的条状微观组织结构. 因此, 很难根据实验过程 X 射线图像序列中的条状微观组织结构判断 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 所处状态, 从而影响了材料凝固过程的分析. 图 2 为 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 实验过程中的 X 射线图像序列, 由于 X 射线透射成像模块距离实验样品较远, X 射线图像整体亮度偏暗, 并且对比度较低, 缺乏明显的微观结构特征. 本研究旨在设计一种针对 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 实验过程中 X 射线图像微观结构的图像分割算法, 较好地分割出凝固过程中的条状微观结构形态变化.

2 GFF-UNet++ 网络结构

UNet 是一种基于编码器-解码器结构的深度学习图像分割网络^[16]. 与 UNet 网络相比, UNet++ 图像分割网络保留了 UNet 的编码器-解码器结构, 并引入密集连接的概念^[17]. 图 3 给出本文提出的全局特征融合 UNet++ (Global Feature Fusion UNet++, GFF-UNet++) 的网络结构, 在 UNet++ 的基础结构上增加了全局特征融合模块进行深度监督. 这里使用

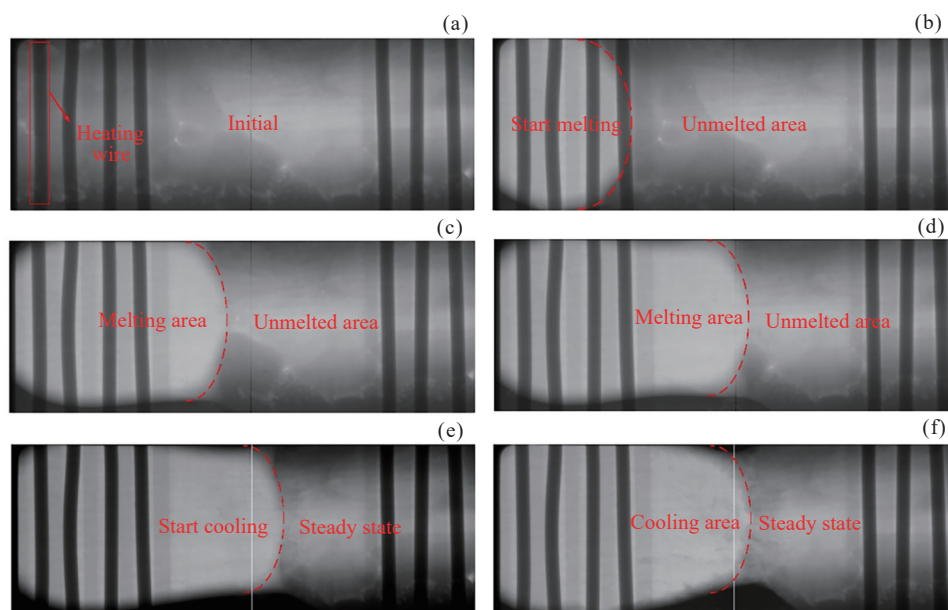


图 2 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 实验过程 X 射线图像的条状微观结构变化

Fig. 2 Strip microstructure change of X-ray images of $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ experimental process

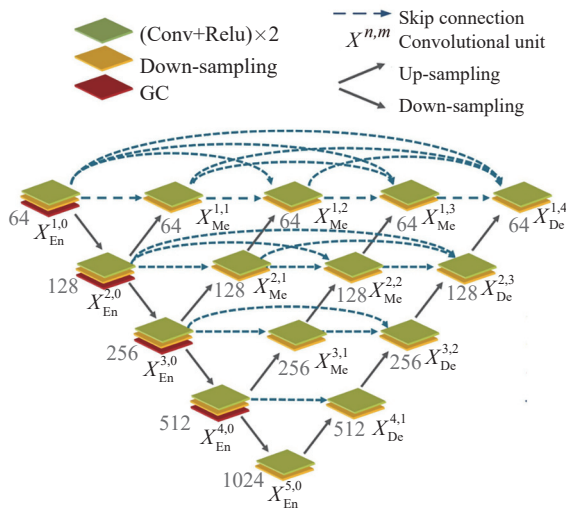


图 3 GFF-UNet++网络结构

Fig. 3 Structure diagram of the GFF-UNet++ network

批处理的方式, 将图 2 中的系列图像作为一个批次输入网络, 每幅图像均经过 Min-Max 归一化, 基于图像自身的最小值和最大值, 将像素值缩放到 $[0,1]$ 区间. 归一化后的图像作为网络输入.

网络编码器逐层进行特征提取, 并在每一层生成不同尺度的特征图, 逐步捕捉多层次的语义信息, 最终通过解码器生成最终的具有高精度的分割结果. GFF-UNet++网络结构如图 3 所示.

UNet++的编码器层在相邻卷积单元之间通过密集跳跃连接进行特征传递. 该连接方式不仅在相邻卷积单元之间建立密集连接, 还通过跨层跳跃连接融合多个前层卷积单元的特征信息, 能够同时利用浅层的细节信息和深层的高级语义信息, 从而提升分割精度. 这种连接方式虽然能够提高网络的表达能力, 但是在处理具有复杂背景和条状微观结构的 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ X 射线图像时, 仍存在不足. 为弥补其不足, GFF-UNet++网络在编码器层引入全局特征融合模块, 该模块被添加在编码器的每一个卷积单元之后, 能够帮助网络在每一层都捕捉到图像的全局信息, 提升对复杂背景和条状微观结构的感知能力.

图 3 给出了 GFF-UNet++网络的结构, 该网络包含编码器部分 X_{En} 、编码器与解码器之间的密集连接 X_{Me} 和解码器部分 X_{De} . 其中, $X^{n,m}$ 表示 GFF-UNet++第 n 层的第 m 个卷积单元. 例如, $X_{\text{En}}^{1,0}$ 表示网络第一层的第一个卷积单元, En 表示该卷积单元

属于编码器部分. 左下角的数字, 例如 64, 128, 256, 512, 1024, 代表了网络不同层的特征图通道数.

GFF-UNet++网络从编码器开始, 首先处理输入图像, 从 $X_{\text{En}}^{1,0}$ 开始, 图像经过卷积单元的卷积处理, 生成初步的分割特征图, 在每个卷积单元之后, GFF 模块被应用, 进一步增强了特征提取的能力. 通过这种方式, 网络能够更好地理解图像中的整体结构和细节, 尤其在处理具有复杂背景图像时, 能够有效提高微观结构的分割精度.

随后, 跳跃连接的方式将 $X_{\text{En}}^{1,0}$ 提取到的特征直接传递到密集连接部分的卷积单元 $X_{\text{Me}}^{1,1}$, $X_{\text{Me}}^{1,2}$, $X_{\text{Me}}^{1,3}$ 和解码器部分的卷积单元 $X_{\text{De}}^{1,4}$. 跳跃连接的方式无需经过中间的卷积层处理, 能够有效利用浅层特征, 帮助网络在训练过程中更好地捕获细节.

图像经过下采样 (最大池化) 操作, 逐步减少空间分辨率. 特征图通过每一层的卷积单元处理, 传递至下层. 在第二层时, 处理过程会继续进行, 特征图从 $X_{\text{En}}^{1,0}$ 传递至 $X_{\text{En}}^{2,0}$, 每一层会逐步减少空间分辨率, 并增加通道数. 在密集连接部分, 每个卷积单元不仅接受来自前一层的上采样操作输出, 还会接收到跳跃连接来自前几层的特征. 这些输入特征不仅有来自上一层的输出, 还包括从之前卷积单元通过跳跃连接传递过来的输出. 每个卷积单元的输出会继续传递给下一层的卷积单元, 同时也与该层前几个卷积单元的输出进行结合, 形成密集连接. 在解码器部分, 特征图通过反卷积操作逐步恢复空间分辨率. 每一层的解码器单元会从编码器的相应层以及中间密集连接部分接收到特征图进行融合, 从而恢复图像的细节.

区别于传统的图像分割算法, GFF-UNet++不依赖于固定网格或大小区域的划分, 通过网络模型逐步提取和融合特征, 进行端到端的像素级预测, 避免了固定网格尺寸的限制. $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ X 射线图像数据集每张图片的大小为 719 pixel $\times$ 588 pixel. $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ X 射线图像的实际物理大小为 12.83 mm $\times$ 5.24 mm, 原始分辨率下每个像素分别对应约 17.85 μm $\times$ 8.91 μm 的物理尺寸. 为了适配深度学习模型的输入要求, 并提高模型计算效率, 这里将所有图像调整为 512 pixel $\times$ 512 pixel 的统一大小.

在 UNet++基础网络的前四层卷积单元后加入 GFF 模块, 旨在优化编码器层下采样中相邻卷积单元特征融合的方式. 其中“全局”指代输入的

$\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ X 射线图像的整体信息, 通过扩大感受野捕捉图像全局特征, 感受野是指网络中某一层的神元能够接收到的输入图像区域范围, 感受野越大, 网络就能整合更多特征, 从而避免仅依赖局部区域的特征表示. GFF 模块使用全局平均池化操作, 对图像的整体特征进行提取; 特征融合指通过加权融合不同层或不同阶段的特征图与图像整体特征, 生成新的特征图, 强调特征融合和增强; 模块表示执行这一全局特征融合操作的独立功能单元. 具体而言, GFF 模块通过全局平均池化提取全局特征, 通过全连接层生成图像的全局特征权重, 然后使用 Sigmoid 激活函数生成加权系数, 通过对图像特征的融合和加权来增强模型的特征感知能力. 计算表达式为

$$z_1 = \mathbf{W}^{(1)} \odot \mathbf{x}^{i,0} + \mathbf{b}^{(1)}, \quad \mathbf{x}^{i,0} \in \mathbb{R}^{H \times W \times C}. \quad (1)$$

式中, z_1 表示第一层卷积的输出, $\mathbf{W}^{(1)}$ 和 $\mathbf{b}^{(1)}$ 分别为第一次卷积操作的卷积核权重和偏置项,

$$\mathbf{x}^{i',0} = f\{\sigma_2[\mathbf{W}^{(1)} \odot \sigma_1(z_1) + \mathbf{b}^{(2)}]\}, \quad \mathbf{x}^{i',0} \in \mathbb{R}^{H \times W \times C}. \quad (2)$$

其中, $\mathbf{W}^{(1)}$ 和 $\mathbf{b}^{(2)}$ 是第二次卷积操作的卷积核权重和偏置项; σ_1 和 σ_2 分别表示第一和第二个卷积操作对应的激活函数操作; f 表示下采样操作.

GFF 模块工作原理如图 4 所示, $\mathbf{x}^{i,0}$ 图像经过初步的特征提取, 包含一定的局部信息. 但是由于还没有经过 GFF 模块的全局特征增强, 尚未捕捉到更大范围的图像特征. GFF 先对 $\mathbf{x}^{i,0}$ 进行平均池化操作, 对特征图的每个区域进行操作, 计算该区域内所有像素值的均值, 以此来降低特征图的空间维度, 并保留区域的平均特征信息. 本文将特征图的空间维度压缩为单个全局向量 $\mathbf{v}_c$, 其表达式为

$$\mathbf{v}_c = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W \mathbf{x}^{i,0}(h, w, c), \quad \mathbf{v}_c \in \mathbb{R}^C. \quad (3)$$

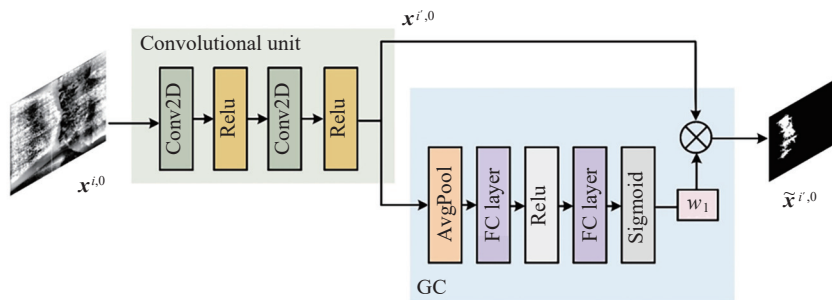


图 4 GFF 模块工作原理

Fig. 4 Schematic diagram of the GFF module

其中, $H \times W$ 是特征图的空间尺寸; h 和 w 遍历特征图的所有空间位置. 将池化得到的全局特征向量经过两个全连接层处理, 并经过激活函数 σ_3 生成一个权重 $\mathbf{w}_1$, 其可以表示为

$$\mathbf{w}_1 = \sigma_3(C_2\{\text{ReLU}[C_1(\mathbf{v}_c)]\}). \quad (4)$$

最后, 使用权重对输入特征图 $\mathbf{x}^{i',0}$ 进行重新标定得到 $\tilde{\mathbf{x}}^{i',0}$, 有

$$\tilde{\mathbf{x}}^{i',0} = \mathbf{w}_1 \odot \mathbf{x}^{i',0}, \quad \tilde{\mathbf{x}}^{i',0} \in \mathbb{R}^{H \times W \times C}. \quad (5)$$

GFF 模块通过全局特征的加权来调整每个特征图通道的权重, 增强对目标区域的响应能力, 并有效抑制背景噪声的影响. 全局信息的引入能够有效减少局部特征相似性导致的误分割, 提高分割的可靠性.

图 3 给出的 GFF-UNet++ 编码器结构突出了 GFF 模块的集成. 在编码器的前 4 个下采样卷积单元之后, 引入设计的 GFF 模块, 该模块通过融合全局特征增强模型对复杂背景和条状微观结构的鲁棒性. 图 5 给出了具有 GFF 模块构造的 GFF-UNet++ 编码器, 图 5 每层最右侧的箭头均指向 GFF-UNet++ 网络结构中的稠密跳跃连接模块, 通过稠密跳跃连接实现层间的高效传递与融合.

3 评价指标

为验证 GFF-UNet++ 方法的有效性, 将通过预测图与标签图的比较, 并使用混淆矩阵评价模型性能. 混淆矩阵是一种区分真正值 (TP)、真负值 (TN)、假正值 (FP)、假负值 (FN) 的方法^[10]. 本文选用 5 个评价指标, 即准确率 (A)、召回率 (R)、精确度 (P)、Dice 系数和交并比 (IOU), 对 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ X 射线图像的条状微观结构分割任务进行全面评估.

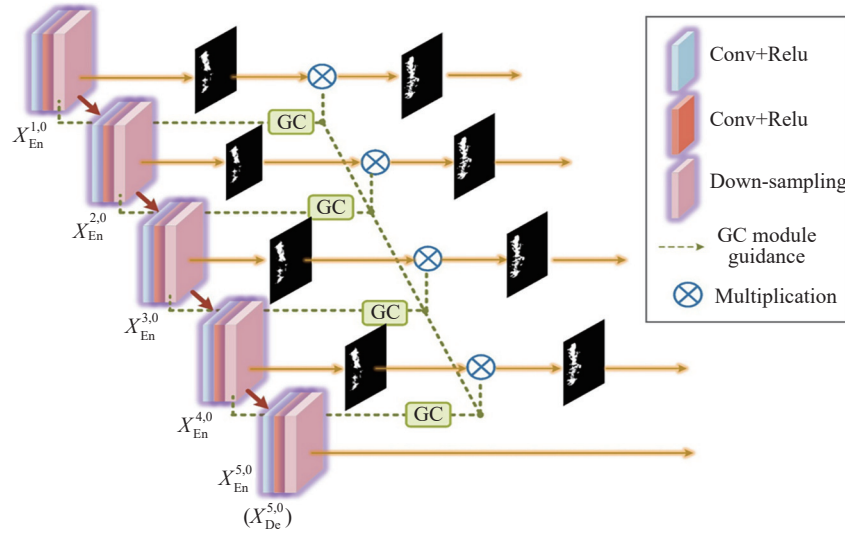


图 5 具有 GFF 模块构造的 GFF-UNet++编码器

Fig. 5 Schematic diagram of the GFF-UNet++ encoder with the construction of the GFF module

表 1 给出了混淆矩阵的构建方法, 其中列代表模型预测的类别结果, 行代表实际的类别标签。

准确率是模型预测正确的像素相对于所有像素的百分比, 衡量模型在所有像素上预测的准确性, 即

$$A = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}}. \quad (6)$$

召回率指的是真实标签为正类的像素中, 被模型正确预测为正类的比例, 即

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}}. \quad (7)$$

召回率衡量模型对目标区域 (正类) 的覆盖能力, 召回率越高, 表示模型更全面地检测出了目标区域。

精确度是模型预测为正类的像素中, 真实标签为正类的比例。其表达式为

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}. \quad (8)$$

模型的精确度越高, 表示模型预测的正类区域更精确, 假正例更少。

Dice 系数是一种用于衡量两个集合相似度的统计指标, 用于衡量预测分割结果与真实分割结果之间的相似度。其计算方法如下:

$$\delta_{Dice} = \frac{2N_{TP}}{2N_{TP} + N_{FN} + N_{FP}}. \quad (9)$$

Dice 系数的值越大, 表示两个集合的相似度越高, 模型的分割效果越好。

表 1 本实验中混淆矩阵的构建方法

Table 1 Construction method of the confusion matrix in this experiment

混淆矩阵	预测结果	
	条状微观结构	背景
真实标签	条状微观结构	TP
	背景	FN
		FP
		TN

IOU 通过计算预测分割与实际分割区域的重叠情况来评估模型的准确性, 即

$$I = \frac{N_{TP}}{N_{FP} + N_{FN} + N_{TP}}. \quad (10)$$

IOU 值越大, 表示预测的分割区域与真实的标注区域之间的重叠程度越高, 即分割的准确度越高。

4 GFF-UNet++模型实验结果

4.1 与其他图像分割模型的比较

表 2 给出了 UNet^[18], UNet++^[19], DC-UNet^[20], UNet3+^[21], PMM^[8] (Pretrained-Microscopy-Models) 的分析结果, GFF-UNet++等图像分割模型在 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ X 射线条状微观结构图像分割中的表现。在表 2 的评估过程中, 所有模型的性能基于 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 数据集的整幅图像像素进行计算。DC-UNet^[20] 是一种用于 X 射线图像分割的新架构。其改进 U-Net, 设计新 CNN (Convolutional Neural Net-

表 2 图像分割评价指标
Table 2 Image segmentation evaluation index

Architecture	Evaluation index				
	Accuracy	Recall	Precision	Dice	IOU
UNet	0.822	0.790	0.982	0.868	0.835
UNet++	0.813	0.788	0.988	0.834	0.822
DC-UNet	0.850	0.776	0.978	0.842	0.849
UNet3+	0.867	0.851	0.990	0.877	0.882
PMM	0.880	0.833	0.983	0.906	0.870
GFF-UNet++	0.898	0.885	0.993	0.894	0.901

注 粗体数值表示每列在所述指标下的最优结果.

work), 用残差模块替代跳跃连接, 在多数数据集上性能优于 U-Net, 为 X 射线图像分割提供了新的方法. UNet 3+^[21] 利用了全面的跳跃连接和深度监督将不同尺度特征图中的低级细节与高级语义相结合, 并从全尺寸聚合特征图中学习分层表示, 提高了图像分割的准确性. PMM 模型在材料显微图像的微观结构分割中表现优异. 将其与 UNet^[18], UNet++^[19], DC-UNet^[22], UNet3+^[21], PMM^[8] 模型进行比较, 进一步验证 GFF-UNet++模型在材料微观结构分割上的优越性能.

从表 2 可以看出, GFF-UNet++模型的各个指标均优于 UNet^[18], UNet++^[19], DC-UNet^[22], UNet3+^[21], PMM^[8] 模型. UNet++集成了 GFF 模块后, 在 Accuracy, Recall, Precision 和 Dice 等指标上均显示出明显的提升, 这一结果证明了 GFF 模块具备出色的全局特征处理能力. 通过表 2 可以得出结论, 在 Na₆Mo₁₁O₃₆ X 射线图像的条状微观结构分割任务中, 相比其他图像分割模型, GFF-UNet++更具有针对性和优势.

4.2 随训练轮次的变化

以在分割结果中表现最好的模型分析为基础, 进一步探讨每个模型的性能随训练轮次 (epoch) 的变化情况. 通过对不同 epoch 下的模型的表现进行对比, 不仅能够揭示模型在训练初期如何快速学习全局特征, 在中期逐步优化细节特征, 以及在后期可能出现的过拟合现象, 还可以帮助确定模型的最佳训练轮次. 在图像分割任务中进行模型性能统计时, 使用 BCELoss 作为损失函数, 其可以表示为

$$L_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^N [y_i \lg(p_i) + (1 - y_i) \lg(1 - p_i)]. \tag{11}$$

其中, N 为总样本数, y_i 为第 i 个样本的所属类别, p_i 为第 i 个样本的预测值.

图 6 给出了训练损失和验证损失随训练次数的变化. 从图 6 可以看出, 使用 UNet, UNet++, DC-UNet, UNet3+模型对 Na₆Mo₁₁O₃₆ 条状微观结构图像进行分割时, 模型的验证损失值波动过大, 并非最优的图像分割方法; 使用 UNet3+, PMM 模型进行图像分割, 模型的训练损失和验证损失相差较大, PMM 模型在训练集上表现较好, 但在验证集上泛化性能较差. GFF-UNet++模型不仅优于其他的对比模型, 标准差也小得多, 该模型在训练过程中取得良好的泛化能力, 避免了过拟合现象, 并能够更好地适应新的数据.

4.3 GFF-UNet++更好地描绘了模糊的边界

由于研究使用的是非公开的 Na₆Mo₁₁O₃₆ X 射线图像数据集, 目前尚无现有的图像分割方法能够直接应用于该数据集进行对比. 因此, 为确保研究结果的有效性和可比性, 这里选择了公开的开源图像分割模型 (UNet^[16], UNet++^[17], DC-UNet^[23], UNet3+^[21], PMM^[8]), 并在 Na₆Mo₁₁O₃₆ X 射线图像数据集上进行训练和测试. 通过这种方式, 能够基于这些已验证的开源模型进行比对, 进而获得相应的实验结果.

图 7 为 Na₆Mo₁₁O₃₆ 的分割结果对比, 图 7(a) 为图 2(f) 经过对比度受限的自适应直方图均衡化 (CLAHE) 预处理的图像. 使用 CLAHE (Contrast-Limited Adaptive Histogram Equalization) 算法进行图像预处理能够自适应地提升图像局部对比度, 为后续分割任务提供更清晰的图像基础. 图 7 还给出了不同模型对条状微观结构的分割结果, 不同颜色显示出分割结果中的不同类型区域. 图 7 (b) 为分割任务

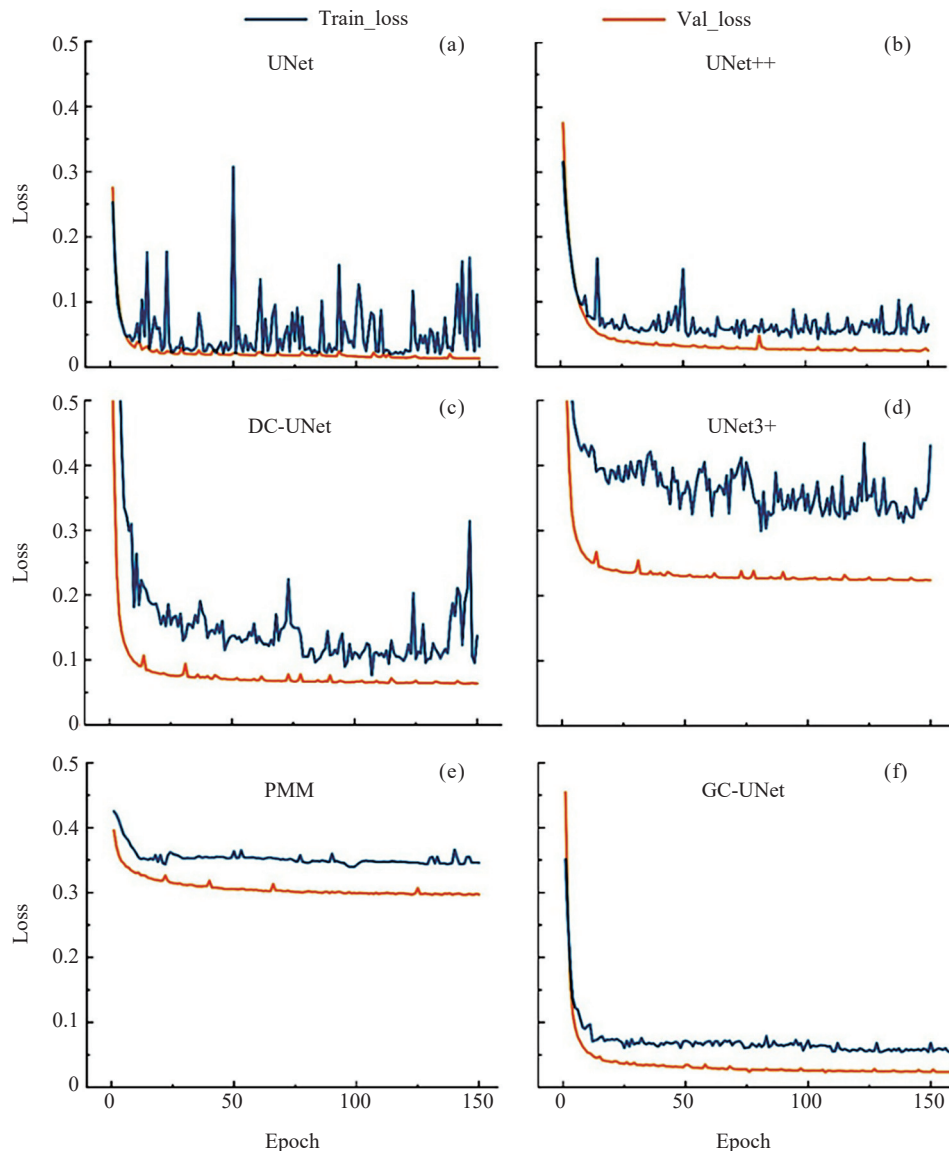


图 6 训练损失和验证损失随训练次数的变化

Fig. 6 Variation of training loss and validation loss with a number of iterations

的真实分割标签 (Ground Truth), 即人工标注的真实 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 图像分割结果. 图 7(c)~(h) 为各个模型的预测结果, 即通过图像分割模型对输入图像进行处理后得到的分割结果, 用于与图 7(a) 对应的各个模型分割结果进行对比, 验证各模型在分割目标结构时的准确性和鲁棒性. GFF-UNet++ 算法在所有算法中表现最佳, 其在图中的绿色和紫色区域面积最少, 同时在白色区域的表现也最为突出, 证明其在正确识别目标对象方面具有很高的准确性.

图 7 (a) 中, UNet 和 UNet++ 模型对部分条状微观结构的分割较为粗糙, 尤其是在微观结构的边

缘区域, 分割结果与 Ground Truth 存在较大的偏差 (绿色和紫色区域较多). DC-UNet 和 UNet3+ 模型虽然在一定程度上改善了分割结果, 但是在复杂背景中仍然存在较多的绿色区域, 影响了分割精度. PMM 模型对背景的误分割更加显著, 显示出其对背景和目标区域的区分能力不足. 通过对比, 可以直观发现 GFF-UNet++ 模型分割结果与 Ground Truth 最为接近, 在 False Positive 和 False Negative 的分布上均显著减少, 边缘区域的细节分割也更加准确, 白色区域的面积最大, 体现了对目标对象的识别能力.

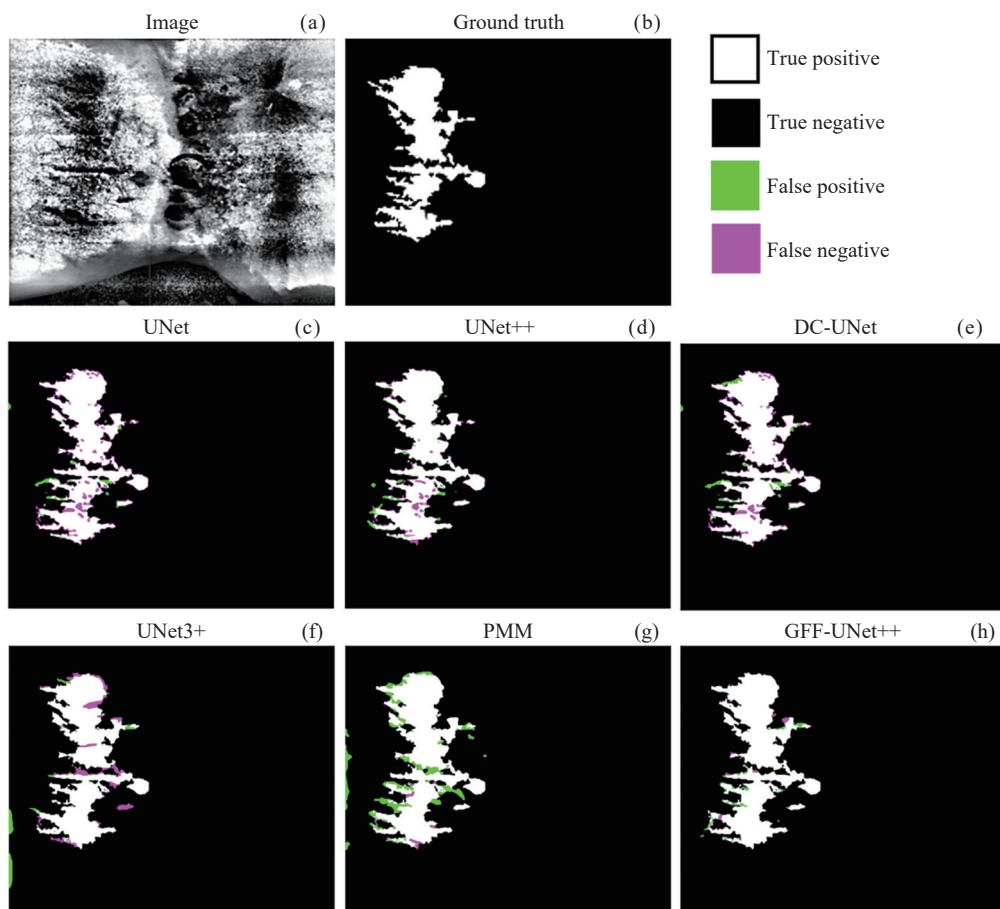


图 7 不同模型对 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 的 X 射线图像条状微观结构分割综合性能的定性比较

Fig. 7 Qualitative comparison of the comprehensive performance of different models in segmenting the striped microstructure of $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ X-ray images

5 结论

为能够清晰地分割出复杂背景下的 X 射线图像中的条状微观结构, 提出了一种针对 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ X 射线图像中条状微观结构的图像分割算法 GFF-UNet++. 通过引入全局特征融合模块, 所提出的 GFF-UNet++ 充分利用编码器学习的边缘特征, 显著提升了条状微观结构与复杂背景的区别精度. 与较为流行的分割模型 Unet, UNet++, DC-UNet++, UNet3+ 和 PMM 等模型相比, GFF-UNet++ 模型具有更强的边缘特征分割能力. GFF-UNet++ 算法证实了条状微观结构分割的可行性, 为空间材料实验过程中 X 射线图像条状微观结构分割提供了新的思路.

致谢 中国科学院上海硅酸盐研究所提供了 $\text{Na}_6\text{Mo}_{11}\text{O}_{36}$ 数据集, 中国航天科技集团公司第五研究院第

五一〇研究所提供了实验设备图像.

参考文献

- [1] WANG Y, LI S, LIU Z, *et al.* Anisotropy-dependent seaweed growth during directional solidification of Al-4.5%Cu single crystal[J]. *Scripta Materialia*, 2020, **186**: 121-126
- [2] AGHALARI M, AGHAGOLZADEH A, EZOJI M. Brain tumor image segmentation via asymmetric/symmetric UNet based on two-pathway-residual blocks[J]. *Biomedical Signal Processing and Control*, 2021, **69**(6): 102841
- [3] WANG E K, CHEN C-M, HASSAN M M, *et al.* A deep learning based medical image segmentation technique in Internet-of-Medical-Things domain[J]. *Future Generation Computer Systems*, 2020, **108**: 135-144
- [4] HUANG H, LIN L, TONG R, *et al.* UNet 3+: a full-scale connected UNet for medical image segmentation[C]// 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Barcelona, Spain: IEEE, 2020: 1055-1059

- [5] SONG Y, QU Z, LIAO H, *et al.* Material twins generation of woven polymer composites based on ResL-U-Net convolutional neural networks[J]. *Composite Structures*, 2023, **307**: 116672
- [6] GUAN Y, PANG Z, SUN H, *et al.* Material strain image segmentation algorithm based on improved UNET network[C/OL]//Proceedings of the 9th International Symposium on Test Automation & Instrumentation (ISTAI 2022), 2022
- [7] SHI P, DUAN M M, YANG L F, *et al.* An improved U-Net image segmentation method and its application for metallic grain size statistics[J]. *MATERIALS*, 2022, **15**(13)
- [8] 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**(1): 200
- [9] IBTEHAZ N, RAHMAN M S. MultiResUNet : Rethinking the U-Net architecture for multimodal biomedical image segmentation[J]. *Neural Networks*, 2019, 121: 74-87
- [10] GAO WEIZE, CHEN SHANXIONG, MO BOFENG, *et al.* R-UNet++: a local segmentation network for the classification of oracle bone materials[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2022, **34**(3): 415-424
- [11] HIRABAYASHI Y, IGA H, OGAWA H, *et al.* Deep learning for three-dimensional segmentation of electron microscopy images of complex ceramic materials[J]. *NPJ Computational Materials*, 2024, **10**(1): 46
- [12] 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**(1): 108
- [13] YANG B, WU M, TEIZER W. Modified UNet++ with attention gate for graphene identification by optical microscopy[J]. *Carbon: An International Journal Sponsored by the American Carbon Society*, 2022, **195**: 246-252
- [14] LIU P, SONG Y, CHAI M, *et al.* Swin-UNet++: a nested swin transformer architecture for location identification and morphology segmentation of dimples on 2.25Cr1-Mo0.25V fractured surface[J]. *Materials*, 2021, **14**(24): 7504
- [15] SHU K Y, CHEN Z X, ZHU B, *et al.* Unsupervised Segmentation for Microstructure Identification of High Strength Steel with Superpixel Segmentation and Texture Feature Clustering[C]//Proceedings of the 14th International Conference on the Technology of Plasticity - Current Trends in the Technology of Plasticity, 2024. Switzerland: Springer
- [16] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[J]. *Springer, Cham*, 2015. DOI:10.1007/978-3-662-54345-0_3
- [17] ZHOU Z, RAHMAN SIDDIQUEE M M, TAJBAKHS N, *et al.* Unet++: a nested U-Net architecture for medical image segmentation[J]. 2018. DOI:10.1007/978-3-030-00889-5_1
- [18] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation; proceedings of the medical image computing and computer-assisted intervention[C]//MICCAI 2015: 18th international conference, 2015. Munich: Springer
- [19] ZHOU Z, RAHMAN SIDDIQUEE M M, TAJBAKHS N, *et al.* UNet++: A Nested U-Net Architecture for Medical Image Segmentation[C]//Proceedings of the Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, 2018. Springer International Publishing. DOI: 10.1007/978-3-030-00889-5_1
- [20] LOU A, GUAN S, LOEW M H. DC-UNet: rethinking the U-Net architecture with dual channel efficient CNN for medical image segmentation[J]. *Proceedings of SPIE*, 2021, 11596(000). DOI:10.1117/12.2582338
- [21] HUANG H, LIN L, TONG R, *et al.* UNet 3+: a full-scale connected UNet for medical image segmentation[J]. *arXiv*, 2020. DOI:10.1109/ICASSP40776.2020.9053405
- [22] ISENSEE F, JAEGER P F, KOHL S A A, *et al.* mnU-Net: a self-configuring method for deep learning-based biomedical image segmentation[J]. *Nature Methods*, 2021, **18**(2): 203-211
- [23] LOU A, GUAN S, LOEW M. DC-UNet: rethinking the U-Net architecture with dual channel efficient CNN for medical image segmentation[C]//Image Processing. SPIE, 2021. DOI:10.1117/12.2582338

作者简介



李欣泽 (第一作者) 女, 2000 年 8 月出生于广西南宁市, 现为中国科学院大学硕士研究生, 主要研究方向为图像对比度增强、图像分割。

E-mail: lixinze221@mailsucas.ac.cn



于 强 (通信作者) 男, 1965 年出生于陕西宝鸡, 现为中国科学院国家空间科学中心博士生导师, 主要研究方向为空间科学实验技术、计算机应用技术、软件工程、人工智能等。

E-mail: yuqiang@nssc.ac.cn