

DOI:10.11784/tdxbz202502007

基于自监督学习的钛合金金相组织智能识别

宋凯¹, 张宗瀚¹, 余京泰¹, 高帅², 王璇², 钟秀堃¹

(1. 天津大学化工学院, 天津 300350;

2. 国家市场监督管理总局重点实验室(特种设备数字孪生共性技术), 天津 300192)

摘要: Ti-6Al-4V (TC4) 合金因其优异的性能而成为各行业应用最广泛的钛合金之一. TC4 合金的金相组织大致可分为 4 类: 等轴组织、双相组织、魏氏组织、网篮组织, 不同组织结构具有显著性能差异. 由于锅炉、管道、压力容器等设备长期处于高温、高压等极端工况会导致合金材料的金相组织特征发生变化, 进而显著影响 TC4 合金的强度、韧性、耐腐蚀性、硬度等关键性能指标. 为了满足 TC4 合金金相组织自动高效识别的需求, 开发了一种基于掩蔽重建的新型自监督分类模型 (MSC-Net), 旨在克服常规深度学习模型依赖大型标注数据集的局限性. 通过引入掩码重建技术, 该模型能在无显式标注的情况下学习金相组织图像的复杂内在特征, 并准确区分等轴组织、双相组织、魏氏组织和网篮组织等不同类型. 实验结果表明, MSC-Net 不仅在实验室数据集中实现了 100.0% 的识别准确率, 且当主流模型 (如 ResNet 和 Inception v3) 在手持光学显微镜图像数据集的泛化能力仅为 60.4% 和 51.5% 时, MSC-Net 的准确率竟高达 87.6%, 展现了 MSC-Net 卓越的泛化性能. 充分证明了本算法能有效迁移到使用便携金相显微镜采集的现场图片中进行无损检测, 并能够实时、准确地识别出当前的金相组织类型. 这一功能是实现服役过程中组织状态监控的基础, 为设备性能评估和预测性健康维护等提供了可靠的技术支持.

关键词: 无损检测; 自监督学习; 金相组织; 钛合金; 掩码自编码器; 计算机视觉; 模式识别

中图分类号: TQ028.8

文献标志码: A

文章编号: 0493-2137(2025)10-1021-10

Intelligent Recognition of Titanium Alloy Microstructures Based on Self-Supervised Learning

Song Kai¹, Zhang Zonghan¹, Yu Jingtai¹, Gao Shuai², Wang Xuan², Zhong Xiukun¹

(1. School of Chemical Engineering and Technology, Tianjin University, Tianjin 300350, China;

2. Key Laboratory of Digital Twin Generic Technology in Special Equipment, State Administration for Market Regulation, Tianjin 300192, China)

Abstract: Ti-6Al-4V (TC4) alloy is one of the most widely used titanium alloys in various industries owing to its outstanding characteristics. The microstructures of TC4 alloy can be broadly classified into four types, i.e., equiaxed, duplex, Widmanstätten and basket-weave structures. Each of these microstructures exhibits distinct characteristics. Prolonged exposure of boilers, pipelines, pressure vessels and other equipment to extreme conditions such as high temperatures and high pressures causes changes in the microstructural characteristics of alloy material. Such microstructural changes significantly affect the key performance indicators of TC4 alloy, such as strength, toughness, corrosion resistance and hardness. To meet the requirements for automated and efficient microstructure identification in TC4 alloys, we propose MSC-Net, a novel self-supervised classification model based on masked autoencoder architecture. The proposed model overcomes the limitations of conventional deep learning models that require large annotated datasets by enabling unsupervised learning of complex intrinsic features in metallographic images. Our model achieves accurate differentiation between equiaxed, duplex, Widmanstätten, and basket-weave

收稿日期: 2025-02-11; 修回日期: 2025-03-27.

作者简介: 宋凯 (1975—), 女, 博士, 副教授, ksong@tju.edu.cn.

通信作者: 钟秀堃, 2023207580@tju.edu.cn.

基金项目: 国家市场监督管理总局科技计划资助项目 (2024MK008).

Supported by the Science and Technology Plan Project of State Administration for Market Regulation (No. 2024MK008).

structures without explicit labeling. The experimental results demonstrate the remarkable performance of MSC-Net, achieving 100.0% recognition accuracy in laboratory datasets. More significantly, compared to mainstream algorithms (ResNet and Inception v3 with accuracies of 60.4% and 51.5%, respectively), it exhibits superior generalization capabilities, achieving 87.6% accuracy on field images captured by portable optical microscopes. These results validate the effective transferability of MSC-Net to real-world non-destructive testing scenarios. The model can identify the current microstructure type in real-time and with high accuracy. This capability is fundamental for monitoring the material's condition during service, providing reliable technical support for equipment performance evaluation and predictive health maintenance.

Keywords: non-destructive testing; self-supervised learning; microstructure; titanium alloy; masked auto-encoder; computer vision; pattern recognition

TC4(Ti-6Al-4V)合金是世界上最广泛使用的钛合金之一,因在室温下显示出 $\alpha + \beta$ 双相结构,被称为 $\alpha + \beta$ 型钛合金^[1].凭借低密度、高强度、优异的耐腐蚀性和抗疲劳性能,钛合金已成为航空航天^[2]、船舶^[3]、化工及医疗设备^[4-5]等行业的关键结构材料.在化工领域,TC4合金依托其卓越的耐腐蚀性、高比强度、重量轻及良好加工性,被广泛用于换热器、压力容器和反应器等设备的管道、板材、泵阀等部件.特别是在海水淡化、强酸、强碱及高氯化物等苛刻工况下,它以优异的耐蚀性、高温强度和较低维护成本展现出不可替代的优势^[6-8].

典型的钛合金主要展现两种晶体结构: α 相和 β 相. α 相钛合金以其优良的延展性和加工性能而闻名,而 β 相钛合金则以其高强度著称.这两种相的化学组成、形态、分布以及其他微观结构特征如位错和织构等形成了钛合金金相组织的4种类型:等轴组织、双相组织、魏氏组织和网篮组织,它们是影响钛合金性能的关键因素.在实际应用中, α 相与 β 相的比例通常根据性能要求进行优化^[9-10].钛合金的微观结构,即金相组织,包括晶粒的尺寸、形状和晶界等因素,对其性能有决定性影响.因此金相组织分析^[11-12],尤其是金属材料的微观结构表征和分析,是科学领域中的关键技术^[13],为材料的性能优化和在工业生产中保证材料具有一致和预期的性能提供理论基础^[14-17].服役过程中,温度、应力、腐蚀、辐射等服役环境和工况的变化会导致金相组织不同相的相互转化、相界面的移动和相分布的变化、分解、沉淀或重新排列等多种微观组织类型的变化,从而影响材料的力学性能、耐腐蚀性和使用寿命等^[1,18].为此,快速识别和分析金相组织对于设备健康检测至关重要.

既有的金相组织分析方法主要包括手工操作或固定算法,这些方法依赖于手动设计的特征提取技术,如边缘检测、纹理分析和阈值处理. Rauch 等^[19]设计了一种用于金相组织图像分析的边缘检测和过

滤方法,通过使用 Canny 边缘检测算法和基于粒子动力学的过滤技术,专门用于检测合金等材料颗粒之间的边界,解决了因图像中存在的划痕和微小杂质等噪声问题而带来的挑战. Zhe 等^[20]开发了一种用于分离钛合金金相组织中复杂接触的等轴相和片状 α 相的方法,该方法通过识别微观结构特征中的凹点,选择凹点对,并确定连接两个凹点的候选分离线是否可以接受.这种方法有效地分离了复杂接触的微观结构特征,为钛合金的定量表征提供了技术支持.然而这些方法涉及一系列手动步骤和主观判断,在面对不同类型的金相组织图像时,往往需要重新设计和调整,因此过程效率低下.尤其是在需要处理大量图像的情况下,这种低自动化程度不适合现代工业应用.

随着人工智能尤其是基于深度学习的计算机视觉技术的迅猛发展,众多研究者已将这些方法引入金相组织分析,以克服传统方法的诸多不足.深度学习算法能够自动从数据中提取特征,无需人工干预,显著提升了分析的精度、效率和可靠性^[11]. Baskaran 等^[21]使用卷积神经网络对钛合金的微观结构进行分类,将其分为层片状、双相和针状3种类型. Nigam 等^[22]探讨了使用卷积神经网络来自动识别和分类金属微结构的方法,首先将微结构分类为树枝状和非树枝状,然后进一步将树枝状微结构分类为纵向和横向视图.该研究表明,深度学习模型能够绕过传统的特征提取步骤,直接从图像中学习以区分不同金相组织的特征,从而提高了分类的准确性和效率. Kuchar 等^[23]使用传统的阈值处理方法来分割和分类 Zn-Al-Mg-Sn 合金的 SEM BSE 微观结构,该研究侧重于通过半自动化 Matlab 算法进行系统化的图像分析,从而有效提取和分析材料的微观组成特征.这些模型能够从复杂的图像数据中自动识别和提取有用特征,并具备更好的泛化能力,能够应对各种不同的金相图像.

尽管深度学习技术在图像分类和分割等模式识别领域已展现出卓越能力,但其在材料科学特别是微

观结构识别中的应用仍面临多重挑战,主要原因是深度学习模型对大量带标注数据的依赖.与自然图像如ImageNet相比,微观结构图像在细节和信息密度上的复杂性更高,这不仅使模型训练变得更加困难,而且也增加了精确标注的难度.这些因素共同导致了高质量、有代表性的带标签数据集的极度匮乏,极大地限制了深度学习在此领域的进一步应用和发展^[11, 24-25].如何减少对大规模带标注数据集的依赖成为了研究者们关注的重点.这一需求催生了对数据获取和处理方法的创新,目标是开发能够从有限或未标注的数据中有效学习的算法. Goetz 等^[24]采用了一种名为无监督域适应的转移学习方法,目的是在源数据(已标注)和目标数据(未标注)之间找到域不变特征,从而优化在目标数据分布上的性能. Chen 等^[26]提出了一种名为 Con2Net 的半监督模型,采用多解码器架构,通过一致性约束和对正确预测像素的锐化操作来改善分割效果;此外,研究者还引入了对比学习模块,以增强模型在特征空间中的类内紧凑性和类间分散性. Kim 等^[27]介绍了一种无监督的机器学习技术,用于对低碳钢的微观结构进行分割,该技术结合了卷积神经网络和超像素算法,在不依赖于大量带标签数据的情况下处理和分析微观结构图像.本实验室团队开发了一种基于掩码自编码器(masked autoencoder, MAE)的AI生成模型 sSEM-Net,利用 MAE 复杂的深度学习模型学习光学显微镜图像与 SEM 图像之间的对比关系,并根据光学显微镜图像在线生成 SEM 图像,用于辅助金相组织分析^[28].

然而,上述这些方法需要图像具有极高的细节丰富性和分辨率,主要以扫描电子显微镜等图像的离线分析为主,当输入降至在线监测或手持终端的低分辨率图像时,模型的泛化能力明显不足,亟需开发以低分辨率图像为输入的金相组织分类识别算法.更为重要的是,对于锅炉、管道和压力容器等大型设备的实时监控,现有检测方法^[29]通常需经历破坏性的样本处理步骤,限制了其在需要快速、无损评估的工业场合中的适用性.因此开发基于计算机视觉技术的在线自动检测技术助力材料宏观性能的预测与解释至关重要.

为此,本文在 sSEM-Net^[28]的基础上,提出了一种新的基于 MAE 的自监督预训练模型 MSC-Net,来完成金相组织的自动分类任务. MAE 通过掩码和重建的策略,使模型能够在没有显式标签的情况下学习到更深入、更鲁棒且更有用的数据表示,从低分辨率的光学显微镜图像呈现的复杂金相组织图像中捕捉

到本质和关键特征,因此这种方法特别适用于金相组织的分类.通过这种自监督学习框架, MSC-Net 能够有效提高金相分类任务的准确性.这种训练策略不仅优化了模型对金相组织图像的解析能力,还提高了其在实际应用中的泛化能力,为未来无损在线检测和材料科学领域的深度学习应用开辟了新的可能性.此外,本技术并不限于钛合金,其应用范围可以拓展至其他类型的合金,为金相分析的非破坏性、高效应用提供了新的视角和技术支持.

1 实验材料和方法

1.1 钛合金金相组织自动识别模型 MSC-Net

图 1 展示了 MSC-Net 模型的整体框架及其两阶段训练流程.在第 1 阶段的自监督预训练阶段中,模型基于 MAE 架构,首先对输入的金相组织图像随机掩蔽 50% 的块状区域,随后将剩余的可见图像块输入编码器进行多层次特征提取;解码器基于编码器输出的潜在特征与掩蔽位置标记,重构被掩蔽区域的像素信息,继而使用遮蔽重建图像与原图像的均方误差作为损失函数驱动模型自主学习晶界分布、相区形态等关键视觉特征,整个预训练过程完全无需人工标注数据,仅通过大规模无标签金相图像的自监督学习即可完成.第 2 阶段迁移至金相组织分类任务,移除预训练阶段的解码器模块,保留编码器作为固定特征提取器,并使用第 1 阶段训练后的模型参数作为初始权重对分类模型的编码器进行权重初始化,后在 ViT 编码器的输出端接入由全连接层构成的分类头,通过少量标注数据对分类层进行微调以执行金相组织的识别任务.此框架可以充分利用无标签数据来增强模型的特征提取能力,显著减少对大量标注数据的依赖,有效降低了数据标注的成本和复杂性.

1.1.1 自监督预训练阶段

如图 1 所示,自监督预训练网络采用编码器-解码器结构.编码器主要负责从输入数据中提取特征,其输入是部分被掩码的图像,输出则是图像的高维特征表示.解码器的任务是利用编码器输出的压缩特征表示,重建图像的完整视觉内容.这不仅包括从已知数据部分复原图像,还需要准确预测由于掩码而缺失的图像信息.这种掩蔽策略迫使编码器学习如何从限定的可见数据中抽取图像的本质和关键特征.通过从随机掩码的图像中预测原始图像,编码器获得了提取有意义信息表示的能力,显著增强了网络对图像全局结构和局部细节的理解能力.

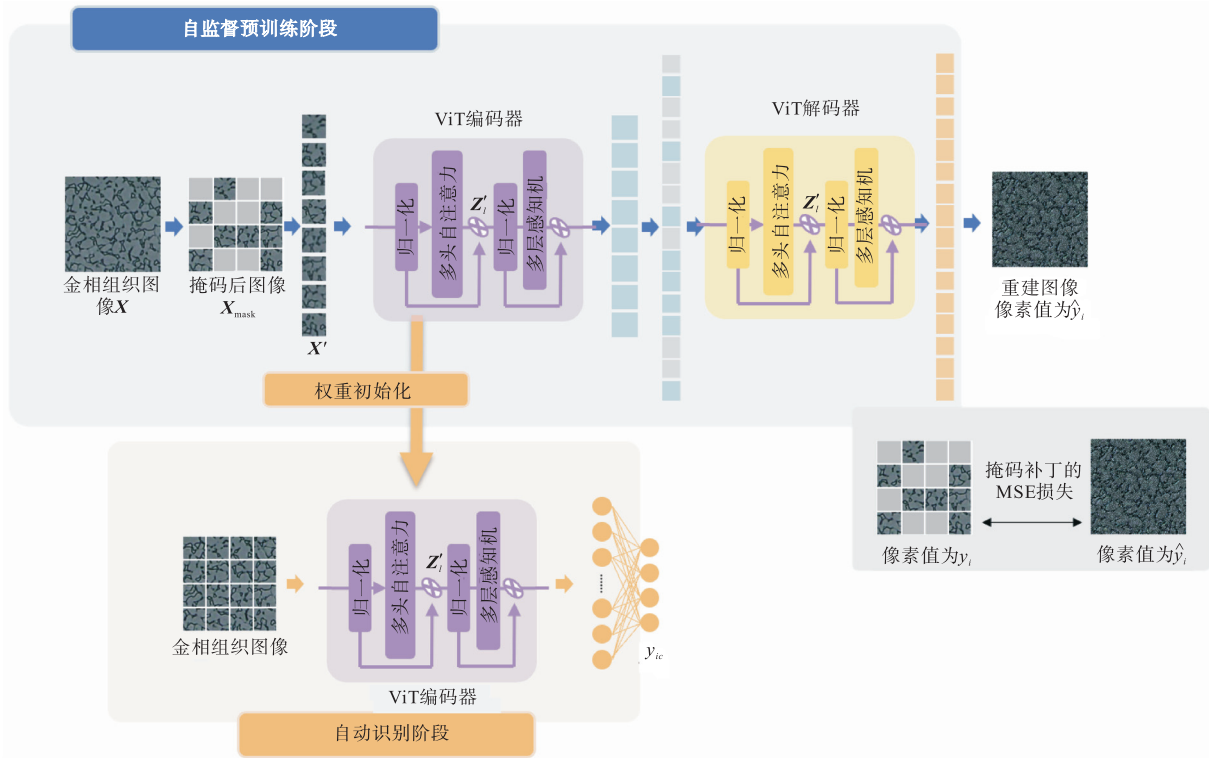


图 1 MSC-Net 模型的整体框架
Fig.1 Overall framework of the MSC-Net model

1) 预处理块

在将图片输入到编解码器之前,需要先对其进行预处理.首先,将输入的二维三通道图像划分为规则的不重叠的块.其次,通过补丁嵌入将图像块转化为高维空间中的点,方便对其进行更高效的特征提取,同时添加位置编码使模型理解补丁在原始图像中的相对位置.之后,对这些补丁进行随机掩模处理.最后,将处理过的补丁嵌入与类标记结合.该过程的公式表达如下:

$$X_{emb} = Emb(X) \quad X \in \mathbf{R}^{H \times W}, X_{emb} \in \mathbf{R}^{\frac{HW}{P^2} \times D} \quad (1)$$

$$X_{mask} = mask(X_{emb} + E_{pos}) \quad X_{mask} \in \mathbf{R}^{(1-r) \left(\frac{HW}{P^2} \right) \times D} \quad (2)$$

$$X' = concat(X_{mask}, E_{cls})$$

$$E_{cls} \in \mathbf{R}^{1 \times D}, X' \in \mathbf{R}^{(1-r) \left(\frac{HW}{P^2} + 1 \right) \times D} \quad (3)$$

式中: X 为原始输入图像; X' 为处理后的图像; P^2 表示将图像划分为 $P \times P$ 的补丁,作为 Transformer 的最小信息处理单元; D 表示嵌入向量维度; r 为对图像的掩码比率; X_{mask} 为掩码后的图像; E_{pos} 和 E_{cls} 分别为位置嵌入和类别嵌入; \mathbf{R} 代表实数域; H 、 W 分别为输入图像的高度和宽度(单位为像素).

2) 编码器

因 Transformer 在复杂图像任务中的性能已显著优于传统 CNN^[30-31],且其特征提取过程不依赖预设模板,适应各种金相图像的特征和变化,因此,MSC-Net 以 Transformer 块作为核心特征提取模块,以充分发挥其捕捉全局上下文信息的优势.在自监督预训练阶段,编码器仅作用于可见的图像块,以此减少模型训练的计算需求.

如图 2 所示,每个 Transformer 块主要由两部分构成:一个多头自注意力(multi-head self-attention,

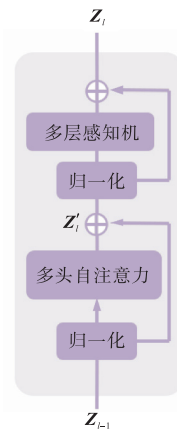


图 2 Transformer 块的详细结构
Fig.2 Detailed structure of the Transformer block

MSA)层和一个多层感知机 (multi-layer perceptron, MLP)层^[31]。Transformer 块的核心设计是多头自注意力机制。多头自注意力层将输入特征划分为多个头并行计算,以捕捉不同子空间的信息,在多种表示空间内建模数据的内在相关性,模型因此能够同步关联图像的多样特征。多层感知机层则作为一个简单的前馈神经网络,用于处理自注意力层的输出,进一步转换特征以提供更深层的非线性映射。此外,每个 Transformer 块内还嵌入了残差连接,不仅加速了训练过程还避免梯度消失问题。Transformer 块的处理过程由式(4)和(5)表示,即

$$\mathbf{Z}'_l = \text{MSA}(\text{LN}(\mathbf{Z}_{l-1})) + \mathbf{Z}_{l-1} \quad l=1,2,\dots,L \quad (4)$$

$$\mathbf{Z}_l = \text{MLP}(\text{LN}(\mathbf{Z}'_l)) + \mathbf{Z}'_l \quad l=1,2,\dots,L \quad (5)$$

式中: $\mathbf{Z}_0 = \mathbf{X}'$; \mathbf{Z}'_l 表示第 l 个自注意力层的输出信息; \mathbf{Z}_l 表示第 l 个多层感知机层的输出信息; L 为 MSC-Net 中 Transformer 块的个数。

3) 重建解码器

重建解码器的主要目标是利用编码器学习到的特征来精确重构原始图像。

在特征输入到解码器之前,添加位置嵌入,帮助解码器理解各补丁在整个图像中的相对关系,这对于正确重建图像的全局结构至关重要。为减轻模型的计算负担和训练时间并有效提高更精准的图像重构, MSE-Net 采用非对称设计:相比于编码器而言,解码器的设计尽量保持简洁,迫使编码器学习到信息密集的压缩表示。解码器处理的输入包括完整的标记集,其中既包含未被掩码的可见块标记,也包含被掩码的不可见块标记。

4) 损失函数

为使重构像素尽量接近真实像素,采用均方误差 (mean square error, MSE, e_{ms}) 作为损失函数,其数学计算如式(6)所示。MSE 的值越小,误差越小,表示预测模型的准确度越高。

$$e_{\text{ms}} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

式中: n 代表图像中被掩蔽的像素总数; \hat{y}_i 和 y_i 分别表示重建图像和原始图像中对应的像素值。损失函数的计算仅包括被掩蔽的像素点,确保模型训练的重点在于提高图像的重建质量。

1.1.2 分类阶段

1) 分类头

由于经过自监督预训练的编码器能够高效提取图像的深层特征,分类阶段的网络结构无需过分复杂。因此,为了在标记数据有限的条件下减少模型参数、规避过拟合风险, MSC-Net 的分类头采用了简洁的设计,仅由一层全连接构成。

2) 损失函数

损失函数 (\mathcal{L}) 采用交叉熵损失 (cross entropy loss) 结合 Softmax 激活函数和对数损失 (logarithmic loss), 其计算式为

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^M y_{ic} \ln p_{ic} \quad (7)$$

式中: N 为样本总数; M 为类别总数; y_{ic} 为真实标签的 one-hot 编码,对于真实类别 $y_{ic} = 1$, 否则 $y_{ic} = 0$; p_{ic} 为模型预测样本 i 属于类别 c 的概率。

1.2 实验制备与数据采集

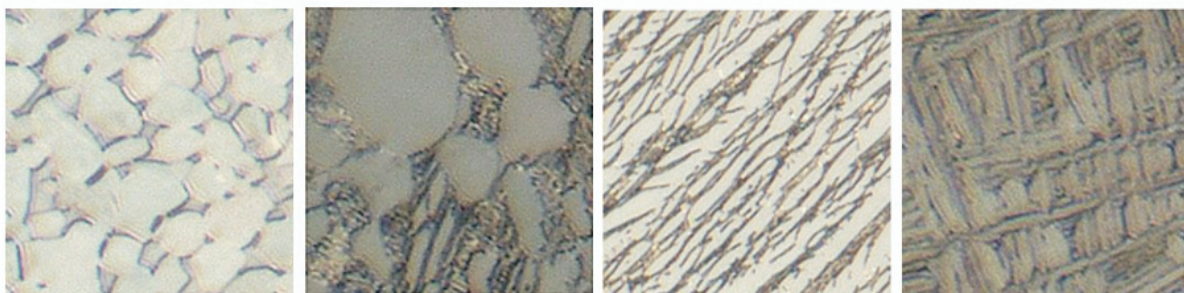
1.2.1 样本材料制备

如图 3 所示,室温下的 TC4 合金金相组织主要分为 4 类:等轴组织、双相组织、魏氏组织和网篮组织。表 1 详细列出了这 4 类典型组织的组织特点以及热处理工艺参数^[32]。

1.2.2 数据集

1) 自监督预训练数据集

采用 3 台光学显微镜 (DSX510、AXIO Scope. A1 和 VHX-7000) 从以上 4 个不同样本的多个区域捕获大量图像,这些图像的放大倍数为 500 倍和 1000 倍。随后,图像被裁剪至 256 像素 \times 256 像素的大



(a) 等轴组织

(b) 双相组织

(c) 魏氏组织

(d) 网篮组织

图 3 TC4合金的金相组织

Fig.3 Microstructures of TC4 alloy

表 1 TC4 合金 4 类金相组织的加工工艺及参数

Tab.1 Processing technology and parameters of four types of microstructure of TC4 alloy

金相组织	组织特点	工艺参数
等轴组织	主要是由等轴状的 α 相和 β 相转变组织构成,晶粒呈近似球形,整个材料中晶粒大小相对均匀,没有明显的晶粒取向	先经过两次自耗电弧炉熔炼获得铸锭,再在 β 单相区(1 000 $^{\circ}\text{C}$)进行开坯,再加热至 $\alpha+\beta$ 相区并锻造为棒材,最终在 940 $^{\circ}\text{C}$ ($\alpha+\beta$ 相区)终锻获得等轴组织
双相组织	由等轴状的 $\alpha+\beta$ 相转变组织构成,其特征是等轴 α 相含量约为 30% 以下,在这种结构中,等轴初生 α 晶粒可能呈现为均匀分布的圆形或近圆形晶粒	TC4 热锻板材经 800 $^{\circ}\text{C}$ 热处理 1 h, 空气冷却,再经 540 $^{\circ}\text{C}$ 时效 4 h, 得到由等轴初生 α 相和 β 相转变组织生成的细小的细片状次生 α 相组成的双态组织
魏氏组织	由 β 相转变组织构成,无等轴 α 相, β 相晶粒内的 α 相以片层状为主,在晶粒内存在集束	TC4 热轧板材经过 1 020 $^{\circ}\text{C}$ 热处理 2 h 空气冷却,得到具有粗大原始 β 晶粒的魏氏组织
网篮组织	细长的片层结构多方向交错排列,具有网篮编织状形态,原始晶界 β 相已被破碎	通过定向能量沉积(DED)技术制备,得到具有针状马氏体组成的网篮组织;功率为 1 400 W,扫描速度为 1 200 mm/min,送粉速率为 1.8 g/min,层厚为 0.5 mm,扫描间距为 0.6 mm,抬升量为 0.3 mm

小. 排除存在对焦不准确、图像框架内有明显污点、边缘模糊或曝光过度/曝光不足等缺陷的不合格图像,预训练数据集最终包含 66 968 张图像.

2) 实验室验证数据集

为验证模型的性能,另外采集整理实验室验证数据集. 同样由 3 台光学显微镜(DSX510、AXIO Scope. A1 和 VHX-7000)从 4 个不同样本的多个区域采集,共包含 4 个类别总共 400 个图像样本,每个类别 100 个样本. 为保证检验的合理性,此数据集中不包含任何属于自监督预训练数据集中的图像.

3) 泛化验证数据集

为了更好地评估模型是否能满足现场金相检验的要求,使用便携式 PTI-5000 光学数字显微镜获取的数据集作为现场验证数据集进行泛化性能的验证. 该设备易于连接到各种材料上,特别适用于锅炉、管道、压力容器以及大型铸造或锻造部件的现场金相检查. 使用该设备,共收集了 4 种不同类型的 192 张钛合金微观结构图像.

1.3 自监督预训练及分类实验

1.3.1 自监督预训练的实验设置

在自监督预训练阶段,模型在包含 66 968 张光学显微镜图像的数据集上进行了训练. 这些图像的分辨率统一为 256 像素 \times 256 像素. 每张输入图像被划分为 16 个小块,每小块的嵌入维度设置为 256. 在训练过程中,采用了 AdamW 优化器,批处理大小为 256,初始学习率设定为 0.001,权重衰减率为 0.05,掩码比例设置为 0.5,模型训练 1 000 个周期.

1.3.2 分类实验的实验设置

在对分类头进行训练的阶段,编码器的权重由预训练权重初始化,全连接层为随机初始化. 数据集按照 9 : 1 的比例划分为训练集和测试集,以确保训练的充分性和模型评估的准确性. 模型训练了 50 个周期,批处理大小为 32,初始学习率设定为 0.001,权

重衰减率为 0.05.

2 实验结果与讨论

2.1 自监督预训练结果与讨论

如图 4 所示,第 1 行是原始图像,第 2 行是掩码后图像,第 3 行则展示了 MSC-Net 模型基于掩码后图像重建的结果. 可视化的结果表明, MSC-Net 模型能够从有限数量的可见图像块中有效地学习金相结构的特征,并恢复大部分图像细节,说明 MSC-Net 模型能够充分提取原始图像的高维特征,为后续分类头的分类任务打下良好的基础. 金相图像相较于自然图像,需要更高层次的细节恢复,且其数据获取相对更具挑战性. 在这种情况下,尽管使用的数据集显著小于处理自然图像时通常使用的数据集,模型仍表现出良好的重建能力. 在自监督学习过程中,为了增加模型处理复杂图像细节的能力,采用高比例的掩码. 这虽然增加了捕获某些细节的复杂度,但同时也为模型提供了学习从部分信息重建完整图像的机会. 由于损失函数只针对被掩码的块进行计算,不涉及可见块,这种策略有助于强化模型对于丢失数据部分的重建性能. 可视化重建结果表明,预训练不仅提升了模型对原始数据的理解,还进一步挖掘了图像的内在信息,对提高下游任务的性能具有重要作用.

2.2 分类结果及模型对比

为了全面评估 MSC-Net 模型的性能,将其与当前主流的深度学习架构进行了比较. 基于 CNN 的 VGGNet^[33]、ResNet^[34]和 Inception v3^[35]模型是图像分类领域的经典模型,在多种图像分类任务中表现优异. LSTM 模型^[36]基于多层感知机,可以对长序列进行建模并捕获长距离依赖关系,作为分类器使用时有助于提取长距离特征. 基于 Transformer 的 ViT 模型^[37]和 Swin-ViT 模型^[38]是当前最主流的图像分类模

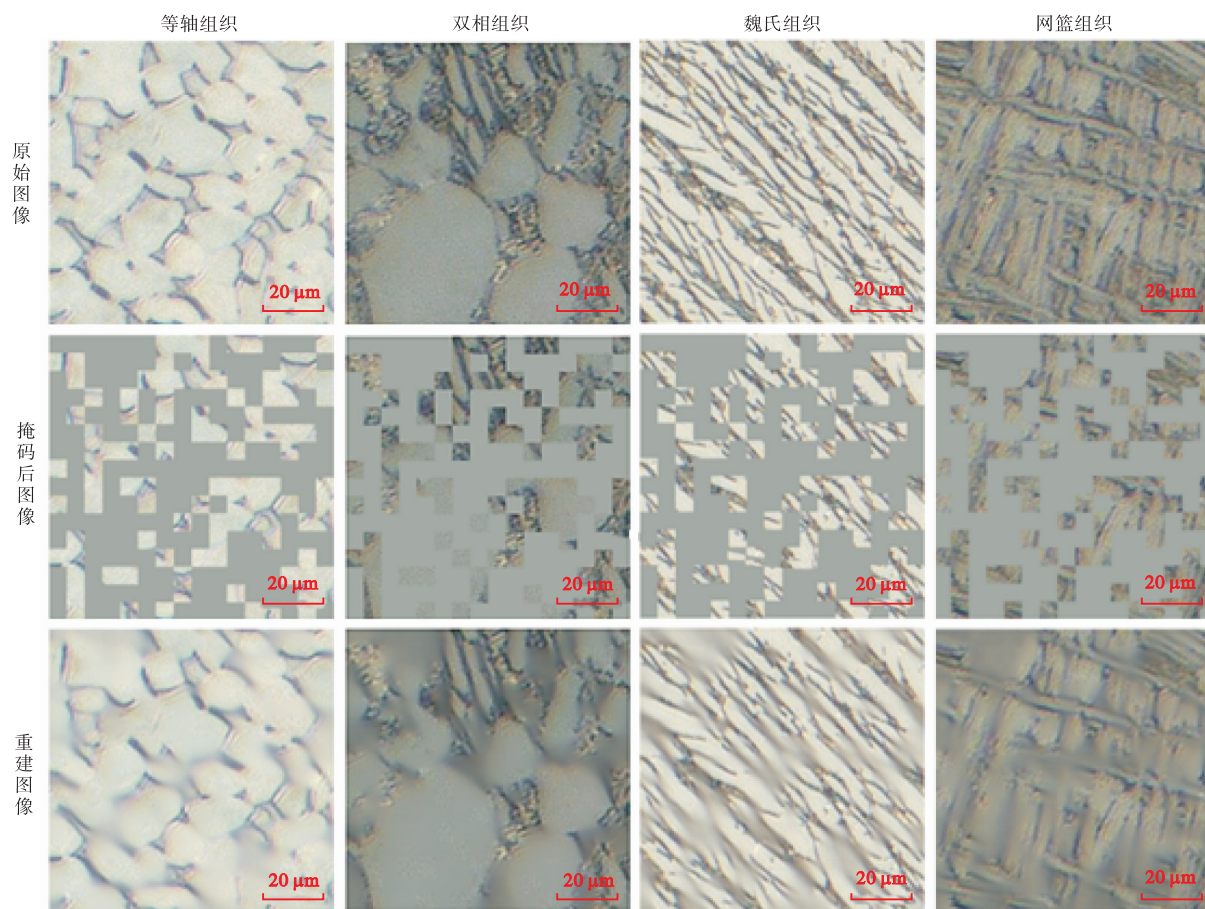


图 4 自监督预训练图像重建的视觉结果

Fig.4 Visual results of self-supervised pre-training image reconstruction

型,使用注意力机制捕获图像的全局依赖关系. 这些模型覆盖了从传统 CNN 到最新 Transformer 架构的深度学习主流方法. 本实验在实验室数据集上进行训练和验证,并且在泛化数据集上进行泛化性能的验证. 实验室数据集通常在受控环境下获得,图像质量和采集条件相对一致,而现场数据集则包括使用便携式 PTI-5000 光学数字显微镜收集的图像,这些图像可能包含各种噪声,更具挑战性.

从表 2 的对比结果可以看出,所使用的 MSC-Net 模型在实验室验证集上达到了 100.0% 的最高准确率,显示了优异的性能. 其他模型如 LSTM 和 Swin-ViT 也展现了很高的准确率,达到了 99.5%. Inception v3 模型性能稍差,但也达到了 98.8% 的准确率. ViT 模型和 VGGNet 模型的准确率相对较低,分别为 92.0% 和 95.0%. 现场金相检测为材料微观结构的评估提供了快速便捷的手段,但与实验室的稳定条件相比仍存在较大的局限性,现场环境中的光照条件以及设备稳定性等参数难以达到实验室的理想水平,在更具挑战性的现场泛化数据集上, MSC-Net 模

型表现出了非常优异的性能,准确率达 87.6%,验证了该方法对现场即时检测场景的适用性. 而其他模型的泛化能力均不尽如人意. 其中 ResNet 模型、Inception v3 模型和 VGGNet 模型的泛化能力相对较好,现场验证集准确率分别为 60.4%、51.5% 和 49.5%. LSTM 模型和 ViT 模型在现场验证集上的准确率仅有 26.0%,难以区分现场条件下钛合金显微组织的金相类型,其原因为 ViT 系列和 VGGNet 作为参数量较大的模型,在数据量有限的情况下难以充分学习,模型很容易陷入过拟合. 大量的可训练参数在数据稀缺的环境下容易关注训练样本中的噪声等不重要信息,捕捉数据中的偶然模式,而无法学习到有效的、泛化能力强的特征表示. 本研究使用的分类数据集中,每个类别仅有 100 张图片,是一个对于分类任务而言较小的数据集,这使得参数量较大的模型无法学习到真正的因果关系,无法充分学习到能够泛化到新数据的特征. 此外,与基于 CNN 的网络架构相比,基于 Transformer 的网络在建模细粒度特征上不如 CNN 高效,同时缺乏对图像局部结构的自然归纳

偏置,在数据量不足时难以进行高效学习. CNN 由于其卷积结构的特点,能够自然地捕捉到图像中的局部连续性和空间层次性,更具有鲁棒性,这对于图像处理任务尤为重要,也使得 ResNet 模型和 Inception v3 模型在现场泛化数据集上达到了相对较好的结果. LSTM 作为序列模型依靠将图像切割为序列以进行特征提取,天然地难以有效建模像素之间的空间局部关系,并且在输入时将二维图片转换为序列破坏了图像的二维结构,使学习图像的多层次特征变得困难,难以泛化到差距过大的数据集.

表 2 不同模型分类结果对比

Tab.2 Comparison of the classification results obtained with different models

模型	参数量/ 10 ⁶	实验室数据集 准确率/%	现场数据集 准确率/%
VGGNet	139.6	95.0	49.5
ResNet	11.2	97.5	60.4
Inception v3	24.3	98.8	51.5
LSTM	33.1	99.5	26.0
ViT	86.6	92.0	26.0
Swin-ViT	86.7	99.5	34.0
MSC-Net	85.8	100.0	87.6

MSC-Net 模型通过自监督预训练机制,有效利用海量无标注图像数据进行特征表示学习,实现了对数据资源的充分利用,显著降低了对人工标注数据的依赖性,从而避免了大规模数据标注所产生的高昂成本,学习到了更为广泛和通用的特征表示,在现场数据集上表现出更好的适应性和泛化能力.

MSC-Net 模型展现出的优异泛化性能表明,该模型能够有效捕获从标准实验室环境到现场无损检测场景数据之间的分布差异,这一特性在实际工业部署中具有重要应用价值. 在现场数据集上的高准确率表明, MSC-Net 模型不仅在实验室环境中表现出色,在更具挑战性的现场环境中也能保持高准确性. 这种泛化能力的提高得益于模型学习到的广泛和通用的特征表达,使得模型能够更好地适应新的、分布外的数据环境. MSC-Net 模型将训练分为 2 个阶段:第 1 阶段进行特征提取器的预训练,自监督学习作为预训练步骤是 MSC-Net 模型建立优势的核心机制,自监督预训练通过掩码重建,引导模型学习 TC4 合金不同组织之间的核心表征,构建领域不变特征空间,而非局限于实验室数据的分类特征;第 2 阶段将预训练完成后的编码器作为特征提取器使用,利用在实验室数据集上进行任务特异性优化,实现了对 TC4 合金金相组织的精确分类.

3 结 语

本研究提出了一种对于金相组织自动识别分类的 MSC-Net 模型,其展示了卓越的性能,在实验室验证集上实现了 100.0% 的准确率,并在现场泛化数据集上也保持了 87.6% 的高准确率,表明其优异的泛化能力. 该模型的成功主要归功于其自监督预训练策略,这种策略使得模型能够学习并提取广泛且通用的特征,使得预训练的编码器在后续作为特征提取器使用时表现出高效和精确的性能,从而有效避免过拟合现象. MSC-Net 模型与其他主流模型如 Inception v3、VGGNet、ResNet 及基于 Transformer 的 ViT 的比较显示,虽然这些模型在实验室验证集上也表现出较高的准确率,但在现场数据集上的泛化能力较 MSC-Net 模型有明显差距. 这进一步验证了 MSC-Net 模型在处理现场数据的变化时的强大能力,对于实际应用场景尤为关键. 通过将训练分为自监督学习和分类决策两个阶段, MSC-Net 模型能够更有针对性地优化性能,提高对特定任务的适应性和准确性. 这种分阶段的方法不仅提高了模型的学习效率,还增强了其在面对分布外数据环境时的应用潜力,为钛合金设备服役过程中的快速无损检测提供了新方法,开辟了新思路.

参考文献:

- [1] Leyens C, Peters M. Titanium and Titanium Alloys: Fundamentals and Applications[M]. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, 2003.
- [2] Zhao Q Y, Sun Q Y, Xin S W, et al. High-strength titanium alloys for aerospace engineering applications: A review on melting-forging process[J]. Materials Science and Engineering: A, 2022, 845: 143260.
- [3] Boyce A, Morton P H, Bell T. Surface Engineering of Titanium and Titanium Alloys[M]. Ohio, USA: ASM International, 1994.
- [4] Anene F A, Aiza Jaafar C N, Zainol I, et al. Biomedical materials: A review of titanium based alloys[J]. Journal of Mechanical Engineering Science, 2021, 235(19): 3792-3805.
- [5] Semiatin S L. An overview of the thermomechanical processing of α/β titanium alloys: Current status and future research opportunities[J]. Metallurgical and Materi-

- als Transactions A, 2020, 51(6): 2593-2625.
- [6] Pushp P, Dasharath S M, Arati C. Classification and applications of titanium and its alloys[J]. Materials Today: Proceedings, 2022, 54: 537-542.
- [7] Veiga C, Davim J P, Loureiro A J R. Properties and applications of titanium alloys: A brief review[J]. Reviews on Advanced Materials Science, 2012, 32(2): 133-148.
- [8] Florkiewicz W, Malina D, Tyliczszak B, et al. Sustainable Production: Novel Trends in Energy, Environment and Material Systems[M]. Berlin/Heidelberg, Germany: Springer International Publishing, 2020.
- [9] Masahashi N, Mizuhara Y, Matsuo M, et al. High temperature deformation behavior of titanium-aluminide based gamma plus beta microduplex alloy[J]. ISIJ International, 1991, 31(7): 728-737.
- [10] Sadeghpour S, Abbasi S M, Morakabati M, et al. Correlation between alpha phase morphology and tensile properties of a new beta titanium alloy[J]. Materials & Design, 2017, 121: 24-35.
- [11] Durmaz A R, Müller M, Lei B, et al. A deep learning approach for complex microstructure inference[J]. Nature Communications, 2021, 12(1): 6272.
- [12] Koyama M, Zhang Z, Wang M, et al. Bone-like crack resistance in hierarchical metastable nanolaminate steels[J]. Science, 2017, 355(6329): 1055-1057.
- [13] Scott D A, Schwab R. Introduction to Metallography [M]. Cham, Germany: Springer, 2019.
- [14] Holm E A, Cohn R, Gao N, et al. Overview: Computer vision and machine learning for microstructural characterization and analysis[J]. Metallurgical and Materials Transactions A, 2020, 51: 5985-5999.
- [15] 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.
- [16] Zubayer M H, Zhang C, Wang Y. Deep learning-based automatic defect detection of additive manufactured stainless steel[J]. Metals, 2023, 13(12): 1987.
- [17] Bartlett J L, Jarama A, Jones J, et al. Prediction of microstructural defects in additive manufacturing from powder bed quality using digital image correlation[J]. Materials Science and Engineering: A, 2020, 794: 140002.
- [18] Lu W Q, Liu Y J, Wu X, et al. Corrosion and passivation behavior of Ti-6Al-4V surfaces treated with high-energy pulsed laser: A comparative study of cast and 3D-printed specimens in a NaCl solution[J]. Surface and Coatings Technology, 2023, 470: 129849.
- [19] Rauch L, Kusiak J. Edge detection and filtering approach dedicated to microstructure images analysis[J]. Computer Methods in Materials Science, 2007, 7(2): 305-310.
- [20] Ji Z, Yang H, Li H W, et al. A new method for separating complex touching equiaxed and lamellar alpha phases in microstructure of titanium alloy[J]. Transactions of Nonferrous Metals Society of China, 2013, 23(8): 2265-2269.
- [21] Baskaran A, Kane G, Biggs K, et al. Adaptive characterization of microstructure dataset using a two stage machine learning approach[J]. Computational Materials Science, 2020, 177: 109593.
- [22] Nigam R, Khavala V B, Dash K, et al. Image-driven deep-learning-enabled automatic microstructural recognition[J]. Emerging Materials Research, 2022, 12(1): 47-51.
- [23] Kuchar D, Gogola P, Gabalcova Z, et al. Segmentation and classification of Zn-Al-Mg-Sn SEM BSE microstructure[J]. Applied Sciences, 2023, 13(2): 1045.
- [24] Goetz A, Durmaz A R, Müller M, et al. Addressing materials' microstructure diversity using transfer learning[J]. Npj Computational Materials, 2022, 8(1): 27.
- [25] Jung J, Na J, Park H K, et al. Super-resolving material microstructure image via deep learning for microstructure characterization and mechanical behavior analysis[J]. Npj Computational Materials, 2021, 7(1): 96.
- [26] 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.
- [27] Kim H, Inoue J, Kasuya T. Unsupervised microstructure segmentation by mimicking metallurgists' approach to pattern recognition[J]. Scientific Reports, 2020, 10(1): 17835.
- [28] 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.
- [29] Gou J, Shen J Q, Hu S S, et al. Microstructure and

- mechanical properties of as-built and heat-treated Ti-6Al-4V alloy prepared by cold metal transfer additive manufacturing[J]. *Journal of Manufacturing Processes*, 2019, 42: 41-50.
- [30] Wei X R, Zhou K, Feng S W, et al. Exploring fast-infering in transformer backbone model for fatigue crack detection and propagation tracking for proton exchange membrane[J]. *Journal of Power Sources*, 2023, 573: 233129.
- [31] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//*Advances in Neural Information Processing Systems 30*. Red Hook, USA: Curran Associates Inc., 2017: 5998-6008.
- [32] Yadav P, Saxena K K. Effect of heat-treatment on microstructure and mechanical properties of Ti alloys: An overview[J]. *Materials Today: Proceedings*, 2020, 26: 2546-2557.
- [33] Traore B B, Kamsu-Foguem B, Tangara F. Deep convolution neural network for image recognition[J]. *Ecological Informatics*, 2018, 48: 257-268.
- [34] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA: IEEE, 2016: 770-778.
- [35] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston, USA: IEEE, 2015: 1-9.
- [36] Graves A. *Supervised Sequence Labelling with Recurrent Neural Networks*[M]. Berlin, Germany: Springer, 2012.
- [37] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16×16 words: Transformers for image recognition at scale[C]//*Proceedings of the 9th International Conference on Learning Representations (ICLR)*. Vienna, Austria: PMLR, 2021: 1-21.
- [38] Liu Z, Lin Y T, Cao Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*. Montreal, Canada: IEEE, 2021: 10012-10022.

(责任编辑: 田 军)