附件 1:

浙江大学多学科交叉优秀博士学位论文申报表

论文题目	基于原位 SEM 的材料变形行为三维表征方法研究		
论文英文题目	Research on 3D Characterization Methods for Material Deformation Behavior Based on In-situ SEM		
作者学号	作者姓名	获得博士学位日期	现工作或学习单位
12360001	张雪成	2026 年 6 月 30 日	浙江大学材料学院硅及先进半导体材料全国重点实验室
作者电话	作者邮箱	论文涉及的研究方向	申报交叉组别①
13395710427	zhangxczju@zju.edu.cn	微观原位三维表征与测量	工程信息
作者主学科/专业代码	作者主学科/专业名称	作者交叉学科/专业代码	作者交叉学科/专业名称
085400	电子信息	080500	材料科学与工程
主导师工号	主导师姓名	主导师所在学院（系）	主导师研究方向
0019838	孟文超	控制科学与工程学院	工业具身智能、能源电力系统
交叉导师工号	交叉导师姓名②	交叉导师所在学院（系）	交叉导师研究方向
0021166	张跃飞	材料科学与工程学院	高温原位电镜智能表征
交叉导师工号	交叉导师姓名②	交叉导师所在学院（系）	交叉导师研究方向
0010292	张泽	材料科学与工程学院	电子显微学
申请学位时论文评阅结果		优秀（ 4 ）人；良好（ 1 ）人；其他（ 0 ）人	

作者攻博期间及获得博士学位后至申报通知发布之日，与博士学位论文密切相关的**浙江大学为第一单位**的代表性**交叉创新性成果**③

（1）论文④

发表学术论文题目	刊物/会议名称	发表或录用时间	作者排序（本人排序/总人数/导师排序）	论文级别/ 5 年平均影响因子/它引次数	学位论文对应章节
基于扫描电子显微镜的三维表面重建研究进展	电子显微学报	2025 年	1/4/3	CSCD/0.98/1 次	第一章

In Situ 3-D Surface Reconstruction for Metal Deformation Evolution via SEM	IEEE Transactions on Instrumentation and Measurement	2025 年	1/8/4	SCI/7.2/1 次	第二章
ETDMS: Efficient two-stage diffusion model for accelerated SEM image super-resolution	Ultramicroscopy	2025 年	1/8/4	SCI/2.7/2 次	第三章
Knowledge-distilled swin transformer for efficient vibration artifact removal in SEM imaging	Micron	2025 年	1/6/3	SCI/2.7/1 次	第三章
Superalloys fracture process inference based on overlap analysis of 3D models	Communications Engineering	2024 年	1/11/4	ESCI/7.1/9 次	第五章

（2）著作[5]

著作名称	出版社	出版时间	作者排序（本人排序/总人数/导师排序）	著作类型

（3）专利[6]

专利名称	专利类型	专利状态	申请时间	发明人排序（本人排序/总人数/导师排序）	可转化应用交易额（人民币）
一种合金断口三维重建与尺度校准方法	发明专利	授权	2023 年 10 月 16 日	1/7/2	参考价 20 万元
一种基于扫描电镜的原位三维表面重建方法	发明专利	授权	2025 年 3 月 11 日	1/6/3	参考价 30 万元

（4）承担或参与交叉科研项目[7]

参与科研项目名称	项目类型	项目起止时间	参与人排名	参与完成的任务

论文的交叉性和主要创新点	（请从论文选题、研究内容、解决问题的技术手段等方面阐述论文的交叉性） 本论文面向航空发动机等高端装备关键材料在极端服役条件下的微观变形与断裂机理研究需求，围绕原位扫描电子显微镜仅能进行二维表面观测、复杂环境下成像质量不稳定、大变形测量易失效以及材料内部损伤过程难以直接观测等问题，开展材料科学与实验力学、电子显微成像、控制工程、计算机视觉、人工智能及高性能计算的深度交叉研究。论文并非简单叠加不同学科方法，而是以材料损伤机理分析为目标，以原位 SEM 实验提供真实物理场景和数据，以三维视觉与数字图像相关实现定量测量，以深度学习和高性能计算解决复杂成像与极端变形条件下的可靠性和效率问题，构建了“三维形貌获取—动态伪影消除—三维变形测量—内部断裂反演”的全链条表征体系。 主要创新点包括：（1）提出基于四象限背散射电子信号的单视角光度立体重建方法，引入阈值触发的三次多项式补偿和抗噪多尺度残差金字塔积分策略，有效抑制高斜率形貌、阴影遮挡及低频漂移造成的重建误差，实现了从亚微米级滑移迹线到百微米级表面起伏的跨尺度三维形貌测量。（2）揭示低频环境振动与高频电子束行扫描共同造成 SEM 条纹伪影的时空耦合机制，构建真实物理激振数据集，提出融合方向性物理约束、知识蒸馏与通道剪枝的轻量化 SwinIR-KD 复原网络，在模型参数量减少 79%的同时将推理速度提升 2.4 倍，并形成可处理大视场图像的工程化软件。（3）提出物理驱动深度学习与传统高精度优化相融合的 E-DIC 方法，将晶体塑性有限元产生的物理变形场、SEM 真实纹理、深度特征匹配及 CPU/GPU 并行亚像素求解相结合，提高了极端大变形和局部非连续条件下的测量鲁棒性；进一步耦合二维位移场与三维高度场，实现了原位 SEM 条件下全场离面位移和微观有限应变的定量计算。（4）提出基于互补断口三维模型重叠分析的断裂过程逆向推演方法，通过多视角三维重建、无损绝对尺度标定和模型逐步分离，将断裂后的空间几何信息转化为损伤演化时序信息，揭示了高温条件下内部夹杂或孔洞诱发开裂早于表面晶界裂纹萌生的演化机制。 上述研究突破了传统原位 SEM 二维表面观测的限制，实现了从表面形貌、全场三维变形到内部损伤时序的连续定量表征，为高温合金等关键材料的损伤机理研究和高端装备安全评价提供了新的交叉研究方法与工程化工具。
学位论文作者或导师承诺意见	本人是该学位论文作者☑本人|☑导师，愿意推荐该学位论文参评浙江大学多学科交叉优秀博士学位论文，并承诺：以上所填全部信息及“代表性成果”证明材料准确无误、真实可靠；本学位论文不涉密，可在互联网上公开评审并全文公示；**该学位论文作者无学术失范或学术不端问题**。如信息不实，愿承担由此带来的一切后果和法律责任。 承诺人签名： 2026 年 10 月 16 日

所在学院（系）推荐意见	经学院（系）审查，以上承诺信息属实，同意推荐。 学院（系）负责人签字： 单位公章 2016 年 11 月 6 日
研究生院意见	分管领导签字： 年　月　日

填表要求：

①申报组别，分为以下五组，根据研究内容选填组别：

基础科学：基于数学、物理、化学等基础学科的交叉科学研究，面向国际科学前沿和国家重大需求，解决信息、生命、材料、能源、环境等领域的核心基础科学问题，取得重大突破或形成新的学科增长点。

工程信息：基于先进材料、现代工程、信息技术、农业等领域的交叉科学研究，面向国际科学前沿、国家重大需求和经济主战场，解决我国社会发展过程中相关的工程控制、先进制造等领域的关键科学和技术问题

生命科学：面向人民生命健康，基于理学、工学、医学、农学等领域的交叉科学研究，发展生物医学前沿技术方法，阐明生命相关复杂系统的多层次跨尺度相互作用与调控机制，揭示生命现象背后的科学规律和共性原理，探索生命健康研究新模式，应对人类健康与疾病防治中的重大挑战。

社会治理：基于自然科学的理论和方法，采用多学科会聚、多工具融合的交叉科学研究手段，研究人类社会发展过程中的资源开发利用、全球变化应对、人类文明演化、自然-社会耦合互馈等宏观复杂系统，解决人类可持续发展中的重大科学问题。

文化文明：采用多学科会聚、多工具融合的交叉科学研究手段，研究人类文化文明的重大难题和新兴问题，构建融合文化文明和自然科学的研究范式，形成与科学技术发展相协调的文化文明理论体系。

②如导师组涉及多位导师，可自行在表格添加交叉导师信息。

③“代表性成果”限填作者攻博期间及获得博士学位后一年内与博士学位论文密切相关、能反映学位论文水平、体现多学科交叉研究的成果。

④论文中论文级别分别为 SCI/SSCI/AHCI/ESCI/ CSCD/CSSCI，具体详见学校图书馆链接：https://libweb.zju.edu.cn/55981/list.htm

⑤发表著作中的著作类型分别为，专著/编著/编译/校注。

⑥技术专利的专利类型分别为，发明专利/实用新型和外观设计专利/计算机软件著作权。

⑦承担或参与科研项目的项目类型分别为，国家级/省部级/国际科技合作项目/横向科研项目；项目状态为，在研/结题。

文章编号: 1000-6281(2025)06-0714-12

基于扫描电子显微镜的三维表面重建研究进展

张雪成[1], 张跃飞[2*], 孟文超[3], 张 泽[2]
(1.浙江大学工程师学院,浙江 杭州 310015; 2.浙江大学材料科学与工程学院,浙江 杭州 310027;
3.浙江大学控制科学与工程学院,浙江 杭州 310027)

摘 要 扫描电子显微镜(SEM)凭借高分辨率、大景深、快速成像等优势,广泛应用于材料学和生物学研究中的微观表征。尽管 SEM 图像呈现一定的立体感,其本质仍是二维图像,缺乏直接的三维信息。为实现更直观的视觉表达并精确测量样品表面特性,基于二维 SEM 图像的三维表面重建已成为显微视觉领域的核心课题。本文系统综述该领域近年来的研究进展,涵盖设备开发、算法优化及应用探索,重点分析单视角与多视角两种三维重建方法的特点与进展。同时,剖析当前技术瓶颈,并展望未来研究方向,以期推动该领域的进一步发展。

关键词 扫描电子显微镜; 三维重建; 光度立体法; 运动恢复结构; 原位三维成像
中图分类号: TB302.3; TP391.4; O463; TG115.21 **文献标识码:** A
Doi: 10.3969/j.issn.1000-6281.2025.06.10

扫描电子显微镜(scanning electron microscope, SEM)因其高空间分辨率、大景深、快速成像和灵活的成像能力,在材料学[1-3]和生命科学[4-5]等领域得到广泛运用。其成像原理基于电子束与样品表面相互作用,激发二次电子(secondary electron, SE)、背散射电子(back-scattered electron, BSE)等物理信号,并通过探测器捕获这些信号生成图像[6]。然而,由于成像机制的特性,SEM 图像仅为散射电子强度的二维映射,而非样品表面的真实高度图。尽管图像呈现一定立体感,其本质仍不含三维信息。因此,将 SEM 图像与真实几何形状对比时,需谨慎解释结果并充分考虑其局限性。

为深入观察和精确测量微观样品的表面结构与形貌,从二维 SEM 图像重建三维模型已成为显微学领域的重要研究方向。该方法不仅能提供更直观的视觉信息,还可实现对表面特征的定量分析[7]。此外,SEM 的大样品腔室允许集成电子背散射衍射(electron backscatter diffraction, EBSD)、能谱仪(energy dispersive spectrometer, EDS)以及原位测试(*in-situ* test)设备[8]等,通过融合多源数据显著丰富三维模型的信息含量,这一优势是其他常规表征技术难以企及的。

在 SEM 中,三维模型的获取方法依据是否破坏样品可分为有损和无损两类[9]。有损方法包括连续块面法(serial block-face, SBF)[9]、金相连续抛光(metallographic serial polishing, MSP)[10]和聚焦离子束(focused ion beam, FIB)层析成像[11]等。这些方法均通过逐层切片对样品进行三维重建,但过程不可逆,区别在于切片方式的不同。此类技术适用于样品内部结构的重建,但耗时较长,需根据研究目标选择合适的切片方式,同时在切片时避免额外损伤,确保成本、时间、重建范围与精度的平衡。

无损方法利用三维计算机视觉算法分析处理二维图像[12],从而重建样品表面三维模型。其优点在于不破坏样品,便于后续分析,且重建范围广、速度快。然而,该方法无法获取内部信息,仅适用于表面重建,目前主要用于分析表面形貌或内部结构与表面关联较弱的样品。

本文重点探讨 SEM 无损表面重建技术的应用与发展前景[13]。近年来,随着软硬件升级和算法优化,基于 SEM 图像的三维表面重建在精度、纹理质量以及数据采集与重建效率上均取得显著进展[14]。本文将分析单视角与多视角两种 SEM 三维重建方法的原理、研究现状及应用场景。表 1 对比了两者在设备需求、图像获取和适用性等方面的异同。此外,针对当前的技术挑战,本文提出未来研究展望。

收稿日期: 2025-04-15; **修订日期:** 2025-06-04
基金项目: 国家科技重大专项资助项目(No.J2019-III-0008-0051).
作者简介: 张雪成(1999—),男(汉族),浙江人,博士研究生. E-mail: zhangxuecheng@zju.edu.cn
* **通讯作者:** 张跃飞(1974—),男(汉族),山西人,教授. E-mail: yfzhang76@zju.edu.cn

中国知网 https://www.cnki.net

表 1　基于单视角与多视角表面重建对比

Table 1　Comparison of surface reconstruction based on single-view versus multi-view

特　性	单视角重建	多视角重建
重建原理	光度立体法[15-16]	运动恢复结构[17]
探测器配置	多个探测器同时接收电子信号	单个探测器通过旋转与倾转样品台接收多角度信号
图像采集数量	与探测器数量相同(通常为 4 幅)	与表面形貌复杂度以及所需重建范围有关(通常多于 20 幅)[18-19]
图像采集时间	较短,可原位采集	较长,仅离位操作
重建范围	限于单一视角表面	通过多图像扩展范围(范围由探测器位置及样品台最大倾转角度决定)
图像对齐(匹配)	无需对齐(无需进行图像匹配)	需要对齐(需要检测图像特征点进行匹配以及探测器位置估计)
适用样品	光滑均匀表面	粗糙或具纹理表面

1　基于单视角图像的三维表面重建

1.1　光度立体法基本原理

单视角 SEM 三维表面重建基于光度立体法(photometric stereo, PS)的基本思想与流程,该方法由 Woodham 教授于 20 世纪 70 年代末提出[15-16]。在处理光学图像时,该方法通过同一视角下不同光照条件的图像,计算物体表面每个像素的法向向量,进而重建三维高度场。其前提包括:(1)相机与物体相对位置固定;(2)相机成像无畸变(如使用长焦或远心镜头);(3)光照方向和强度可控且均匀;(4)物体表面满足朗伯(Lambertian)反射特性。立体光度法的实现包含五个步骤[13]:(1)采集多光照条件下的单视角图像;(2)标定光照方向;(3)计算表面法向量;(4)估计反照率;(5)重建深度信息。

与光学成像不同,扫描电子显微镜(SEM)利用电子束与样品相互作用生成的高分辨率图像表征表面形貌。电子源发射的高能电子束经透镜聚焦至样品表面,激发二次电子(SE)和背散射电子(BSE)等信号。这些信号经探测器收集、处理和放大后,形成反映微观形貌的图像。由于 SEM 使用电子束而非光源,无法通过改变光照角度获取不同明暗图像。为此,可在 SEM 腔室内沿电子束方向对称布置多个背散射电子(BSE)探测器(如图 1 所示为四象限 BSE 探测器系统实物图)。图 2 所示在极坐标下四象限 BSE 探测器系统的几何形状[20]。该系统由环绕电子束路径的四个象限探测器组成,每个探测器捕获来自不同方向的背散射电子(BSE)。在扫描过程中,四个探测器同时或依次获取图像,每个图像反映了样品表面不同方向的 BSE 信号强度。通过比较对角探测器之间的信号差异,可以精确确定样品表面的斜率。随后,通过对这些斜率信息进行积分,可以重建样品的三维表面高度信息。这种方法不仅提高了重建的精确度,而且由于样品在采集过程中没有发生变化与移动,支持了原位时序数据的采集,允许对动态过程(如材料变化或生物样本的实时观察)进行实时三维成像。合理组合不同权重的图像数据(例如通过加权平均或其他算法)能进一步减小误差并提升重建的可靠性。需注意的是,探测器需精确对称布置,保持一致的灵敏度和增益,以确保信号差异准确反映样品表面的真实斜率[21]。否则,探测器对称性偏差会导致不同象限探测器的信号响应不一致,从而在计算斜率时引入误差,进而影响三维重建的准确性,例如可能出现表面扭曲或伪影。为解决这个问题,通常需要进行校准,使用三维校准标准(如球形段和阶梯金字塔结构)来调整探测器的增益、偏置和几何参数。除此之外,在实际计算过程中,背散射电子(BSE)的发射特性与传统光度立体法中的朗伯反射假设存在显著差异。BSE 信号的强度分布受样品材料属性(如原子序数和密度)、电子束入射角度以及表面几何形状的影响,且在复杂形貌下常伴随明显的阴影效应,导致其分布呈现各向异性,而非朗伯反射假设中的各向同性。因此,SEM 中应用单视角三维表面重建时,需对光度立体法进行调整以适应这些非理想条件。

1.2　相关研究进展

图像质量、空间分辨率、放大倍率以及探测器的位置与倾斜角度是决定 SEM 三维表面重建精度的关键因素。为提升光度立体法在 SEM 中的应用效果,研究人员通过算法优化与设备改进,开发了一系列技术和工具,推动了其在科学研究与工业领域的广泛应用。

在算法优化与设备开发方面,单视角 SEM 三维重建技术取得了诸多进展。1991 年,Beil 等[22]将传统立体学与基于阴影的最小二乘法相结合,采用动

中国知网　https://www.cnki.net

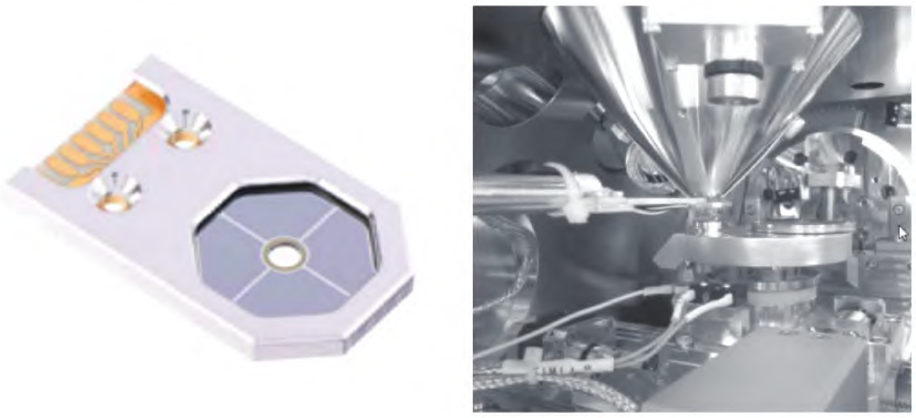


图 1 四象限 BSE 探测器系统。

Fig.1 Four-quadrant BSE detector system.

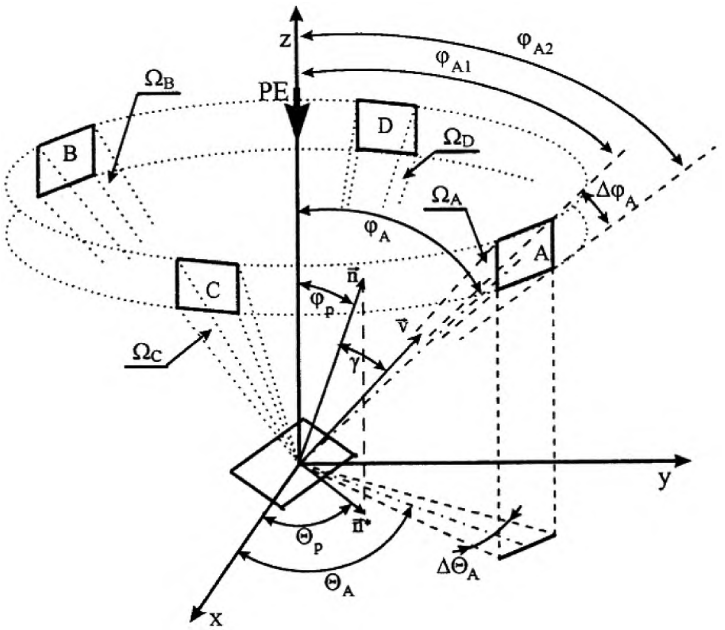


图 2 极坐标下四象限 BSE 探测器系统的几何形状[20]。

Fig.2 Geometry of the four-region BSE detector system in polar coordinates[20].

态编程技术和图像金字塔技术重建表面轮廓，有效缩短计算时间。2005 年，Drzazga 和 Paluszyński 等[20,23]同步采集硅板蚀刻痕迹图像，结合信号处理消除误差，基于阴影形状重建光滑三维形貌，将垂直误差降至视场尺寸的 10%以下。2008 年，Pintus 等[24-25]提出一种四源光度立体自动对准算法，充分考虑浮雕效应、阴影干扰及探测器非线性等光学与 SEM 成像特性，使图像更适配标准光度立体法。2010 年，Vynnyk 等[26]开发了一种高效数学模型，仅需两个对称探测器即可完成重建，优化了系统效率。在国内，赵以恒等[27]基于 STM32 微控制器设计了多路信号采集与生成模块，并利用 Qt(一种跨平台 C++图形用户界面开发框架) 开发 SEM 图像采集软件，显著提升了多通道信号处理能力，但在采用光栅样品进行精度验证中表现不佳。光栅样品高度为 1.5 μm，而在三维模型中为 1.1 μm，误差高达 26.7%。其采集图像及重建结果如图 3 所示，高误差主要源于计算过程中并没有针对解决光栅样品边界接近垂直的形貌特征所产生的严重阴影效应。

随着技术日趋成熟，SEM 三维重建的研究重点已从系统开发转向优化重建质量。值得注意的是，BSE 信号的发射特性受到材料属性、入射角度和几何形状的复杂影响，且阴影效应显著增强了信号的各向异性。当前研究中，2017 年，Miyamoto 等[28]提出一种引导去阴影与自校准方法，可自动校准梯度

中国知网 https://www.cnki.net

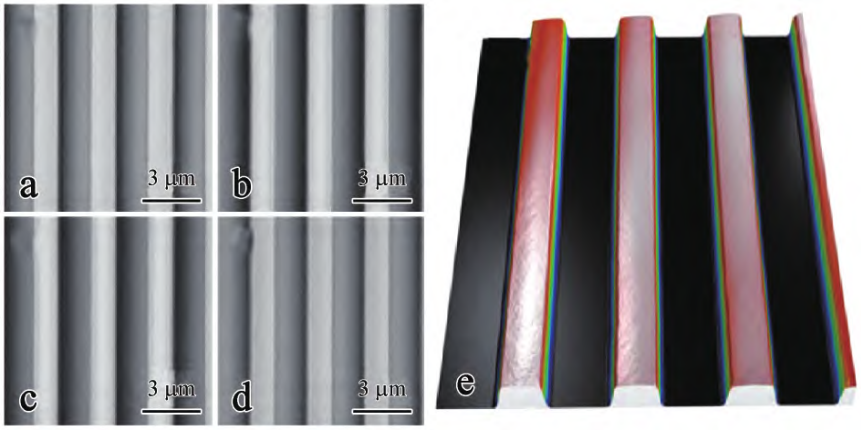


图 3　四象限 BSE 探测器采集图像及三维成像图[21]。

Fig.3　Four-quadrant BSE detector acquisition images and 3D imaging result[21].

测量参数,消除阴影误差,精确重建三维形状及无影图像。随后,该团队进一步将几何约束融入目标函数,解决复杂形状重建中的变形与不确定性问题[29]。2019 年,Borzunov 等[21] 基于探测器对 BSE 信号高度的响应特性,优化了光度立体法对表面梯度的灵敏度,提出了新型 BSE 探测器排列方式(如图 4a 所示)。该设计降低电子辐照对材料的影响,保持高信噪比,梯度灵敏度较传统四象限 BSE 探测器提升 1.7 倍。此外,他们采用简易最小二乘法优化图像数据,有效应对逆向问题,图 4b 为利用该探测器分布系统采集到的图像重建金字塔形样品的三维模型[30]。然而,对这些因素如何影响 BSE 三维重建精度的系统性分析仍较缺乏。例如,材料属性的变化可能导致反射强度的非线性分布,而阴影效应在复杂几何形状下的作用机制尚待深入研究。这些问题为后续算法优化提供了方向。

2020 年,Bayazid 等[31] 以 Ti-6Al-4V 球形颗粒为研究对象,分析放大倍率(M)、加速电压(V)和工作距离(WD)对重建效果的影响。结果表明,工作距离为 9~10 mm 时重建效果最佳;低倍率下提高放大倍率可显著提升图像质量,但超过阈值后质量改善有限,且牺牲扫描范围;加速电压升高则可能影响小特征暴露。WD 对通过 SEM 图像进行的三维重建过程是一个重要的影响参数。文中的最佳参数为放大倍率 1 000 倍、加速电压 3 kV,工作距离 9 mm。2021 年,Neggers 等[32] 开发了一种基于主成分分析的模态分解方法,将探测器数量与方法测定融入算法,省去校准步骤。文中选用六种具有不同形貌的样品作为示例案例以突出方法的适应性,其重建结果如图 5 所示。2023 年,Giardino 等[33] 引入正则化技术,解决图像噪声引发的表面伪影与误差,大幅提升重建质量。

光度立体法在多领域展现了广泛应用。例如,Wzorek 等[34] 利用该方法重建熔融 KOH-NaOH 共晶混合物中蚀刻坑的形状与深度分布,确定位错类型及位置,并与原子力显微镜结果对比验证。

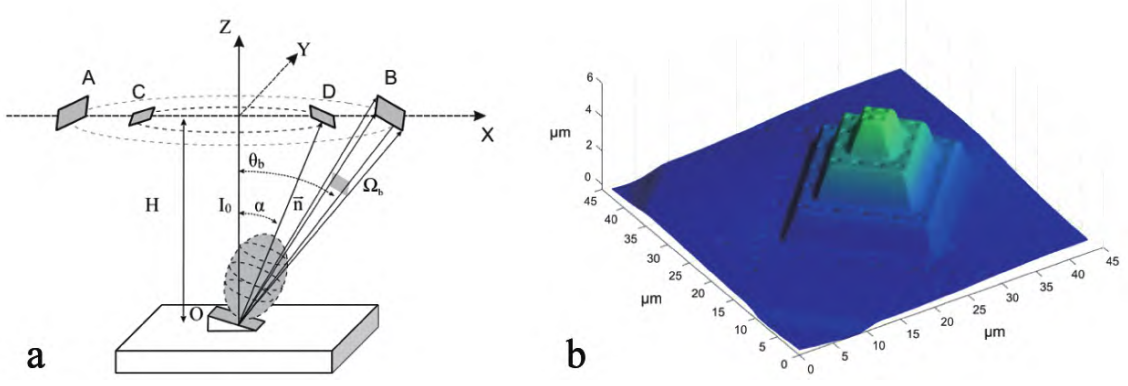


图 4　a. 针对表面梯度灵敏度改进的 BSE 探测器排列方式; b. 金字塔样品三维重建结果[21]。

Fig.4　a. Improved BSE detector arrangement for surface gradient sensitivity; b. Pyramid sample 3D reconstruction result[21].

中国知网　https://www.cnki.net

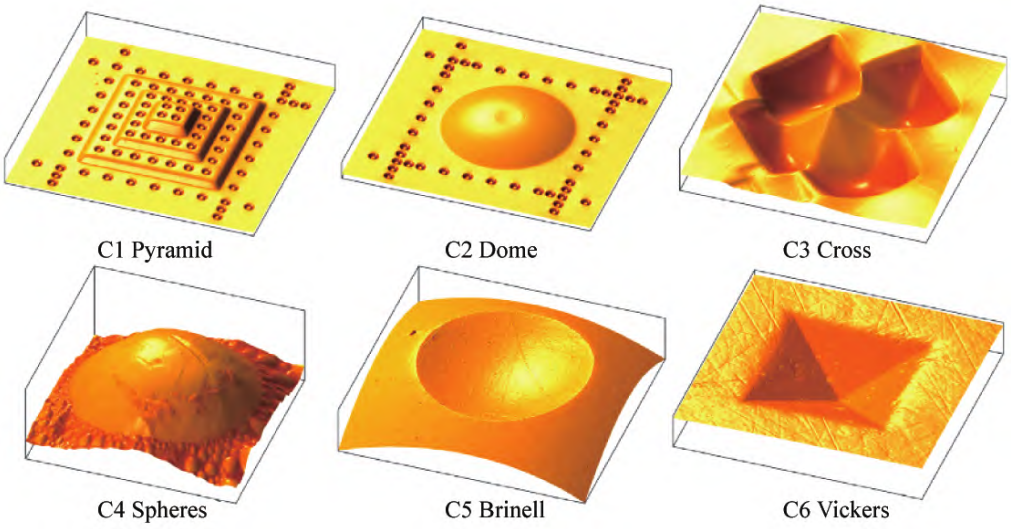


图 5 六种样品的重建地形图[32]。

Fig.5 Reconstructed topographies of the six samples[32].

Estellers 等[35]通过引入形状和平滑度先验知识优化重建流程,重现印刷电路硅晶圆表面结构,模拟制造过程变异性,用于检测故障与关键区域。2016 年,Hemmleb 等[36]开发 4Q-BSE 探测器,通过校准增益、偏移及探测器几何参数(半径、高度、面积和转角),分析合金断裂过程。Slówko 等[37]将多探测器技术应用于湿生物样品(如花瓣和猪关节软骨)的三维成像(如图 6 所示)。然而,其探测器系统厚度(12 毫米)过大,限制了高倍率成像时的工作距离。为此,他们设计了微型电离多探测器单元[38],作为低通量电子能量过滤器,支持表面下微观细节成像,满足生物医学领域的迫切需求。该系统可嵌入物镜孔内,不受工作距离限制,实现高倍率成像。

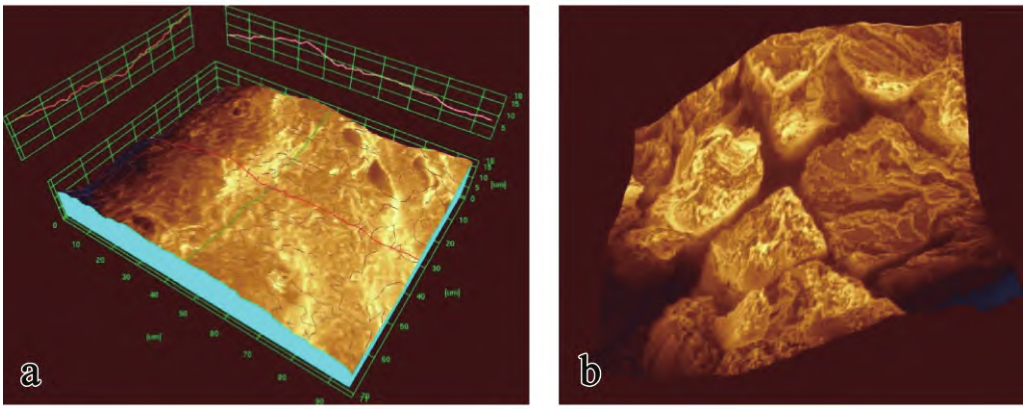


图 6 用柠檬酸消化的家猪关节软骨[37]。

Fig.6 Articular cartilage of the domestic pig digested with citric acid[37].

2 基于立体视觉的多视角表面重建

2.1 多视角重建方法基本原理

SEM 三维表面重建的另一重要方法是基于运动恢复结构(structure from motion, SfM)的多视角立体表面重建,运动恢复结构流程如图 7 所示。

相较于光度立体法,该方法能生成更完整的表面三维模型。通过移动探测器或旋转、倾斜样品,可有效获取多视角图像。对于只需重建复杂表面形貌的应用,SEM 多视角表面重建因成本低、操作简便,有望在某些场景下替代微米级断层扫描[17]。以下为其三维重建的主要流程。

(1) 获取 SEM 图像:通过 SEM 对样品表面进行多视角拍摄。由于电子束角度调节范围有限,通常通过旋转或倾斜样品台调整样品与探测器的相对角度,以采集重建所需的图像,确保每帧图像清晰。

中国知网 https://www.cnki.net

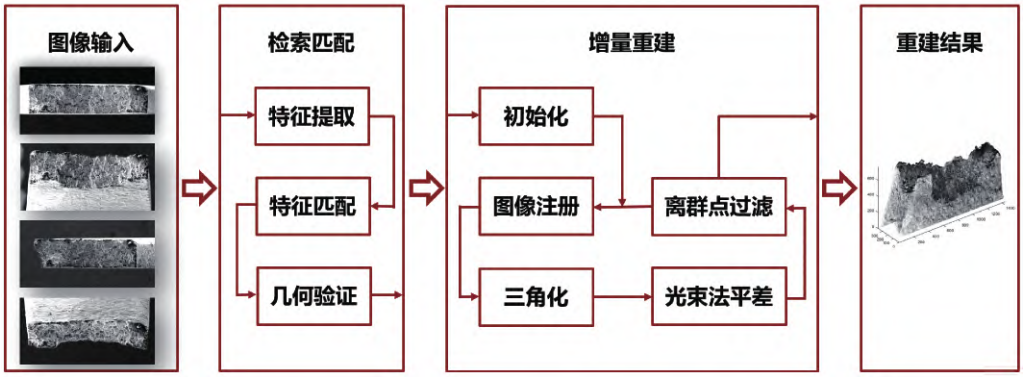


图 7　运动恢复结构流程。

Fig.7　Structure-from-Motion pipeline.

(2) 特征点识别与匹配: 采用尺度不变特征转换(scale-invariant feature transform, SIFT)等算法,识别并匹配图像对中的特征点。

(3) 点云计算: 基于特征点位置及图像中心坐标生成稀疏点云,再通过算法优化生成高密度点云,以捕获更多表面细节。

(4) 表面模型生成: 利用稠密点云数据构建表面网格模型,重现样品完整的三维形状。

(5) 后处理: 对模型进行去噪、平滑和修复等处理,以提升精度,并可添加纹理贴图,使模型更真实直观。

多视角表面重建在 SEM 三维重建中仍面临若干挑战。首先,原始图像中的噪声、过亮区域和多重阴影会干扰特征点识别与匹配,降低精度和可靠性。其次,该方法需依赖图像细节特征,因此不适合光滑表面的样品。常用解决方法是采用数字图像相关(digital image correlation, DIC)技术,通过散斑喷涂标记特征,但可能改变样品原始形貌。此外,过于复杂或明暗对比过强的表面可能导致探测器信息捕获不足,造成重建失败或结果失准。最后,图像采集过程中,样品需保持稳定,任何变形均可能引入时序误差。

2.2　相关研究进展

多视角表面重建的研究主要通过多图像特征点匹配,重建 SEM 样品表面的三维微观结构。SEM 图像为灰度图,研究中常使用成熟的光学摄影测量软件,如 123D Catch[18,39]、PhotoScan[40]、Geomagic Studio[17]、Metashape[41],生成高分辨率三维模型。Tondare 等[42]利用商业软件分析虚拟样品,发现该方法对粗糙表面表现优异,但在均匀低对比度区域高度误差高达 46%,需进一步优化。Tafti 等[13]于 2015 年综述相关算法,并开发了 3DSEM++方法[14],该方法通过整合机器学习、基于统计假设检验的 contrario 方法和极线几何,构建了一个自适应且智能的框架。特别值得一提的是,该方法采用了全局优化技术(差分进化算法),显著提高了 SEM 外参估计和 3D 表面建模的准确性,从而实现了高精度的重建结果。2017 年 Baghaie 等[43]引入能量最小化因子图和循环信念传播,提升微观样品重建质量。2019 年 Barbieri 等[17]以断层扫描为基准,比较了不同采集步长(9°和 27°)的误差,分别为 0.3%和 1.4%。2020 年,Kozikowski[41]利用鼠标宏控制样品台旋转、倾斜和图像采集,启发了高效采集策略,虽未详述实现细节。Toberg 等[44]对比正交与仿射相机模型对重建效果的影响。2022 年,Sihvonen 等[19]分析免费软件重建质量对模型纹理的影响,提出评估模型准确性的新方法。

多视角重建方法在生物[40]、化学[45-46]和材料[18,47]领域应用广泛。Omrani 等[48]采用多视角几何与全局优化方法,开展微观表面摩擦学分析。Gontard 等[39]量化图像采集与重建耗时,以 $LiTi_2(PO_4)_3$ 颗粒为例验证重建效果。次年,该团队优化图像采集策略[18](如图 8 所示),将传统的旋转样品台为主调整为倾转样品台为主,以适应粗糙表面高斜率形貌的重建与分析,增强梯度信息获取,并基于三维模型评估样品工程表面粗糙度(如图 9 所示)。值得注意的是,在后续工作中 Gontard 等成功融合了多视角重建与 EDS 元素分布图[49](如图 10 所示),实现了化学成分与三维形貌的同步可视化。

Raudoniene 等[46]开发了一种简便的湿化学合成方法,制备形状均一的 Ag_3PO_4 晶体,并通过三维重建生成晶体网格模型,辅助分析并提供直观的视觉表达。Tondare 等[50]利用 SfM 方法重建水族馆砾石颗粒的三维结构并分析体积,探索 SEM 在颗粒计

中国知网　https://www.cnki.net

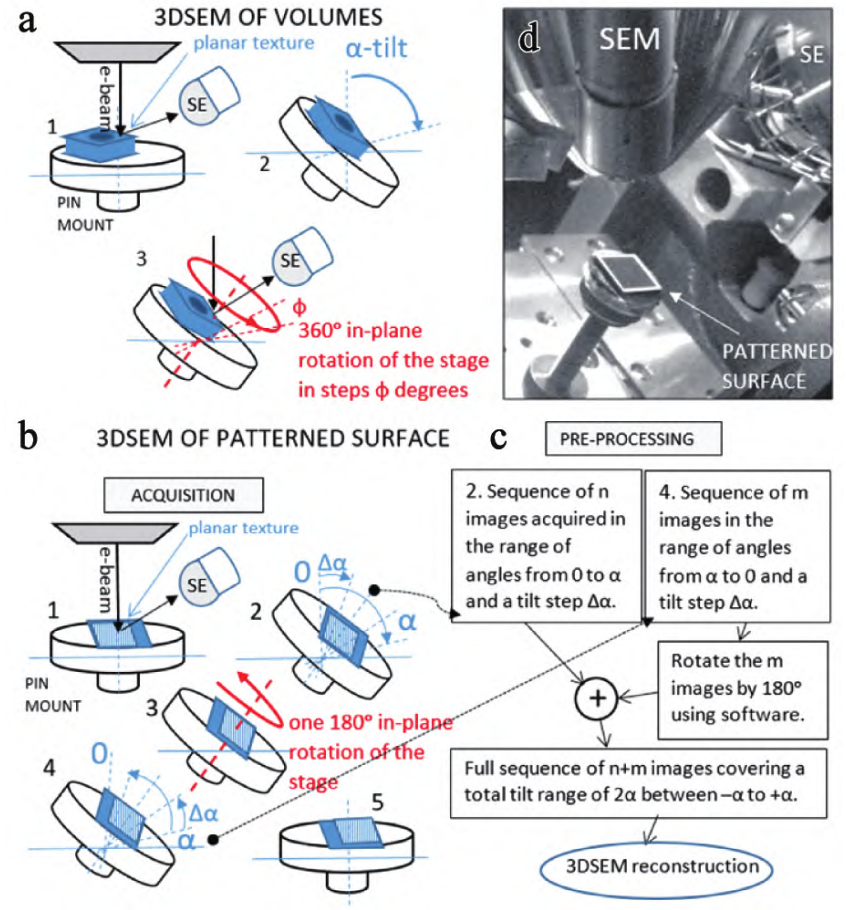


图 8 a. **旋转样品台策略**; b. **优化的倾转样品台策略**; c. **优化的倾转样品台策略流程图**; d. SEM **内部采集视角**[18]。

Fig.8 a. Rotation stage strategy; b. Optimized tilt stage strategy; c. Optimized tilt stage workflow; d. SEM internal view[18].

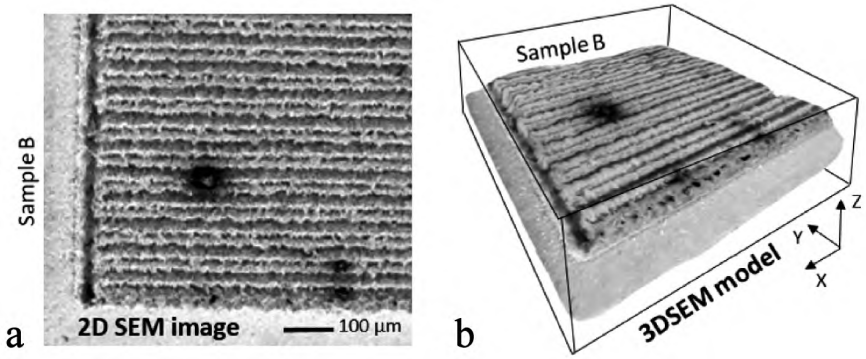


图 9 AISI 304L **不锈钢二维** SEM **图像及其三维重建结果**[18]。

Fig.9 2D SEM image and 3D reconstruction of AISI 304L stainless steel[18].

量学的潜力。Podor **等**[51] **在** SEM **中集成加热台，支持光束摇动模式，采集连续倾斜图像序列，实现高温下样品形貌的动态测量，适用于研究金属氧化、玻璃陶瓷结晶及高温机械测试中的裂纹扩展等过程。近期，**Zhang **等**[52] **提出一种集成框架，首次将多视角三维重建应用于合金断口分析，显著增强断面视觉信息并提供精准尺度测量，重建结果如图** 11 **所示。该框架通过三维模型重叠分析，推断内部断裂过程，精确识别裂纹起始点及扩展路径(如图** 12 **所示)，为材料研究开辟了新的观察视角。**

中国知网 https://www.cnki.net

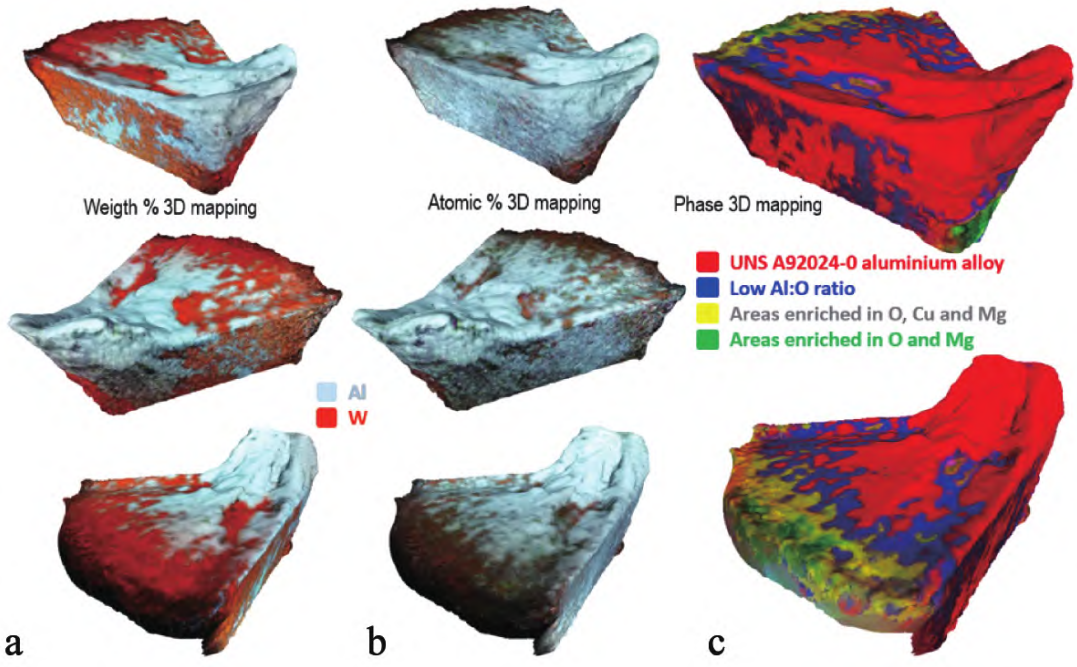


图 10　WC-Co 刀片表面的高分辨率的三维模型和化学信息[49]。

Fig.10　High-resolution 3D models and chemical information of the surface of a WC-Co insert[49].

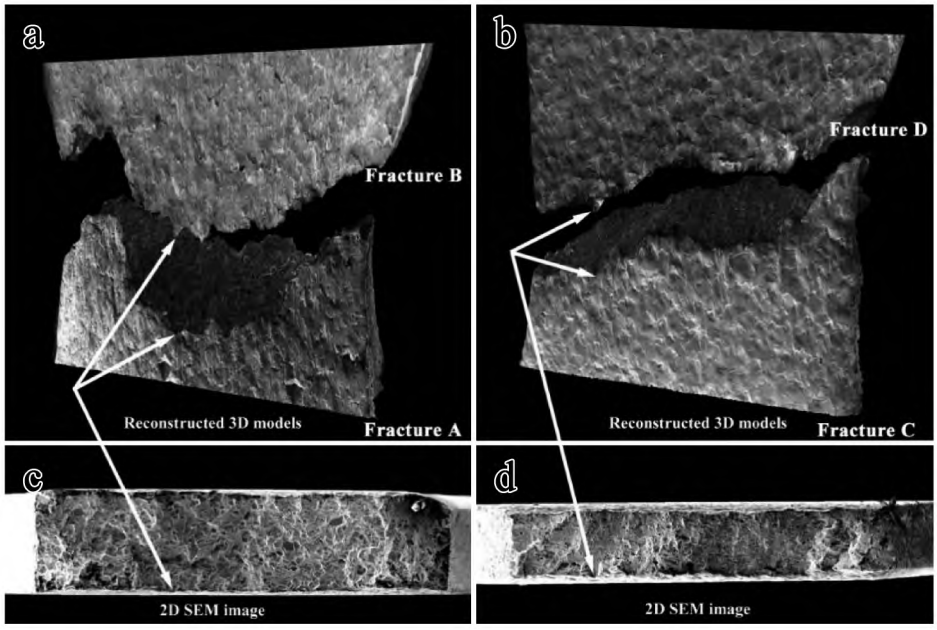


图 11　合金断口三维模型以及与二维图像对比结果[52]。

Fig.11　3D models of the alloy fracture and comparison with 2D images[52].

3　总结与展望

三维表面重建技术凭借其独特的优势，已成为扫描电子显微学领域的重要研究方向。从设备层面看，多探测器系统与智能化算法的结合显著提升了数据采集效率与重建精度；在信息维度上，该技术突破了传统二维成像的局限，实现了表面形貌的量化分析与动态信息捕获；算法优化进一步增强了复杂结构解析能力，实现高保真度模型重建。目前，该技术已广泛应用于材料科学（如金属断口分析、半导体器件检测）、生物医学（如细胞动态观测、组织形貌表征）以及工业检测（如表面粗糙度评估、微观缺陷识别）等领域。然而，尽管已取得显著进展，当前技术仍面临诸多挑战，亟待进一步突破。

中国知网　https://www.cnki.net

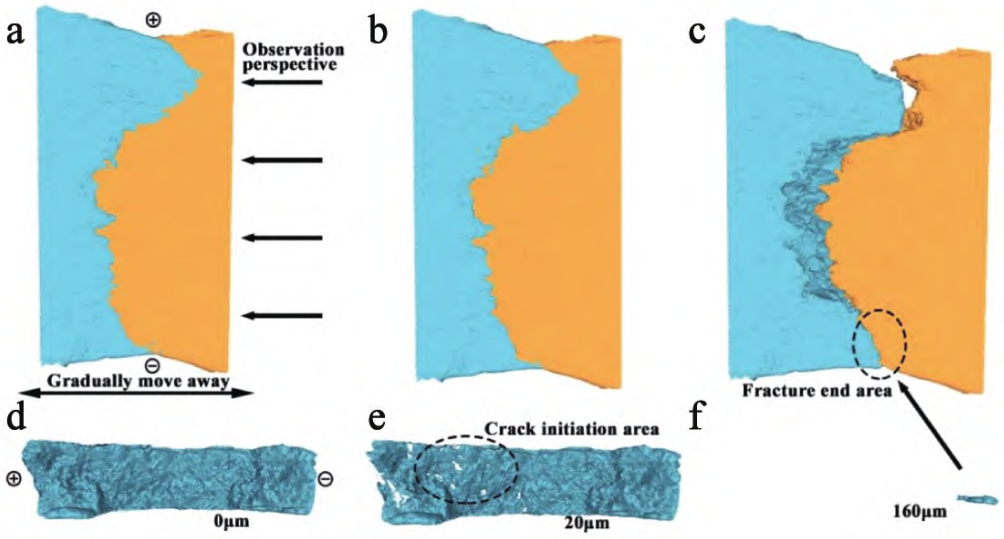


图 12 三维模型重叠分析的实施过程[53]。

Fig.12 Implementation process of 3D models overlap analysis[53].

3.1 仍存在的问题和挑战

尽管 SEM 三维表面重建技术取得显著进展，仍面临诸多挑战。首先，成像质量限制了重建效果。无论是单视角方法中的多探测器设计，还是多视角方法中的样品台旋转与倾转，重建算法均仅适用于样品完全可见的区域。对于复杂样品，遮挡或不可见区域会导致视野缺失，降低模型精度与完整性。其次，图像质量受多种干扰源（漂移、振动、噪声）影响的问题依然存在。对于广泛使用的 BSE 探测器（单视角），其信号的非理想特性（受材料、角度、几何形状影响）和各向异性阴影效应仍是提升精度的核心障碍，相关物理机制对重建误差的量化研究仍需深入。

在图像数据采集方面，多视角重建通过增加图像数量可提升精度[17]，但受限于样品台角度范围与采集时间，难以无限扩展数据量。因此，需优化采集策略，以最少图像实现最优重建效果。此外，当前数据采集效率较低，亟需高效精准的自动化方案。另一方面，用于三维重建的 SEM 图像数据集资源匮乏。Tafti 等[53]虽公开了 3DSEM 数据集，但其图像数量与种类有限，难以满足算法优化需求。为进一步提升重建算法的准确性、可靠性和效率，亟需构建更丰富的高质量数据集。

值得一提的是，在 SEM 三维表面重建方法发展的过程中，似乎都缺乏对于重建分辨率以及精度验证的详细描写，不同的研究工作中对于误差的判断方法存在较大差异性，缺乏一套行之有效的评判标准与测试流程。此外，当前 SEM 三维表面重建研究主要聚焦静态形貌，无法满足特定实验中对动态形貌变化的实时表征需求，限制了其在动态过程分析中的应用。

3.2 未来研究展望

3.2.1 多视角图像自动采集方案

多视角重建需采集大量不同角度的图像，开发自动化采集系统尤为关键。通过宏程序或机器人控制样品台运动，可大幅降低手动操作负担。例如，Kozikowski 等[41]利用鼠标宏实现图像自动采集，验证了其可行性。未来，可进一步研发智能化采集策略，优化图像数量与角度选择，提升效率与精度。

3.2.2 原位三维成像

原位热机械实验中的动态三维重建可实现样品在模拟服役条件下的时空多维分析。其实现需克服动态加载与高温等环境下图像质量下降的难题，同时针对材料变形、断裂等具体场景，深入探索相关科学问题。

3.2.3 SEM-3D-DIC 的应用

三维数字图像相关技术（three-dimensional digital image correlation，3D-DIC）在宏观场景中已广泛应用并持续发展，但在 SEM 微观观测中仍局限于二维分析。开发结合 SEM 三维模型的 SEM-3D-DIC 技术，可实现微观尺度下材料变形与应变的精准测量，为材料科学提供广阔应用前景。

3.2.4 基于三维模型的深度学习预测

利用深度学习分析 SEM 三维模型，结合原位实

中国知网 https://www.cnki.net

验的时序数据,可预测材料性能或失效模式,如基于表面拓扑估计抗拉强度或疲劳寿命。但这依赖于大量高质量三维 SEM 数据集的支撑。未来,可通过开源社区合作,构建丰富的数据集,拓展其在材料设计与失效分析中的应用。

参考文献:

[1] KABOLI S, NOEL P, CLÉMENT D, et al. On high-temperature evolution of passivation layer in Li-10 wt.% Mg alloy via *in situ* SEM-EBSD [J]. Science Advances, 2020, 6(50): eabd5708.

[2] **张昊翔, 丁青青, 姚霞, 等. 热处理对一种新型析出强化型奥氏体不锈钢显微组织与性能的影响** [J]. **电子显微学报**, 2024, 43(5): 540-548.

[3] **李菲祺, 吕俊霞, 程晓鹏, 等. 不同取向镍基单晶高温合金** 650 ℃ **原位拉伸** [J]. **电子显微学报**, 2024, 43(3): 295-302.

[4] TATE M L K. A new age in scanning electron microscopy: Applications in the life sciences [J]. Science, 2017, 358(6368): 1344.

[5] BILEN B, SENER L T, ALBENIZ I, et al. Determination of ultrastructural properties of human carotid atherosclerotic plaques by scanning acoustic microscopy, Micro-computer tomography, scanning electron microscopy and energy dispersive X-ray spectroscopy [J]. Scientific Reports, 2019, 9: 679.

[6] VERNON-PARRY K D. Scanning electron microscopy: An introduction [J]. III-Vs Review, 2000, 13(4): 40-44.

[7] LIMANDRI S, GALVÁN JOSA V, VALENTINUZZI M C, et al. 3D scanning electron microscopy applied to surface characterization of fluorosed dental enamel [J]. Micron, 2016, 84: 54-60.

[8] BOEHLERT C J, LI H, WANG L, et al. Slip system characterization of inconel 718: Using *in-situ* scanning electron microscopy [J]. Advanced Materials & Processes, 2010, 168(11/12): 41-46.

[9] LU J, CHEN B, LIU X, et al. 3D microstructure reconstruction of casting aluminum alloy based on serial block-face scanning electron microscopy [J]. Journal of Alloys and Compounds, 2019, 778: 721-730.

[10] JUNG J M, YOO J H, JEONG H J, et al. Three-dimensional characterization of SiC particle-reinforced Al composites using serial sectioning tomography and thermo-mechanical finite element simulation [J]. Metallurgical and Materials Transactions A, 2014, 45(12): 5679-5690.

[11] YUAN H, DE MOORTÈLE B V, EPICIER T. Accurate post-mortem alignment for focused ion beam and scanning electron microscopy (FIB-SEM) tomography [J]. Ultramicroscopy, 2021, 228: 113265.

[12] VERHOEVEN G. Taking computer vision aloft-archaeological three-dimensional reconstructions from aerial photographs with photoscan [J]. Archaeological Prospection, 2011, 18(1): 67-73.

[13] TAFTI A P, KIRKPATRICK A B, ALAVI Z, et al. Recent advances in 3D SEM surface reconstruction [J]. Micron, 2015, 78: 54-66.

[14] TAFTI A P, HOLZ J D, BAGHAIE A, et al. 3DSEM++: Adaptive and intelligent 3D SEM surface reconstruction [J]. Micron, 2016, 87: 33-45.

[15] WOODHAM R J. Photometric stereo: A reflectance map technique for determining surface orientation from image intensity [C]//Image Understanding Systems and Industrial Applications I: vol. 0155. SPIE, 1979: 136-143.

[16] WOODHAM R J. Photometric method for determining surface orientation from multiple images [J]. Optical Engineering, 1980, 19(1): 139-144.

[17] BARBIERI G, DA SILVA F P. Acquisition of 3D models with submillimeter-sized features from SEM images by use of photogrammetry: A dimensional comparison to microtomography [J]. Micron, 2019, 121: 26-32.

[18] GONTARD L C, LÓPEZ-CASTRO J D, GONZÁLEZ-ROVIRA L, et al. Assessment of engineered surfaces roughness by high-resolution 3D SEM photogrammetry [J]. Ultramicroscopy, 2017, 177: 106-114.

[19] SIHVONEN T, REINIKAINEN S P. Image based evaluation of textured 3DSEM models [J]. Ultramicroscopy, 2022, 238: 113518.

[20] DRZAZGA W, PALUSZYNSKI J, SLOWKO W. Three-dimensional characterization of microstructures in a SEM [J]. Measurement Science and Technology, 2006, 17(1): 28-31.

[21] BORZUNOV A A, KARAULOV V Y, KOSHEV N A, et al. 3D surface topography imaging in SEM with improved backscattered electron detector: Arrangement and reconstruction algorithm [J]. Ultramicroscopy, 2019, 207: 112830.

[22] BEIL W, CARLSEN I C. Surface reconstruction from stereoscopy and "shape from shading" in SEM images [J]. Machine Vision and Applications, 1991, 4(4): 271-285.

[23] PALUSZYŃSKI J, SLÓWKO W. Surface reconstruction with the photometric method in SEM [J]. Vacuum, 2005, 78(2): 533-537.

中国知网 https://www.cnki.net

[24] PINTUS R, PODDA S, VANZI M. An automatic alignment procedure for a four-source photometric stereo technique applied to scanning electron microscopy [J]. IEEE Transactions on Instrumentation and Measurement, 2008, 57(5): 989-996.

[25] PINTUS R, PODDA S, VANZI M. Improvements in automated photometric stereo 3D SEM [J]. Microscopy and Microanalysis, 2008, 14(S2): 608-609.

[26] VYNNYK T, SCHULTHEIS T, FAHLBUSCH T, et al. 3D-measurement with the stereo scanning electron microscope on sub-micrometer structures [J]. Journal of the European Optical Society - Rapid Publications, 2010, 5(10038s). DOI: 10.2971/jeos.2010.10038s.

[27] **赵以恒，张宜旭，刘陵恩，等. 扫描电镜多角度图像采集系统研究** [J]. **电子显微学报**, 2024, 43(6): 701-710.

[28] MIYAMOTO A, CHEN D, KANEKO S. Bootstrapping de-shadowing and self-calibration for scanning electron microscope photometric stereo [J]. Measurement Science and Technology, 2014, 25(10): 105402.

[29] MIYAMOTO A, MATSUSE H, KOUTAKI G. Robust surface reconstruction by design-guided SEM photometric stereo [J]. Measurement Science and Technology, 2017, 28(4): 045405.

[30] BORZUNOV A A, LUKYANENKO D V, RAU E I, et al. Reconstruction algorithm of 3D surface in scanning electron microscopy with backscattered electron detector [J]. Journal of Inverse and Ill-posed Problems, 2021, 29(5): 753-758.

[31] BAYAZID S M, BRODUSCH N, GAUVIN R. Investigation of the effect of magnification, accelerating voltage, and working distance on the 3D digital reconstruction techniques [J]. Scanning, 2020, 2020: e3743267.

[32] NEGGERS J, HÉRIPRÉ E, BONNET M, et al. Principal image decomposition for multi-detector backscatter electron topography reconstruction [J]. Ultramicroscopy, 2021, 227: 113200.

[33] GIARDINO M, MENON D M N, JANNER D L. Regularization techniques for 3D surface reconstruction from four quadrant backscattered electron detector images [J]. Ultramicroscopy, 2023, 250: 113746.

[34] WZOREK M, CZERWINSKI A, RATAJCZAK J, et al. Depth measurements of etch-pits in GaN with shape reconstruction from SEM images [J]. Journal of Microscopy, 2010, 237(3): 242-245.

[35] ESTELLERS V, THIRAN J P, GABRANI M. Surface reconstruction from microscopic images in optical lithography [J]. IEEE Transactions on Image Processing, 2014, 23(8): 3560-3573.

[36] HEMMLEB M, BETTGE D, DRIEHORST I, et al. 3D surface reconstruction with segmented BSE detector: New improvements and application for fracture analysis in SEM [C]//European Microscopy Congress 2016: Proceedings. John Wiley & Sons, Ltd, 2016: 489-490.

[37] SLÓWKO W, WIATROWSKI A, KRYSZTOF M. Detection of secondary and backscattered electrons for 3D imaging with multi-detector method in VP/ESEM [J]. Micron, 2018, 104: 45-60.

[38] SLÓWKO W. Specific features of the miniature ionisation BSE multi-detector unit for 3D imaging in environmental conditions [J]. Micron, 2019, 126: 102752.

[39] GONTARD L C, SCHIERHOLZ R, YU S, et al. Photogrammetry of the three-dimensional shape and texture of a nanoscale particle using scanning electron microscopy and free software [J]. Ultramicroscopy, 2016, 169: 80-88.

[40] EULITZ M, REISS G. 3D reconstruction of SEM images by use of optical photogrammetry software [J]. Journal of Structural Biology, 2015, 191(2): 190-196.

[41] KOZIKOWSKI P. Extracting Three-dimensional information from SEM images by means of photogrammetry [J]. Micron, 2020, 134: 102873.

[42] TONDARE V N, VILLARRUBIA J S, VLADÁR A E. Three-dimensional (3D) nanometrology based on scanning electron microscope (SEM) stereophotogrammetry [J]. Microscopy and Microanalysis, 2017, 23(5): 967-977.

[43] BAGHAIE A, TAFTI A P, OWEN H A, et al. SD-SEM: Sparse-dense correspondence for 3D reconstruction of microscopic samples [J]. Micron, 2017, 97: 41-55.

[44] TOBERG S, REITHMEIER E. Quantitative 3D reconstruction from scanning electron microscope images based on affine camera models [J]. Sensors, 2020, 20(12): 3598.

[45] CEGUERRA A V, DAY A C, RINGER S P. Assessing the spatial accuracy of the reconstruction in atom probe tomography and a new calibratable adaptive reconstruction [J]. Microscopy and Microanalysis, 2019, 25(2): 309-319.

[46] RAUDONIENE J, SKAUDZIUS R, ZARKOV A, et al. Wet-chemistry synthesis of shape-controlled Ag_3PO_4 crystals and their 3D surface reconstruction from SEM imagery [J]. Powder Technology, 2019, 345: 26-34.

[47] KANG K W, PEREDA M D, CANAFOGLIA M E, et al. Uncertainty studies of topographical measurements on steel surface corrosion by 3D scanning electron microscopy [J]. Micron, 2012, 43(2): 387-395.

中国知网 https://www.cnki.net

[48] OMRANI E, TAFTI A P, FATHI M F, et al. Tribological study in microscale using 3D SEM surface reconstruction [J]. Tribology International, 2016, 103: 309-315.

[49] GONTARD L C, BATISTA M, SALGUERO J, et al. Three-dimensional chemical mapping using non-destructive SEM and photogrammetry [J]. Scientific Reports, 2018, 8(1): 11000.

[50] TONDARE V N. A concept for three-dimensional particle metrology based on scanning electron microscopy and structure-from-motion photogrammetry [J]. Journal of Research of the National Institute of Standards and Technology, 2020, 125: 125014.

[51] PODOR R, GOFF X L, LAUTRU J, et al. Direct observation of the surface topography at high temperature with SEM [J]. Microscopy and Microanalysis, 2020, 26(3): 397-402.

[52] ZHANG X, GUO G, LI Z, et al. Superalloys fracture process inference based on overlap analysis of 3D models [J]. Communications Engineering, 2024, 3(1): 1-10.

[53] TAFTI A P, KIRKPATRICK A B, HOLZ J D, et al. 3DSEM: A 3D microscopy dataset [J]. Data in Brief, 2016, 6: 112-116.

Review of 3D surface reconstruction based on scanning electron microscopy

ZHANG Xuecheng[1], ZHANG Yuefei[2*], MENG Wenchao[3], ZHANG Ze[2]

(1.Polytechnic Institute, Zhejiang University, Hangzhou Zhejiang 310015; 2.School of Materials Science and Engineering, Zhejiang University, Hangzhou Zhejiang 310027; 3.College of Control Science and Engineering, Zhejiang University, Hangzhou Zhejiang 310027, China)

Abstract The scanning electron microscope (SEM) is widely used in materials science and biology research due to its high resolution, large depth of field, and fast imaging capabilities. Although SEM images exhibit a sense of three-dimensionality, they are inherently two-dimensional and lack direct three-dimensional information. To achieve more intuitive visualization and accurate measurement of surface features, three-dimensional surface reconstruction from two-dimensional SEM images has become a key research focus in microscopic imaging. This paper presents a comprehensive review of recent advances in this field, including developments in instrumentation, algorithm optimizations, and emerging applications. Particular emphasis is placed on the feature and progress of two three-dimensional reconstruction methods: Single-view and multi-view. Current technical challenges are also discussed, along with future research prospects, to support continued progress in this field.

Keywords scanning electron microscope; three-dimensional reconstruction; photometric stereo; structure from motion; *in-situ* three-dimensional imaging

* Corresponding author

中国知网 https://www.cnki.net

In Situ 3-D Surface Reconstruction for Metal Deformation Evolution via SEM

Xuecheng Zhang, Bin Zhang, Guanghao Guo, Wenchao Meng, *Senior Member, IEEE*, Xianjue Ye, Yuefei Zhang, Ze Zhang, and Jiming Chen, *Fellow, IEEE*

***Abstract*—This article proposes an enhanced submicrometer-resolution 3-D surface reconstruction framework for in situ scanning electron microscopy (SEM) characterizing metal deformation evolution, establishing the first practical solution for spatiotemporal 3-D surface analysis in dynamic polycrystalline systems. The methodology integrates an optimized gradient computation algorithm with a multiscale reconstruction (MSR) strategy, specifically addressing high-slope surface recovery challenges under mechanical loading. Key innovations include a cubic compensation term for mitigating steep-region signal occlusion and a hierarchical Gaussian pyramid architecture that preserves more microscopic features through residual propagation while suppressing measurement noise. Experimental validation on IN718 superalloy demonstrates dual-scale characterization capabilities: submicrometer-resolution reconstructions during early plastic deformation capture slip trace activation dynamics and cross-slip system interactions. In contrast, large-area reconstructions in later stages quantitatively reveal grain boundary collapse patterns and strain gradient distributions preceding crack nucleation. The demonstrated approach bridges slip behavior and dual-scale deformation mechanisms, offering significant implications for material design and failure prediction.**

***Index Terms*—3-D surface reconstruction, cross-scale characterization, in situ scanning electron microscopy (SEM) experiments, metal deformation behavior, slip trace analysis.**

I. Introduction

IN RECENT years, in situ scanning electron microscopy (SEM) experiments have gained significant attention for their ability to observe the dynamic behavior of materials under simulated service conditions such as mechanical loading and thermal variations [1], [2]. Based on the powerful characterization capabilities and extensive extensibility of SEM, this technique provides direct visualization across scales from nanometers to millimeters, offering valuable insights into the fundamental mechanisms of material performance. Particularly in metal material research, in situ SEM is widely used to study mechanical properties and deformation behaviors, such as observing microstructural evolution [3], crack initiation and propagation [4], and phase transformation [5]. These studies provide important foundations for material design and optimization in engineering applications.

Due to the imaging mechanism of SEM, it can only generate 2-D information, making it challenging to capture the dynamic changes in the 3-D morphology of the sample surface [6], [7]. For instance, it cannot quantitatively describe the height variations of slip traces or the strain distribution at grain boundaries. This is especially critical in in situ experiments, where the surface morphology undergoes significant changes during deformation, such as the appearance of slip traces in the early stages of plastic deformation and overall undulation due to extensive slip in the later stages. Therefore, there is an urgent need to develop methods for reconstructing the 3-D surface morphology from SEM images to supplement spatial information and enhance analytical capabilities [8].

SEM-based 3-D reconstruction methods include multiview and single-view approaches [9]. Multiview 3-D reconstruction requires tilting and rotating the sample stage to obtain images from different perspectives, enabling large-scale 3-D surface reconstruction [7]. However, this method is limited by the size of the SEM chamber and changes in loading conditions during sample stage tilting, making it less suitable for in situ experiments. An alternative solution is single-view 3-D reconstruction using a four-quadrant backscattered electron (4Q-BSE) detector [10], [11]. The 4Q-BSE detector captures backscattered electron (BSE) signals from different angles, deriving surface gradients and thus reconstructing the sample's 3-D surface height field. This method, often called shape from shading (SFS), has made progress in instrument equipment and algorithm improvement [6], [11]. Nevertheless, the existing application-oriented research [12] still reveals critical limitations when handling complex morphological features.

Specifically, single-view 3-D reconstruction methods based on the 4Q-BSE detector face persistent challenges in high-slope morphology scenarios, where signal occlusion significantly degrades reconstruction accuracy [13]. Signal occlusion occurs when parts of the sample block the path of BSEs, leading to incomplete or inaccurate data in some areas of the detector, which can result in artifacts and distortions in the reconstructed 3-D surface [14]. Additionally, research in the SEM single-view 3-D reconstruction area is limited, and literature on such methods often consists of method-

Received 31 March 2025; revised 9 July 2025; accepted 3 August 2025. This work was supported in part by the National Natural Science Foundation of China under Grant U24A20258 and Grant 92367205. The Associate Editor coordinating the review process was Dr. Maryam Shamgholi. *(Corresponding authors: Wenchao Meng; Yuefei Zhang.)*

Xuecheng Zhang, Wenchao Meng, and Jiming Chen are with the College of Control Science and Engineering, Zhejiang University, Hangzhou, Zhejiang 310027, China (e-mail: zhangxuecheng@zju.edu.cn; wmengzju@zju.edu.cn; cjm@zju.edu.cn).

Bin Zhang, Guanghao Guo, Xianjue Ye, Yuefei Zhang, and Ze Zhang are with the Institute of Superalloys Science and Technology, School of Materials Science and Engineering, Zhejiang University, Hangzhou, Zhejiang 310027, China (e-mail: msezhangbin@zju.edu.cn; guogh@zju.edu.cn; yexianjue@zju.edu.cn; yfzhang76@zju.edu.cn; zezhang@zju.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/TIM.2025.3604113, provided by the authors.

Digital Object Identifier 10.1109/TIM.2025.3604113

1557-9662 © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

ological articles lacking precision validation and with limited application scope. This highlights the necessity for systematic accuracy assessment and practical implementations in critical domains like in situ SEM thermomechanical experiments.

This article proposes a 3-D surface reconstruction method for in situ SEM experiments, integrating an optimized gradient calculation algorithm with multiscale reconstruction strategies to address critical challenges in metallic deformation analysis. This framework provides a versatile tool for elucidating complex deformation mechanisms in a wide range of metallic materials, thereby contributing significantly to the rational design and optimization of advanced engineering alloys and enhancing the predictability of their long-term performance. The main contributions of this work are as follows.

1) We establish the first practical framework for 3-D surface reconstruction in in situ SEM experiments of metallic materials, enabling spatiotemporal analysis of metal deformation behavior.
2) We propose an enhanced 4Q-BSE-based 3-D reconstruction algorithm that effectively resolves signal occlusion issues caused by high-slope surfaces, significantly improving reconstruction accuracy compared to conventional methods.
3) Our methodology achieves cross-scale observation capabilities (from sub-micrometer slip traces to 100-μm grain clusters), allowing systematic investigation of microstructure evolution across different deformation stages.

II. Related Works

A. In Situ SEM Experiments on Metallic Materials

In situ SEM integrates multifield coupling modules to accurately simulate service conditions (e.g., thermal cycling and mechanical loading) within vacuum environments, enabling real-time observation of dynamic deformation and microstructural evolution in metallic materials. By correlating macroscopic mechanical responses with microscale structural features, this technique effectively captures transient behaviors under extreme thermomechanical conditions, such as crack initiation and propagation, providing critical experimental support for multiscale material modeling.

Recent studies exemplify in situ SEM's pivotal role in metallic material research. Zhang et al. [2] demonstrated its efficacy in revealing crack propagation dynamics through fatigue testing nickel-based superalloys and additive-manufactured metals. Advancements now enable comprehensive phase transformation analysis across temperature regimes. Wang et al. [15] tracked fatigue crack growth and microstructural changes in additive-manufactured Al–Si10–Mg alloys under thermal–mechanical coupling, while reactor steel studies [16] quantified microscale strain fields during deformation. Cutting-edge high-temperature imaging systems capture dynamic processes like metal melting and recrystallization [17]. These breakthroughs establish in situ SEM as an indispensable tool for decoding material deformation behavior, driving innovation in metallic systems.

These studies have established benchmark datasets critical for optimizing material performance under extreme operational conditions. Nevertheless, current methodologies are predominantly reliant on 2-D imaging modalities, including advanced quantitative strain mapping techniques like digital image correlation (DIC), which provide valuable insights into plastic localization and crystallographic slip [18] and inherently lack direct spatially resolved 3-D height information. This limitation introduces systematic errors in comprehensive 3-D morphology characterization, particularly for features with significant out-of-plane displacement components such as slip steps and surface wrinkling during plastic deformation.

B. 3-D Reconstruction With 4Q-BSE Detector

Three-dimensional surface reconstruction techniques are crucial in materials science and engineering, particularly for revealing complex surface topographies. Photometric Stereo, as a noncontact 3-D imaging technique, infers surface normals by analyzing intensity variations under different lighting conditions to reconstruct surface geometry. Recently, deep learning methods have significantly enhanced the robustness and accuracy of photometric stereo for challenging surfaces, especially for reflective metallic surfaces [19], leading to confidence-aware networks for more precise end-to-end normal and depth estimation in smart metrology [20]. Furthermore, to overcome the limitations of individual techniques, research has also explored combining multiple 3-D reconstruction principles, such as fusing photogrammetry with photometric stereo to obtain more complete and accurate 3-D data [21].

Against this background, the 4Q-BSE detector integration in SEM enables tilt-free 3-D surface reconstruction by extracting topographic gradients from BSE signals via SFS algorithms. Pintus et al. [10] presented an automatic alignment procedure for a four-source photometric stereo technique applied to SEM, which is crucial for accurate height field reconstruction. Hardware advancements, exemplified by Borzunov et al.'s [6] optimized detector arrays for enhanced quantum efficiency, synergize with computational breakthroughs, including Giardino et al. 's [11] regularization-based distortion correction, Miyamoto et al.'s [14], [22] shadow-robust photometric stereo, and Neggers et al.'s [23] multidetector decomposition framework. These innovations underpin transformative industrial applications. Notably, Agashe et al.'s reference-free defect metrology achieves 3% height estimation errors [24].

Despite progress, existing methodologies predominantly address static nonmetallic systems, assuming small surface slopes. They have not been optimized for signal occlusion under dynamic loading conditions. They lack temporal resolution to resolve dynamic deformation mechanisms like slip traces evolution in metals—a critical gap for aerospace materials science. Our research bridges this divide through an occlusion-resistant computational framework experimentally validated via standardized protocols and demonstrates its first implementation in in situ tensile testing of IN718 superalloy. This approach enables concurrent micrometer-scale spatial resolution and millisecond-level temporal tracking, revealing previously undetectable fracture nucleation processes under operational stress conditions.

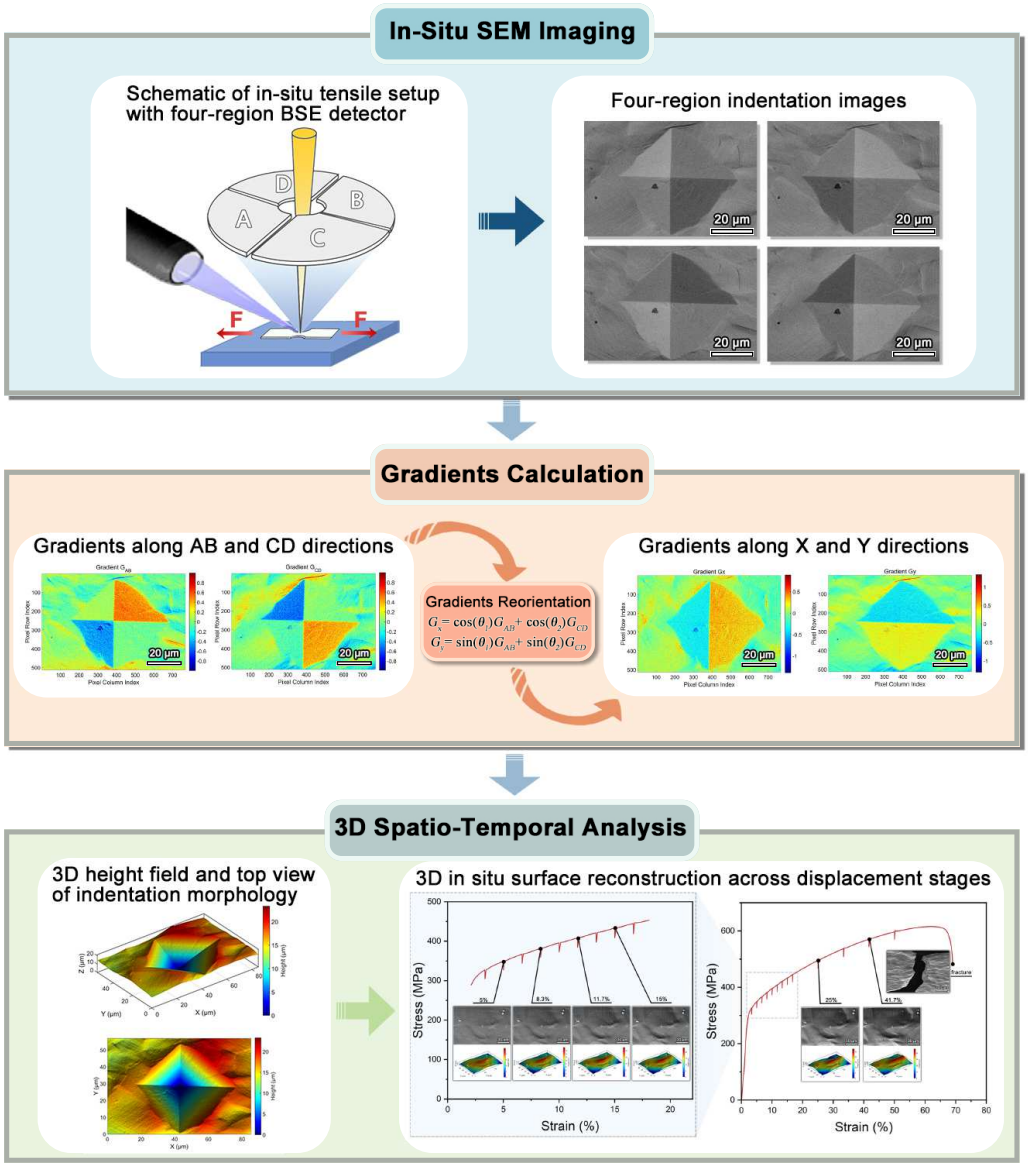


Fig. 1. Framework for in situ 3-D surface reconstruction and deformation evolution analysis.

III. Framework Overview and Experimental Setting

A. Framework Implementation

The proposed framework combines in situ SEM imaging with computational reconstruction algorithms through three coordinated modules: 1) synchronized mechanical loading and 4Q-BSE signal acquisition; 2) adaptive gradient calculation with occlusion compensation; and 3) multiscale surface reconstruction with 3-D spatiotemporal analysis of material deformation behavior. As illustrated in Fig. 1, this architecture ensures closed-loop validation from raw signal acquisition to quantitative 3-D morphology analysis.

B. In Situ SEM Tensile Experiment and Images Acquisition

This study conducted in situ tensile testing on IN718 nickel-based superalloy, a polycrystalline material widely adopted in aerospace and petrochemical applications due to its high-temperature stability, superior strength, and corrosion resistance. The specimen preparation protocols, including chemical composition, dimensions, heat treatment, and electrochemical polishing parameters, are detailed in [25].

The custom-developed tensile testing stage [26] was integrated with a CIQTEK 5000 Pro SEM equipped with a precalibrated 4Q-BSE detector to establish the in situ characterization system. Quasi-static tensile loading was applied at 1 μm/s under ambient conditions. This displacement-controlled loading strategy allowed for precise incremental deformation, enabling intermittent image acquisition at specific strain intervals to capture the dynamic evolution of surface features such as slip traces and overall morphology. As illustrated in Fig. 2(a), the detector quadrants (A, C, B, and D) arranged counterclockwise generated corresponding intensity images (I_A, I_C, I_B, and I_D). Synchronized mechanical loading and imaging enabled intermittent microstructure observation at user-defined strain intervals, with the physical setup shown in Fig. 2(b).

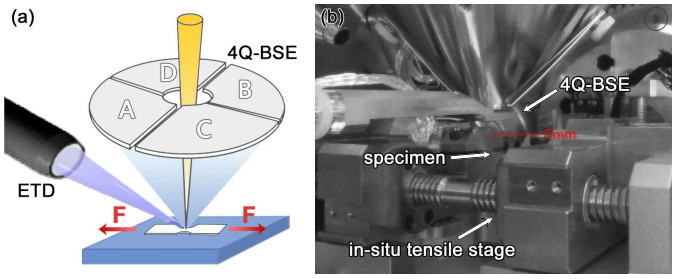


Fig. 2. In situ tensile testing configuration with synchronized mechanical loading and 4Q-BSE detector imaging. (a) Schematic of the load-controlled deformation coordinated with the 4Q-BSE detector and the Everhart–Thornley detector (ETD). (b) Photograph of the experimental setup, highlighting the specimen, in situ tensile stage, and the 4Q-BSE detector.

SEM imaging employed 15-kV acceleration voltage and 15-mm working distance with spot size maintained at 60 μm. Multiscale observation was achieved through progressive magnification adjustments from 250× to 1200× across six designated gauge regions, corresponding to horizontal field widths (HFWs) ranging from 510 μm at 250× to 110 μm at 1200×. All BSE images were captured at 3072 × 2048 pixel resolution, with HFW metadata embedded for spatial calibration during 3-D reconstruction. This pixel resolution, combined with the varying HFWs, yields an effective pixel size ranging from approximately 36 nm/pixel (at 1200×) to 166 nm/pixel (at 250×), thereby supporting the capability for submicrometer-resolution 3-D reconstruction.

IV. Methodology

This section comprehensively details our 3-D surface reconstruction framework, tailored for characterizing the microstructural evolution of IN718 nickel-based superalloy during in situ SEM tensile testing. The framework's core relies on an optimized gradient calculation algorithm and a multiscale reconstruction (MSR) strategy, crucial for accurately capturing dynamic surface morphology under various deformation conditions by mitigating signal occlusion and suppressing noise. Unlike conventional single-scale least-squares approach with regularization, which often suffers from error accumulation and noise sensitivity, the MSR method leverages a Gaussian pyramid-based strategy to enhance reconstruction accuracy and suppress high-frequency noise, enabling precise capture of surface morphology under tensile deformation. This section is structured as follows. First, we outline the gradient calculation methodology, addressing signal disappearance and occlusion challenges. Then, we introduce the MSR algorithm for surface reconstruction. Finally, we validate its performance using a custom 3-D standard specimen.

A. Gradient Calculation

During SEM imaging, the interaction between the electron beam and the specimen surface constitutes the fundamental mechanism for image formation. When high-energy primary electrons impinge on the sample, they may undergo either inelastic or elastic scattering. The inelastic scattering process generates secondary electrons (SEs)—low-energy particles (<50 eV) predominantly utilized for high-resolution topographic imaging. Conversely, elastic scattering produces BSEs

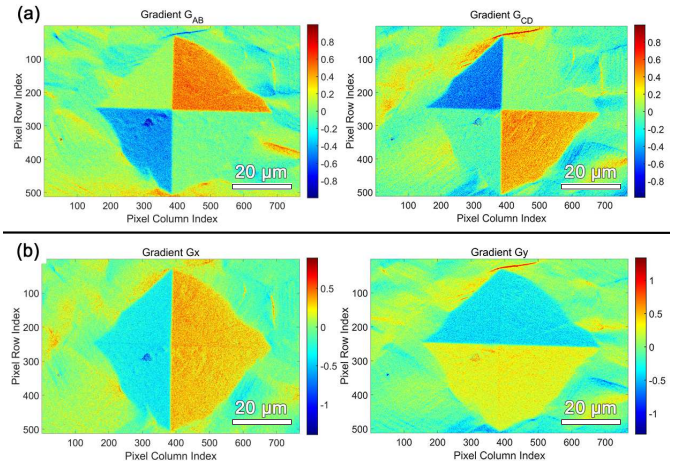


Fig. 3. Surface gradient calculation derived from cross-shaped indentation morphology characterization. (a) Gradient maps along the initial acquisition directions, AB and CD (G_{AB} and G_{CD}). (b) Reoriented gradients along the principal X and Y axes (G_X and G_Y).

characterized by higher kinetic energies (typically >50 eV), approaching the incident beam energy.

The BSE detector is strategically positioned directly below the objective lens pole piece, with two principal configurations: single-area detector and 4Q-BSE detector. The 4Q-BSE detector enables flexible signal combinations from four distinct segments (A, B, C, and D) to extract compositional and morphological information. As BSE emission efficiency correlates with surface inclination, differential signal analysis from opposing quadrants permits surface gradient computation through geometric considerations. Based on the quadrant configuration illustrated in Fig. 2(a), the signal intensities from detector regions A and B can be mathematically expressed as follows [27]:

$$\begin{cases} i_A = \dfrac{i_P \eta_{\mathrm{BSD}}}{\pi} \left[d_A \tan(\Phi_S) \cos(\theta_A - \theta_S) + c_A \right] \\ i_B = \dfrac{i_P \eta_{\mathrm{BSD}}}{\pi} \left[d_B \tan(\Phi_S) \cos(\theta_B - \theta_S) + c_B \right] \end{cases} \tag{1}$$

where Φ_S and θ_S denote the surface orientation parameters in the polar-spherical coordinate system, while d_A, d_B, c_A, c_B, θ_A, and θ_B represent geometry-dependent constants determined by the detector configuration.

Furthermore, assuming identical geometric coefficients across all four quadrants ($d_A = d_B$, $c_A = c_B$, and $\theta_A = 180° - \theta_B$) with matched sensitivity/gain settings, the surface slope can be derived from normalized differential intensities between opposing quadrant pairs. Let I_A and I_B denote the signal intensities from quadrants A and B, respectively. The surface height gradient along the AB direction can be expressed as [11]

$$G'_{AB} = \tan(\Phi_S) \cos(\theta_S) = k \left(\frac{I_A - I_B}{I_A + I_B} \right) \tag{2}$$

where k is a proportional coefficient with its value given by $k = c_a/d_a$.

This gradient calculation formula is theoretically correct but presents two practical computational issues. The first is the signal disappearance problem; for instance, electrons excited in the crack interior region cannot be received by any detector

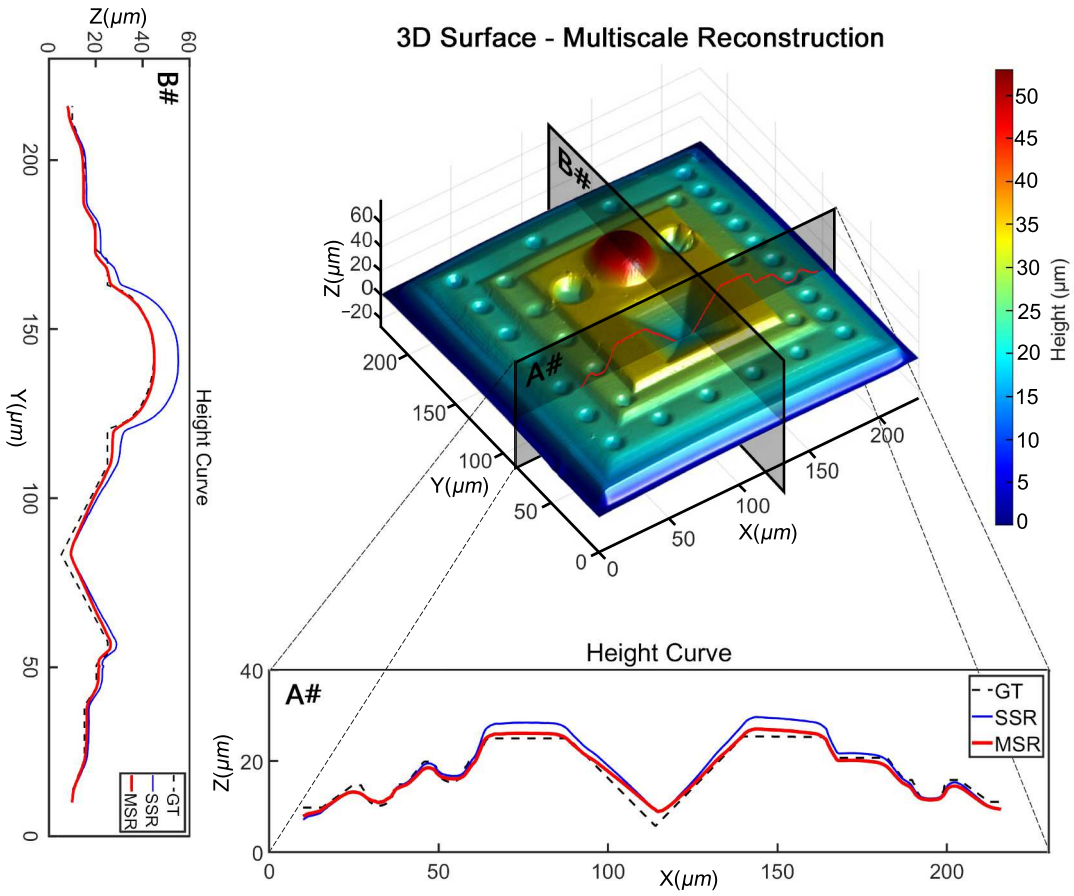


Fig. 4. Accuracy validation system for MSR 3-D reconstruction algorithm.

(i.e., $I_A = I_B = 0$), leading to computational errors in (2). This issue is resolved by adding a minimal positive constant ε to the denominator.

The second is the signal occlusion problem: in regions with steep surface slopes, BSEs may be blocked by higher sample features [14], creating significant shadow effects that amplify signal intensity differences between quadrant detectors. This phenomenon is challenging to model accurately; thus, we adopt a simple yet practical compensation method by introducing a cubic term in the gradient calculation. This cubic compensation term is specifically designed to counteract the nonlinear overestimation of surface gradients that occurs due to signal loss or amplification in steep topographical regions. The parameter α serves as a critical threshold: the compensation is only activated when the observed normalized signal difference, indicative of a steep slope, exceeds this predefined value. Conversely, b_0 is a negative constant that applies the actual dampening effect, proportionally reducing the calculated gradient to compensate for the occlusion-induced signal amplification. The coefficient b of this term remains zero for small surface slopes and becomes a negative constant b_0 only when the absolute value of the normalized signal difference, which is proportional to the surface slope, exceeds a predefined threshold α. In practice, b_0 and α are empirically determined through a calibration process, often involving samples with known steep geometries. The successful reconstruction of multitier pyramidal structures with 30°, 45°, and 60° sidewall slopes, as demonstrated in our accuracy validation (Section IV-C), implicitly validates the robustness of our chosen parameters for handling high-slope morphologies. Consequently, the final gradient result for AB-direction is expressed as

$$G_{AB} = k\left(\frac{I_A - I_B}{I_A + I_B + \varepsilon}\right) + b\left(\frac{I_A - I_B}{I_A + I_B + \varepsilon}\right)^3 \tag{3}$$

$$b = \begin{cases} b_0, & \text{if } \left|\frac{I_A - I_B}{I_A + I_B + \varepsilon}\right| > \alpha \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Following the same methodology, the gradient calculation result for the CD-direction is given by:

$$G_{CD} = k\left(\frac{I_C - I_D}{I_C + I_D + \varepsilon}\right) + b\left(\frac{I_C - I_D}{I_C + I_D + \varepsilon}\right)^3. \tag{5}$$

Due to the potential misalignment between the detectors' actual configuration and the principal orientation of the in situ tensile stage, the aforementioned gradients necessitate systematic reorientation along the X and Y coordinate axes

$$\begin{cases} G_x = \cos(\theta_1)\, G_{AB} + \cos(\theta_2)\, G_{CD} \\ G_y = \sin(\theta_1)\, G_{AB} + \sin(\theta_2)\, G_{CD} \end{cases} \tag{6}$$

where θ_1 and θ_2 denote the angular deviations between the AB/CD orientations and the X-axis coordinate system. These angular parameters can be experimentally determined during detector installation procedures or refined through Radon

transform analysis combined with standardized calibration specimens for enhanced metrological precision [23]. Fig. 3 demonstrates the computed gradient distributions derived from cross-shaped indentation morphology characterization using a microhardness tester. Specifically, Fig. 3(a) illustrates the gradient profile aligned with the principal detector orientation, while Fig. 3(b) presents the orthogonally decomposed gradient mappings along the X- and Y-axes. For the corresponding 4Q-BSE images and the reconstructed 3-D morphology of this indentation, please refer to the Supplemental Material (Fig. S1).

B. Surface Reconstruction

Reconstructing 3-D surface height fields from surface gradients constitutes a classical problem in computer vision, typically addressed through integration-based approaches. However, due to the imaging principle of SEM employing point-by-point scanning and line-by-line imaging, error accumulation progressively intensifies during the reconstruction process through continuous integration. Previous research on 4Q-BSE-based reconstruction has achieved promising results by applying the least-squares method combined with regularization techniques [6], [11]. The procedure is as follows. Initially, precomputed surface gradient matrices along the X- and Y-axes are retrieved, with dimensions corresponding to the image size (W, H). Subsequently, a height value of 0 is assigned to the designated starting point

$$\begin{cases} \vec{\nabla} z\left(x, y\right) = \left(G_x, G_y\right), (x, y) \in I \\ z\left(x_0, y_0\right) = 0. \end{cases} \tag{7}$$

Subsequently, the surface gradient matrix G_x along the X-axis direction is flattened column-wise into a column vector G_{xc}

$$G_x \equiv \begin{pmatrix} G^x_{1,1} & G^x_{1,2} & \cdots & G^x_{1,H} \\ G^x_{2,1} & G^x_{2,2} & \cdots & G^x_{2,H} \\ \vdots & \vdots & \ddots & \vdots \\ G^x_{W,1} & G^x_{W,2} & \cdots & G^x_{W,H} \end{pmatrix} \sim \begin{pmatrix} G^x_{1,1} \\ G^x_{2,1} \\ \vdots \\ G^x_{W,1} \\ G^x_{1,2} \\ G^x_{2,2} \\ \vdots \\ G^x_{W,2} \\ \vdots \\ G^x_{1,H} \\ G^x_{2,H} \\ \vdots \\ G^x_{W,H} \end{pmatrix} \equiv G_{xc}. \tag{8}$$

The surface gradient matrix G_y along the Y-axis direction undergoes identical row-wise flattening to form the column vector G_{yc}. The mathematical formulation for surface reconstruction using the least-squares method can be expressed as

$$P\hat{Z} = \begin{pmatrix} G_{xc} \\ G_{yc} \end{pmatrix} \equiv G \tag{9}$$

where P denotes the forward difference operator matrix with dimensions $(2HW \times HW)$, first HW rows correspond to finite difference approximation of operator $(\partial/\partial x)$, and the last HW rows correspond to $(\partial/\partial y)$. $\hat{Z}$ represents the surface height field vector to be solved, and G constitutes the complete gradient column vector formed by vertically concatenating G_{xc} and G_{yc}. The construction methodology of matrix P is explicitly delineated in Algorithm 1.

Algorithm 1 Forward Difference Operator P Matrix Construction

Input: Image height H, width W.
Output: Forward difference matrix P.

```
Initialize sparse matrix P ∈ ℝ^{2HW×HW}
Compute total pixels len ← H × W
for i = 1 to len − H do
    P[i, i] ← −1
    P[i, i + H] ← 1
end for
for col_start = 1 to len−H + 1 step H
    for j = 0 to H − 2
      row ← len + col_start + j
      col ← col_start + j
      P[row, col] ← −1
      P[row, col + 1] ← 1
    end for
end for
```

Building upon this theoretical framework, Giardino et al. [11] incorporated regularization techniques to mitigate interquadrant sensitivity disparities in the detector and suppress image noise interference, achieving significantly enhanced 3-D surface reconstruction fidelity. Following the Tikhonov regularization paradigm, the surface height determination problem was reformulated as an optimization formulation with a regularization term:

$$\min \left\{ \left\| P\hat{Z} - G \right\|^2 + \lambda \left\| L\hat{Z} \right\|^2 \right\} \tag{10}$$

$$\left(P^T P + \lambda L^T L\right) \hat{Z} = P^T G \tag{11}$$

where λ denotes the regularization strength and L is a regularization operator matrix of size $(2HW \times HW)$ constructed based on the second-order central difference operator. The detailed implementation is described in Algorithm 2.

The regularized 3-D surface height field $\hat{Z}$ can be obtained by solving (11). However, the reconstruction process fails to integrate the accurate distance information encoded in HFW, resulting in deviations from the actual morphological dimensions of the sample. Furthermore, this method exhibits sensitivity to noise in initial images, where local optima are readily induced by noise interference. Increasing the regularization strength λ would lead to excessive global smoothing at the expense of fine details. We propose a MSR method to address these limitations. By constructing a Gaussian pyramid of gradient matrices, MSR provides globally optimal initial guesses at coarse scales to accelerate convergence while demonstrating superior high-frequency noise suppression capabilities [28], [29]. The cross-scale residual compensation mechanism also effectively enhances high-frequency detail preservation [30].

Suppose that the Gaussian pyramid consists of S scales. The initial gradient fields are designated as $(G_x^{(1)}, G_y^{(1)})$, corre-

Algorithm 2 Second-Order Difference Operator L Matrix Construction

Input: Image height H, width W.
Output: Second-order difference matrix L.

```
Initialize sparse matrix L ∈ R^{2HW×HW}
Compute total pixels len ← H × W
for i = 1 to len − 2H do
    L[i, i] ← 1
    L[i, i + H] ← −2
    L[i, i + 2H] ← 1
end for
for col_start = 1 to len−H + 1 step H
    for j = 0 to H − 3
      col ← col_start + j
      row ← len + col
      L[row, col] ← 1
      L[row, col + 1] ← −2
      L[row, col + 2] ← 1
    end for
end for
```

sponding to the first layer (finest scale). The gradient fields of subsequent layers are then defined as follows:

$$\begin{cases} G_x^{(s)} = \left(G_x^{(s-1)} * g_\sigma\right)_{\downarrow 2} \\ G_y^{(s)} = \left(G_y^{(s-1)} * g_\sigma\right)_{\downarrow 2} \end{cases} \tag{12}$$

where the Gaussian kernel $g_\sigma(x, y) = (1/2\pi\sigma^2)e^{-(x^2+y^2)/(2\sigma^2)}$ is the unique filter satisfying linear scale-space axioms, capable of effectively suppressing aliasing artifacts. The surface height field $Z^{(s)}$ at scale s is computed through the residual propagation scheme; it is obtained by upsampling the solution from the coarser $s+1$ scale and subsequently adding the interscale residual

$$Z^{(s)} = \mathcal{U}\left(Z^{(s+1)}\right) + \Delta Z^{(s)} \tag{13}$$

$$\mathcal{U}(Z)(i, j) = \sum\nolimits_{p,q} h(p, q)\, Z(|i/2| + p, |j/2| + q). \tag{14}$$

By incorporating the HFW recorded during image acquisition and dividing it by the gradient field width $W^{(s)}$ at each scale, we compute the actual scaling factor $\gamma^{(s)} = \mathrm{HFW}/W^{(s)}$ for each layer. Furthermore, the differential operators and regularization strength parameters for each scale are defined as

$$\begin{cases} P'^{(s)} = P^{(s)}/\gamma^{(s)} \\ L'^{(s)} = L^{(s)}/\gamma^{(s)} \\ \lambda'^{(s)} = \lambda^{(s)} \cdot \left(\gamma^{(s)}\right)^2 \cdot 2^{(S-s)}. \end{cases} \tag{15}$$

Specifically, the residual field $\Delta Z^{(s)}$ is obtained by solving the following optimization equation:

$$\min_{\Delta Z^{(s)}} \left\| P'^{(s)}\left(\mathcal{U}\left(Z^{(s+1)}\right) + \Delta Z^{(s)}\right) - G^{(s)} \right\|_2^2 + \lambda'^{(s)} \left\| L'^{(s)} \Delta Z^{(s)} \right\|_2^2. \tag{16}$$

Through iterative computation, the height field $Z^{(1)}$ obtained at the finest scale ($s = 1$) is thus designated as the final reconstruction result Z_{final}. Compared with the single-scale method (SSR), MSR significantly reduces noise sensitivity and error accumulation through coarse-scale global optimization and fine-scale residual compensation. The complete computational workflow of this multiscale reconstruction framework is summarized in Algorithm 3.

Algorithm 3 Multiscale Surface Reconstruction

Input: Initial gradient fields $G_x^{(1)}$ and $G_y^{(1)}$, maximum number of scales S, regularization coefficient $\lambda^{(1)}$, horizontal field of view width HFW.
Output: Height map Z_{final}.

```
for s = 2 to S do
    G_x^(s) ← GaussianBlur(G_x^(s−1))↓2
    G_y^(s) ← GaussianBlur(G_y^(s−1))↓2
end for
for s = S down to 1 do
    γ^(s) = HFW/W^(s)
    P'^(s) = P^(s)/γ^(s)
    L'^(s) = L^(s)/γ^(s)
    λ'^(s) = λ^(s) · (γ^(s))^2 · 2^(S−s)
    if s = S
    Z^(S) = arg min ||P'^(S) Z − G^(S)||^2
    +λ'^(S) ||L'^(S) Z||^2
    else
    Z_up ← U(Z^(s+1))
    ΔG_x = G_x^(s) − ∇_x Z_up, ΔG_y = G_y^(s) − ∇_y Z_up
    ΔZ^(s) = arg min ||P'^(s) ΔZ − ΔG^(s)||^2
    +λ'^(s) ||L'^(s) ΔZ||^2
    Z^(s) ← Z_up + ΔZ^(s)
end for
Z_final ← Z^(1)
return Z_final
```

Remark 1: In the gradient calculation process, we incorporated a gradient compensation term to address signal occlusion caused by high-slope planes. Furthermore, we combined the MSR method through multiscale gradient decomposition and residual compensation mechanisms, effectively suppressing reconstruction errors induced by high-frequency noise in SEM images.

C. Accuracy Analysis

To validate the accuracy of 3-D reconstruction algorithms, we developed a custom 3-D standard specimen with multilayer topographic features. The pyramidal structure, designed in SolidWorks, features three tiers with sidewall slope angles of 30°, 45°, and 60° from bottom to top. The lower two tiers contain multiple hemispherical indentations, while the top tier incorporates larger hemispherical indentations, hemispherical protrusions, and cruciform grooves. Fabricated on a silicon substrate using high-precision two-photon polymerization 3-D printing (0.5-μm resolution), the specimen has a base edge length of 200 μm. A 100-nm gold layer was deposited via magnetron sputtering (40-mA current and 100-s duration) to ensure conductivity for SEM imaging. Comparative reconstruction was performed using the regularized SSR from [11] and our proposed MSR algorithm. The regularized SSR method [11] represents the state-of-the-art approach for 4Q-BSE-based 3-D surface reconstruction, making it the most relevant baseline for this comparative analysis. Fig. 4 demonstrates the submicrometer-resolution 3-D reconstruction achieved by

TABLE I
PERFORMANCE COMPARISON OF SURFACE RECONSTRUCTION METHODS

Plane	Method	RMSE (μm)	MAE (μm)	MaxAE (μm)	R^2
Plane **A#**	SSR	2.1952	1.8211	4.2708	0.8520
	MSR	**1.2192** (44.5% ↓)	**1.0342** (43.2% ↓)	**3.5660** (16.5% ↓)	**0.9543** (+0.102 ↑)
Plane **B#**	SSR	5.0623	3.8472	10.8892	0.7782
	MSR	**1.4226** (71.9% ↓)	**1.1232** (70.8% ↓)	**4.0875** (62.5% ↓)	**0.9825** (+0.204 ↑)

MSR. Height profile comparisons along characteristic planes **A#** and **B#** reveal that the MSR results (red curves) align more closely with the ground truth (black dashed curves) than SSR results (blue curves). The MSR method's incorporation of third-order gradient terms for shadow mitigation and multi-scale reconstruction strategy effectively reduces error accumulation in high-slope regions (e.g., the hemispherical protrusion in plane **B#**) while producing smoother surface morphology.

To quantitatively assess the reconstruction accuracy, we employed several standard error metrics: root-mean-square error (RMSE), mean absolute error (MAE), maximum error (MaxE), and the coefficient of determination (R^2). RMSE and MAE provide measures of the average magnitude of the errors between the reconstructed height field and the ground truth. MaxE identifies the largest individual deviation, indicating potential outliers or regions with significant error. R^2 quantifies the proportion of variance in the ground truth that is predictable from the reconstructed height field, with values closer to 1 indicating a better fit. As detailed in Table I, our MSR method consistently demonstrates superior performance across all metrics compared to the SSR approach. Specifically, MSR achieves a substantial reduction in RMSE and MAE for both Plane **A#** (low-gradient region) and Plane **B#** (high-gradient region), with improvements ranging from 43.2% to 71.9%. The significantly higher R^2 values (up to +0.204 for Plane **B#**) further underscore MSR's improved fidelity in capturing the true surface topography. These quantitative results robustly confirm the enhanced accuracy and robustness of our proposed method, particularly in challenging high-slope and complex topographical regions.

V. RESULTS

In this section, we present our in situ SEM tensile testing findings coupled with advanced 3-D reconstruction analysis of IN718 nickel-based superalloy, utilizing the MSR method detailed in Section IV. Our approach combines high-precision 3-D surface morphology reconstruction with mechanical deformation data to provide a comprehensive spatiotemporal characterization of material behavior under tensile loading. The results are structured into three main subsections: 1) in situ spatiotemporal analysis, which correlates the evolution of 3-D surface morphology with the stress–strain response across various deformation stages; 2) submicrometer-resolution characterization of slip trace dynamics during the early-to-midstages of plastic deformation; and 3) large-field 3-D reconstruction capturing polycrystalline deformation characteristics in the middle-to-late stages of strain. Compared to traditional 2-D imaging techniques, the MSR method enables the simultaneous capture of fine-scale slip features and macroscale surface relief phenomena with exceptional accuracy, even under severe plastic deformation. This section will first explore the relationship between 3-D surface evolution and mechanical behavior, followed by detailed insights into slip trace development and large-scale polycrystalline deformation patterns.

A. In Situ Spatiotemporal Analysis

The 3-D surface reconstruction and analysis of in situ experimental images were performed using the MSR method described in Section III. Fig. 5 demonstrates the correlation between high-precision 3-D surface morphology reconstruction (characterized by 3-D height fields) and the stress–strain curve at different deformation stages during in situ tensile testing of IN718 alloy, achieving spatiotemporal analysis of material deformation behavior. The observed stress relaxation phenomena in the figure correspond to image acquisition processes during load-holding periods, with each stress relaxation plateau representing a single imaging event. The left panel illustrates intensive image sampling during the initial plastic deformation stage (3.3%–16.7% strain range), designed to capture the dynamic evolution of slip traces in microdeformation. As plastic deformation progresses to the intermediate and final stages, the sampling frequency decreases while slip traces approach saturation, with macroscale morphology features becoming more pronounced in the height fields.

High-precision 3-D reconstruction was consistently achieved throughout all plastic deformation stages, with maintained accuracy as comprehensively illustrated in Fig. 6, demonstrating the method's superior robustness even under severe plastic deformation. As shown in Fig. 6(a), the SEM images of the observed region clearly depict the overall surface evolution from the initial undeformed state to the point nearing fracture. Fig. 6(b) and (c) further presents the dynamic evolution of the 3-D morphology across the entire strain range, with unified color bars within each panel effectively highlighting the overall changes and emphasizing the progressive nature of the deformation. Specifically, Fig. 6(c) reveals distinct local bulging phenomena emerging at the material surface when strain exceeds 16.7%, a characteristic precisely captured by the 3-D height fields. This deformation characteristic, resulting from grain rotation-mediated strain accommodation, is precisely captured by the 3-D height field. For detailed individual 4Q-BSE images and their corresponding 3-D reconstructions at each specific strain step, please refer to the Supplemental Material (Figs. S2–S13).

B. Slip Trace Evolution Characterization With Submicrometer 3-D Resolution

Submicrometer-resolution 3-D reconstruction of slip traces was performed on IN718 alloy during in situ tensile testing

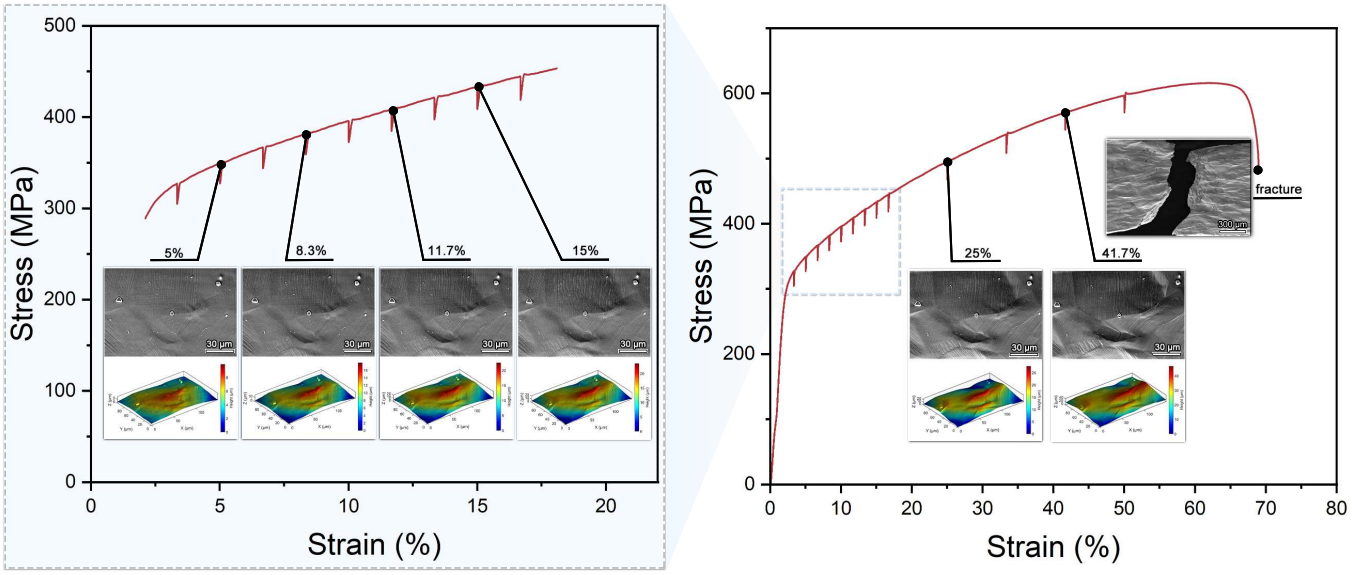


Fig. 5. Correlation between 3-D morphology evolution and mechanical response in IN718 alloy during tensile deformation.

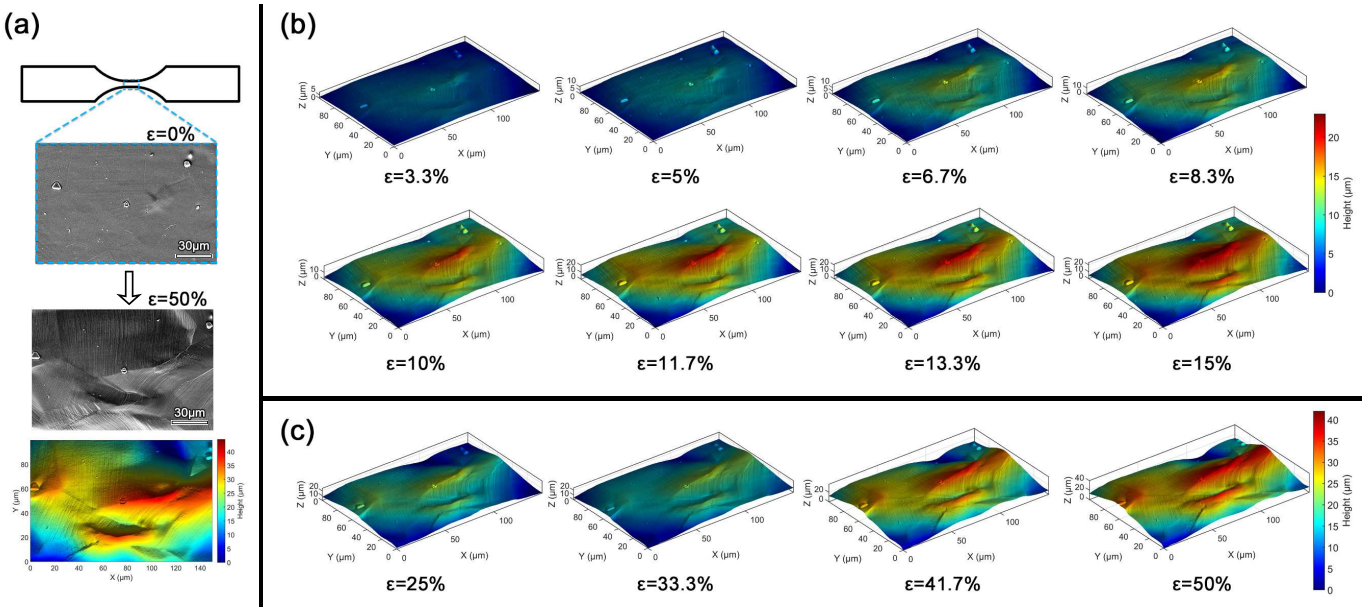


Fig. 6. Dynamic evolution of the 3-D morphology and strain distribution correlation in the central gauge area of IN718 alloy during tensile deformation. (a) Schematic of the observation region on the specimen and the corresponding SE images showing surface evolution from 0% to 50% engineering strain. (b) Reconstructed 3-D height fields illustrating surface morphology changes during early plastic deformation (ϵ = 3.3% to ϵ = 15%), with a unified color scale emphasizing progressive evolution. (c) Reconstructed 3-D height fields depicting surface morphology changes during late plastic deformation (ϵ = 25% to ϵ = 50%), also with a unified color scale to highlight overall transformations.

at engineering strain levels of 5%, 8.3%, 11.7%, and 15% [Fig. 7(a)–(d)], covering a 110-μm-wide field of view. The stress–strain curve confirms that all acquisition points reside in the early-to-mid transition stage of plastic deformation [Fig. 7(e)]. Height profile evolution along identical spatial coordinates [arrows in Fig. 7(a)–(d)] reveals submicrometer-scale slip step characteristics across strain stages (Fig. 7(f), color-coded arrows/curves). A systematic increase in profile slope with strain originates from grain rotation-induced measurement plane wrinkling during plastic deformation.

Detailed analysis of characteristic regions shows that region **1#** exhibits newly activated slip steps maintaining identical Burgers vector orientation as primary systems (evidenced by consistent profile slopes). Region **2#** demonstrates activated secondary slip systems with distinct vectors. Although 2-D morphology suggests parallel slip traces within a single grain (misinterpretable as a single slip system), 3-D height profiles reveal these traces correspond to dislocation motions along different directions within the same slip plane, manifested as differential slope characteristics in height profiles. This quantitative characterization of slip trace morphology at submicrometer resolution, a detail often undetectable by conventional 2-D imaging, provides crucial insights into the dynamic evolution of dislocation motion during incipient

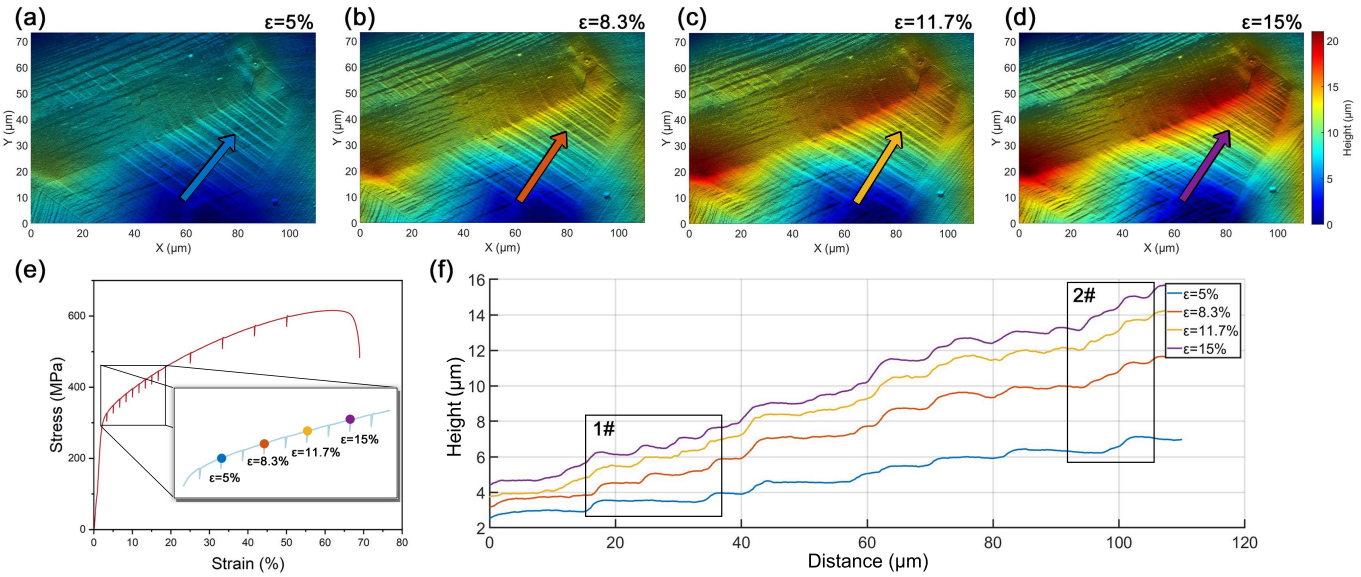


Fig. 7. Height variations and dynamic evolution of surface slip traces on IN718 alloy under different strain levels during the early stages of plastic deformation. (a)–(d) Submicrometer-resolution 3-D reconstructions of slip traces at engineering strain levels of 5%, 8.3%, 11.7%, and 15%, all scaled to a unified height range for direct comparison. The colored arrows indicate the spatial coordinates for height profile analysis. (e) Macroscopic stress–strain curve of IN718 alloy, highlighting the specific strain acquisition points (colored circles) corresponding to the detailed analysis presented in (a)–(d) and (f). (f) Quantitative height profile evolution along the indicated spatial coordinates, demonstrating submicrometer-scale slip step characteristics across strain stages.

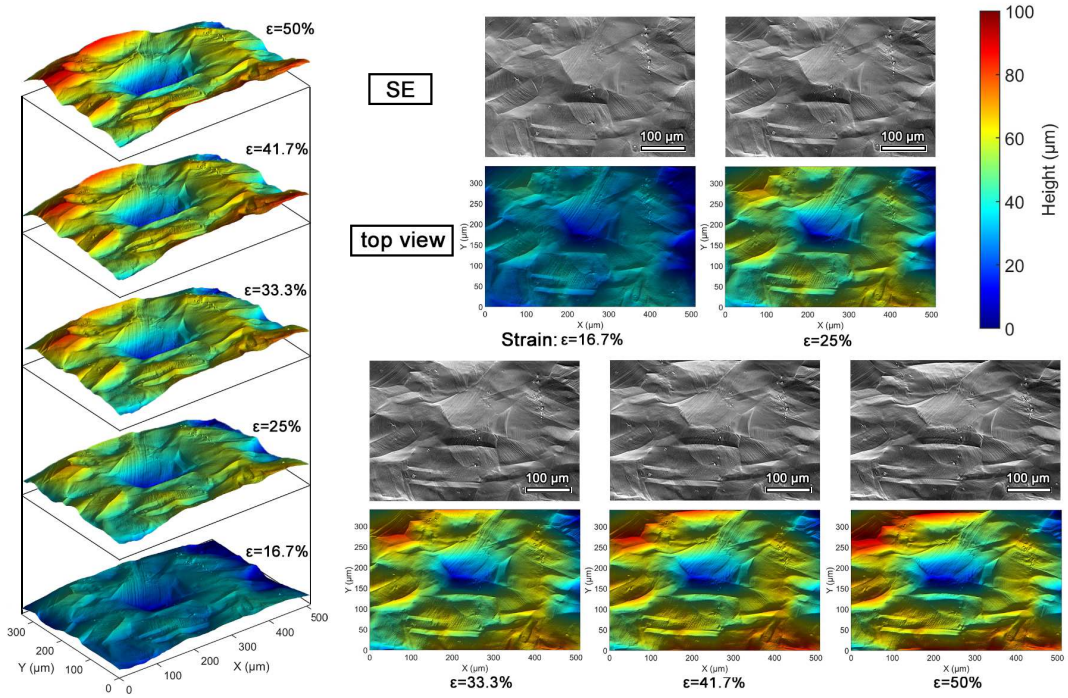


Fig. 8. Variation in the 3-D height field due to the coordinated deformation of multiple grains under a large field of view.

plastic deformation. Such understanding is fundamental for deciphering the onset of plasticity, predicting work hardening behavior, and identifying potential sites for early-stage fatigue crack initiation, thereby guiding the design of alloys with improved damage tolerance.

Remark 2: Through high-resolution 3-D reconstruction, this study has quantitatively characterized the morphological undulations of slip traces undetectable by conventional 2-D imaging techniques, providing novel quantitative evidence for deciphering the dynamic evolution of dislocation motion during incipient plastic deformation.

C. Large-Field 3-D Reconstruction of Polycrystalline Deformation Characteristics

During the middle-to-late stages of plastic strain (16.7%–50%), the surface wrinkling and relief phenomena of grains intensify significantly with increasing strain, as demonstrated in Fig. 8. The horizontal field of view reaches 510 μm, approximately five times that of the region shown in Fig. 7, enabling comprehensive observation of coordinated deformation among multiple grains. The left-side 3-D reconstruction results, normalized with a unified

color scale, reveal collapse at the central grain boundary region and pronounced undulations in peripheral grains, indicating substantial deformation gradients near grain boundaries. Comparative analysis between SEM images and 3-D morphology on the right side confirms their consistent deformation patterns. The surface relief and wrinkling originate from heterogeneous deformation caused by grain orientation variations. In particular, enhanced dislocation slip activity within specific grains is constrained by adjacent grain boundaries, leading to ineffective strain transfer and consequent intergranular shear deformation. Progressive strain accumulation exacerbates shear deformation, ultimately initiating intergranular crack nucleation.

The ability to capture such large-field 3-D morphology with submicrometer resolution is invaluable for identifying macroscopic strain localization hot spots, understanding the critical role of grain boundaries in accommodating heterogeneous deformation, and accurately predicting failure mechanisms like intergranular fracture in engineering components. This cross-scale correlative analysis provides a comprehensive view from grain-level interactions to macroscopic material behavior, crucial for optimizing material design and enhancing reliability.

Remark 3: The large-field-of-view 3-D reconstruction not only maintains submicrometer spatial resolution but also achieves macroscopic characterization of full-field strain distribution in polycrystalline materials, demonstrating its unique advantages in cross-scale correlative analysis.

VI. Conclusion

This article developed an innovative 3-D surface reconstruction technology for in situ SEM experiments, aiming to elucidate the deformation evolution mechanisms of metallic materials systematically. By integrating an improved gradient calculation method with a multiscale reconstruction algorithm, we effectively resolved the signal occlusion problem encountered by conventional methods when processing high-gradient surfaces, achieving significant enhancement in reconstruction accuracy and robustness. Experimental validation demonstrated the superior performance of this technique in both standard specimens and in situ tensile tests of IN718 nickel-based superalloy, accomplishing cross-scale correlation analysis from microscopic slip traces to macroscopic polycrystalline deformation. During the initial plastic deformation stage, intragranular slip systems' sequential activation and dynamic evolution were systematically captured, revealing interaction mechanisms between different slip systems. Large-field 3-D reconstruction unveiled grain boundary accommodation behavior in the advanced plastic deformation stage, particularly the correlation between strain gradient distribution and crack nucleation mechanisms near grain boundaries. These findings deepen the fundamental understanding of metallic material deformation physics and provide critical experimental evidence for multiscale material design and performance optimization. The methodology can be further extended to diverse material systems and complex loading conditions, paving new avenues for advancing in situ experimental technologies.

References

[1] M. A. Haque and M. T. A. Saif, "In-situ tensile testing of nano-scale specimens in SEM and TEM," *Experim. Mech.*, vol. 42, no. 1, pp. 123–128, Mar. 2002.

[2] B. Zhang et al., "In situ SEM fatigue testing technology for metallic materials: A review," *Nanoscale*, vol. 16, no. 28, pp. 13300–13330, 2024.

[3] X. Liu, Y. Liu, B. Jin, Y. Lu, and J. Lu, "Microstructure evolution and mechanical properties of a SMATed mg alloy under in situ SEM tensile testing," *J. Mater. Sci. Technol.*, vol. 33, no. 3, pp. 224–230, Mar. 2017.

[4] W. Wang, T. G. Liu, X. Y. Cao, Y. H. Lu, and T. Shoji, "In-situ SEM study of crack initiation and propagation behavior in a dissimilar metal welded joint," *Mater. Sci. Eng., A*, vol. 729, pp. 331–339, Jun. 2018.

[5] M. Wang, Z. Li, and D. Raabe, "In-situ SEM observation of phase transformation and twinning mechanisms in an interstitial high-entropy alloy," *Acta Mater.*, vol. 147, pp. 236–246, Apr. 2018.

[6] A. A. Borzunov et al., "3D surface topography imaging in SEM with improved backscattered electron detector: Arrangement and reconstruction algorithm," *Ultramicroscopy*, vol. 207, Dec. 2019, Art. no. 112830.

[7] X. Zhang et al., "Superalloys fracture process inference based on overlap analysis of 3D models," *Commun. Eng.*, vol. 3, no. 1, p. 108, Aug. 2024.

[8] A. P. Tafti, A. B. Kirkpatrick, Z. Alavi, H. A. Owen, and Z. Yu, "Recent advances in 3D SEM surface reconstruction," *Micron*, vol. 78, pp. 54–66, Nov. 2015.

[9] A. P. Tafti, J. D. Holz, A. Baghaie, H. A. Owen, M. M. He, and Z. Yu, "3DSEM++: Adaptive and intelligent 3D SEM surface reconstruction," *Micron*, vol. 87, pp. 33–45, Aug. 2016.

[10] R. Pintus, S. Podda, and M. Vanzi, "An automatic alignment procedure for a four-source photometric stereo technique applied to scanning electron microscopy," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 5, pp. 989–996, May 2008.

[11] M. Giardino, D. M. N. Menon, and D. L. Janner, "Regularization techniques for 3D surface reconstruction from four quadrant backscattered electron detector images," *Ultramicroscopy*, vol. 250, Aug. 2023, Art. no. 113746.

[12] J. Rosenberger, J. Tlatlik, C. Beckmann, B. Rohrmüller, and S. Münstermann, "Extracting ductile cast iron microstructure parameters from fracture surfaces: A deep learning based instance segmentation approach," *Eng. Fract. Mech.*, vol. 311, Nov. 2024, Art. no. 110586.

[13] J. D. Woodward and R. A. Wepf, "Macromolecular 3D SEM reconstruction strategies: Signal to noise ratio and resolution," *Ultramicroscopy*, vol. 144, pp. 43–49, Sep. 2014.

[14] A. Miyamoto, D. Chen, and S. Kaneko, "Bootstrapping de-shadowing and self-calibration for scanning electron microscope photometric stereo," *Meas. Sci. Technol.*, vol. 25, no. 10, Oct. 2014, Art. no. 105402.

[15] Z. Wang, W. Wu, G. Qian, L. Sun, X. Li, and J. A. F. O. Correia, "In-situ SEM investigation on fatigue behaviors of additive manufactured Al-Si10-Mg alloy at elevated temperature," *Eng. Fract. Mech.*, vol. 214, pp. 149–163, Jun. 2019.

[16] F. Latourte, T. Salez, A. Guery, N. Rupin, and M. Mahé, "Deformation studies from in situ SEM experiments of a reactor pressure vessel steel at room and low temperatures," *J. Nucl. Mater.*, vol. 454, nos. 1–3, pp. 373–380, Nov. 2014.

[17] Y. Zhang et al., "Development and application of a high-temperature imaging system for in-situ scanning electron microscope," *Mater. Today Commun.*, vol. 38, Mar. 2024, Art. no. 107782.

[18] J. C. Stinville et al., "Insights into plastic localization by crystallographic slip from emerging experimental and numerical approaches," *Annu. Rev. Mater. Res.*, vol. 53, no. 1, pp. 275–317, Jul. 2023.

[19] R. Yang, Y. Wang, S. Liao, and P. Guo, "DPPS: A deep-learning based point-light photometric stereo method for 3D reconstruction of metallic surfaces," *Measurement*, vol. 210, Mar. 2023, Art. no. 112543.

[20] Y. Zhang, R. Yang, and P. Guo, "Confidence-aware photometric stereo networks enabling end-to-end normal and depth estimation for smart metrology," *IEEE/ASME Trans. Mechatronics*, vol. 30, no. 2, pp. 910–920, Apr. 2025.

[21] A. Karami, F. Menna, and F. Remondino, "Combining photogrammetry and photometric stereo to achieve precise and complete 3D reconstruction," *Sensors*, vol. 22, no. 21, p. 8172, Oct. 2022.

[22] A. Miyamoto, H. Matsuse, and G. Koutaki, "Robust surface reconstruction by design-guided SEM photometric stereo," *Meas. Sci. Technol.*, vol. 28, no. 4, Apr. 2017, Art. no. 045405.

[23] J. Neggers et al., "Principal image decomposition for multi-detector backscatter electron topography reconstruction," *Ultramicroscopy*, vol. 227, Aug. 2021, Art. no. 113200.

[24] S. S. Agashe, G. Kumar, S. Kulasekaran, and Y. Cho, "Defect 3-D profile estimation using SEM images," *IEEE Trans. Semicond. Manuf.*, vol. 36, no. 3, pp. 351–358, Aug. 2023.
[25] G. Guo et al., "Effect of precipitate phase on the plastic deformation behavior of alloy 718: In-situ tensile experiment and crystal plasticity simulation," *Int. J. Plasticity*, vol. 187, Apr. 2025, Art. no. 104286.
[26] W. Jiang et al., "In-situ EBSD investigation of the effect of orientation on plastic deformation behavior of a single crystal superalloy," *Mater. Sci. Eng., A*, vol. 849, Aug. 2022, Art. no. 143453.
[27] W. Slówko, "Directional detection of secondary electrons for electron beam profilography," *Vacuum*, vol. 52, no. 4, pp. 441–449, Apr. 1999.
[28] T. Lindeberg, *Scale-Space Theory in Computer Vision*, vol. 256. Cham, Switzerland: Springer, 2013.
[29] J. Nocedal and S. J. Wright, *Numerical Optimization*. Cham, Switzerland: Springer, 1999.
[30] S. Mallat, *A Wavelet Tour of Signal Processing*. Amsterdam, The Netherlands: Elsevier, 1999.



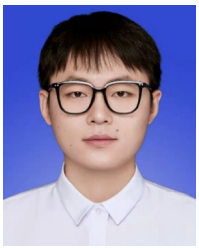
Xuecheng Zhang received the B.S. degree from Central South University, Changsha, China, in 2021. He is currently pursuing the Ph.D. degree in control science and engineering with the Polytechnic Institute, Zhejiang University, Hangzhou, China.

His current research is focused on deep learning and 3-D characterization in microscopy.



Bin Zhang received the M.A.Eng. degree from Nanchang University, Nanchang, China, in 2022. He is currently pursuing the Ph.D. degree with the School of Materials Science and Engineering, Zhejiang University, Hangzhou, China.

His research field is the application of DIC method in in situ SEM mechanical testing environment.



Guanghao Guo received the B.S. degree from Nanjing Tech University, Nanjing, China, in 2019. He is currently pursuing the Ph.D. degree in materials science and engineering with the School of Materials Science and Engineering, Zhejiang University, Hangzhou, China.

His current research is focused on in situ characterization and mechanical deformation of superalloys.



Wenchao Meng (Senior Member, IEEE) received the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2015.

He is currently with the College of Control Science and Engineering, Zhejiang University. His current research interests include adaptive intelligent control, cyber-physical systems, renewable energy systems, and smart grids.



Xianjue Ye received the Ph.D. degree from Central South University, Changsha, China, in 2023.

He is currently a Post-Doctoral Researcher with the School of Materials Science and Engineering, Zhejiang University, Hangzhou, China. His current research is focused on the design and multiscale characterization of superalloys.



Yuefei Zhang received the Ph.D. degree from Beijing University of Technology, Beijing, China, in 2008.

He is currently a Professor with the School of Materials Science and Engineering, Zhejiang University, Hangzhou, China. His research interests include nanomechanics, in situ electron microscopy, superalloys, and the application of in situ high-throughput characterization with big data and machine learning.



Ze Zhang received the Ph.D. degree from Jilin University, Changchun, China, in 1980, and the M.S. and Ph.D. degrees from the Institute of Metal Research, Chinese Academy of Sciences, Shenyang, China, in 1983 and 1987, respectively.

He is currently a Professor with the Department of Materials Science and Engineering, Zhejiang University, Hangzhou, China. His research interests include quasicrystals, low-dimensional nanomaterials, and electron microscopy with an emphasis on AI and big data integration for materials science.

Dr. Zhang has received numerous awards, including the National Natural Science First Class Award. He was an Academician at Chinese Academy of Sciences.



Jiming Chen (Fellow, IEEE) received the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2005.

He is currently a Professor with the Department of Control Science and Engineering, Zhejiang University. His research interests include the Internet of Things, networked control, and wireless networks.

Dr. Chen is a fellow of the CAA. He was a recipient of the Seventh IEEE ComSoc Asia/Pacific Outstanding Paper Award, Japan Society for the Promotion of Science Invitation Fellowship, and IEEE ComSoc AP Outstanding Young Researcher Award. He was on the editorial boards of multiple IEEE TRANSACTIONS and the General Co-Chair for IEEE RTCSA'19, IEEE Datacom'19, and IEEE PST'20. He is also the Distinguished Lecturer of the IEEE Vehicular Technology Society Distinguished Lecturer.

Ultramicroscopy 278 (2025) 114226

Contents lists available at ScienceDirect

Ultramicroscopy

journal homepage: www.elsevier.com/locate/ultramic

ETDMS: Efficient two-stage diffusion model for accelerated SEM image super-resolution

Xuecheng Zhang [a,b,1], Zixin Li [b,1], Bin Zhang [a], Wenchao Meng [b], Yuefei Zhang [a,*], Chaojie Gu [b], Xianjue Ye [a], Ze Zhang [a]

[a] Institute of Superalloys Science and Technology, School of Materials Science and Engineering, Zhejiang University, Hangzhou, 310027, China
[b] College of Control Science and Engineering, Zhejiang University, Hangzhou, 310027, China

ARTICLE INFO

Dataset link: https://github.com/Elton0427/ETDMS

Keywords:
Scanning electron microscopy (SEM)
Denoising diffusion probabilistic model
SEM image super-resolution
Efficient two-stage diffusion model

ABSTRACT

The scanning electron microscope (SEM) is a crucial tool for characterizing material microstructures, and it is renowned for its high resolution and depth of field. However, SEM image quality is affected by the scanning speed and resolution settings. When using SEM to capture fast-changing dynamic processes and other specific tasks, maintaining high image quality while using fast scanning mode is often tricky. To address these challenges, this paper introduces introduce an efficient two-stage diffusion model for accelerated SEM image super-resolution named ETDMS. The image denoising and super-resolution were divided into independent tasks based on SEM imaging principles and completed in two separate stages. Specifically, in Stage 2, a conditional lightweight encoder–decoder architecture SR network is proposed to replace the large U-Net in the traditional diffusion model and combine it with accelerated sampling technology to improve image generation efficiency. Experimental results prove that compared with previous super-resolution methods, the images generated by ETDMS significantly improve evaluation parameters, subjective visual quality, and detail generation.

1. Introduction

In the research of material science and other physical sciences, it is crucial to conduct in-depth analysis and exploration of the sample's microstructure [1]. The microstructure determines the material's macroscopic properties and is crucial to understanding material science phenomena. Scanning Electron Microscopy (SEM) has become an indispensable instrumentation in material science research due to its outstanding characteristics such as nanometer-level resolution and high depth of field [2,3]. SEM scans the surface of the sample with an electron beam, and the signals excited are captured and amplified by detectors, ultimately generating high-contrast two-dimensional images that reveal the detailed characteristics of the sample's surface, providing researchers with a window to observe, analyze and measure the micro-world of samples.

Despite the many advantages of SEM, its image acquisition process is inevitably subject to noise interference. These noises may originate from the instability of the interaction between the electron beam and the sample, the noise of the detector, and other electronic components of the image acquisition system [4,5]. The presence of noise reduces the image's quality and affects the accuracy of subsequent image analysis and processing. To improve the signal-to-noise ratio of the image, one can increase the dwell time of the electron beam on each pixel point by adopting a slow scan mode. However, while this method can enhance image quality, it also significantly increases the image acquisition time. Moreover, under specific experimental conditions, such as in situ observation of catalytic and other chemical reactions, the rapidity of the reaction makes it impossible for the slow scan mode to capture the fast-changing dynamic processes [6,7]. Therefore, fast scan mode is a necessary choice. Additionally, the prolonged focusing of a high-energy electron beam can cause damage to specific samples, especially biological samples [8,9] and thermosensitive materials [10,11]. In pursuing high-resolution images (HR), the increased acquisition time makes finding a balance between image quality and acquisition efficiency even more challenging.

In recent years, many deep learning-based methods for low-resolution image (LR) denoising and super-resolution (SR) have emerged to address the issues above. These methods can be mainly divided into three categories: Methods oriented towards Peak Signal-to-Noise Ratio (PSNR) [12–14], methods based on Generative Adversarial Networks (GAN) [15–17], and methods based on Denoising Diffusion Probabilistic Models (DDPM) [18–20]. PSNR, as a widely used image quality evaluation metric, can directly reflect the signal-to-noise ratio of

* Corresponding author.
E-mail address: yfzhang76@zju.edu.cn (Y. Zhang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ultramic.2025.114226
Received 25 March 2025; Received in revised form 22 July 2025; Accepted 25 August 2025
Available online 13 September 2025
0304-3991/© 2025 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

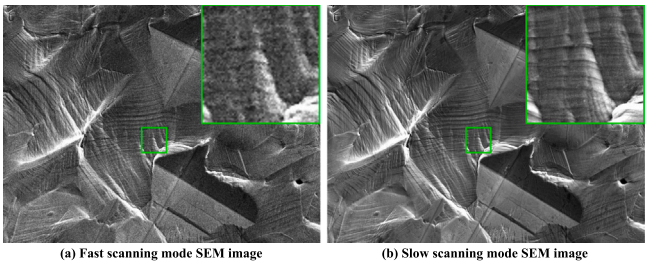


Fig. 1. Comparison of SEM images captured at different scanning speeds and resolution settings. The left image was captured at a scanning speed of 3 (1 μs/pixel) and upsampled 4× using bicubic interpolation to align the resolution with the right image. The right image was captured at 5 (10 μs/pixel). The difference between the images acquired at the two settings can be viewed by zooming in on the details.

an image. Methods guided by PSNR train neural networks by minimizing the Euclidean or Manhattan distance between LR and HR images. These methods are easy to quantify and optimize during training and have a lower computational cost. However, PSNR-oriented methods tend to generate smooth images, which may lose high-frequency details, thus not being entirely satisfactory in visual quality. Methods based on GAN introduce an adversarial training mechanism, which can generate images rich in detail. However, the training process of GAN is usually more complex, requiring careful design of network architecture and loss functions, and is prone to mode collapse. Methods based on DDPM simulate the reverse process of data generation, gradually removing noise from the image and effectively capturing the statistical characteristics of the image. Images generated by this method usually perform better in visual performance and detail preservation. However, the training and inference process of DDPM could be more convenient, demanding on computational resources, and generally time-consuming. Therefore, it is necessary to design corresponding lightweight strategies to improve this problem.

In some specific applications of SEM, such as capturing the fast-changing dynamic process, it is difficult to balance high image quality while using fast scanning modes, as these images often come with problems of high noise and insufficient detail information. Low quality images prevent experimenters from accurately analyzing and measuring microscopic features. Existing methods cannot address this issue, highlighting a significant research gap. To close this gap, inspired by the multi-stage concept in SRFS [20], this paper proposes a novel and efficient two-stage diffusion model named ETDMS. This model is specifically designed to address the unique challenges of SEM image SR by leveraging the distinct characteristics of image degradation in SEM. The core innovation of ETDMS lies in its strategic decomposition of the complex SEM image SR problem into two independent and optimized stages: dedicated denoising followed by conditional SR. Based on DDPM, this model is designed explicitly for denoising and SR of SEM images. In the SEM imaging process, the signal-to-noise ratio (image quality) is mainly determined by the scanning speed (i.e., the dwell time of the electron beam). In contrast, the diameter of the incident electron beam spot primarily determines the image's resolution. In practical applications, the fast scan mode produces low-quality, low-resolution images (LQLR), while the slow scan mode produces high-quality, high-resolution images (HQHR). Fig. 1 compares image details under the two modes. In addition to the set acquisition resolution, scanning speed influences the acquisition process. In SRFS, LQLR is first transformed into low-quality, high-resolution images (LQHR) through a shallow PSNR-oriented SR network and then serves as a conditional input to the subsequent SR network. However, in this process, the upsampling without denoising leads to the diffusion of noise, introducing new noise and artifacts, making the high-frequency details of the image blurry and subsequently losing important information.

In the ETDMS model, the first stage employs a robust denoising network (DN) to transform LQLR images into high-quality low-resolution (HQLR) images. This crucial preprocessing step is designed to effectively remove noise and enhance image details and local contrast, thereby preventing the propagation and diffusion of noise during subsequent upsampling. The second stage then takes these cleaned HQLR images from the first stage as conditional input, allowing a conditional lightweight SR network to focus exclusively on the super-resolution task and the generation of high-frequency details through a diffusion model. During this process, noise is gradually introduced through a Markov chain until the image distribution approaches the standard Gaussian distribution. In the inference phase, the complex U-Net network structure used in SRFS is abandoned. Instead, a cross-attention module is adopted to reduce the computational burden of many convolutional layers, thereby improving computational efficiency. Additionally, combining denoising diffusion implicit models (DDIM) reduces the number of sampling steps and further accelerates the inference process. Experimental results demonstrate that this two-stage approach of ETDMS significantly enhances the quality and resolution of SEM images, particularly in the performance of high-frequency details, while maintaining high computational efficiency.

Experiments are conducted to evaluate the performance of ETDMS. Specifically, LQLR-HQLR-HQHR SEM image datasets are captured from multiple superalloy samples, aiming to provide more accurate and reliable data support for the microstructural analysis of specimens in material science by enhancing the quality of SEM images. The experimental results demonstrate that ETDMS outperforms existing algorithms in multiple aspects, including higher evaluation metrics, more excellent visual quality, and faster inference runtime. Under the default settings of ETDMS, the Learned Perceptual Image Patch Similarity (LPIPS), which best simulates human subjective perception, is reduced to below 0.2, representing an 8.5% improvement over the state-of-the-art SR3. The inference time for a single image is reduced from 25 s to less than 1 s. The main contributions are as follows:

- The difference of images in SEM fast and slow scanning modes regarding quality and resolution is clarified, and a training data acquisition strategy is proposed for LQLR-HQLR-HQHR. The image pairs are captured under actual working conditions, not synthetically generated, which offers better authenticity and reliability, and helps to enhance the robustness of the trained models.
- A novel two-stage SR network method called ETDMS is proposed based on DDPM, which can effectively improve the quality of generated images, especially the recovery of high-frequency details.
- The practicality of this research is demonstrated by effectively reducing inference time without compromising image quality. This is achieved through the use of an efficient network and DDIM sampling.

2. Related work

2.1. Diffusion model for image super-resolution

Many learning-based methods have been applied to vision-based measurement [21,22], including multi-source image fusion and knowledge distillation [23], phase retrieval [24], and dynamic 3D measurement [25]. Subsequently, in pursuit of superior viewing effects and enhanced measurement accuracy, researchers have introduced further innovations and conducted extensive trials in methodology.

DDPM is a powerful generative model first proposed in [18]. It generates clear images from noisy ones through an iterative denoising process and is flexible and applicable to many visual tasks, including video generation [26,27], semantic segmentation [28,29], and so on. DDPM is also widely used in image denoising and SR tasks across various industries and fields [19,30]. DDPM can recover clear images

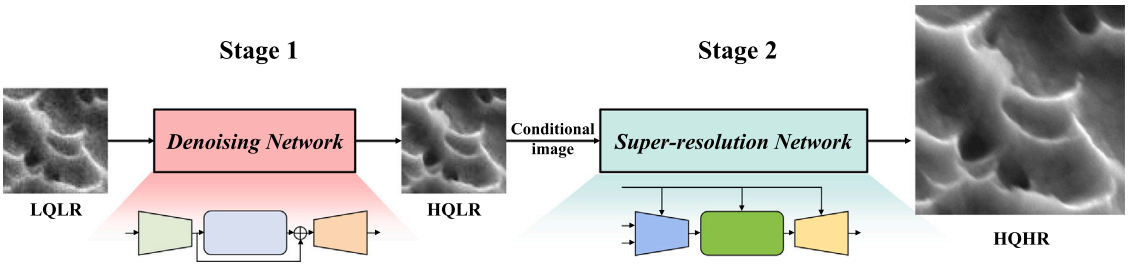


Fig. 2. The complete process of proposed ETDMS.

from images containing Gaussian noise, especially in terms of stability when dealing with complex scenes and details. By taking low-resolution images as input, DDPM can generate high-resolution images while preserving the details and texture information of the images. Recent research, such as the work by Liu et al. [31] and An et al. [32], has applied DDPM to SR tasks for remote sensing images. Wu et al. [33] incorporating self-attention mechanisms to generate high-resolution brain images, further proving the effectiveness of DDPM in SR tasks.

Although DDPM has shown excellent performance in image SR tasks, it also has some limitations. For instance, the iterative process requires noise and sample data to maintain the same size, leading to high computational costs, especially when dealing with high-dimensional data such as images and videos. Additionally, the sampling efficiency of DDPM is relatively low, necessitating a large number of time steps to ensure the stable generation of high-quality samples.

To address these issues, researchers have proposed various methods to reduce the complexity of DDPM and improve efficiency. For example, the Latent Diffusion Model (LDM) proposed by Rombach et al. [34] significantly reduces computational costs by training diffusion models in latent space. The Denoising Diffusion Implicit Models (DDIM) introduced by Song et al. [35] accelerate the generation process by using smaller sampling steps. The Come-Closer-Diffuse-Faster (CCDF) strategy proposed by Chung et al. [36] utilizes the reconstruction of pre-trained networks for initialization, allowing for smaller time steps and further accelerating the generation process of conditional diffusion models.

The application of DDPM in image denoising and super-resolution tasks has demonstrated its strong potential as a generative model. Despite some challenges, with continuous research and improvement, the prospects for the application of DDPM and its variant models in these fields are very promising.

2.2. Denoising and super-resolution of SEM images

Numerous learning-based methods have been applied to the SR tasks of SEM images. In earlier times, Tsiper et al. [37] employed sparse coding and dictionary learning algorithms to enhance the low-resolution SEM images of microelectronic chips, achieving the level of high-resolution images obtained through slow-scan SEM while significantly reducing noise. Haan et al. [38] and Wang et al. [39] respectively utilized GAN-based and attention mechanism-based methods for the super-resolution of SEM images. In [20], the authors proposed a novel two-stage approach, first super-resolving the SEM images and then generating details based on a diffusion model, achieving excellent results.

These methods aim to improve the resolution of the image and generate a certain amount of detail but do not take into account problems such as the low signal-to-noise ratio caused by the short electron beam dwell time in the fast scanning mode, which makes it difficult to generate sufficient detail.

3. Methodology

In this paper, an efficient two-stage diffusion model for the SEM images SR task named ETDMS is presented. ETDMS is an innovative two-stage network comprising a denoising network (DN) and a conditional lightweight SR network (CLSR), which is shown in Fig. 2. In stage 1, the LQLR images are denoised by DN into cleaned HQLR images. In stage 2, the cleaned HQLR images are upsampled as the conditional image and use the conditional diffusion model to predict the difference between the cleaned HQLR images and HQHR images. In Stage 1, the denoising process effectively removes the noise part from the image while enhancing the detailed information of the image. This preprocessing method ensures that the image is in a relatively clean state before undergoing SR, and it also eliminates the noise diffusion effect caused by direct upsampling. In Stage 2, CLSR focuses more on reconstructing the details of the image rather than diverting its learning capability to deal with noise. This strategy allows the network to precisely and reasonably address the loss of quality and resolution in SEM images under fast-scanning mode, thereby generating images with better visual quality for satisfactory analysis and measurement. In this section, the mainstream conditional diffusion-based SR network structures will first be introduced. Then, the proposed efficient two-stage network, ETDMS, including its DN and CLSR components, will be presented.

3.1. Conditional DDPM

The implementation process of DDPM is composed of two processes: the forward diffusion process and the reverse denoising process. In the forward diffusion process, data is subjected to a fixed Markov chain. This process gradually adds Gaussian noise generated according to a preset noise variable schedule $\{\beta_1, \beta_2, \ldots, \beta_T\}$ into the initial data x_0, thereby continuously generating a sequence of latent samples, that is $\{x_1, x_2, \ldots, x_T\}$, identical in dimension to x_0. The data distribution can be set as

$$q(x_{1:T}|x_0) = \prod_{t=1}^{T} q(x_t|x_{t-1}) \tag{1}$$

$$q(x_t|x_{t-1}) = \mathcal{N}\left(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I\right) \tag{2}$$

where T is the diffusion time step and the noise variable $\{\beta_t \in (0,1)\}_{t=1}^{T}$ with $\{\beta_{t+1} \geq \beta_t\}_{t=1}^{T-1}$. x_T will be conformed to standard Gaussian distribution, $x_T \sim \mathcal{N}(0, I)$ under the premise that β_T is sufficiently large.

The forward diffusion process allows for sampling the state x_t at any given time t, we have

$$q(x_t|x_0) = \mathcal{N}\left(x_t; \sqrt{\bar{\alpha}_t}x_0, (1-\bar{\alpha}_t)I\right) \tag{3}$$

where $\bar{\alpha}_t = \prod_{s=1}^{t}(1-\beta_s)$. Reparameterization trick allows for the simplification of Markov chains during the diffusion process. The relationship between x_t and x_0 at any time step t can be expressed as

$$\mathrm{x}_t = \sqrt{\bar{\alpha}_t}x_0 + \sqrt{1-\bar{\alpha}_t}\epsilon, \epsilon \sim \mathcal{N}(0, I) \tag{4}$$

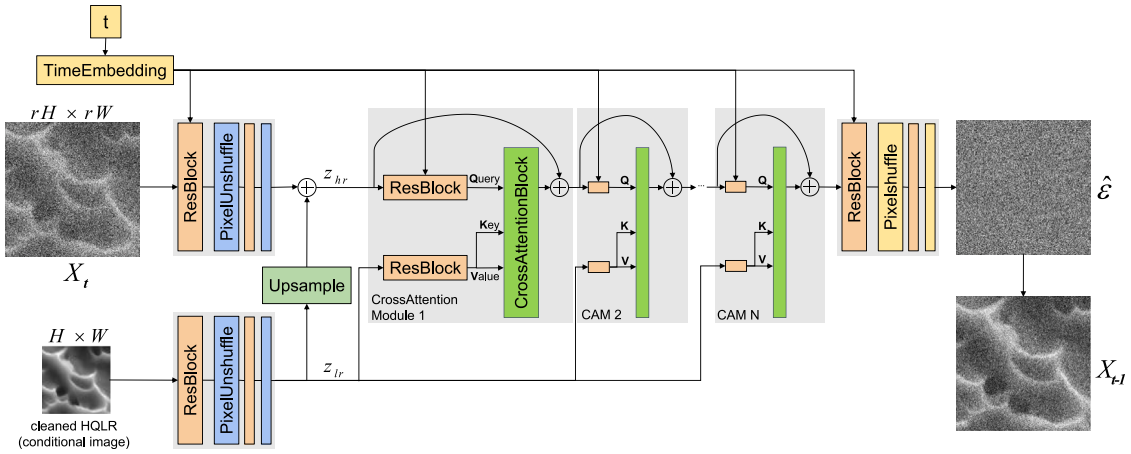


Fig. 3. Network structure of the conditional lightweight SR network (CLSR).

The posterior probability $q(x_{t-1}|x_t, x_0)$ follows a Gaussian distribution, which can be written as

$$\begin{cases} q(x_{t-1}|x_t, x_0) = \mathcal{N}\left(x_{t-1}; \tilde{\mu}(x_t, x_0), \tilde{\beta}_t I\right) \\ \tilde{\mu}(x_t, x_0) = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t} x_t + \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t} x_0 \\ \tilde{\beta}_t = \frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}\beta_t \end{cases} \tag{5}$$

In the reverse denoising process, $\hat{x}_0^t$ is predicted by a conditional denoising model $\hat{\epsilon}_\theta$ parameterized by θ with the input of time step t, sampling data x_0 and Gaussian noise ϵ at time t. In addition, we set y as the conditional input.

$$q(x_{0:T-1}|x_T, \hat{x}_0^t) = p(x_T)\prod_{t=1}^{T} q(x_{t-1}|x_t, \hat{x}_0^t) \tag{6}$$

$$\hat{x}_0^t = \frac{1}{\sqrt{\bar{\alpha}_t}}(x_t - \sqrt{1-\bar{\alpha}_t})\hat{\epsilon}_\theta(x_0, y, t) \tag{7}$$

The object of the training process is to minimize the Kullback–Leibler (KL) divergence between the noise added during the diffusion process $q(x_{t-1}|x_t, x_0)$ and the noise predicted denoising model ϵ_θ in $p_\theta(x_{t-1}|x_t)$, and the loss function is

$$\min_\theta L_{simple}(\theta) = E_{x_0,\epsilon,t}\left\|\epsilon - \hat{\epsilon}_\theta(\sqrt{\bar{\alpha}_t}x_{hr} + \sqrt{1-\bar{\alpha}_t}\epsilon, x_{lr}, t)\right\|^2 \tag{8}$$

In the context of image SR tasks, there is still a problem in the aforementioned process. The high-resolution images $x_{hr} \in \mathbb{R}^{rw\times rh\times c}$ used for training are continuously degraded by the addition of noise during the diffusion process, but their dimensions remain unchanged. Here, w, h, and c represent the width, height, and color channels of the image, respectively, and r represents the scale factor. If we want to incorporate $x_{lr} \in \mathbb{R}^{w\times h\times c}$ as a conditional input into the prediction network to generate a high-resolution image $x_{sr} \in \mathbb{R}^{rw\times rh\times c}$, we need to optimize the network architecture to address the issue of dimension mismatch. In SR3 [19], Saharia et al. used bicubic interpolation to upscale x_{lr} and then combine it with x_{sr} and ϵ to obtain the merged information

$$x_{merged} = Concat((x_{lr})^{bicubic}, \sqrt{\bar{\alpha}_t}x_{sr} + \sqrt{1-\bar{\alpha}_t}\epsilon) \tag{9}$$

In the work of SRFS [20], Wang et al. proposed an innovative approach that divides the SR and denoising processes into two stages. Initially, a simple Shallow PSNR-oriented Super-Resolution Network (SPSN) is trained, which consists of nearest interpolation upsample layer and three convolutional layers, to achieve low-resolution SR and obtain SPSN(LR), which is then used as a conditional input. Subsequently, based on the conditional DDPM with a large U-Net, noise prediction is carried out to realize the denoising process. Following this, temporal information t is processed through a transformer sinusoidal positional encoding [40] to obtain temporal encoding information.

$$t_e = TimeEmbedding(t) \tag{10}$$

t_e is then fed into the conditional denoising network ϵ (typically a large U-Net), which is trained to generate a high-resolution image x_{sr} corresponding to the provided low-resolution image x_{lr}.

3.2. Proposed ETDMS

Departing from the prior methods, denoising and SR are considered as two separate tasks. As illustrated in Fig. 2, DN has been integrated with CLSR to propose ETDMS. In Stage 1, LQLR image is processed by the denoising procedure DN to obtain the HQLR image, which serves as the conditional input for Stage 2 to guide the CLSR in continuously denoising the Gaussian noise image and predicting HQHR $\hat{x}_0^t$. By employing a pre-denoising process with DN, CLSR can focus more on the detail recovery in the SR task. This not only prevents the spread of noise but also significantly enhances the overall quality of the generated images.

The SwinIR network [14] was employed as DN for image denoising. SwinIR is an image restoration network based on the Swin Transformer [41], which effectively restores image details through hierarchical feature extraction and a high-quality image reconstruction module. The LQLR image obtained from the fast scanning mode is processed by a trained DN to generate cleaned HQLR image, denoted by

$$x_{HQLR} = DN(x_{LQLR}) \tag{11}$$

In Stage 2, inspired by LDM [34] and LWTDM [32], the large U-Net network previously employed in the reverse denoising process of the diffusion model has been discarded, opting instead for a conditional lightweight end-to-end encoder–decoder architecture that incorporates cross-attention modules, named CLSR. The network structure of CLSR is illustrated in Fig. 3.

In the encoder segment, temporal encoding information t_e is obtained based on (10). The coded features are divided into two separate branches. The cleaned HQLR image generated in Stage 1, serving as conditional image, is introduced as input to the low-resolution branch, while the noisy sample X_t and t_e serve as inputs for the high-resolution branch. Both branches are subjected to identical sequences of ResBlocks(Res(·)) and PixelUnshuffle(PUS(·)) operations. The output from the low-resolution branch, upscaled through a bicubic interpolation upsample module(ups(·)), is integrated into the high-resolution branch. This process utilizes elemental reconstruction, which is able to reduce the dimensionality of the features by a factor q, and the whole

Table 1
Specific information of the dataset.

Dataset	LQLR images (pixels/image)	HQLR images (pixels/image)	HQHR images (pixels/image)
Scanning speed	3(1 μs/pixel)	5(10 μs/pixel)	5(10 μs/pixel)
Raw data size	1024 × 768	1024 × 768	4096 × 3072
Cropped data size	64 × 64	64 × 64	256 × 256

process is still nonparametric. After pixelunshuffle, the feature maps' length and width are reduced to $1/q$ of their original sizes, and the number of channels d becomes $d' = d \times q^2$. The encoded features from the low-resolution and high-resolution branches can be respectively expressed as:

$$z_{lr} = PUS(Res(x_{HQLR})) \tag{12}$$

$$z_{hr} = PUS(Res(X_t, t_e)) + ups(z_{lr}) \tag{13}$$

In the middle feature extraction layer, we introduce the lightweight embedding module named CrossAttention Module (CAM) that utilizes cross-attention mechanisms to incorporate conditional information into the high-resolution branch. N identical CAM modules are instantiated, denoted as CAM_n for $n = 1, 2, \ldots, N$. The representation for each module CAM_n is delineated as follows:

$$CrossAttn(\boldsymbol{q}_n, \boldsymbol{k}_n, \boldsymbol{v}_n) = softmax(\frac{\boldsymbol{q}_n \boldsymbol{k}_n^T}{\sqrt{d'}}) \cdot \boldsymbol{v}_n \tag{14}$$

where $\boldsymbol{q}_n = W_q^n \cdot Res(z_{hr})$, $\boldsymbol{k}_n = W_k^n \cdot Res(z_{lr})$ and $\boldsymbol{v}_n = W_v^n \cdot Res(z_{lr})$. n represents the module index and W_*^n is a learnable projection matrix. During training, each CAM is tasked with updating the high-resolution branch's z_{hr} through the forward propagation process

$$z_{hr}^n \leftarrow z_{hr}^{n-1} + CrossAttn(\boldsymbol{q}_n, \boldsymbol{k}_n, \boldsymbol{v}_n) \tag{15}$$

where z_{hr}^n is taken as the output of CAM_n.

In the final decoder section, a design opposite to that in the encoder is employed, efficiently decoding z_{hr}^N through ResBlocks and PixelShuffle(PS(·)) to obtain the predicted noise $\widetilde{\epsilon}$

$$\widetilde{\epsilon} = PS(Res(z_{hr}^N, t_e)) \tag{16}$$

3.3. Distinction from SRFS

While SRFS [20] also proposes a two-stage DDPM super-resolution strategy, ETDMS introduces fundamental distinctions, particularly regarding noise handling and efficiency:

- First Stage Design & Noise Handling: SRFS's first stage upsamples noisy LQLR without prior denoising, fundamentally leading to noise diffusion across the higher resolution, introducing artifacts. In contrast, ETDMS's first stage performs dedicated pre-denoising, providing a clean input to prevent noise propagation.
- Second Stage Efficiency: ETDMS employs a lightweight CLSR with cross-attention and integrates DDIM sampling, significantly reducing computational burden and accelerating inference compared to SRFS's larger U-Net.

4. Experiments

4.1. Implementation details

4.1.1. Datasets

In this paper, images with various morphological features were captured from multiple superalloy samples to construct the dataset. These samples include single-crystal superalloy SX and polycrystalline superalloy IN718 under high and room-temperature tensile fracture conditions. Details regarding the composition, dimensions, and fabrication techniques of these samples can be found in our previous work [42,43]. All images were captured using the TESCAN Clara field emission scanning electron microscope. We captured images from 60 distinct locations on multiple superalloy samples. For each of these 60 locations, three corresponding images were acquired, all based on specific scanning speeds and resolution settings. The image quality was influenced by the scanning speed, which was set to 3 and 5, while the resolutions were 1024 × 768 pixels and 4096 × 3072 pixels, respectively, to meet the requirements of a fourfold SR task. Subsequently, these images were cropped, and 8,000 pairs were obtained for training, 2,000 for validation, and 1,520 for testing. Detailed information about the dataset is shown in Table 1.

4.1.2. Evaluation metrics

For evaluation metrics, we employ general PSNR(↑) and structural similarity index (SSIM)(↑) to assessing image accuracy and perceptual quality. To comprehensively assess the quality of the generated SR SEM images, we employed the Fréchet Inception Distance (FID)(↓) to evaluate the stylistic differences between the generated images and the HQHR images, as well as LPIPS(↓) to calculate the perceptual similarity between the generated images and the HQHR images, thereby simulating human subjective perception. Finally, we compared the time required for different methods to SR LQLR (64 × 64 pixels) to HQHR (256 × 256 pixels). All image quality tests and inference time calculations were performed on an NVIDIA GeForce RTX4090 D Graphics Processing Unit (GPU).

4.1.3. Training settings

In this paper, the scale factor r is set to 4. In Stage 1, we train DN for 200 epochs (1000 iterations per epoch) with Adam optimizer ($\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$) and the learning rate α is 2×10^{-4}. In stage 2, we train CLSR for 1000 epochs (2000 iterations per epoch) with Adam optimizer ($\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-4}$), and network parameters $N = 2$, $q = 4$. During the training step, the time step is fixed at 2000, and the initial learning rate α is 10^{-4}. The source code for ETDMS is openly available on GitHub at https://github.com/Elton0427/ETDMS under the MIT license.

4.2. ETDMS analysis

To conduct a detailed analysis of ETDMS, we compared the generated images and evaluation metrics at two stages under different settings, as shown in Table 2. The HQLR images output in Stage 1 have excellent evaluation metrics, mainly because the deep feature extraction module in the SwinIR network can effectively extract the detailed features of the input SEM images, eliminate image noise while preserving as much valid information as possible, and form cleaned HQLR images. As shown in Table 2, Stage 1 (DN) delivers HQLR images with a high PSNR of 32.71, as it is specifically optimized for image denoising and signal-to-noise ratio improvement. It is important to clarify that Stage 2 (CLSR), as a generative diffusion model for super-resolution, primarily focuses on reconstructing high-frequency details and achieving superior perceptual quality (reflected by lower FID and LPIPS). While Stage 2 also involves a denoising process (reversing the diffusion process noise), it is not designed to further remove original image acquisition noise. Stage 1's pre-denoising is thus crucial, providing a clean foundation and enabling Stage 2 to concentrate on realistic and perceptually high-quality SR image generation. Stage 2 is implemented based on diffusion models, which are generative SR models that can provide more similar image styles and better visual quality. In practical implementation, two sets of models were trained separately according to different training data, namely captured HQLR-captured HQHR and cleaned HQLR-capture HQHR. The results show that the model trained solely from captured images can achieve better PSNR and SSIM. However, the model trained with the cleaned images

Table 2
Quantitative results of generated HQHR images with different settings and sampling steps.

Method	Train input	Sampling steps T_s	PSNR(↑)	SSIM(↑)	FID(↓)	LPIPS(↓)	Infertime(↓)
ETDMS Stage 1: DN	captured LQLR	–	32.71	0.883	44.05	0.1012	0.0421
ETDMS Stage 2: CLSR	captured HQLR	50	23.76	0.704	63.38	0.3900	0.234
		100	23.82	0.698	53.53	0.3391	0.453
		200	23.97	0.694	48.53	0.2954	0.894
		500	23.87	0.687	45.19	0.2564	2.26
		1000	23.86	0.683	44.79	0.2437	4.44
		2000	25.63	0.691	44.06	0.2427	8.9
ETDMS Stage 2: CLSR	cleaned HQLR	50	22.44	0.680	45.08	0.2684	0.234
		100	22.56	0.665	42.75	0.2218	0.453
		200	**22.54**	**0.652**	**42.79**	**0.1960**	**0.894**
		500	22.54	0.638	42.85	0.1840	2.26
		1000	22.56	0.632	42.52	0.1832	4.44
		2000	**24.96**	**0.643**	**39.03**	**0.1760**	**8.9**
ETDMS Single stage	captured LQLR	2000	22.73	0.593	51.62	0.2282	8.9

generated in Stage 1 as input has a significant advantage in FID and LIPIS metrics, restoring high-frequency detail information and achieving better visual quality. This is believed to be due to the significant differences between cleaned HQLR and captured HQLR. The cleaned HQLR is smoother and has a higher signal-to-noise ratio. Therefore, this configuration is set as the default setting for ETDMS, under which CLSR can focus on the generation of structural and high-frequency information. A control experiment was also conducted with a single-stage method, training CLSR with captured LQLR-captured HQHR, and the results indicate that the two-stage method has a clear advantage over this single-stage approach. In this control experiment, the CLSR network was trained directly using captured LQLR images as input and captured HQHR images as the target, thereby bypassing the initial denoising stage. As shown in Table 2, the two-stage ETDMS consistently demonstrates superior performance across all evaluation metrics compared to this single-stage approach, particularly in terms of perceptual quality metrics (FID and LPIPS), which are crucial for realistic detail generation. This clear advantage validates our design choice to separate denoising and super-resolution into distinct stages.

Diffusion models typically require thousands of sampling steps during the inference process, which results in a longer time for image generation. In CLSR, this situation was effectively improved by introducing a lightweight encoder–decoder architecture network with cross-attention modules. In addition, the DDIM sampling method was adopted to accelerate the process further. Ablation experiments were conducted to determine the optimal number of sampling steps, with the results summarized in Table 2. Fig. 4 compares four evaluation metrics under different sampling steps, emphasizing the inherent trade-off between the perceptual accuracy of the generated model and the generation time. The visualization results of CLSR with different sampling steps are shown in Fig. 5. By balancing image quality and generation efficiency, T_s=200 was selected as the default setting for ETDMS while retaining the T_s=2000 highest-quality output mode, thereby increasing the practicality of the method under real-world conditions.

In addition, ablation studies were conducted on the network parameter N to analyze its impact on experimental outcomes. We examined the effects of various N configurations on model performance and computational complexity, as shown in Table 3. Based on the findings in Table 3, the N=2 configuration was ultimately selected as the default setting for ETDMS, establishing an optimal balance between computational efficiency and output quality.

4.3. Model comparison

In this section, ETDMS is compared with several existing SR methods, including the primary interpolation-based method Bicubic, PSNR-oriented network SwinIR, enhanced super-resolution generative adversarial network ESRGAN, and the conditional diffusion model-based

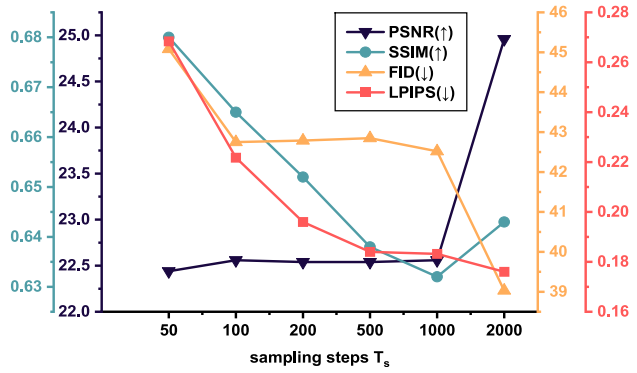


Fig. 4. Influence of sampling steps T_s on evaluation metrics of ETDMS.

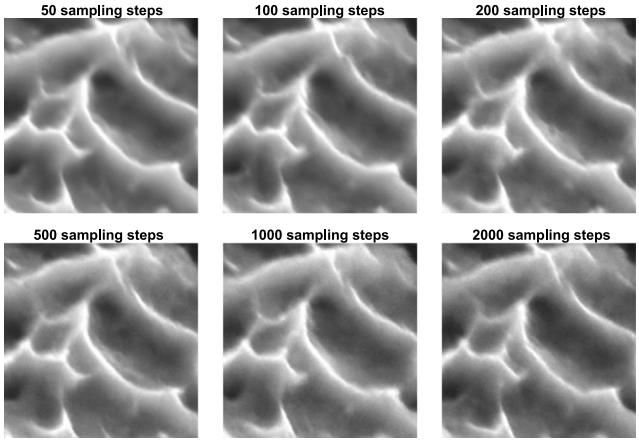


Fig. 5. Visualization results of different sampling steps of ETDMS.

network SR3. The results are shown in Table 4, where SR3 used a sampling step count of T_s=2,000. Additionally, in Table 5, statistics on FLOPs (G), Params (M), and Infertime (s) are provided so that these methods can be delved into in terms of their complexity and efficiency. The visualization results are shown in Fig. 6. The experimental results were analyzed, where Bicubic performed poorly as a primary interpolation method, which is not discussed in detail. Analysis of the remaining experimental results revealed that ESRGAN, which introduced adversarial and perceptual losses, had a good FID value but performed average in other metrics and visual effects. SwinIR, as a PSNR-oriented method, incorporated pixel loss during training and thus

Table 3
Quantitative results of generated HQHR images with different settings and sampling steps.

Method	N	Sampling steps T_s	PSNR(↑)	SSIM(↑)	FID(↓)	LPIPS(↓)	Parameters (M)	FLOPs (G)	Infertime (s)
ETDMS Stage 2: CLSR	4	200	23.86	0.657	39.27	0.1991	17.214	80.821	1.1197
		2000	25.81	0.659	36.33	0.1749			11.8
	2	200	22.54	0.652	42.79	0.1960	11.583	56.700	0.894
		2000	24.96	0.643	39.03	0.1760			8.9
	1	200	25.08	0.664	43.65	0.2070	8.768	44.639	0.786
		2000	25.55	0.640	39.50	0.1810			7.6

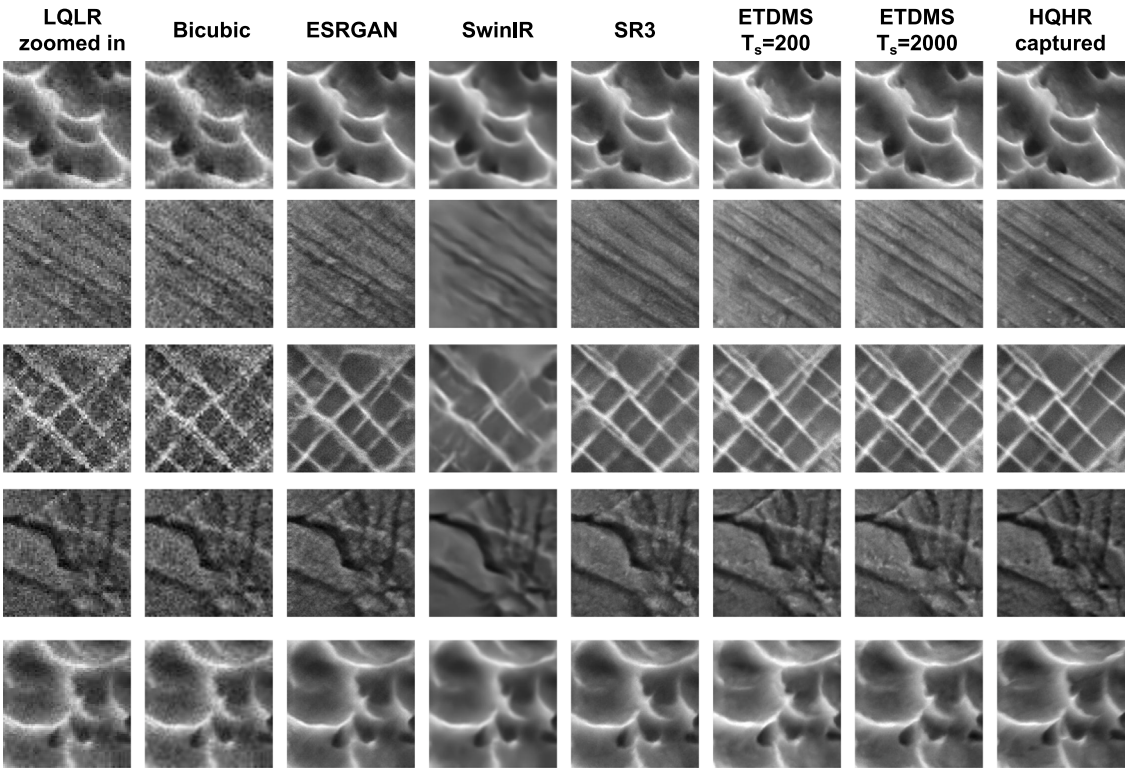


Fig. 6. Visualization results of output SR images generated by different SR methods at 4× scale factor.

had the best PSNR and SSIM. However, this also led to over-smoothing in the images generated by SwinIR, with excessive FID and LPIPS values and noticeable detail loss in the visualization results. The diffusion-based SR3 network provided satisfactory FID and LPIPS and visually was closer to captured HQHR but had lower PSNR and SSIM values, as SR3 does not calculate pixel loss. In our SR task, previous methods could not achieve satisfactory results. Our ETDMS, through its innovative two-stage structure, first effectively removes the interference caused by noise in LQLR images and then focuses on high-frequency detail recovery in SR using a diffusion model. This approach enables ETDMS to achieve the best FID and LPIPS, which are critical metrics for simulating human subjective perception and evaluating the realism of generated details, while also remaining competitive in PSNR and SSIM. In the visualization results, the SR images generated by ETDMS are also stylistically and in detail closer to the captured HQHR. Compared to the two-stage method in SRFS, ETDMS fundamentally avoids the problem of noise diffusion by pre-denoising. Furthermore, without compromising generation quality, it utilizes a lightweight model and accelerated sampling techniques to significantly enhance the efficiency of SEM SR tasks. Running the two-stage method on an RTX4090 D to achieve a 4× super-resolution task for 64 × 64 images takes less than 1 s, while SR3 requires about 25 s.

Table 4
Comparison of quantitative results by different aforementioned SR methods.

Method	PSNR(↑)	SSIM(↑)	FID(↓)	LPIPS(↓)
Bicubic	21.73	0.420	188.87	0.5837
ESRGAN	22.41	0.524	57.676	0.617
SwinIR	24.21	0.675	157.58	0.5872
SR3	23.85	0.607	41.25	0.2143
ETDMS(T_s=200)	22.54	0.652	42.79	0.1960
ETDMS(T_s=2000)	24.96	0.643	39.03	0.1760

Table 5
Complexity of the aforementioned methods.

Method	Parameters (M)	FLOPs (G)	Infertime (s)
ESRGAN	16.735	90.680	0.0249
SwinIR	11.852	50.546	0.0723
SR3	92.514	177.993	25.0039
ETDMS	11.583	56.700	0.894

5. Conclusion

In this paper, an efficient two-stage diffusion model for accelerated SEM image super-resolution named ETDMS is proposed, which enhances instrumentation characterization capability for vision-based analysis and measurement. Previous SR methods for SEM images, due to the lack of analysis and optimization of the SEM imaging principles, often result in overly smooth generated images and a lack of high-frequency details. The differences in image quality between SEM's fast and slow scanning modes have been clarified, and separated noise

removal and SR into two independent tasks, thus proposing a two-stage model. Stage 1 of ETDMS focuses on removing image noise, while Stage 2 focuses on generating high-frequency details. Unlike previous SR methods that use image synthesis or acquisition methods, a new training data acquisition strategy is proposed, and LQLR-HQLR-HQHR datasets have been captured under actual working conditions. In Stage 2, by using an encoder–decoder architecture and cross-attention modules, combined with accelerated sampling techniques, we significantly improve the generation efficiency. ETDMS has achieved a significant improvement in both perceptual image quality and generation speed compared to previous methods. In future work, the application of generative models in SEM images will continue to be explored, including the development of SEM image detail enhancement methods for specific scenarios.

CRediT authorship contribution statement

Xuecheng Zhang: Writing – original draft, Visualization, Validation, Software, Resources, Methodology. **Zixin Li:** Writing – original draft, Visualization, Validation, Software, Resources, Methodology. **Bin Zhang:** Writing – review & editing, Validation, Supervision, Data curation, Conceptualization. **Wenchao Meng:** Writing – review & editing, Methodology, Conceptualization. **Yuefei Zhang:** Writing – review & editing, Investigation, Funding acquisition. **Chaojie Gu:** Writing – review & editing, Writing – original draft, Validation, Methodology. **Xianjue Ye:** Writing – review & editing, Supervision, Funding acquisition. **Ze Zhang:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by National Science and Technology Major Project of China [grant numbers J2019-III-0008-0051], and National Natural Science Foundation of China [grant number 52401025].

Data availability

The source code for ETDMS and the corresponding dataset are openly available on GitHub under the MIT license. URL: https://github.com/Elton0427/ETDMS.

References

[1] Kenneth Kanayo Alaneme, Eloho Anita Okotete, Recrystallization mechanisms and microstructure development in emerging metallic materials: A review, J. Sci.: Adv. Mater. Devices 4 (1) (2019) 19–33.
[2] Ruiqi Liang, Yazhen Xue, Xiaowei Fu, An N. Le, Qingliang Song, Yicheng Qiang, Qiong Xie, Ruiqi Dong, Zehao Sun, Chinedum O. Osuji, et al., Hierarchically engineered nanostructures from compositionally anisotropic molecular building blocks, Nat. Mater. 21 (12) (2022) 1434–1440.
[3] Ahmad R. Kirmani, Todd A. Byers, Zhenyi Ni, Kaitlyn VanSant, Darshpreet K. Saini, Rebecca Scheidt, Xiaopeng Zheng, Tatchen Buh Kum, Ian R. Sellers, Lyndsey McMillon-Brown, et al., Unraveling radiation damage and healing mechanisms in halide perovskites using energy-tuned dual irradiation dosing, Nat. Commun. 15 (1) (2024) 696.
[4] F. Timischl, M. Date, S. Nemoto, A statistical model of signal–noise in scanning electron microscopy, Scanning 34 (3) (2012) 137–144.
[5] Dominic Chee Yong Ong, Kok Swee Sim, Single image signal-to-noise ratio (SNR) estimation techniques for scanning electron microscope–A review, IEEE Access (2024).
[6] Zhiheng Li, Gaoxin Lin, Linqin Wang, Husileng Lee, Jian Du, Tang Tang, Guoheng Ding, Rong Ren, Wenlong Li, Xing Cao, et al., Seed-assisted formation of NiFe anode catalysts for anion exchange membrane water electrolysis at industrial-scale current density, Nat. Catal. (2024) 1–9.
[7] Maryam Abdinejad, Amirhossein Farzi, Robin Möller-Gulland, Fokko Mulder, Chengyu Liu, Junming Shao, Jasper Biemolt, Marc Robert, Ali Seifitokaldani, Thomas Burdyny, Eliminating redox-mediated electron transfer mechanisms on a supported molecular catalyst enables CO2 conversion to ethanol, Nat. Catal. (2024) 1–11.
[8] K.L. Small, N.T. Henthorn, Deepa Angal-Kalinin, A.L. Chadwick, Elham Santina, Adam Aitkenhead, Karen J. Kirkby, R.J. Smith, M. Surman, J. Jones, et al., Evaluating very high energy electron RBE from nanodosimetric pBR322 plasmid DNA damage, Sci. Rep. 11 (1) (2021) 3341.
[9] Toma Susi, Jannik C. Meyer, Jani Kotakoski, Quantifying transmission electron microscopy irradiation effects using two-dimensional materials, Nat. Rev. Phys. 1 (6) (2019) 397–405.
[10] Dj Gheysari, A. Behjat, M. Haji-Saeid, The effect of high-energy electron beam on mechanical and thermal properties of LDPE and HDPE, Eur. Polym. J. 37 (2) (2001) 295–302.
[11] Sandra Stangebye, Yin Zhang, Saurabh Gupta, Ting Zhu, Olivier Pierron, Josh Kacher, Understanding and quantifying electron beam effects during in situ TEM nanomechanical tensile testing on metal thin films, Acta Mater. 222 (2022) 117441.
[12] Chao Dong, Chen Change Loy, Kaiming He, Xiaoou Tang, Learning a deep convolutional network for image super-resolution, in: Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13, Springer, 2014, pp. 184–199.
[13] Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, Yun Fu, Image super-resolution using very deep residual channel attention networks, in: Proceedings of the European Conference on Computer Vision, ECCV, 2018, pp. 286–301.
[14] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, Radu Timofte, Swinir: Image restoration using swin transformer, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 1833–1844.
[15] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al., Photo-realistic single image super-resolution using a generative adversarial network, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4681–4690.
[16] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, Chen Change Loy, Esrgan: Enhanced super-resolution generative adversarial networks, in: Proceedings of the European Conference on Computer Vision (ECCV) Workshops, 2018.
[17] Xintao Wang, Liangbin Xie, Chao Dong, Ying Shan, Real-esrgan: Training real-world blind super-resolution with pure synthetic data, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 1905–1914.
[18] Jonathan Ho, Ajay Jain, Pieter Abbeel, Denoising diffusion probabilistic models, Adv. Neural Inf. Process. Syst. 33 (2020) 6840–6851.
[19] Chitwan Saharia, Jonathan Ho, William Chan, Tim Salimans, David J. Fleet, Mohammad Norouzi, Image super-resolution via iterative refinement, IEEE Trans. Pattern Anal. Mach. Intell. 45 (4) (2022) 4713–4726.
[20] Zelin Wang, Rui Feng, Jikai Mao, Xingyu Tao, Deao Zhang, Xuliang Shan, Jianguang Zhou, SRFS: Super resolution in fast-slow scanning mode of SEM based on denoising diffusion probability model, IEEE Trans. Instrum. Meas. (2024).
[21] Shervin Shirmohammadi, Alessandro Ferrero, Camera as the instrument: The rising trend of vision based measurement, IEEE Instrum. Meas. Mag. 17 (3) (2014) 41–47.
[22] Di Tian, Dangjun Zhao, Dongyang Cheng, Junchao Zhang, Lidar super-resolution based on segmentation and geometric analysis, IEEE Trans. Instrum. Meas. 71 (2022) 1–17.
[23] Wanxin Xiao, Yafei Zhang, Hongbin Wang, Fan Li, Hua Jin, Heterogeneous knowledge distillation for simultaneous infrared-visible image fusion and super-resolution, IEEE Trans. Instrum. Meas. 71 (2022) 1–15.
[24] Xiaodong Li, Ning Fu, Pinjun Zheng, Liyan Qiao, A compact super resolution phase retrieval method for Fourier measurement, IEEE Trans. Instrum. Meas. 71 (2022) 1–15.
[25] Yang Zhao, Rongbiao Zhu, Kai Zhang, Haotian Yu, Lianfa Bai, Dongliang Zheng, Jing Han, Accurate dynamic 3-D shape measurement based on the fringe pattern super-reconstruction technique, Measurement 200 (2022) 111575.
[26] Shitong Luo, Wei Hu, Diffusion probabilistic models for 3d point cloud generation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 2837–2845.
[27] William Harvey, Saeid Naderiparizi, Vaden Masrani, Christian Weilbach, Frank Wood, Flexible diffusion modeling of long videos, Adv. Neural Inf. Process. Syst. 35 (2022) 27953–27965.
[28] Dmitry Baranchuk, Ivan Rubachev, Andrey Voynov, Valentin Khrulkov, Artem Babenko, Label-efficient semantic segmentation with diffusion models, 2021, arXiv preprint arXiv:2112.03126.
[29] Emmanuel Asiedu Brempong, Simon Kornblith, Ting Chen, Niki Parmar, Matthias Minderer, Mohammad Norouzi, Denoising pretraining for semantic segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 4175–4186.
[30] Haoying Li, Yifan Yang, Meng Chang, Shiqi Chen, Huajun Feng, Zhihai Xu, Qi Li, Yueting Chen, Srdiff: Single image super-resolution with diffusion probabilistic models, Neurocomputing 479 (2022) 47–59.

[31] Jinzhe Liu, Zhiqiang Yuan, Zhaoying Pan, Yiqun Fu, Li Liu, Bin Lu, Diffusion model with detail complement for super-resolution of remote sensing, Remote. Sens. 14 (19) (2022) 4834.
[32] Tai An, Bin Xue, Chunlei Huo, Shiming Xiang, Chunhong Pan, Efficient remote sensing image super-resolution via lightweight diffusion models, IEEE Geosci. Remote. Sens. Lett. (2023).
[33] Zhanxiong Wu, Xuanheng Chen, Sangma Xie, Jian Shen, Yu Zeng, Super-resolution of brain MRI images based on denoising diffusion probabilistic model, Biomed. Signal Process. Control. 85 (2023) 104901.
[34] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, Björn Ommer, High-resolution image synthesis with latent diffusion models, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 10684–10695.
[35] Jiaming Song, Chenlin Meng, Stefano Ermon, Denoising diffusion implicit models, 2020, arXiv preprint arXiv:2010.02502.
[36] Hyungjin Chung, Byeongsu Sim, Jong Chul Ye, Come-closer-diffuse-faster: Accelerating conditional diffusion models for inverse problems through stochastic contraction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 12413–12422.
[37] Shahar Tsiper, Or Dicker, Idan Kaizerman, Zeev Zohar, Mordechai Segev, Yonina C. Eldar, Sparsity-based super resolution for SEM images, Nano Lett. 17 (9) (2017) 5437–5445.
[38] Kevin de Haan, Zachary S. Ballard, Yair Rivenson, Yichen Wu, Aydogan Ozcan, Resolution enhancement in scanning electron microscopy using deep learning, Sci. Rep. 9 (1) (2019) 12050.
[39] Jia Wang, Chuwen Lan, Caiyong Wang, Zehua Gao, Deep learning super-resolution electron microscopy based on deep residual attention network, Int. J. Imaging Syst. Technol. 31 (4) (2021) 2158–2169.
[40] Ashish Vaswani, Attention is all you need, 2017, arXiv preprint arXiv:1706.03762.
[41] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, Baining Guo, Swin transformer: Hierarchical vision transformer using shifted windows, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 10012–10022.
[42] Wenqiang Zhang, Shuang Chen, Guanghao Guo, Xuan Liu, Bin Zhang, Xianjue Ye, Jin Wang, Yuefei Zhang, Xinbao Zhao, Ze Zhang, Study on plastic deformation mechanism of a fourth-generation nickel-based single crystal superalloy: in-situ SEM characterization and dislocation density-based model, Mater. Sci. Eng.: A (2025) 148132.
[43] Xuecheng Zhang, Guanghao Guo, Zixin Li, Wenchao Meng, Yuefei Zhang, Qing Ye, Jin Wang, Shibo He, Xinbao Zhao, Jiming Chen, et al., Superalloys fracture process inference based on overlap analysis of 3D models, Commun. Eng. 3 (1) (2024) 108.

Micron 199 (2025) 103897

Contents lists available at ScienceDirect

Micron

journal homepage: www.elsevier.com/locate/micron

Full length article

Knowledge-distilled swin transformer for efficient vibration artifact removal in SEM imaging

Xuecheng Zhang [a,b], Bin Zhang [a], Wenchao Meng [b], Yuefei Zhang [a,*], Xianjue Ye [a], Ze Zhang [a]

[a] Institute of Superalloys Science and Technology, School of Materials Science and Engineering, Zhejiang University, Hangzhou, 310027, China
[b] College of Control Science and Engineering, Zhejiang University, Hangzhou, 310027, China

ARTICLE INFO

Keywords:
Scanning electron microscopy (SEM)
Image restoration
Vibration artifact removal
Swin transformer
Knowledge distillation

ABSTRACT

This study addresses the challenge of removing vibration-induced horizontal stripe artifacts in Scanning Electron Microscopy (SEM) images, which degrade image fidelity and compromise quantitative analysis. The main contribution lies in the development of SwinIR-KD, a computationally efficient image restoration framework. This framework synergistically integrates the Swin Transformer architecture with knowledge distillation and a novel Horizontal Stripe Suppress Loss. Specifically, SwinIR-KD employs a lightweight student model, distilled from a larger pre-trained SwinIR teacher model, and incorporates architectural modifications to reduce complexity. A specialized loss function, combining reconstruction, distillation, and stripe suppression components, guides the training. Experimental results on an SEM speckle image dataset demonstrate that SwinIR-KD significantly reduces model parameters and computational complexity by approximately 79% compared to the baseline SwinIR, while achieving comparable or even superior image restoration performance in terms of PSNR, SSIM, FID, and LPIPS. Furthermore, SwinIR-KD effectively processes large-scale SEM images and is complemented by a user-friendly interface for practical application.

1. Introduction

Scanning Electron Microscopy (SEM) stands as an indispensable characterization tool within the field of materials science. Its capacity for high-resolution imaging is pivotal for elucidating the intrinsic correlations among the microstructure, morphology, and macroscopic properties of diverse materials (Panwar et al., 2020; Zhang et al., 2024; Guo et al., 2025). For example, SEM is extensively employed to scrutinize the surface characteristics of alloys (Zhang et al., 2025a), nanomaterials (Suga et al., 2014), and composite materials (Chowdhury et al., 2025). This exceptional imaging prowess establishes SEM as the preferred technique for investigating nanoscale structures. Nevertheless, the image acquisition process in SEM is notably susceptible to external environmental disturbances, especially mechanical vibrations originating from other laboratory equipment or personnel activities (Jung et al., 2012; Shin et al., 2020). Such vibrations typically manifest as horizontal stripe artifacts in the acquired images (Fig. 1), which fundamentally degrade image fidelity and can severely compromise the accuracy of subsequent quantitative analyses in materials science. For instance, such distortions can lead to erroneous measurements of critical microstructural features like grain size, phase distribution, and defect density, thereby affecting the reliability of downstream analysis pipelines, such as automated segmentation or classification.

Addressing these vibration-induced artifacts is crucial for maintaining the integrity of SEM data and ensuring reliable scientific conclusions. Traditional strategies for mitigating these artifacts have primarily relied on hardware solutions, such as high-performance vibration isolation tables. While effective, their substantial cost often limits widespread adoption in resource-constrained environments. Software-based correction methods have also been explored. For instance, signal processing techniques that incorporate frequency analysis and phase control are particularly suited for periodic artifacts like vibration stripes (Jung et al., 2012). Other algorithmic approaches, such as those based on Fourier filtering and variational methods, have been developed specifically for stripe removal by targeting frequency coefficients or optimizing image smoothness while preserving structures (Rottmayer et al., 2024). General digital image processing methods, encompassing denoising and distortion correction, may also indirectly mitigate the impact of vibrations (Płuska et al., 2009). However, these conventional software approaches frequently encounter limitations such as low computational efficiency, inadequate capacity for handling complex or large-scale images, and difficulties in achieving real-time processing, especially when dealing with diverse and complex real-world degradations (Appalanaidu et al., 2023). These inherent

* Corresponding author.
E-mail address: yfzhang76@zju.edu.cn (Y. Zhang).

https://doi.org/10.1016/j.micron.2025.103897
Received 17 June 2025; Received in revised form 7 August 2025; Accepted 12 August 2025
Available online 21 August 2025
0968-4328/© 2025 Published by Elsevier Ltd.

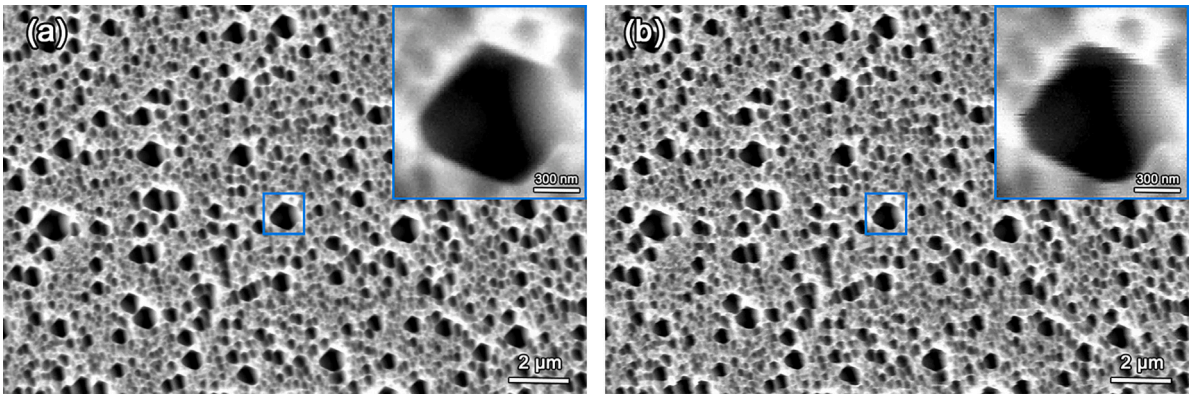


Fig. 1. (a) Ground truth image; (b) Image with vibration artifacts. The difference between the images can be viewed by zooming in on the details.

limitations underscore an urgent demand for more efficient and cost-effective computational techniques to effectively eliminate vibration artifacts in SEM images.

In recent years, deep learning techniques have emerged as a powerful paradigm in electron microscopy image processing, demonstrating remarkable performance across a range of tasks including image segmentation (Wang et al., 2022), feature recognition (Jain, 2024), and denoising (Oho et al., 2019; Weigert et al., 2018). Specifically, Convolutional Neural Networks (CNNs) have shown success in SEM image denoising and resolution enhancement (De Haan et al., 2019). More recently, Vision Transformers (ViTs), particularly the Swin Transformer, have introduced a transformative impact by effectively capturing long-range dependencies with superior computational efficiency and performance (Liu et al., 2021). Building upon this, the SwinIR model has achieved further improvements in various image restoration tasks, such as denoising, deblurring, and super-resolution (Liang et al., 2021). Despite their outstanding performance, these advanced deep learning models, especially Transformer-based architectures, often come with a large number of parameters and high computational complexity, which can hinder their practical deployment in resource-constrained environments, particularly when processing large-sized SEM images (Jiang et al., 2024). To address this bottleneck, Knowledge Distillation (KD) offers an effective model compression paradigm, training a compact student model to mimic the behavior and knowledge of a larger, more accurate teacher model (Hinton et al., 2015; Gou et al., 2021). KD has been successfully applied in image restoration to develop lightweight denoising models (Chen et al., 2021; Wang et al., 2023).

This paper introduces SwinIR-KD, a computationally efficient framework designed to eliminate vibration artifacts in SEM images. SwinIR-KD synergistically integrates the Swin Transformer architecture with knowledge distillation techniques and incorporates a horizontal stripe suppress loss specifically targeting these artifacts. The efficacy of this framework has been rigorously validated on an experimentally acquired SEM speckle image dataset. Notably, SwinIR-KD achieves a substantial reduction in computational complexity, evidenced by approximately 79% fewer parameters, while concurrently maintaining or even enhancing key image restoration performance metrics. Furthermore, to augment its practical utility, we have developed a user-friendly interface that facilitates direct comparison of processing outcomes from both the SwinIR and SwinIR-KD models. The main contributions of this work are subsequently outlined as follows:

- We propose a novel adaptation of the SwinIR model specifically tailored for vibration stripe removal in SEM images, addressing the unique characteristics of these artifacts. We introduce a customized Horizontal Stripe Suppress Loss that effectively targets the directional nature of vibration stripes, significantly enhancing the model's capability to suppress these specific artifacts.
- We develop SwinIR-KD, a significantly lightweight variant of SwinIR through a carefully designed knowledge distillation framework. Our approach achieves remarkable model compression (78.9% reduction in parameters and FLOPs) while maintaining competitive performance, with processing efficiency improvements of up to 58.2% for large-scale image processing tasks.
- We design an intelligent tiling strategy to effectively process high-resolution SEM images that exceed typical GPU memory constraints. This approach incorporates adaptive patch-based processing with weighted blending mechanisms, enabling seamless artifact removal across tile boundaries while maintaining global image consistency.

2. Related work

2.1. Electronic microscopy image restoration

SEM is a crucial characterization tool, but its image acquisition is highly susceptible to external mechanical vibrations, which typically manifest as horizontal stripe artifacts (Jung et al., 2012; Shin et al., 2020). These artifacts severely degrade image fidelity and can compromise the accuracy of subsequent quantitative analyses, such as measurements of grain size, phase distribution, and defect density, thereby affecting downstream analysis pipelines like automated segmentation or classification (Yi and Voyles, 2011).

Traditional strategies for mitigating vibration artifacts primarily rely on hardware solutions, such as high-performance vibration isolation tables. Although effective, their substantial cost often limits widespread adoption, particularly in resource-constrained environments. Software-based correction methods have also been explored. For stripe artifact removal, methods often fall into categories like Fourier filtering and variational methods. Fourier filtering approaches assume stripe artifacts are directional and repetitive, encoding their information in specific frequency coefficients. By appropriately damping or removing these coefficients in the Fourier domain, artifacts can be significantly reduced, though careful selection is needed to preserve structural information (Rottmayer et al., 2024). Variational methods, on the other hand, define denoising as an optimization problem, often using regularization techniques like Total Variation (TV) to preserve edges while smoothing noise (Tawfik et al., 2022). General digital image processing methods, encompassing denoising and distortion correction, may also indirectly mitigate the impact of vibrations (Płuska et al., 2009). More broadly, traditional digital image denoising methods can be categorized into filter techniques (e.g., bilateral filtering), sparse coding (e.g., K-SVD), and model-based denoising (e.g., Block Matching and 3D filtering (BM3D), Wiener filter) (Yi and Voyles, 2011; Mao et al., 2025). While these methods can effectively remove noise and preserve features, they often suffer from limitations such as low computational efficiency, high computational complexity for high-resolution

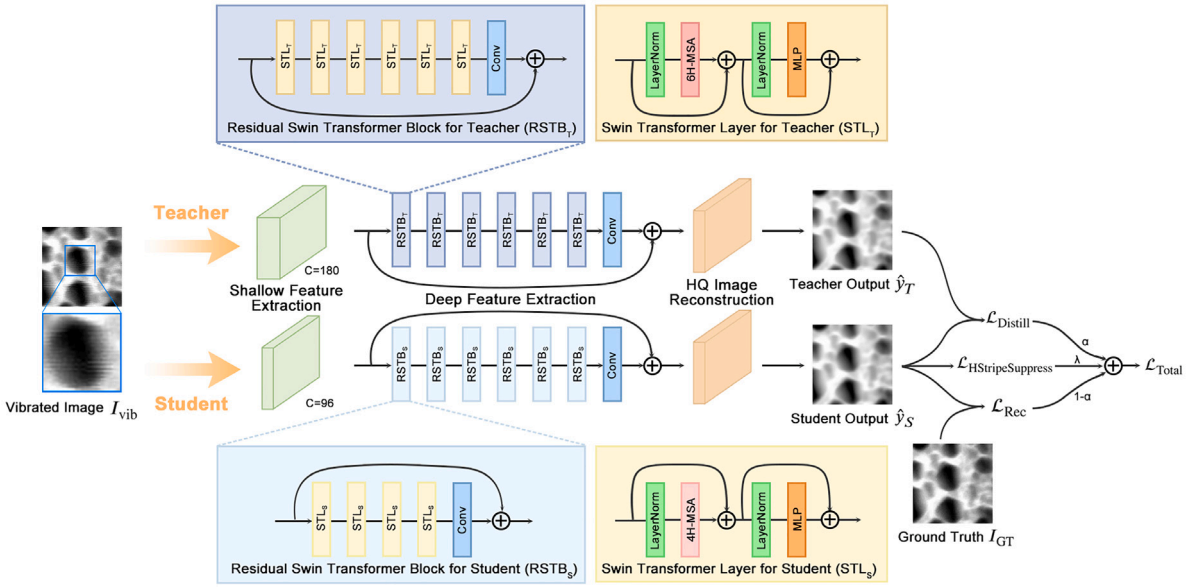


Fig. 2. Overview of the proposed SwinIR-KD framework. The vibrated image I_{vib} is processed by both the pre-trained SwinIR teacher model and the lightweight SwinIR-KD student model. The student model is trained using a composite loss function, including reconstruction loss ($\mathcal{L}_{Rec}$), distillation loss ($\mathcal{L}_{Distill}$) from the teacher's output ($\hat{y}_T$), and a horizontal stripe suppression loss ($\mathcal{L}_{HStripeSuppress}$), to produce the restored student output ($\hat{y}_S$). Key differences between teacher (RSTB$_T$/STL$_T$) and student (RSTB$_S$/STL$_S$) modules: channel reduction (C).

images, and difficulties in achieving real-time processing, particularly when generalizing to various real-world scenarios or complex noise types (Appalanaidu et al., 2023).

In recent years, deep learning techniques have achieved significant advancements in electron microscopy image processing, demonstrating outstanding performance in tasks like image segmentation (Wang et al., 2022), feature recognition (Jain, 2024), and denoising (Oho et al., 2019). Specifically, Convolutional Neural Networks (CNNs) have been successfully employed for SEM image denoising and resolution enhancement (De Haan et al., 2019). Content-Aware Image Restoration (CARE) networks, for instance, utilize CNNs trained on noisy and clean image pairs to restore microscopy data, often outperforming traditional methods even under low signal-to-noise ratio conditions (Weigert et al., 2018). This data-driven approach allows for learning complex noise characteristics and structural information.

2.2. Advanced deep learning image restoration: Focusing on transformer

More recently, Vision Transformers (ViTs) have introduced a transformative impact. Notably, the Swin Transformer, characterized by its hierarchical window-based multi-head self-attention mechanism, effectively overcomes the limitations of traditional CNNs in capturing long-range dependencies, thereby offering superior computational efficiency and performance (Liu et al., 2021). Building upon this architecture, the SwinIR model has realized further improvements in diverse image restoration tasks, such as denoising, deblurring, and super-resolution (Liang et al., 2021).

The field of Transformer-based image restoration continues to evolve rapidly (Jiang et al., 2024; Zhang et al., 2025b). Recent models like HINT (Hierarchical multi-head atteNtion driven Transformer) address redundancy issues in standard Multi-Head Attention (MHA) through Hierarchical Multi-Head Attention (HMHA) and Query-Key Cache Updating (QKCU) mechanisms, learning more diverse contextual features and enhancing interactions between attention heads (Zhou et al., 2025). Additionally, dual-Transformer models such as Dformer and Rformer propose an all-in-one image restoration approach by leveraging frequency decomposition to understand how degradations affect different frequency bands, enabling frequency-aware degradation estimation and degradation-adaptive restoration with a single model for multiple degradation types (Chu et al., 2024).

Despite the excellent performance of advanced Transformer models like SwinIR, their large parameter size and high computational complexity significantly hinder their deployment in practical resource-constrained environments, especially when processing large-sized images (Jiang et al., 2024). This computational demand poses a significant barrier to real-world application, making model compression techniques indispensable.

2.3. Knowledge distillation for image restoration model compression

Knowledge Distillation (KD) provides an effective model compression paradigm to address the computational bottleneck of large models. KD trains a compact and efficient student model to mimic the behavior and knowledge encapsulated by a larger and more accurate teacher model (Hinton et al., 2015; Gou et al., 2021). This process typically involves transferring soft targets (e.g., probability distributions) or intermediate feature representations from the teacher model to guide the student's learning (Li et al., 2023).

Various model compression techniques exist, including pruning, quantization, low-rank decomposition, and lightweight model design. Pruning reduces parameters by eliminating unnecessary weights or neurons, though it can lead to irregular sparsity or performance degradation. Quantization reduces numerical precision to decrease memory usage and improve efficiency, but may result in accuracy loss if not handled properly. Low-rank decomposition approximates weight matrices to reduce redundancy, but the decomposition itself can be computationally expensive and require extensive retraining. Lightweight model design involves creating inherently smaller, efficient architectures, which may have poorer generalization or be difficult to combine with other compression methods. In contrast to these methods, knowledge distillation guides the student model to learn more robust and generalizable feature representations, which is crucial for preserving fine details in image restoration tasks (Li et al., 2023). This knowledge transfer aims to maintain or even enhance the student's performance by leveraging the powerful representations of the teacher model, making it a preferred choice for complex tasks where perceptual quality is paramount.

Knowledge distillation is not merely a model compression technique; it is a mechanism for transferring robust and generalizable

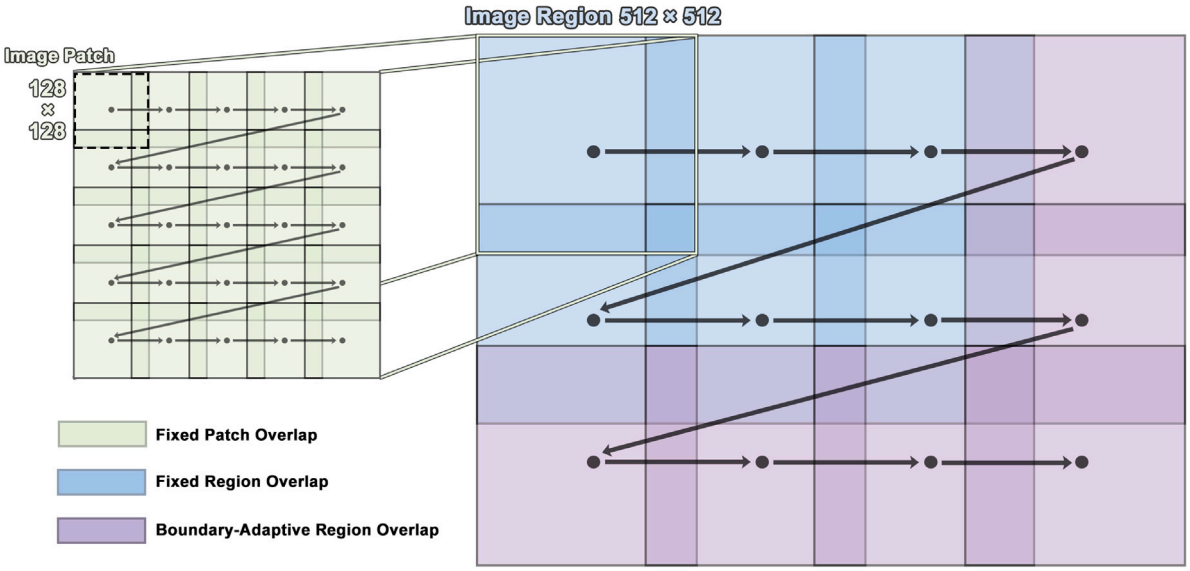


Fig. 3. Illustration of the Large Image Processing Strategy. High-resolution SEM images are processed using a hierarchical approach involving fixed region overlap, fixed patch overlap within regions, and boundary-adaptive region overlap to ensure seamless stitching and efficient handling of images exceeding GPU memory limits. An example image region size of 512 × 512 pixels is shown.

feature representations, leading to perceptually superior results that can even surpass larger teacher models. This is achieved by guiding the student model to learn a better-regularized function, thereby mitigating overfitting and yielding perceptually more pleasing and accurate results.

In the field of image restoration, the KD framework has been successfully applied to train lightweight denoising models (such as U-Net variants (Chen et al., 2021)) and to handle tasks like denoising unpaired low-dose CT images (Wang et al., 2023). Recently, frameworks like Simultaneous Learning Knowledge Distillation (SLKD) have been proposed, employing a dual-teacher, single-student architecture with distinct learning strategies to significantly reduce FLOPs and parameters while maintaining strong performance in tasks like deraining, deblurring, and denoising (Zhang and Yan, 2025).

3. Methodology

This section provides a comprehensive description of the methodologies used in this study to develop and evaluate the SwinIR-KD model for vibration artifact removal in SEM images. We begin by detailing the preparation of our specialized SEM image dataset, covering sample fabrication, image acquisition, and artifact induction. Subsequently, we present the architecture of the proposed SwinIR-KD model, explaining the knowledge distillation framework, the specific lightweight modifications to the student network, and the composite loss function designed to optimize reconstruction fidelity and artifact suppression. Finally, we describe the strategy for processing large-scale SEM images, which addresses typical memory limitations while maintaining processing integrity.

3.1. Dataset preparation

To evaluate the impact of vibration artifacts on microstructural characterization, two distinct nickel-based superalloy samples were selected: a polycrystalline IN718 alloy with artificial speckle patterns and a proprietary fourth-generation nickel-based single-crystal superalloy developed by the Institute of Superalloys Science and Technology, Zhejiang University (Zhang et al., 2025a). Both samples were mechanically ground to 4000 grit followed by meticulous polishing. Electrolytic polishing was performed using a 10 vol.% $HClO_4$ solution at 18 V for 15 s. For the IN718 alloy, speckle patterns (diameter: 0.5–1.5 μm) were generated by chemical etching in a solution of 5 g $CuCl_2$, 100 mL HCl, and 100 mL C_2H_5OH for 40 s. For the nickel-based single-crystal sample, gold nanoparticle speckles (diameter: 100 nm) were deposited via nanoscale thin-film sputtering (40 mA current, 40 s duration) using a film remodeling technique, following methodologies detailed in (Di Gioacchino and Quinta da Fonseca, 2013).

Image acquisition was conducted using a CIQTEK SEM 5000 Pro system under 15 kV accelerating voltage, 6 μs beam dwell time, and an 11 mm working distance. Vibration artifacts were induced during imaging by introducing controlled mechanical disturbances: (1) operation of a mechanical vacuum pump to simulate equipment-induced vibrations, and (2) floor-impact excitations from human activity to replicate common laboratory environmental noise. A total of 43 high-resolution image pairs (1536 × 1024 pixels) of vibrated images and artifact-free ground truth (GT) images were captured. Of these, 40 image pairs were cropped into 128 × 128 patches for model training and evaluation, then partitioned into training (60%), validation (20%), and testing (20%) sets. The remaining 3 full-size image pairs were reserved for large-scale vibration removal evaluation.

3.2. SwinIR-KD model architecture

The SwinIR (Swin Transformer for Image Restoration) model, which is particularly well-suited for removing SEM vibration artifacts, operates in three distinct stages. The SwinIR architecture was selected as our base model for several reasons. Its hierarchical window-based multi-head self-attention mechanism is particularly well-suited for SEM image restoration because it can effectively capture both local fine details and long-range dependencies, which is crucial for preserving intricate microstructures while simultaneously removing global stripe patterns.

First, shallow feature extraction is performed using convolutional layers to capture low-level image characteristics. Second, deep feature extraction and transformation are accomplished by leveraging Residual Swin Transformer Blocks (RSTBs) and Swin Transformer Layers (STLs); RSTBs enhance the learning process through residual connections, while STLs employ a shifted window self-attention mechanism for efficient feature processing. Finally, a reconstruction stage generates the restored high-quality image. While various model compression techniques exist, such as quantization, pruning, and low-rank decomposition, KD was chosen for SwinIR-KD due to its unique advantage

in transferring knowledge and maintaining performance. Unlike other methods that may sacrifice performance for size reduction, KD guides the student model to learn a more robust and generalizable feature representation, which is critical for the fine-grained details in image restoration tasks. Building upon this foundation, in this paper, we present the SwinIR-KD model, a lightweight image restoration framework designed to effectively remove vibration artifacts from SEM images while minimizing computational complexity and parameter count (see Fig. 2). The model leverages knowledge distillation, learning from a pre-trained SwinIR network (teacher branch) to guide a streamlined student branch, SwinIR-KD.

The input to the model is a vibrated image, processed through both branches, with the teacher branch producing an output denoted as $\hat{y}_T$ and the student branch yielding $\hat{y}_S$. The teacher branch employs a pre-trained SwinIR network, which serves as our enhanced baseline. This model was trained on our collected SEM dataset using a composite loss function that included both the Charbonnier loss and the horizontal stripe suppress loss with a weight of $\lambda = 0.1$. This ensures a fair performance benchmark for our proposed student model, which is engineered for efficiency through targeted lightweight modifications and is also trained using a composite loss function that balances reconstruction, distillation, and stripe suppression components. Specifically, the shallow feature extraction module reduces output channels from 180 to 96, the number of STL_S layers within each RSTB_S module decreases from 6 to 4, and the multi-head self-attention mechanism (MSA) is simplified from 6H-MSA to 4H-MSA. These adjustments substantially lower resource demands while striving to preserve restoration performance.

The training of SwinIR-KD employs a composite loss function to balance reconstruction accuracy, knowledge transfer, and artifact suppression. The total loss is defined as:

$$\mathcal{L}_{\text{Total}} = (1-\alpha)\mathcal{L}_{\text{Rec}} + \alpha\mathcal{L}_{\text{Distill}} + \lambda\mathcal{L}_{\text{HStripeSuppress}} \tag{1}$$

where $\mathcal{L}_{\text{Rec}}(\hat{y}_S, I_{\text{GT}})$ measures the reconstruction error between the student output $\hat{y}_S$ and the ground truth image I_{GT}, $\mathcal{L}_{\text{Distill}}(\hat{y}_S, \hat{y}_T)$ enforces similarity between the student and teacher outputs for knowledge distillation, $\mathcal{L}_{\text{HStripeSuppress}}(\hat{y}_S)$ specifically targets vibration-induced horizontal stripe artifacts in SEM images.

The stripe suppression loss is formulated as:

$$\mathcal{L}_{\text{HStripeSuppress}}(\hat{y}) = \frac{1}{N_h}\sum_{i=1}^{H-1}\sum_{j=1}^{W}|\hat{y}_{i+1,j} - \hat{y}_{i,j}| \tag{2}$$

where H and W represent the image height and width, respectively, and N_h is a normalization factor. This term penalizes differences between adjacent row pixels, effectively reducing horizontal stripes. Through this distillation-driven approach, SwinIR-KD achieves significant reductions in computational overhead while maintaining or even enhancing critical performance metrics, making it highly suitable for resource-constrained applications.

The training procedure for SwinIR-KD is outlined in Algorithm 1, which details the iterative optimization process. The algorithm takes as input the vibrated image I_{vib}, the ground truth image I_{GT}, a pre-trained teacher model F_{teacher}, the student model F_{student}, and hyperparameters α and λ. In each iteration, the vibrated image is processed by both the teacher and student branches to generate $\hat{y}_T$ and $\hat{y}_S$, respectively. The reconstruction loss is computed using the Charbonnier loss, the distillation loss via mean squared error (MSE), and the stripe suppression loss based on pixel differences. These components are combined into the total loss, which is then used to update the student model parameters through backpropagation. This process ensures that SwinIR-KD efficiently learns to restore images while suppressing artifacts, leveraging the teacher's knowledge to enhance performance.

Algorithm 1 SwinIR-KD Training Procedure

Input: Vibrated Image I_{vib}, Ground Truth Image I_{GT}, Pre-trained Teacher Model F_{teacher}, Student Model F_{student}, Hyperparameters α, λ.
Output: Trained Student Model F_{student}.
Initialize: Set learning rate, optimizer, and other training parameters.
for each training iteration do
  $\hat{y}_T \leftarrow F_{\text{teacher}}(I_{\text{vib}})$
  $\hat{y}_S \leftarrow F_{\text{student}}(I_{\text{vib}})$
  $\mathcal{L}_{\text{Rec}} \leftarrow \text{CharbonnierLoss}(\hat{y}_S, I_{\text{GT}})$
  $\mathcal{L}_{\text{Distill}} \leftarrow \text{MSELoss}(\hat{y}_S, \hat{y}_T)$
  $\mathcal{L}_{\text{HStripeSuppress}} \leftarrow \frac{1}{N_h}\sum_{i=1}^{H-1}\sum_{j=1}^{W}|\hat{y}_{S_{i+1,j}} - \hat{y}_{S_{i,j}}|$
  $\mathcal{L}_{\text{Total}} \leftarrow (1-\alpha)\cdot\mathcal{L}_{\text{Rec}} + \alpha\cdot\mathcal{L}_{\text{Distill}} + \lambda\cdot\mathcal{L}_{\text{HStripeSuppress}}$
  Compute gradients of $\mathcal{L}_{\text{Total}}$ w.r.t. F_{student} parameters
  Update F_{student} parameters using optimizer
end for
return Trained F_{student}

3.3. Large image processing strategy

Processing large SEM images efficiently requires overcoming GPU memory limitations and computational complexity. To address this, we propose a hierarchical strategy that divides the image into overlapping regions and smaller patches for processing (see Fig. 3). The large image is first segmented into regions of size $R \times R$ with an overlap of O_r pixels, and each region is further divided into patches of size $P \times P$ with an overlap of O_p pixels. These patches are individually denoised using the model, and the results are seamlessly stitched back together using weight maps to ensure smooth transitions across overlapping areas.

Weight maps $W(x, y)$ are defined with linear decay at the edges as:

$$W(x) = \begin{cases} \frac{x}{O} & \text{if } 0 \le x < O, \\ 1 & \text{if } O \le x < S - O, \\ \frac{S-x}{O} & \text{if } S - O \le x < S \end{cases} \tag{3}$$

where S is the size (region or patch), and O is the overlap.

The 2D weight map is computed as $W(x, y) = W(x) \cdot W(y)$, and the final pixel value in overlapping areas is:

$$I_{\text{out}}(x, y) = \frac{\sum(I_{\text{processed}}(x, y) \cdot W(x, y))}{\sum W(x, y) + \epsilon} \tag{4}$$

where $\epsilon = 10^{-8}$ for numerical stability.

Region and patch positions are adjusted with strides $R - O_r$ and $P - O_p$, respectively, ensuring full coverage even at image boundaries.

The simplified process is outlined in Algorithm 2. This approach balances memory usage and computational efficiency while preserving denoising quality, making it practical for large SEM images in resource-constrained settings.

3.4. Evaluation metrics

For quantitative evaluation of vibration artifact removal, we employ the Peak Signal-to-Noise Ratio (PSNR)(↑) and Structural Similarity Index (SSIM)(↑) to measure the fidelity of restored images relative to their artifact-free ground truth. To holistically assess the restoration quality, we further utilize Fréchet Inception Distance (FID)(↓) to quantify the distributional divergence between processed outputs and high-quality reference images, and Learned Perceptual Image Patch Similarity (LPIPS)(↓) to gauge perceptual similarity, thereby emulating human visual judgment of artifact suppression. Finally, we benchmark computational efficiency by recording the processing time required to eliminate vibration artifacts from input images of size 128 × 128

Table 1
Performance comparison of SwinIR and SwinIR-KD on 128 × 128 patches.

Method	PSNR(↑)	SSIM(↑)	FID(↓)	LPIPS(↓)	Parameters (M)	FLOPs (G)	Time (s)
SwinIR	34.06	0.9057	84.5632	0.2570	11.556	189.681	0.1343
SwinIR-KD	32.56	0.8737	79.2449	0.2575	2.436	40.095	0.0906

Algorithm 2 Large Image Vibration Artifact Removal

Input: Image I_{in}, region size R, overlap O_r, patch size P, overlap O_p, model F_{denoise}, ϵ.
Output: Processed image I_{out}.
Initialize: $I_{\text{out}} \leftarrow \mathbf{0}$, $W_{\text{sum}} \leftarrow \mathbf{0}$.
for each region position with stride $R - O_r$ do
  Extract region R_{current} from I_{in}
  Initialize region output $R_{\text{out}} \leftarrow \mathbf{0}$, $W_{\text{sum_region}} \leftarrow \mathbf{0}$
  for each patch position with stride $P - O_p$ do
    Extract patch P_{current} from R_{current}
    Denoise $P_{\text{denoised}} \leftarrow F_{\text{denoise}}(P_{\text{current}})$
    Generate weight map W_p for P_{denoised}
    Update $R_{\text{out}} \leftarrow R_{\text{out}} + P_{\text{denoised}} \cdot W_p$
    Update $W_{\text{sum_region}} \leftarrow W_{\text{sum_region}} + W_p$
  end for
  Normalize $R_{\text{out}} \leftarrow R_{\text{out}}/(W_{\text{sum_region}} + \epsilon)$
  Generate weight map W_r for R_{out}
  Update $I_{\text{out}} \leftarrow I_{\text{out}} + R_{\text{out}} \cdot W_r$
  Update $W_{\text{sum}} \leftarrow W_{\text{sum}} + W_r$
end for
Normalize $I_{\text{out}} \leftarrow I_{\text{out}}/(W_{\text{sum}} + \epsilon)$
return I_{out}

pixels, restoring them to clean outputs of the same resolution. All quality assessments and inference timing were performed on an NVIDIA GeForce RTX 4090 D Graphics Processing Unit (GPU).

3.5. Training settings

We train the SwinIR-KD model on 128 × 128 patches with a batch size of 4. The Adam optimizer is used with an initial learning rate of 2×10^{-4} and zero weight decay. The learning rate is reduced by a factor of 0.5 at 800k, 1.2M, 1.4M, 1.5M, and 1.6M iterations. The training is guided by a composite loss function, where the Charbonnier loss ($\epsilon = 1 \times 10^{-9}$) is balanced with the distillation loss using a weight $\alpha = 0.5$, and the horizontal stripe suppress loss with a coefficient $\lambda = 0.1$.

4. Results and discussion

This section will comprehensively evaluate the proposed SwinIR-KD model and compare its performance against the baseline SwinIR model. The analysis will cover quantitative metrics, model efficiency, and qualitative visual results, with a particular focus on vibration artifact removal.

4.1. Performance comparison of models

To evaluate the proposed SwinIR-KD model, we conducted a rigorous quantitative comparison with the baseline SwinIR model. The detailed results are presented in Table 1, which systematically outlines the performance differences between the two models across four key evaluation metrics.

The data in Table 1 unequivocally demonstrate that SwinIR-KD consistently outperforms the baseline SwinIR model across all reported evaluation metrics. This enhancement in performance significantly validates the efficacy of the implemented knowledge distillation strategy. Particularly noteworthy is the observation that the student model, SwinIR-KD, not only achieves performance comparable to the teacher model (SwinIR) but surpasses it on several critical metrics. This phenomenon suggests that the knowledge distillation process transcends mere model parameter compression. More profoundly, it appears to facilitate the student model's acquisition of a more robust and generalizable feature representation. Furthermore, distillation may guide the student model towards learning a better-regularized function, thereby mitigating overfitting to specific noise patterns or redundant features within the training data that might have been captured by the larger teacher model.

In terms of model efficiency, the application of knowledge distillation has enabled SwinIR-KD to achieve a substantial reduction of approximately 79% in both parameters and computational complexity (measured in FLOPs) when compared to the original SwinIR. This optimization directly translates to an enhanced inference speed, with the processing time for a single 128 × 128 pixel image being reduced to roughly half that required by the SwinIR model.

For the qualitative assessment, Fig. 4 visually presents the restoration outcomes of both SwinIR and SwinIR-KD on two representative image samples exhibiting vibration artifacts, alongside their corresponding GT images. Visually, both models demonstrate a commendable capability in effectively eliminating these artifacts, yielding images of high clarity and perceptual quality. A common characteristic observed is that images restored by both SwinIR and SwinIR-KD tend to appear smoother when compared to their respective GT counterparts. This smoothing effect is a widely recognized phenomenon in the field of image restoration, typically associated with models that are primarily trained using PSNR as the core optimization objective.

4.2. Ablation studies

4.2.1. Impact of horizontal stripe loss weight (λ)

We conducted a comprehensive ablation study on the horizontal stripe loss weight λ to determine its optimal value for balancing reconstruction quality and stripe artifact suppression. This study was performed on the SwinIR model to justify the effectiveness of this specialized loss component. As shown in Table 2, $\lambda = 0.1$ achieves the best performance across all evaluation metrics, demonstrating superior PSNR (34.06 dB), SSIM (0.9057), and perceptual quality metrics with the lowest FID (84.5632) and LPIPS (0.2570) scores.

When $\lambda = 0$ (without horizontal stripe loss), the model shows significantly degraded performance across all metrics, particularly in perceptual quality measures where FID increases to 87.7521 and LPIPS deteriorates to 0.2794. This confirms the necessity of the horizontal stripe loss component in addressing the inherent stripe artifacts in vibration-affected images. The absence of this specialized loss leads to insufficient suppression of horizontal patterns, resulting in visually noticeable artifacts that compromise both objective and perceptual image quality.

However, excessive weight assignment ($\lambda \geq 0.2$) leads to systematic performance deterioration across all evaluation dimensions. At $\lambda = 0.2$, we observe a decline in PSNR to 33.85 dB and SSIM to 0.8989, while perceptual metrics show substantial degradation with FID rising to 96.8769 and LPIPS increasing to 0.3034. Further increasing to $\lambda = 0.4$ exacerbates this, with FID reaching 110.357 and LPIPS 0.3215. This indicates that over-emphasizing horizontal stripe suppression can harm overall image quality by forcing the model to prioritize stripe removal at the expense of fine detail preservation and natural texture reconstruction.

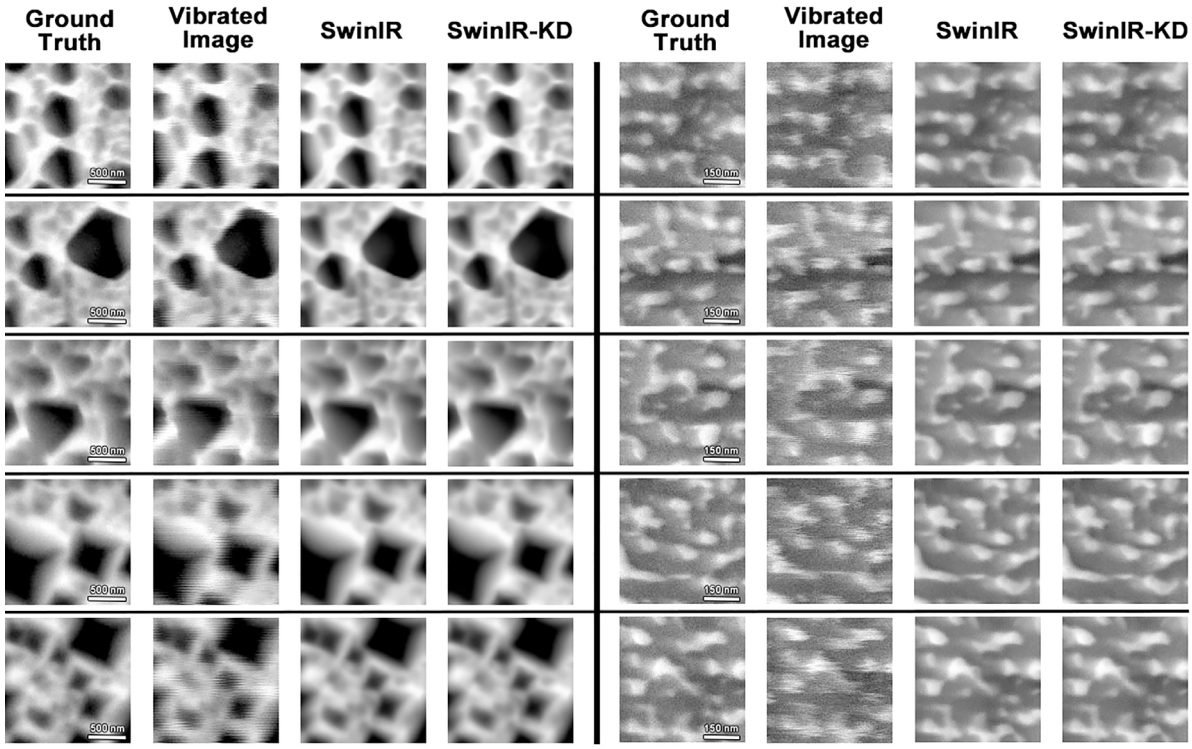


Fig. 4. Qualitative comparison of vibration artifact removal on representative SEM image patches. Results from the baseline SwinIR model and proposed SwinIR-KD model are shown alongside the Ground Truth (GT) and the input Vibrated Image for two different samples. Both models effectively remove artifacts, with SwinIR-KD achieving comparable visual quality to SwinIR.

Table 2
Ablation study on horizontal stripe loss weight λ for SwinIR.

λ	PSNR(↑)	SSIM(↑)	FID(↓)	LPIPS(↓)
0	33.76	0.8997	87.7521	0.2794
0.05	34.00	0.9052	84.6644	0.2656
0.1	**34.06**	**0.9057**	**84.5632**	**0.2570**
0.2	33.85	0.8989	96.8769	0.3034
0.4	33.72	0.8940	110.357	0.3215

$\lambda = 0$: without horizontal stripe loss; $\lambda = 0.1$: optimal setting.

The optimal $\lambda = 0.1$ represents a critical balance point where the model effectively suppresses horizontal artifacts while maintaining high-fidelity reconstruction capabilities. This finding suggests that the horizontal stripe loss should complement rather than dominate the primary reconstruction objective, ensuring that artifact suppression enhances rather than compromises the overall denoising performance.

4.2.2. Impact of channel number (C) in SwinIR-KD

To determine the optimal channel number for the student model, we conducted an ablation study on the channel dimension C of SwinIR-KD. This experiment was performed after fixing the horizontal stripe loss weight at $\lambda = 0.1$. As shown in Table 3, we evaluated three different channel configurations: $C = 64$, $C = 96$, and $C = 128$.

The results reveal an interesting phenomenon: $C = 96$ achieves the best performance across most metrics, particularly excelling in FID (79.2449) and LPIPS (0.2575) compared to both smaller ($C = 64$) and larger ($C = 128$) configurations. This suggests that there exists an optimal model capacity for knowledge distillation frameworks. While $C = 64$ results in a model with the lowest parameters (1.103M) and FLOPs (18.283G), it suffers from insufficient capacity to learn complex features, leading to degraded perceptual quality with higher FID (153.3660) and LPIPS (0.2997).

Conversely, when $C = 128$, despite having a larger capacity with 4.292M parameters, the model shows performance degradation with FID increasing to 156.6725 and LPIPS to 0.3197. This counterintuitive result can be attributed to several factors inherent in the knowledge distillation process:

- **Gradient Dilution:** Increased parameters in a larger student model may lead to diluted effective gradients during the distillation process, making it harder to converge efficiently.
- **Optimization Difficulty:** Larger models are more difficult to optimize and may struggle to find an optimal solution under the dual constraints of the reconstruction loss and the distillation loss.
- **Feature Redundancy:** Excessive channels may be prone to learning redundant or noisy features from the teacher model, rather than a more concise and robust representation.

This finding demonstrates that in knowledge distillation frameworks, there exists an optimal model capacity where $C = 96$ achieves the best trade-off between representational power and training complexity. This validates our architectural modification of reducing the channel number from the original SwinIR, as it maintains high restoration quality while achieving substantial computational efficiency.

4.2.3. Critical role of distillation weight (α) for knowledge distillation

To investigate the impact of knowledge distillation strength, we conducted an ablation study on the distillation weight α in our loss function (Eq. (1)). This parameter is crucial for balancing the knowledge transfer from the teacher model with the direct reconstruction from the ground truth. As shown in Table 4, we evaluated four different α values: 0, 0.2, 0.5, and 0.8.

The results clearly demonstrate a *Balance Hypothesis*: $\alpha = 0.5$ achieves the optimal trade-off between knowledge transfer and autonomous learning. When $\alpha = 0$ (no distillation), the model shows decent PSNR (32.61 dB) and SSIM (0.8751) performance but poor perceptual quality with FID = 154.5667 and LPIPS = 0.3124. This indicates that while the model learns to perform good pixel-level reconstruction, it lacks the high-level feature guidance provided by the teacher, which

Table 3
Ablation study on channel number C for SwinIR-KD on 128 × 128 patches.

C	PSNR(↑)	SSIM(↑)	FID(↓)	LPIPS(↓)	Parameters (M)	FLOPs (G)	Time (s)
64	32.59	0.8742	153.3660	0.2997	1.103	18.283	0.0898
96	**32.56**	**0.8737**	**79.2449**	**0.2575**	**2.436**	**40.095**	**0.0906**
128	32.59	0.8740	156.6725	0.3197	4.292	70.464	0.0939

$C = 96$ achieves optimal balance between model capacity and training complexity.

Table 4
Ablation study on distillation weight α for SwinIR-KD on 128 × 128 patches.

α	PSNR(↑)	SSIM(↑)	FID(↓)	LPIPS(↓)
0	32.61	0.8751	154.5667	0.3124
0.2	32.58	0.8742	154.7504	0.3178
0.5	**32.56**	**0.8737**	**79.2449**	**0.2575**
0.8	32.60	0.8751	156.3405	0.3159

$\alpha = 0.5$ achieves optimal balance between distillation and reconstruction losses.

Table 5
Performance on full-scale SEM images (1536 × 1024 pixels).

Method	Image	PSNR(↑)	SSIM(↑)	FID(↓)	LPIPS(↓)	Time (s)
SwinIR	#1	22.43	0.8014	11.6755	0.2653	21.40
	#2	27.85	0.9055	31.2641	0.2268	
	#3	19.98	0.5733	32.1418	0.7409	
SwinIR-KD	#1	22.52	0.7969	16.6443	0.2195	8.95
	#2	27.57	0.8864	34.5908	0.1669	
	#3	20.06	0.5620	38.7039	0.7433	

is crucial for perceptual quality. With light distillation ($\alpha = 0.2$), the teacher guidance remains insufficient, showing minimal improvement in perceptual metrics.

Remarkably, $\alpha = 0.5$ achieves dramatic improvements in perceptual quality: FID drops from 154.5667 to 79.2449 (a 48.7% reduction) and LPIPS decreases from 0.3124 to 0.2575 (a 17.6% improvement), while maintaining comparable PSNR/SSIM scores. This validates the effectiveness of balanced distillation in preserving pixel-level accuracy while significantly enhancing perceptual quality.

However, excessive distillation ($\alpha = 0.8$) leads to a performance degradation. The student model, with its limited capacity, is forced to excessively mimic high-dimensional teacher features. This can create optimization conflicts between the distillation and reconstruction losses, resulting in performance deterioration with FID increasing to 156.3405 and LPIPS to 0.3159.

Our findings confirm the Balance Hypothesis in knowledge distillation: optimal performance requires careful calibration of the distillation-reconstruction trade-off. The value of $\alpha = 0.5$ represents the sweet spot for achieving superior perceptual quality without sacrificing reconstruction accuracy, thereby justifying our design choice.

4.3. Large image vibration removal results

To rigorously assess the practical applicability of the models, particularly in real-world scenarios that frequently involve high-resolution imagery, we conducted comprehensive evaluations of both SwinIR and SwinIR-KD on three complete, large-scale test images. Recognizing the inherent GPU memory constraints typically encountered when processing such large inputs with deep learning architectures, we employed the previously delineated "Large Image Processing Strategy". This strategy, commonly referred to as a crop-and-stitch or patch-based approach, involves subdividing the large image into smaller, manageable patches. Each patch is then processed independently by the model, and the resulting enhanced patches are subsequently seamlessly reassembled to reconstruct the final high-resolution output image.

Table 5 provides a detailed quantitative comparison of SwinIR and SwinIR-KD when applied to three full-scale test images, each with dimensions of 1536 × 1024 pixels, utilizing the aforementioned crop-and-stitch methodology. The table reports the performance across the same four key evaluation metrics used previously, alongside a critical measure of computational efficiency: the total processing time required for a single large-scale image.

The empirical data presented in Table 5 reveal that both SwinIR and SwinIR-KD achieve comparable quantitative performance levels on these large-scale images processed via the crop-and-stitch technique. However, a stark and significant divergence emerges when examining their computational efficiency. Specifically, the time required to process a single 1536 × 1024 pixel image was dramatically reduced from 21.40 seconds for the original SwinIR model to a mere 8.95 seconds for the distilled SwinIR-KD model. This substantial improvement represents an acceleration factor of approximately 2.4 times.

Although the quantitative image quality metrics for the large, stitched images are similar between the two models, the pronounced reduction in processing time endows SwinIR-KD with considerably greater practical viability for real-world deployment. For applications necessitating the processing of extensive volumes of high-resolution imagery, or those demanding near real-time performance, a 2.4-fold increase in processing speed constitutes a critical operational advantage. This underscores the practical benefits of knowledge distillation beyond mere metric improvement, highlighting its role in bridging the gap between research models and deployable solutions.

Fig. 5 offers a visual testament to the efficacy of SwinIR-KD in processing full-scale test imagery, with particular emphasis on magnified regions to scrutinize fine details. The visual results presented in Fig. 5 indicate that SwinIR-KD effectively handles large-dimension images. Close inspection of the magnified insets reveals that vibration-induced stripe artifacts have been successfully and thoroughly eliminated. This observation serves as compelling evidence of the model's robust restoration capabilities, even when operating under a patch-based processing paradigm. The ability of SwinIR-KD to maintain high visual quality and artifact removal efficacy when using a crop-and-stitch approach suggests that the distilled features are robust and generalize well, preventing the introduction of significant blockiness or inconsistencies at patch boundaries.

Considering its demonstrated proficiency in artifact mitigation coupled with its significant gains in computational efficiency, SwinIR-KD emerges as a highly promising and practical solution for demanding image restoration tasks.

4.4. Application interface showcase

To enhance the practical applicability and streamline the evaluation of the proposed SwinIR-KD model for vibration artifact removal, as well as to offer an efficient interactive platform for subsequent research in this domain, we have developed a user-friendly Graphical User Interface (GUI). The core functionalities and operational workflow of this interface are depicted in Fig. 6. This software tool enables users to conveniently import images designated for processing. It further offers the flexibility to select between the original SwinIR model and the lightweight SwinIR-KD variant, thereby accommodating diverse requirements related to processing speed or specific performance analysis objectives. After the model executes the vibration artifact removal process, users can intuitively select any Region of Interest (ROI) within the processed image via the interface. The software then promptly presents a side-by-side visual comparison of the selected ROI before and after processing, thereby clearly demonstrating the restoration efficacy of the model. Moreover, the fully processed high-resolution

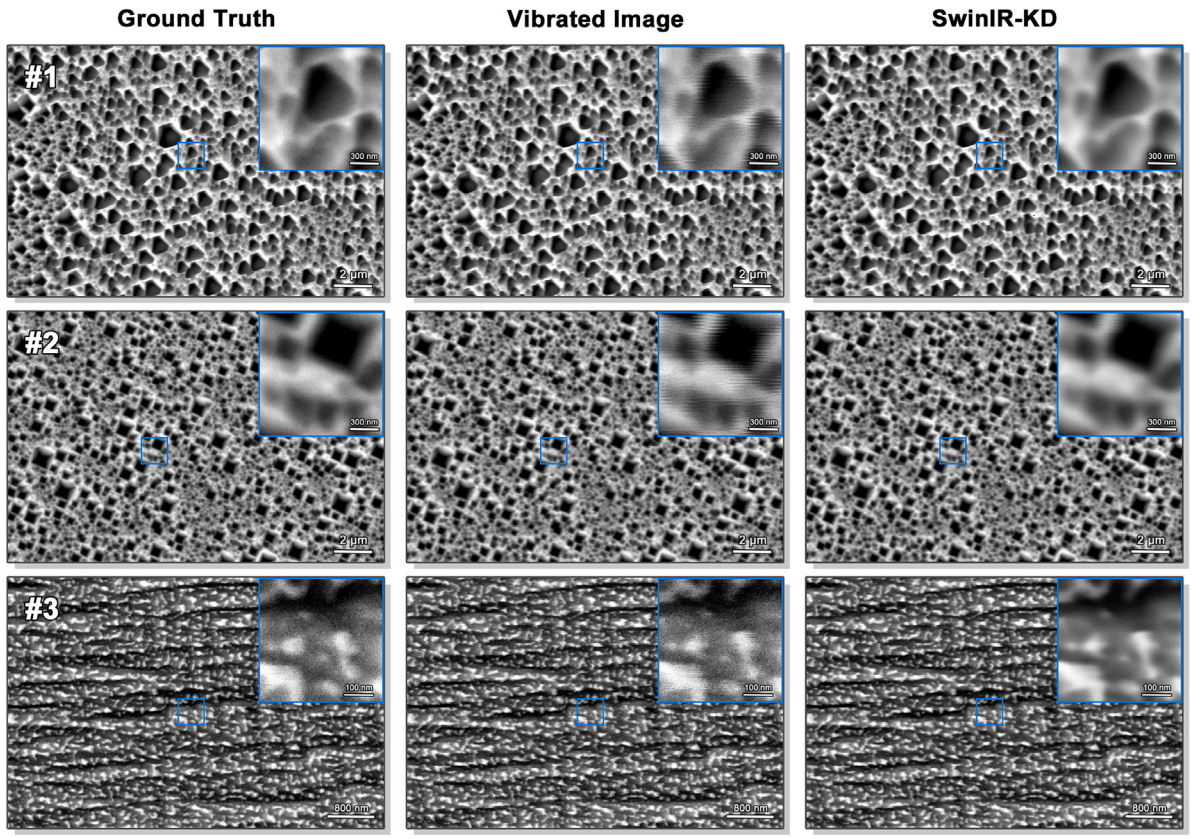


Fig. 5. Visual results of SwinIR-KD on a full-scale SEM test image processed using the large image processing strategy. Magnified insets highlight the effective removal of vibration-induced stripe artifacts, demonstrating the model's robust restoration capabilities on large images.

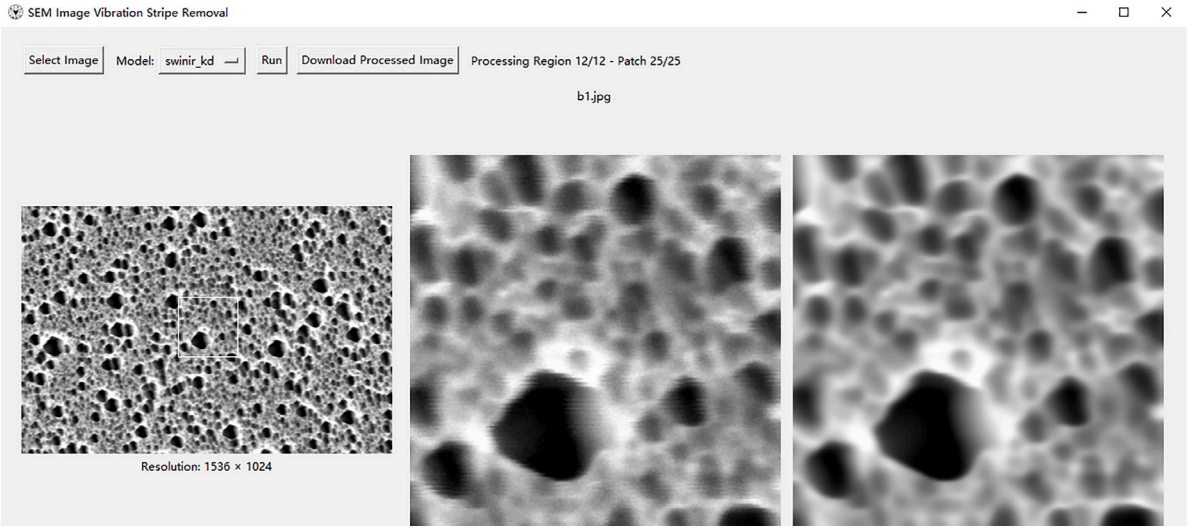


Fig. 6. Screenshot of the developed graphical user interface (GUI) for SEM image vibration stripe removal. The interface allows users to select an input image, choose between the SwinIR and SwinIR-KD models, initiate processing, visualize before/after comparisons of selected regions, and download the processed high-resolution image. The example shows the processing status for an IN718 image with a resolution of 1536 × 1024.

images can be conveniently downloaded and saved for further use or archival purposes. We contend that this interactive software, which synergistically integrates advanced algorithms with an intuitive operational framework, not only effectively showcases the real-world performance of SwinIR-KD but also holds significant potential to serve as an invaluable auxiliary tool in the future research and development of image restoration and enhancement algorithms, markedly enhancing research productivity.

5. Conclusion

This study detailed the development and evaluation of SwinIR-KD, a lightweight framework for removing horizontal stripe artifacts in SEM images. By integrating the Swin Transformer architecture with a tailored knowledge distillation process and a novel Horizontal Stripe Suppress Loss, SwinIR-KD achieved a significant reduction in model parameters and computational load — approximately 79% less than the baseline SwinIR — while delivering comparable or improved image restoration quality based on key evaluation metrics. The research also presented an effective strategy for processing large-scale SEM images and a user-friendly GUI to facilitate practical use. The findings indicate that SwinIR-KD offers an efficient and effective approach for enhancing SEM image quality by mitigating common vibration-induced artifacts.

Future work could focus on enhancing the model's generalization capabilities across a broader range of SEM imaging conditions and sample types. Investigating the model's adaptability to other forms of image degradation in electron microscopy also presents a valuable research direction. Additionally, we plan to integrate our denoising framework into downstream analysis pipelines to quantitatively evaluate its impact on tasks such as microstructure segmentation and classification, thereby providing a more comprehensive assessment of its practical utility. Exploring advanced distillation techniques or alternative loss functions might further refine restoration outcomes. Continued development in this area, including potential exploration of unsupervised learning methods, could further advance the practical application of deep learning for robust and efficient SEM image restoration. We also plan to conduct a comprehensive comparative analysis against a wider range of state-of-the-art image restoration algorithms to further contextualize SwinIR-KD's performance within the broader field. Overall, SwinIR-KD provides a notable contribution towards addressing the challenge of horizontal stripe artifacts in SEM imaging, offering a practical tool for the materials science community.

CRediT authorship contribution statement

Xuecheng Zhang: Writing – original draft, Visualization, Software, Methodology, Data curation. **Bin Zhang:** Writing – original draft, Software, Methodology. **Wenchao Meng:** Writing – review & editing, Supervision, Methodology. **Yuefei Zhang:** Writing – review & editing, Validation, Supervision, Methodology. **Xianjue Ye:** Writing – review & editing, Supervision. **Ze Zhang:** Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by National Science and Technology Major Project of China [grant numbers J2019-III-0008-0051], and National Natural Science Foundation of China [grant number 52401025].

Data availability

Data will be made available on request.

References

Appalanaidu, Botcha, Rathore, Ranjeet Singh, Raj, Manyala Anjani, Dvivedi, Akshay, 2023. Enhancement of electrochemical discharge machining process with pressurized tool feeding mechanism. Mater. Today: Proc..

Chen, Wenshu, Peng, Liyuan, Huang, Yujie, Jing, Minge, Zeng, Xiaoyang, 2021. Knowledge distillation for u-net based image denoising. In: 2021 IEEE 14th International Conference on ASIC. ASICON, IEEE, pp. 1–4.

Chowdhury, Indraneel R, Rao, Pavan S, O'Dowd, Noel P, Comer, Anthony J, 2025. Hygrothermal ageing effects on failure behaviour of fibre-reinforced polymer composite materials under in-situ SEM testing. Compos. Part B: Eng. 294, 112148.

Chu, Jie, Su, Tong, Liu, Pei, Wu, Yunpeng, Zhang, Le, Shi, Zenglin, Wang, Meng, 2024. Learning dual transformers for all-in-one image restoration from a frequency perspective. arXiv preprint arXiv:2407.01636.

De Haan, Kevin, Ballard, Zachary S, Rivenson, Yair, Wu, Yichen, Ozcan, Aydogan, 2019. Resolution enhancement in scanning electron microscopy using deep learning. Sci. Rep. 9 (1), 12050.

Di Gioacchino, F., Quinta da Fonseca, J., 2013. Plastic strain mapping with sub-micron resolution using digital image correlation. Exp. Mech. 53, 743–754.

Gou, Jianping, Yu, Baosheng, Maybank, Stephen J, Tao, Dacheng, 2021. Knowledge distillation: A survey. Int. J. Comput. Vis. 129 (6), 1789–1819.

Guo, Guanghao, Zhang, Wenqiang, Zhang, Bin, Xu, Jiachen, Chen, Shuang, Ye, Xianjue, Zhang, Yuefei, Zhang, Ze, 2025. Effect of precipitate phase on the plastic deformation behavior of alloy 718: In-situ tensile experiment and crystal plasticity simulation. Int. J. Plast. 187, 104286.

Hinton, Geoffrey, Vinyals, Oriol, Dean, Jeff, 2015. Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531.

Jain, Anubhav, 2024. Machine learning in materials research: developments over the last decade and challenges for the future. Curr. Opin. Solid State Mater. Sci. 33, 101189.

Jiang, Junjun, Zuo, Zengyuan, Wu, Gang, Jiang, Kui, Liu, Xianming, 2024. A survey on all-in-one image restoration: Taxonomy, evaluation and future trends. arXiv preprint arXiv:2410.15067.

Jung, Kwang Oh, Kim, Seung Jae, Kim, Dong Hwan, 2012. An approach to reducing the distortion caused by vibration in scanning electron microscope images. Nucl. Instrum. Methods Phys. Res. Sect. A: Accel. Spectrometers Detect. Assoc. Equip. 676, 5–17.

Li, Zhuo, Li, Hengyi, Meng, Lin, 2023. Model compression for deep neural networks: A survey. Computers 12 (3), 60.

Liang, Jingyun, Cao, Jiezhang, Sun, Guolei, Zhang, Kai, Van Gool, Luc, Timofte, Radu, 2021. Swinir: Image restoration using swin transformer. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1833–1844.

Liu, Ze, Lin, Yutong, Cao, Yue, Hu, Han, Wei, Yixuan, Zhang, Zheng, Lin, Stephen, Guo, Baining, 2021. Swin transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 10012–10022.

Mao, Jing, Sun, Lianming, Chen, Jie, Yu, Shunyuan, 2025. Overview of research on digital image denoising methods. Sensors 25 (8), 2615.

Oho, Eisaku, Yamazaki, Sadao, Suzuki, Kazuhiko, 2019. PM-17 noise removal for SEM images using a convolutional neural network. Microscopy 68, i43–i43.

Panwar, Aditya Singh, Singh, Armender, Sehgal, Shankar, 2020. Material characterization techniques in engineering applications: A review. Mater. Today: Proc. 28, 1932–1937.

Płuska, Mariusz, Czerwinski, Andrzej, Ratajczak, Jacek, Kątcki, Jerzy, Oskwarek, Łukasz, Rak, Remigiusz, 2009. Separation of image-distortion sources and magnetic-field measurement in scanning electron microscope (SEM). Micron 40 (1), 46–50.

Rottmayer, Niklas, Redenbach, Claudia, Fahrbach, Florian, 2024. Effective stripe artefact removal by a variational method: application to light-sheet microscopy, FIB-sem and remote sensing images. arXiv preprint arXiv:2401.14220.

Shin, Yun-Ho, Moon, Seok-Jun, Kim, Yong-Ju, Oh, Ki-Yong, 2020. Vibration control of scanning electron microscopes with experimental approaches for performance enhancement. Sensors 20 (8), 2277.

Suga, Mitsuo, Asahina, Shunsuke, Sakuda, Yusuke, Kazumori, Hiroyoshi, Nishiyama, Hidetoshi, Nokuo, Takeshi, Alfredsson, Viveka, Kjellman, Tomas, Stevens, Sam M, Cho, Hae Sung, et al., 2014. Recent progress in scanning electron microscopy for the characterization of fine structural details of nano materials. Prog. Solid State Chem. 42 (1–2), 1–21.

Tawfik, Miral S, Adishesha, Amogh Subbakrishna, Hsi, Yuhan, Purswani, Prakash, Johns, Russell T, Shokouhi, Parisa, Huang, Xiaolei, Karpyn, Zuleima T, 2022. Comparative study of traditional and deep-learning denoising approaches for image-based petrophysical characterization of porous media. Front. Water 3, 800369.

Wang, Ni, Guan, Hao, Wang, Jin, Zhou, Jianli, Gao, Wenjie, Jiang, Wenxiang, Zhang, Yuefei, Zhang, Ze, 2022. A deep learning-based approach for segmentation and identification of δ phase for inconel 718 alloy with different compression deformation. Mater. Today Commun. 33, 104954.

Wang, Jiping, Tang, Yufei, Wu, Zhongyi, Du, Qiang, Yao, Libing, Yang, Xiaodong, Li, Ming, Zheng, Jian, 2023. A self-supervised guided knowledge distillation framework for unpaired low-dose CT image denoising. Comput. Med. Imaging Graph. 107, 102237.

Weigert, Martin, Schmidt, Uwe, Boothe, Tobias, Müller, Andreas, Dibrov, Alexandr, Jain, Akanksha, Wilhelm, Benjamin, Schmidt, Deborah, Broaddus, Coleman, Culley, Siân, et al., 2018. Content-aware image restoration: pushing the limits of fluorescence microscopy. Nature Methods 15 (12), 1090–1097.

Yi, Feng, Voyles, P.M., 2011. Effect of sample thickness, energy filtering, and probe coherence on fluctuation electron microscopy experiments. Ultramicroscopy 111 (8), 1375–1380.

Zhang, Wenqiang, Chen, Shuang, Guo, Guanghao, Liu, Xuan, Zhang, Bin, Ye, Xianjue, Wang, Jin, Zhang, Yuefei, Zhao, Xinbao, Zhang, Ze, 2025a. Study on plastic deformation mechanism of a fourth-generation nickel-based single crystal superalloy: In-situ SEM characterization and dislocation density-based model. Mater. Sci. Eng.: A 930, 148132.

Zhang, Xuecheng, Guo, Guanghao, Li, Zixin, Meng, Wenchao, Zhang, Yuefei, Ye, Qing, Wang, Jin, He, Shibo, Zhao, Xinbao, Chen, Jiming, et al., 2024. Superalloys fracture process inference based on overlap analysis of 3D models. Commun. Eng. 3 (1), 108.

Zhang, Shihua, Li, Zizhuo, Zhang, Kaining, Lu, Yifan, Deng, Yuxin, Tang, Linfeng, Jiang, Xingyu, Ma, Jiayi, 2025b. Deep learning reforms image matching: A survey and outlook. arXiv preprint arXiv:2506.04619.

Zhang, Yongheng, Yan, Danfeng, 2025. Knowledge distillation for image restoration: Simultaneous learning from degraded and clean images. In: ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing. ICASSP, IEEE, pp. 1–5.

Zhou, Shihao, Li, Dayu, Pan, Jinshan, Zhou, Juncheng, Shi, Jinglei, Yang, Jufeng, 2025. Devil is in the uniformity: Exploring diverse learners within transformer for image restoration. arXiv preprint arXiv:2503.20174.

communications engineering

Article

https://doi.org/10.1038/s44172-024-00257-6

Superalloys fracture process inference based on overlap analysis of 3D models

Check for updates

Xuecheng Zhang[1,2,4], Guanghao Guo[3,4], Zixin Li[1,2], Wenchao Meng[1] ✉, Yuefei Zhang[3] ✉, Qing Ye[1], Jin Wang[3], Shibo He[1], Xinbao Zhao[3], Jiming Chen[1] & Ze Zhang[3]

Superalloy materials exhibit susceptibility to fracture failures stemming from the influence of thermomechanical factors. To comprehensively understand the fracture mechanisms, material properties, root causes of failure, and the subsequent optimization of alloys, a detailed analysis of the internal fracture process and the morphological traits of the fracture surface is imperative. Traditional analysis of fracture surfaces solely relies on 2D images, thus lacking crucial 3D information. Although in situ experiments can capture the fracture process, their effectiveness is confined to the specimen's surface, precluding insight into internal changes. Here we introduce an integrated framework encompassing the process of 3D reconstruction of fracture surfaces, aiming to enhance the visual information obtained with micron-level accuracy, visual intuitiveness and sense of depth. Additionally, this framework also facilitates the scrutiny and inference of internal fracture processes. These results demonstrate that under specific service conditions, material deformation fracture probably stems from a combination of surface cracking and internal cracking rather than exclusively one or the other. Overall, our description and analysis of internally initiated cracking due to defects within the specimens can be beneficial in guiding future alloy design and optimization efforts.

Superalloys have gained notable prominence in the aerospace and energy fields[1,2] due to their advantageous properties such as high-temperature strength, oxidation and corrosion resistance, fatigue properties, and notable fracture toughness[3,4]. Analysis of fracture processes in superalloy specimens is of critical importance[5–7] in the field of superalloy research. When an alloy specimen fractures, it results in two matching fracture surfaces, and the study of their appearance[8] and location holds great value. The fracture phenomenon typically occurs in regions of the metal with the least strength, and the analysis of the fracture process is useful[9,10] for gaining insights into mechanical and thermal properties and corrosion resistance of superalloys in extreme environments[11,12], such as high temperature and high pressure[9]. Deep insights can be obtained through detailed analysis of the fracture morphology, structure, and composition of superalloys. Researchers use this as a basis to guide the design of superalloy materials to enhance their strength and plasticity, further applying them to manufacturing and applications[13].

Typical fracture analysis methods for superalloys include optical microscopy (OM)[14,15], scanning electron microscopy (SEM)[16,17], x-ray diffraction (XRD)[18,19], energy dispersive spectroscopy (EDS)[20,21], et al. Appropriate methods and equipment selection depend on the actual needs and research objectives. Among all these methods, SEM has become an essential instrument for analyzing fracture surface morphology and microstructure due to its high resolution, considerable depth of field, and good stereoscopic effect[22]. SEM can directly observe the very rough surface of the specimen as well as the raw fracture of the material. The imaging principle of SEM is that the electron beam excites various physical signals on the specimen surface, such as secondary electrons(SE) and backscattered electrons(BSE), and these signals are acquired through the corresponding detectors to form an image[23]. SEM images are, however, only two-dimensional (2D) intensity maps of the scattered electrons, and lack inherent three-dimensional (3D) information[24]. This may lead to misinterpretations about the surface morphology of the specimen, as the 2D images do not accurately represent the true 3D nature of the specimen[25]. For instance, in metallic materials, 2D images of grain and grain boundary morphologies are commonly depicted as polygons, while their 3D representations exhibit complex and irregular polyhedral geometries. To address this limitation, a suitable method for supplementing the 3D information of the specimen fracture surface is required[13,26].

[1]State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering, Zhejiang University, Hangzhou, 310027, China. [2]Polytechnic Institute, Zhejiang University, Hangzhou, 310015, China. [3]Institute of Superalloys Science and Technology, School of Materials Science and Engineering, Zhejiang University, Hangzhou, 310027, China. [4]These authors contributed equally: Xuecheng Zhang, Guanghao Guo. ✉e-mail: wmengzju@zju.edu.cn; yfzhang76@zju.edu.cn

Reconstructing 3D models from 2D SEM images has emerged as a prominent research area in microscopy. This approach improves the visual representation of the analyzed micro-objects and increases the accessibility of measurable data. The acquisition of a 3D model can be broadly categorized into two types, namely, destructive, and non-destructive, depending on the extent of potential damage caused to the specimen.

Destructive techniques such as serial block-face imaging (SBF)[27], metallographic serial polishing (MSP)[28] and focused ion beam (FIB)[29] capture layer-by-layer images of the specimen and thus such a process is irreversible. These techniques share the standard approach of capturing layer-by-layer images of the specimens, yet the process is irreversible. Careful selection of appropriate slicing methods based on the specimen and the research criteria is crucial, ensuring minimal damage while managing cost, time, reconstruction scope, and accuracy considerations. Furthermore, these techniques are better suited for reconstructing within cuboid regions, but they are not ideal for objects such as fractures with substantial surface profiles.

Standard non-destructive testing techniques in structural analysis encompass X-ray computed tomography (CT) and the utilization of 3D computer vision algorithms to analyze and process 2D images. CT technology allows for acquiring a sequence of cross-sectional images of the specimen through non-invasive tomographic scans, which are then amalgamated to form a 3D structural model for internal assessment[30]. Nevertheless, CT technology is constricted by radiation exposure, high expenses, and limited resolution. In contrast to in situ SEM observations, CT outcomes lack morphological details, rendering them more applicable for quantitative evaluations of structural volume but less convenient for multi-scale observations. Techniques utilizing 3D computer vision algorithms to analyze and process 2D images[31] offer the notable advantage of preserving the integrity of the specimens and facilitating subsequent analysis. Furthermore, it enables a broader reconstruction scope and faster speed[13,32–34].

Among these methods, 3DSEM, based on the photogrammetric reconstruction of SEM images, is widely adopted due to its low cost and ease of use[35]. The technique involves capturing images from multiple perspectives by moving the detector around the specimen or rotating and tilting the specimen within the field of view[36,37]. This capability is particularly suitable for reconstructing the surfaces of specimens with marked variations, such as fractured specimens. Identifying and matching features in these images enables the calculation of a point cloud, which is further used to generate a surface model. Post-processing techniques, including noise reduction, smoothing, and repair, are applied to the surface model to improve the accuracy of the 3D model[38,39]. However, this method has certain limitations. For instance, the model's completeness directly relies on the availability of comprehensive feature information on the fracture surface and the specimen surface in the original images. Therefore, enhancing the data acquisition method is necessary to obtain additional height feature data for complex features. In addition, unlike traditional photogrammetric techniques, the electronic microscope images on which this method is based lack information about the camera's internal and external parameters during the imaging process. Consequently, the overall dimensions of the resulting 3D model may need to be more accurately scaled up or down in proportion.

In this work, we introduce an innovative framework to infer the superalloy fracture process, utilizing 3D model overlap analysis to examine the deformation failure progression in alloy specimens. Notably, our 3D fracture morphology models are reconstructed exclusively from SEM images, thus effectively controlling expenses, and simplifying the operational procedure. These 3D fracture morphology models provide a comprehensive representation of the morphology and intricate details of the specimens in 3D space, overcome difficulties related to cross-section matching and boundary information observation, and include precise dimensional information. Different fracture regions exhibit variational extension during tensile deformation due to differing stress durations. Overlap analysis of the individual 3D fracture morphology models is carried out to elucidate the comprehensive internal fracture mechanism of the specimens. Our results reveal a characteristic material deformation pattern, highlighting the interplay between surface cracking and internal fissures, challenging the notion of these phenomena as isolated occurrences. These conclusions have been subsequently verified and detailed through in situ experiments, showcasing practical and innovative advantages in real-world applications.

Results

Framework of SFPI

We develop a superalloys fracture process inference (SFPI) framework consisting of multiple steps (Fig. 1). The framework aims to reconstruct 3D models of the fracture surface of experimental specimens after tensile testing and provide a comprehensive observation and inference of the fracture process. We conducted in situ tensile experiments on two IN718 specimens at room temperature (RT) and 650 °C high temperature (HT). After conducting the in situ experiments, we proceeded with the 3D reconstruction and scale calibration of the fracture surfaces, which yielded 3D models named Fracture A and Fracture B after the HT in situ tensile testing, as well as 3D models named Fracture C and Fracture D after the RT in situ tensile testing.

After obtaining 3D models featuring precise dimensional discontinuities, the next step involves aligning these models to achieve a multi-surface stereoscopic observation, enriching the visual data and ensuring accurate scale information. To address the challenges posed by varying extension rates in different regions during the stretching deformation process, a periodic overlap and separation strategy is employed for the corresponding 3D models. The periodic overlapping and separation of these models reveal internal regions of the specimen, offering insights into the chronological sequence of events during the fracture process. In Table 1, we provide a comparative analysis of our proposed method with current standard materials characterization and analysis techniques. As a result, our innovative research involves previously inaccessible details about the specimen, including temporal evolution scales, a feat unattainable through other existing methodologies.

Multi-surface stereoscopic observation

For traditional material fracture analysis, matching and seamlessly integrating fracture surface and edge morphology is often necessary to comprehensively examine fracture processes and mechanisms. However, this practice gives rise to two distinct challenges. The first challenge is that the researcher manually matches the SEM images of the fracture surface and the specimen surface to accurately compare and integrate the morphological information. This task is laborious and subjective, leading to biased evaluation results. The second challenge is that analyzing the microstructure at the interface between the specimen and fracture surfaces requires the researcher to rely on their experience and knowledge. Furthermore, the microstructure information, vital for the fracture analysis, must be reconstructed from 2D images, leading to inherently subjective judgments and prone to substantial errors, ultimately lacking in objective credibility. These challenges primarily stem from the inherent limitations of 2D images, which represent planar projections of 3D materials, thereby necessitating a richer spatial dataset for a more comprehensive analysis.

The fracture models offer a more intuitive depiction of the specimen's overall morphology and specific details in 3D space (Fig. 2). Each fracture model comprises a cross-section and four side surfaces. We can observe the information from ten surfaces by aligning the corresponding fracture models. This integration of multidimensional material features provides a more comprehensive and precise visual depiction, particularly emphasizing the morphological attributes at the interface and the elevation within the plane. Fracture A and Fracture C in Fig. 2 demonstrate that our model, as it incorporates both surface and fracture morphology, enables researchers to analyze these characteristics without tedious comparisons efficiently, ultimately enhancing the accuracy of the analysis process. The model exhibits distinct microscopic organizational structures at the interface junction, as denoted by the white arrows, which are not observable in conventional 2D SEM images.

By leveraging these inherent characteristics, it is possible to provide a more comprehensive and accurate description of the typical positions of deformation and the mechanisms of fracture that occur during the process of material fracture. For instance, in the case of the IN718 HT tensile test

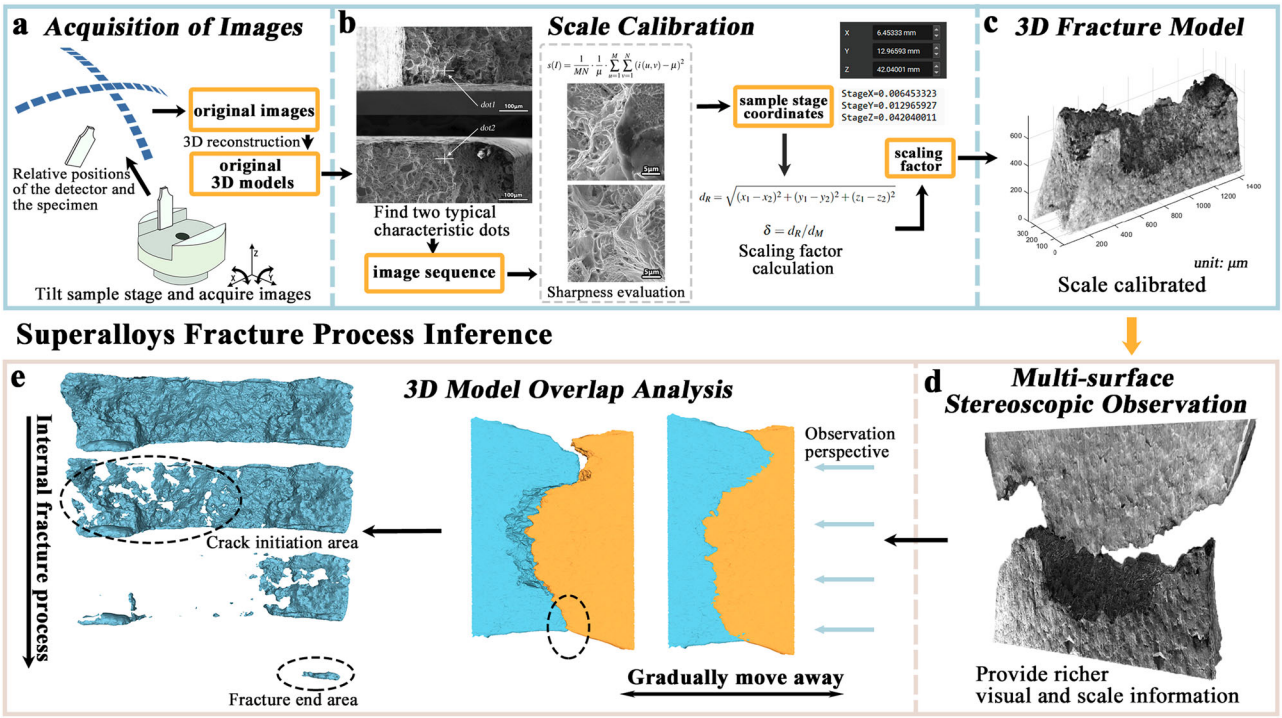


Fig. 1 | Overview of the proposed superalloys fracture process inference framework. a Improved image acquisition process. Images of the fracture specimens from different viewpoints were obtained by tilting the sample stage, and 3D reconstruction was performed to obtain the original 3D models. **b** Calculate the scaling factor to complete the 3D model scale calibration through image sharpness evaluation. **c** Calibrated 3D fracture morphology model in 3D coordinate system. **d** Multi-surface stereoscopic observation. Simultaneous observation of multiple surfaces through a pair of 3D fracture morphology models to solve problems such as difficulty in matching boundary features. **e** The internal fracture process of the sample was obtained by inferring from the 3D model overlap analysis. The figure shows the area where the internal cracking starts and the area where the fracture ends.

Table 1 | Proposed method compared with other methods of material characterization and analysis

Method	Proposed	In situ SEM observation[2,3,45]	In situ X-ray diffraction[19,53]	Layer-by-layer slice[27–29]
Non-destructive	✓	✓	✓	
3D feature	✓		✓	✓
Scale information	✓	✓	✓	✓
Time dimension information	✓	✓	✓	
Internal observation	✓		✓	✓
Morphology information	✓	✓		✓
Multi-scale observation	✓	✓		

specimens, the surface patterns of Fracture A and Fracture B consist of interlocking zig-zag shapes. The cross-section analysis reveals a higher presence of intact grains, indicating that the primary fracture mechanism is intergranular fracture (Fig. 2e). Conversely, the surfaces of Fracture C and Fracture D of the IN718 RT tensile test specimens exhibit a smoother appearance, suggesting that the dominant fracture mechanism in these cases is transgranular fracture (Fig. 2f). The proposed method offers an improved spatial representation of material properties by providing 3D information. This renders it a valuable research tool for the analysis of material fractures.

Reconstruction performance analysis of 3D models

The 3D fracture morphology models allow for precise visual observation of the cross-sectional profiles of fractures and the morphology and roughness of the four side surfaces. Furthermore, it facilitates the precise measurement of grain size, ligamentous fossa depth, and other critical parameters. In order to validate the accuracy of the 3D fracture reconstruction model, we employed a hardness tester to create two indentations spaced 500 μm apart along the horizontal direction on one side surface of Fracture A prior to reconstruction. After reconstruction, we measured the distance between the centers of the two indentations using the measurement tool in Meshlab software in the reconstructed 3D model and compared it with the standard value, obtaining an error of less than 1% (Fig. 2a).

The above method can verify the model's accuracy. In addition, two specific cases have been highlighted to better demonstrate the effectiveness of the three-dimensional cross-sectional model in achieving precise measurement and comprehensive observation. The first figure in Fig. 2g shows the distance between two diagonals of grain in the original SEM image of Fracture A, measured using the distance measurement function in Tescan Essence S9200X Demo software, which is 29.16 μm. In the second figure in Fig. 2g, the distance between two diagonals of the exact grain in the 3D model within the same perspective was measured using the measurement tool in the Meshlab software(v2022.02), which is 30.1396 μm, slightly larger

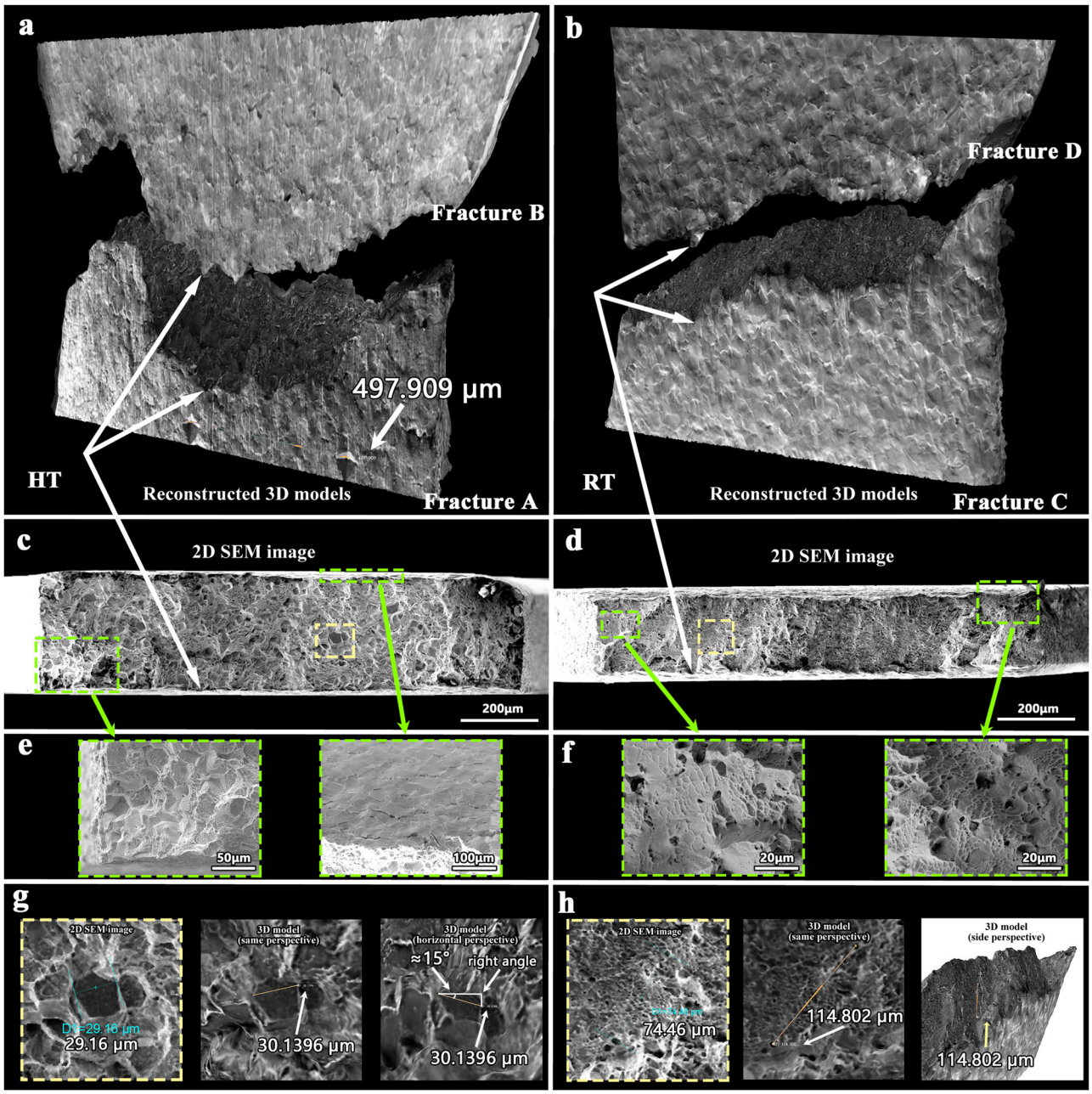


Fig. 2 | Comparison of 2D images and 3D stereoscopic observation and verification of scale calibration results. a, **b** 3D fracture surface models of IN718 after 650 °C high temperature (HT) and room temperature (RT) tensile experiments with an indentation spacing in the model. **c**, **d** Corresponding 2D Scanning Electron Microscope (SEM) images of the fracture surface. **e** Intergranular fracture at 650 °C. **f** Transgranular fracture at RT **g** The distance between two diagonals of grain in an SEM image and a 3D fracture morphology model. **h** The distance between two void defects in an SEM image and a 3D fracture morphology model.

than the value in measured in SEM image. The difference in the number of decimal places in the two sets of measurement results is due to the use of different measurement software. By changing the viewing angle of the 3D model, it can be determined that the reason for this phenomenon is the surface of the grain is not horizontal but has a certain inclination angle. The distance measured in the 2D SEM image is the projection of the actual distance in the horizontal direction. The actual distance between the two characteristic points should be calculated by dividing the distance measured in the 2D SEM image by the cosine of the angle θ formed by the line connecting the two points and the horizontal plane. For the selected grain, the angle θ between the two diagonal lines and the horizontal direction is approximately 15°. The measurement error induced by this angle is not apparent in SEM images but is relatively more precise in the 3D model. Figure 2h provides a more definitive example employing the same methodology. In the SEM image of Fracture C, the measured distance between the two void defects is significantly less than the distance measured from the 3D model. This discrepancy is attributed to the substantial angle between the fracture surface of the sample and the viewing perspective, further demonstrating the superiority of 3D stereoscopic observation.

The model effectively reconstructs the microstructure, reflecting the actual physical information of the material on the fracture surface. The currently obtained 3D model accurately represents the spatial and physical scales, offering precise references for further research. In order to obtain more detailed information, the resolution and clarity of the original SEM images can be enhanced, although it will also lead to increased computation time.

Inference of the fracture process

Many experimental studies on IN718 tensile properties have revealed that the fracture process is usually in transgranular mode at RT due to the stable crystal structure and uniform strain distribution[40,41]. At HT, a more significant number of grain boundary cracks appear as a result of reduced interfacial strength. In addition, as the experimental temperature rises, more slip and other localized deformations will occur within the specimens, leading to uneven deformation in the local regions of polycrystalline materials. However, this analysis is based solely on inferred material fracture morphology analysis, lacking direct visualization and objectivity.

In this work, we have reconstructed and calibrated fracture models for IN718 at both RT and 650 °C. It is important to note that the two fracture surfaces were obtained from in situ tensile experiments, which allows for observing surface morphology evolution during the tensile process. During the tensile deformation process, local stress concentrations arise in the material due to the incomplete uniformity of grain orientation and size and the presence of second-phase particles[42,43]. This results in the preferential cracking of certain grains characterized by reduced elongation. The non-cracked regions experience axial load and further elongate during subsequent stretching. As a result, the fracture surface exhibits varying elongation rates across its local features, rendering complete macroscopic alignment unachievable. Ultimately, the portion that fractures finally exhibits the highest elongation rate, resulting in its initial contact during the alignment process, while the initially cracked parts remain the farthest apart. We inferred the tensile deformation process of the specimens by constructing 3D fracture morphology models and considering the perspective of material fracture.

Our inference method, combined with the model morphology from a fractured perspective, allows for a detailed analysis of the dynamic process of IN718 tensile deformation at both RT and HT. Figure 3 illustrates the inference results of the two sets of experiments, with each image showing a projection of the overlapping portion of the fracture models at different tensile displacements. This overcomes the limitations of traditional fracture analysis and facilitates a transition from traditional inferential analysis to a

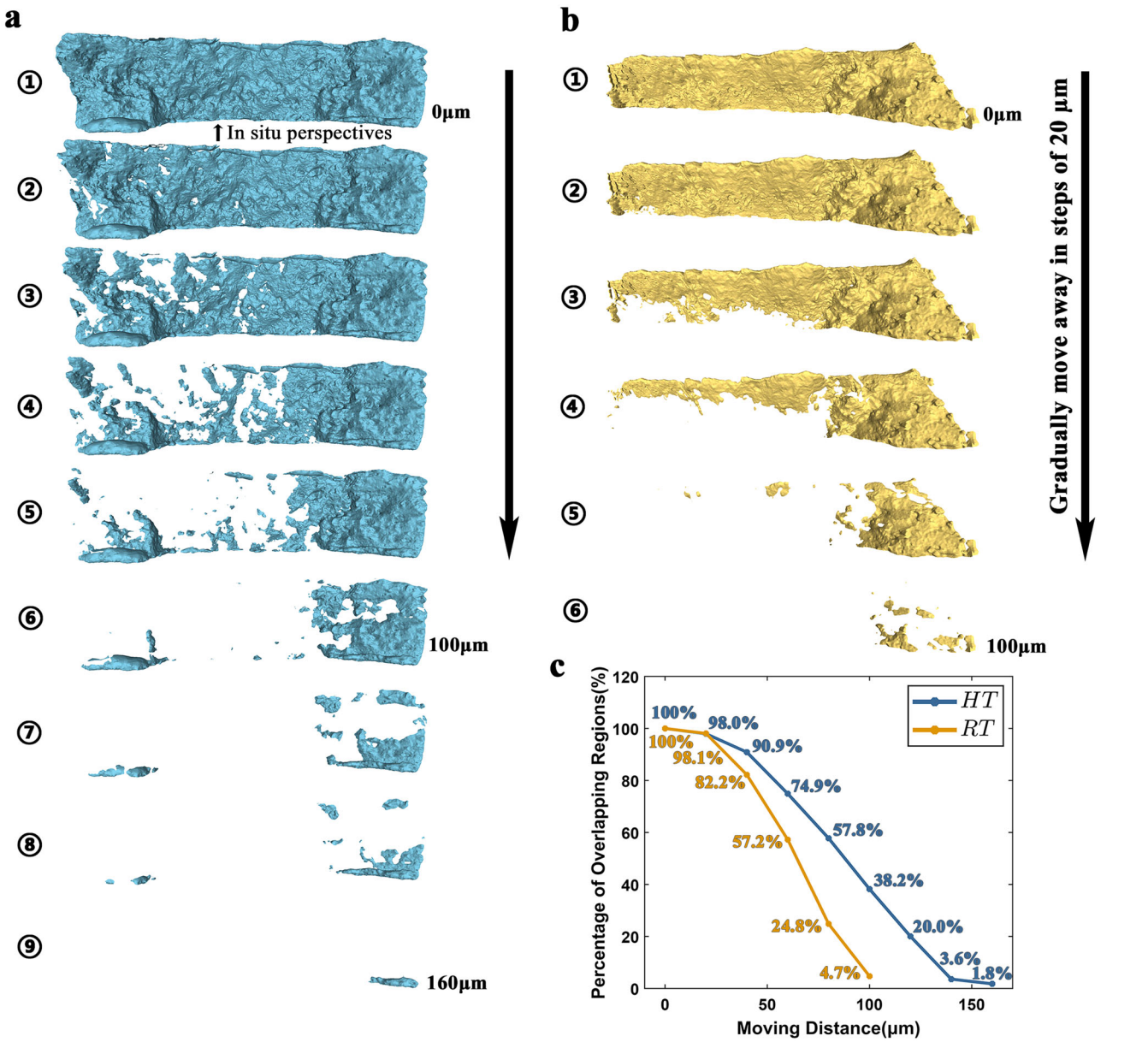


Fig. 3 | Inferred results of specimens fracture processes in IN718 650 °C high temperature (HT) and room temperature (RT) tensile experiments. a Results of HT fracture process inference. **b** Results of RT fracture process inference. **c** Statistical analysis of the ratio of crack area between the results of HT and RT fracture process inference under different stretching displacements.

more intuitive visual observation. By comparing Fig. 3a and b, it becomes evident that there are substantial differences in material cracking between RT and HT tensile deformation. There are more localized plastic cracking regions with smaller areas in HT conditions, indicating that the local non-uniform deformation phenomenon is more evident. This observation is also consistent with traditional theory.

In RT fracture process inference results, the transition from localized cracking to final fracture occurs at a tensile displacement of 100 μm. Conversely, in the HT fracture process inference results, this transition occurs at a more extended tensile displacement of 160 μm. Furthermore, Fig. 3c comparison reveals that RT tensile deformation results in a larger cracking area (overall statistical area) under the same strain, indicating poorer plasticity. This phenomenon cannot be directly observed through traditional fracture analysis techniques. Through the overlap analysis method of the 3D model, the observation and analysis of the fracture process become more intuitive and accurate, effectively solving the limitations of the traditional theory. With further improvements in spatial resolution, it can also be widely used for more microscopic observations and analysis.

Description of the internal fracture process and comparison with in situ experiments

When grain deformation occurs in polycrystalline materials, it will coordinate deformation with surrounding grains. Researchers categorize grains into internal and external grains based on their positions as they experience distinct mechanical constraints[44]. Surface grains possess certain free surfaces. Comparatively, fewer grains deform in coordination. Thus, the contributions of these two-grain types to mechanical deformation are different.

Based on the discussion in the previous section, it has been identified that the tensile fracture of the material is not limited to simple surface or internal cracking but rather involves a comprehensive combination of both. Traditional fracture offset analysis fails to provide a direct understanding of the deformation process, while in situ SEM characterization solely captures the evolution of surface morphology. The method proposed in this paper not only focuses on surface evolution but also simulates the occurrence of internal cracking. This allows for precise determination of the fracture location, whether it emerges internally or on the surface. Notably, previous methodologies have been unable to accomplish this.

According to the results of the HT fracture process inference, we have selected four regions within the specimen that experienced cracking first to describe further. Figure 4 displays a scanning image of Fracture A, highlighting the four selected regions using different colored circles. It can be observed that Region 1 does not show any typical defects. However, a crack is visible at the corresponding position on the opposite side of the fracture surface, potentially leading to cracking within the specimen. Similarly, Region 4 exhibits a comparable scenario, wherein multiple void defects can be observed at the corresponding position on the opposite side of the fracture surface. An inclusion is detected within Region 2. Further analysis using EDS reveals it is a titanium compound. Furthermore, Region 3 directly

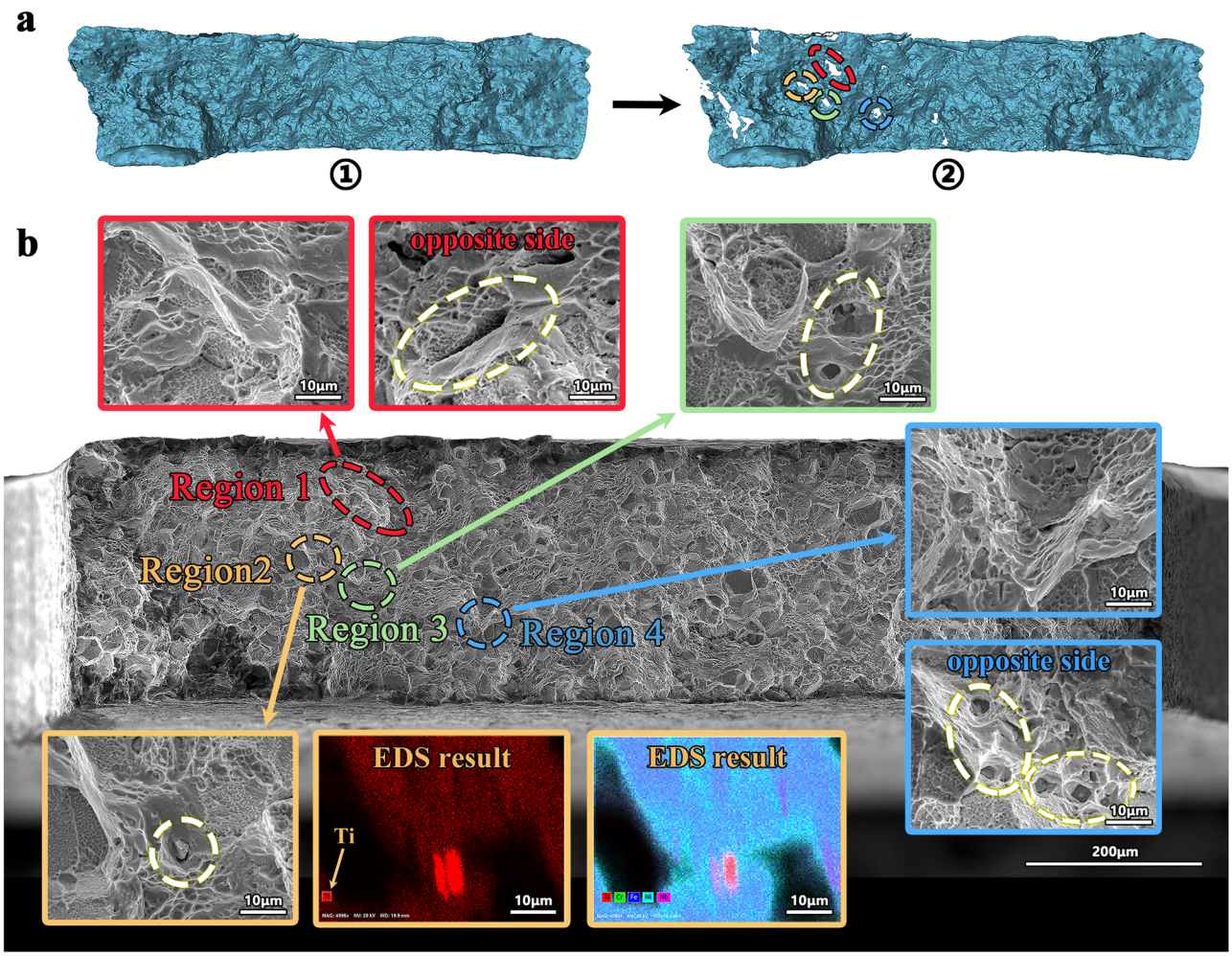


Fig. 4 | Description of internal cracks. a Based on the results of 650 ˚C high temperature (HT) fracture process inference, four regions were randomly selected for further analysis. The four regions are distinguished by different colors. **b** Scanning electron microscope (SEM) images of fracture A and four regions are shown along with additional images to illustrate internal defects. The opposite side in the additional images of Region 1 and Region 4 indicates the corresponding position of the images acquired in Fracture B. The additional images of Region 2 indicate the results of the energy dispersive spectrometer (EDS) analysis.

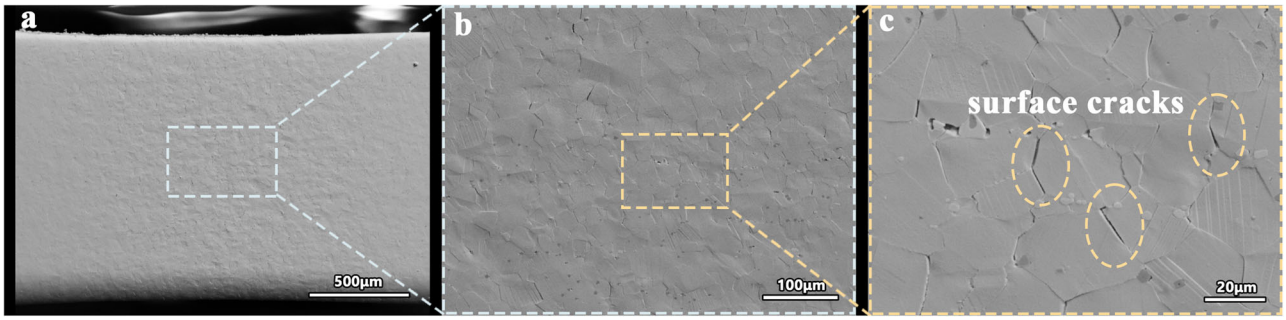


Fig. 5 | The morphology of surface cracks occurring during the in situ experiment under 650 °C high temperature (HT) tensile conditions. a 200x. **b** 500x. **c** 2kx.

displays two void defects. The dimples in the fracture surface images indicate significant plastic deformation within these regions, which is manifested macroscopically as excellent ductility. These defects may precipitate the initiation of cracks in these regions during tensile testing. Notably, the observed defects correspond to the regions where initial cracking was identified during the fracture process analysis, thereby enhancing the credibility of the overall inferential process. These findings, which rely on this study's fracture model and analysis method, offer compelling evidence of their distinct advantages and crucial significance in material fracture research.

In addition, we provide a detailed analysis of the distinct deformation processes of surface grains and internal during the tensile deformation of IN718 by comparing the results with the in situ experiments, with the results of HT fracture process inference serving as an illustrative example. The fracture process inference results depicted in Fig. 3 demonstrate that cracks occur at a strain of 140–160 μm before the specimen's final fracture strain. During the in situ experiment, images were captured to document the deformation length of the specimen. Figure 5a–c depicts the morphology of surface cracks occurring during the in situ experiment under HT tensile conditions. Notably, surface cracks (precisely grain boundary cracks) were observed to occur at 110 μm of strain before the strain reached the final fracture strain of the specimen under HT conditions. This suggests that internal cracking occurs before the manifestation of surface cracks, a phenomenon not previously observed in direct material observations. These empirical findings demonstrate the substantial distinction between surface and internal grains. Additionally, it implies that the surface phenomena observed through in situ SEM exhibit a delay compared to the actual material deformation.

Discussion

This study delves into advanced methodologies dedicated to unraveling the intricate superalloy fracture processes through the comprehensive analysis of overlapping 3D models. We introduce and implement an optimized image acquisition scheme tailored for the three-dimensional reconstruction of SEM images. The resulting calibrated 3D models not only endow our alloy fracture analysis with a wealth of visual details but also ensure a high degree of precision in scale information. By meticulously analyzing these overlapping 3D models, we can infer the initial position of specimen cracking and the evolution of crack paths during the tensile process. Our findings reveal a notable insight: under HT tensile experimental conditions, it is probable that internal defects within the specimens lead to internal cracking before surface cracking. This discovery aligns seamlessly with the outcomes of our in situ experiments, corroborating its validity. Looking ahead, this speculative methodology holds great promise for enriching our understanding of alloy fracture mechanisms, consequently facilitating the design and optimization of alloy performance. Furthermore, we acknowledge the challenges and opportunities ahead, particularly in the realms of enhancing the clarity of 3D fracture morphology models and advancing in situ 3D reconstruction techniques, both of which will considerably shape the trajectory of our future research efforts.

Methods

In situ tensile testing and reconstruction of 3D fracture morphology models

Two sets of in situ tensile experiments on IN718 were conducted to address specific issues. The in situ tensile specimens were designed in a dog-bone shape with a unique structure to enable easy observation during the experiments. Comprehensive information about the composition of the materials (Supplementary Tables 1–2), processing techniques, dimensions and structure of the specimens (Supplementary Fig. 1), and corrosion treatment can be found in the Supplementary Information.

The equipment used for conducting in situ tensile testing was developed by our research group and has been extensively described in previous work[2,45]. The concrete figure and the schematic cross-section of the working parts of the experimental equipment are shown in Supplementary Fig. 3–4. The testing equipment facilitates in situ tensile testing ranging from RT to 1200 °C, meeting high-resolution SEM image acquisition requirements. The experimental conditions and data results strongly support the findings of this research. Two specific sets of in situ tensile experiments were conducted: RT testing and on 650 °C IN718. Each set of in situ uniaxial tensile experiments was conducted at a constant tensile rate of 1 μm/s until fracture was evident. Real-time changes in the morphology of the SEM images are continuously monitored during the tensile process. Photographs are captured at distinctive intervals to record the corresponding displacement in the tensile curve. HT tensile testing at 650 °C necessitates a gradual temperature increase to a specific level, followed by a 30-minute dwell time to minimize temperature fluctuations and non-uniformity in the pre-marked section before the actual tensile testing. These details have been thoroughly elaborated on in previous studies and have reached a high level of applicability[3,45].

In the 3DSEM[35], images are obtained by tilting the sample stage to a fixed angle and capturing images at different angles by rotating the stage. However, this approach results in information loss due to surface features on the fractured surface, which resemble peaks and valleys and are more sensitive to gradient information. To address this concern, we propose an optimized method for image acquisition (Supplementary Fig. 5a), where tilting is employed instead of rotating the sample stage. In this method, the specimen is fixed on the stage inside the SEM chamber, and the tilt angle and field of view are adjusted to achieve optimal focus and ensure a clear and centered view of the specimen cross-section.

Image acquisition was realized with a Tescan Mira4 SEM. Considering the reconstruction effectiveness and data collection efficiency, the image acquisition scheme applied in this article is as follows: the sample stage is tilted along the X-axis and the Y-axis, capturing images at 5° intervals within the −30° to 30° range. The relative positions of the detector and the specimen throughout the image acquisition process are shown in Supplementary Fig. 5b. We used an SE detector to obtain 25 sets of raw image data at different angles, with each image being 2048 × 1024 pixels, and was saved in the TIF format. Throughout the entire image collection and reconstruction process, we only used Photoshop 2020 software to crop out the SEM information from the raw images and saved them in JPG format. The acquisition of the 3D model was achieved by processing the image sequence

using Agisoft Metashape software(v1.8.4) to complete the reconstruction process.

Since accurate camera parameters are not available in advance, the reconstructed 3D models are usually scaled up or down proportionally and thus require scale calibration. In order to obtain the scaling factor, standard measurement methods involve the use of other equipment, such as white light interferometers[46,47], white light chromatic sensors[48], and atomic force microscopes[49]. However, cost and time control are difficult for objects with large volumes and characteristic features like alloy fractures. We propose a scale calibration method for accurate scale calibration and adjustment of 3D models using only SEM images.

We aim to identify two marked points in the cross-section on the fracture surface, *dot*1 and *dot*2 (Supplementary Fig. 6a), each exhibiting distinctive features. By utilizing a 10 kx magnification, we meticulously observe and optimize the focus of *dot*1 to ensure image clarity. While maintaining a fixed focus, we incrementally adjust the height Z of the sample stage in 10 μm steps. This procedure allows us to capture a sequence of images that exhibit an increasing and subsequently decreasing level of sharpness (Supplementary Fig. 6b). We employ an identical method to obtain a sequential set of images for the position of *dot*2. Efficaciously, the entire process is completed in approximately 30 min.

Marturi[50,51] et al. evaluated the efficacy of standard sharpness functions in accurately determining the clarity of SEM images. They identified that the normalized variance equation effectively balances speed and accuracy trade-offs. The formula employed for calculating the normalized variance, denoted as $s(I)$, is presented below:

$$s(I) = \frac{1}{MN} \cdot \frac{1}{\mu} \cdot \sum_{u=1}^{M} \sum_{v=1}^{N} \left(i(u,v) - \mu\right)^2 \quad (1)$$

where M and N are the length and width of the input image, respectively, $i(u, v)$ is the pixel intensity at coordinate (u, v), and μ is the average intensity of all pixels in the input image.

The sharpness score curve for the two markers in Fracture A can be found in Supplementary Fig. 6c. It is important to note that the displacement of the specimen is equivalent to the displacement of the sample stage under a constant focal length, as they are fixed relative to each other. The 3D coordinates of the sample stage were determined at the point of highest clarity. Image sharpness assessment curves for four fracture models are shown in Supplementary Fig. 7. The distance d_R between *dot*1 and *dot*2 is:

$$d_R = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (2)$$

where x_1, y_1, z_1, and x_2, y_2, z_2 represent the XYZ axis coordinates of the sample stage when observing *dot*1 and *dot*2, respectively.

$$\delta = d_R / d_M \quad (3)$$

Locate *dot*1 and *dot*2 in the reconstructed 3D model, and measure and calculate the distance d_M. Calculate the scaling factor δ using equation (3), and use the scaling factor to scale the final calibrated 3D model.

Superalloys fracture process inference method and validation

Using the overlap analysis of 3D fracture morphology models to infer the fracture process of specimens, the specific method proceeds: firstly, the corresponding fracture models are gradually brought together until they overlap entirely (Supplementary Fig. 8a). To observe the projection of the target fracture surface coinciding with the opposite fracture surface, it can be observed from inside the opposite fracture. The complete morphology of the target fracture surface can be seen when two fracture models are fully overlapped (Supplementary Fig. 8d). This stage signifies the connected state of the specimen during the actual tensile process prior to fracture initiation. Subsequently, a gradual distancing from the associated model occurs with a step size of 20 μm (Supplementary Fig. 8b). During the tensile process, the region where cracking initiates displays a notch in the projection figure due to its reduced elongation rate (Supplementary Fig. 8e). With the increasing separation of the corresponding model, the cracking process, encompassing crack initiation, propagation, and ultimate fracture, become evident. After a separation distance of 160 μm, only the region within the black circle exhibits overlapping between the corresponding models (Supplementary Fig. 8c), indicating the location of the fracture termination (Supplementary Fig. 8f). The labeled distances in Fig. 3 represent the displacement at which the corresponding model gradually moves away, equivalent to the tensile displacement observed in actual tensile experiments. It should be noted that our analysis method assumes that all local material points follow a uniaxial tensile state.

To further validate the accuracy and effectiveness of the tensile fracture process inference, we conducted an additional analysis using in situ tensile experiments conducted on a new Ni-based wrought superalloy (NWS)[52] specimen at RT. The material's fracture process becomes increasingly apparent as the deformation progresses. The position of the initial surface cracking coincides with the fracture process inference results, thereby providing robust evidence for the highly accurate fracture process inferred by our framework (Supplementary Fig. 9).

Furthermore, we employed our proposed methodology to reconstruct and analyze the ex-situ tensile fractures of circular cross-section specimens made of IN718 to validate the method's universality. Six typical stages of the fracture process inference are shown (Supplementary Fig. 10a). The original SEM images of the fractures are depicted, revealing the fundamental characteristics of the morphology (Supplementary Fig. 10b). However, they lack precise 3D perspective and quantitative analysis. Subsequently, the results of the 3D reconstruction and registration of the fractures are showcased (Supplementary Fig. 10c). The comparative study demonstrates the practical application of our methodology to circular section specimens' fractures, thereby affirming the universality of our framework.

Data availability

The source data and data folds in the four fractures of two specimens used in this study are provided on the Zenodo at https://zenodo.org/records/10430093.

Code availability

The source code for calculating image sharpness in SFPI is available on the Zenodo at https://zenodo.org/records/10430093.

Received: 22 January 2024; Accepted: 29 July 2024;
Published online: 06 August 2024

References

1. Shahwaz, M., Nath, P. & Sen, I. A critical review on the microstructure and mechanical properties correlation of additively manufactured nickel-based superalloys. *J. Alloy. Compd.* **907**, 164530 (2022).
2. Guo, G. et al. In-situ SEM-EBSD investigation of the low-cycle fatigue deformation behavior of Inconel 718 at grain-scale. *J. Mater. Res. Technol.* **24**, 5007–5023 (2023).
3. Sang, L. et al. In-situ SEM study of temperature-dependent tensile behavior of Inconel 718 superalloy. *J. Mater. Sci.* **56**, 16097–16112 (2021).
4. Jiang, W. et al. Study of pre-precipitated δ phase promoting deformation twinning and recrystallization behavior of Inconel 718 superalloy during hot compression. *Mater. Des.* **226**, 111693 (2023).
5. Hanson, J. P. et al. Crystallographic character of grain boundaries resistant to hydrogen-assisted fracture in nibase alloy 725. *Nat. Comm.* **9**, 3386 (2018).
6. Naragani, D. P., Shade, P. A., Kenesei, P., Sharma, H. & Sangid, M. D. X-ray characterization of the micromechanical response ahead of a propagating small fatigue crack in a Ni-based superalloy. *Acta Mater.* **179**, 342–359 (2019).

7. Griffiths, S. et al. Combining alloy and process modification for microcrack mitigation in an additively manufactured Ni-base superalloy. *Addit. Manuf.* **36**, 101443 (2020).
8. Mandelbrot, B. B., Passoja, D. E. & Paullay, A. J. Fractal character of fracture surfaces of metals. *Nature* **308**, 721–722 (1984).
9. Shyam, A. & Milligan, W. Effects of deformation behavior on fatigue fracture surface morphology in a nickelbase superalloy. *Acta Mater.* **52**, 1503–1513 (2004).
10. Zhu, Q., Wang, C., Qin, H., Chen, G. & Zhang, P. Effect of the grain size on the microtensile deformation and fracture behaviors of a nickel-based superalloy via ebsd and in-situ synchrotron radiation x-ray tomography. *Mater. Charact.* **156**, 109875 (2019).
11. Zhao, Z. et al. A comprehensive study of the anisotropic tensile properties of laser additive manufactured Ni-based superalloy after heat treatment. *Int. J. Plast.* **148**, 103147 (2022).
12. Lv, P. et al. Temperature effects on tensile behaviors and relevant deformation mechanisms of a low-cost nickelbased single crystal superalloy containing 1.5% re. *J. Alloy. Compd.* **926**, 166819 (2022).
13. Gontard, L. et al. Assessment of engineered surfaces roughness by high-resolution 3D SEM photogrammetry. *Ultramicroscopy* **177**, 106–114 (2017).
14. Li, Y. et al. Multi-scale microstructural investigation of a laser 3D printed Ni-based superalloy. *Addit. Manuf.* **34**, 101220 (2020).
15. Eckmann, S. & Schweizer, C. Characterization of fatigue crack growth, damage mechanisms and damage evolution of the nickel-based superalloys mar-m247 cc (hip) and cm-247 lc under thermomechanical fatigue loading using in situ optical microscopy. *Int. J. Fatigue* **99**, 235–241 (2017).
16. Fukunaga, A. Differences between internal and external hydrogen effects on slow strain rate tensile test of ironbased superalloy a286. *Int. J. Hydrog. Energy* **47**, 2723–2734 (2022).
17. Kumar, P. et al. Compressive vs. tensile yield and fracture toughness behavior of a body-centered cubic refractory high-entropy superalloy al0. 5nb1. 25ta1. 25tizr at temperatures from ambient to 1200 °C. *Acta Mater.* **245**, 118620 (2023).
18. Zhu, Q., Chen, G., Wang, C., Qin, H. & Zhang, P. Tensile deformation and fracture behaviors of a nickel-based superalloy via in situ digital image correlation and synchrotron radiation x-ray tomography. *Materials* **12**, 2461 (2019).
19. Liu, K. et al. In-situ X-ray tomography investigation of pore damage effects during a tensile test of a Ni-based single crystal superalloy. *Mater. Charact.* **177**, 111180 (2021).
20. Zhang, L., Sun, Z., Xue, Q., Lei, M. & Tian, X. Transient liquid phase bonding of ic10 single crystal with GH3039 superalloy using bni2 interlayer: microstructure and mechanical properties. *Mater. Des.* **90**, 949–957 (2016).
21. Kim, H., Cong, W., Zhang, H.-C. & Liu, Z. Laser engineered net shaping of nickel-based superalloy Inconel 718 powders onto aisi 4140 alloy steel substrates: interface bond and fracture failure mechanism. *Materials* **10**, 341 (2017).
22. Polcari, D., Dauphin-Ducharme, P. & Mauzeroll, J. Scanning electrochemical microscopy: a comprehensive review of experimental parameters from 1989 to 2015. *Chem. Rev.* **116**, 13234–13278 (2016).
23. Vernon-Parry, K. D. Scanning electron microscopy: an introduction. *III-Vs Rev.* **13**, 40–44 (2000).
24. Ding, W., Zhang, Y., Lu, H., Wan, W. & Shen, Y. Automatic 3D reconstruction of SEM images based on nanorobotic manipulation and epipolar plane images. *Ultramicroscopy* **200**, 149–159 (2019).
25. Eggeman, A. S., Krakow, R. & Midgley, P. A. Scanning precession electron tomography for three-dimensional nanoscale orientation imaging and crystallographic analysis. *Nat. Commun.* **6**, 7267 (2015).
26. Borzunov, A. et al. 3D surface topography imaging in SEM with improved backscattered electron detector: arrangement and reconstruction algorithm. *Ultramicroscopy* **207**, 112830 (2019).
27. Lu, J., Chen, B., Liu, X., Yang, F. & Robinson, I. K. 3D microstructure reconstruction of casting aluminum alloy based on serial block-face scanning electron microscopy. *J. Alloy. Compd.* **778**, 721–730 (2019).
28. Jung, J. M., Yoo, J. H., Jeong, H. J., Lee, S. & Kim, H. S. Three-dimensional characterization of sic particle-reinforced al composites using serial sectioning tomography and thermo-mechanical finite element simulation. *Mater. Trans. A* **45**, 5679–5690 (2014).
29. Yuan, H., Van de Moortele, B. & Epicier, T. Accurate post-mortem alignment for focused ion beam and scanning electron microscopy (FIB-SEM) tomography. *Ultramicroscopy* **228**, 113265 (2021).
30. Liu, K. et al. Effect of cooling rate on carbides in directionally solidified nickel-based single crystal superalloy: X-ray tomography and U-net CNN quantification. *J. Alloy. Compd.* **883**, 160723 (2021).
31. Verhoeven, G. Taking computer vision aloft-archaeological three-dimensional reconstructions from aerial photographs with photoscan. *Archaeol. Prospect.* **18**, 67–73 (2011).
32. Gontard, L. C., Schierholz, R., Yu, S., Cintas, J. & Dunin-Borkowski, R. E. Photogrammetry of the threedimensional shape and texture of a nanoscale particle using scanning electron microscopy and free software. *Ultramicroscopy* **169**, 80–88 (2016).
33. Barbieri, G. & da Silva, F. P. Acquisition of 3D models with submillimeter-sized features from SEM images by use of photogrammetry: a dimensional comparison to microtomography. *Micron* **121**, 26–32 (2019).
34. Sihvonen, T. & Reinikainen, S.-P. Image based evaluation of textured 3DSEM models. *Ultramicroscopy* **238**, 113518 (2022).
35. Tafti, A. P. et al. 3DSEM++: adaptive and intelligent 3D SEM surface reconstruction. *Micron* **87**, 33–45 (2016).
36. Eulitz, M. & Reiss, G. 3D reconstruction of SEM images by use of optical photogrammetry software. *J. Structural Biol.* **191**, 190–196 (2015).
37. Tondare, V. N. A concept for three-dimensional particle metrology based on scanning electron microscopy and structure-from-motion photogrammetry. *J. Res. Natl. Inst. Standards Technol*. **125**, 125014 (2020).
38. Tafti, A. P., Kirkpatrick, A. B., Alavi, Z., Owen, H. A. & Yu, Z. Recent advances in 3D SEM surface reconstruction. *Micron* **78**, 54–66 (2015).
39. Kozikowski, P. Extracting three-dimensional information from SEM images by means of photogrammetry. *Micron* **134**, 102873 (2020).
40. Li, L. et al. The effect of Al and Ti on the microstructure, mechanical properties and oxidation resistance of γ-Ni3 (Al, Ti) strengthened austenitic stainless steels. *J. Mater. Res. Technol.* **24**, 4650–4660 (2023).
41. Zhang, Z. et al. Microstructures and mechanical properties of a L12-structured precipitation strengthened Co-based superalloy. *J. Mater. Res. Technol.* **26**, 7789–7802 (2023).
42. Stinville, J. C. et al. High resolution mapping of strain localization near twin boundaries in a nickel-based superalloy. *Acta Mater.* **98**, 29–42 (2015).
43. Texier, D. et al. Strain localization in the Alloy 718 Ni-based superalloy: from room temperature to 650 °C. *Acta Mater.* **268**, 119759 (2024).
44. Gao, W. et al. Effect of grain size on deformation and fracture of Inconel718: an in-situ SEM-EBSD-DIC investigation. *Mater. Sci. Eng. A* **861**, 144361 (2022).
45. Chen, J. et al. In-situ study of adjacent grains slip transfer of Inconel 718 during tensile process at high temperature. *Int. J. Plast.* **163**, 103554 (2023).
46. Jacobs, S. D. Manipulating mechanics and chemistry in precision optics finishing. *Sci. Technol. Adv. Mater.* **8**, 153–157 (2007).
47. Yan, L., Rong, Y., Jiang, F. & Zhou, Z. X. Threedimension surface characterization of grinding wheel using white light interferometer. *Int. J. Adv. Manuf. Technol.* **55**, 133–141 (2011).
48. Darafon, A., Warkentin, A. & Bauer, R. Characterization of grinding wheel topography using a white chromatic sensor. *Int. J. Mach. Tools Manuf.* **70**, 22–31 (2013).

49. Tian, X. et al. Afm characterization of surface mechanical and electrical properties of some common rocks. *Int. J. Min. Sci. Technol.* **32**, 435–445 (2022).
50. Marturi, N., Tamadazte, B., Dembele, S. & Piat, N. Visual servoing-based approach for efficient autofocusing in scanning electron microscope. *In 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2677–2682 (2013).
51. Marturi, N., Tamadazte, B., Dembele, S. & Piat, N. Visual servoing-based depth-estimation technique for manipulation inside SEM. *IEEE Trans. Instrum. Meas.* **65**, 1847–1855 (2016).
52. Liu, H. et al. Influence of thermal exposure on microstructural stability and tensile properties of a new Ni-base superalloy. *J. Mater. Res. Technol.* **21**, 4462–4472 (2022).
53. Guillaume, R. et al. Impact of solute flow during directional solidification of a Ni-based alloy: in-situ and real-time X-radiography. *Acta Mater.* **177**, 111180 (2021).

Acknowledgements

This work has been supported by National Natural Science Foundation of China (No. 92367205), National Science and Technology Major Project of China (J2019-III-0008-0051), Basic Science Center Program for Multiphase Media Evolution in Hypergravity of the National Natural Science Foundation of China (No. 51988101), Zhejiang Provincial Natural Science Foundation of China under Grant LZ22F030011, Shanxi-Zheda Institute of Advanced Materials and Chemical Engineering (2022SZ-TD007). The supports are gratefully acknowledged.

Author contributions

X. Zhang, W.M., J.C., Z.Z. designed the research study. J.W., X. Zhao prepared and provided the specimens. X. Zhang, J.W. conducted in situ tensile experiments. X. Zhang, Z.L., S.H. implemented the method and wrote the code. X. Zhang, G.G., Y.Z. performed the analysis. X. Zhang, G.G., Z.L., W.M., Q.Y. wrote the paper. All authors read and approved the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44172-024-00257-6.

Correspondence and requests for materials should be addressed to Wenchao Meng or Yuefei Zhang.

Peer review information *Communications Engineering* thanks E.S. Perdahcioglu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Emanuela Bosco, Anastasiia Vasylchenkova, Rosamund Daw. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2024

证书号第8122929号

专利公告信息

发明专利证书

发明名称：一种合金断口三维重建与尺度校准方法

专利权人：浙江大学

地　　址：310058 浙江省杭州市西湖区余杭塘路866号

发 明 人：张雪成;孟文超;贺诗波;王晋;顾超杰;陈积明;陈爽

专　利　号：ZL 2023 1 1344109.2　　授权公告号：CN 117456090 B

专利申请日：2023年10月16日　　授权公告日：2025年08月01日

申请日时申请人：浙江大学

申请日时发明人：张雪成;孟文超;贺诗波;王晋;顾超杰;陈积明;陈爽

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，并予以公告。

专利权自授权公告之日起生效。专利权有效性及专利权人变更等法律信息以专利登记簿记载为准。

局长
申长雨

国家知识产权局
2025年08月01日

证书号第8519674号

专利公告信息

发明专利证书

发明名称：一种基于扫描电镜的原位三维表面重建方法

专利权人：浙江大学

地　　址：310058 浙江省杭州市西湖区余杭塘路866号

发 明 人：张雪成;张跃飞;孟文超;张斌;郭广浩;叶显爵

专　利　号：ZL 2025 1 0283025.5　　授权公告号：CN 120121657 B

专利申请日：2025年03月11日　　授权公告日：2025年11月25日

申请日时申请人：浙江大学

申请日时发明人：张雪成;张跃飞;孟文超;张斌;郭广浩;叶显爵

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，并予以公告。

专利权自授权公告之日起生效。专利权有效性及专利权人变更等法律信息以专利登记簿记载为准。

局长
申长雨

国家知识产权局
2025年11月25日