

文章编号 1004-924X(2025)15-2441-14

MScU-Net 全息相位解包裹与像差同步求解

严利平, 苗小喆, 黄 柳*

(浙江理工大学 信息科学与工程学院(网络空间安全学院), 浙江 杭州 310018)

摘要: 数字全息显微术中相位恢复面临噪声导致的解包裹误差与像差引入的相位畸变双重挑战, 现有分步处理方法存在计算效率低与误差传递的固有局限, 亟需发展同步求解新范式。为此, 提出了一种基于多光谱通道 U-Net(MScU-Net) 的数字全息显微相位恢复方法。该网络在 ResU-Net 架构中引入 Transformer 增强的多光谱通道注意力机制, 通过端到端非线性映射, 直接从含噪、失真的包裹相位数据中实现相位解包裹与像差补偿的同步求解。实验结果表明, MScU-Net 在相位恢复中表现出显著优势, 其中, 峰值信噪比 PSNR 达 29.82 dB, 结构相似度 SSIM 保持 0.96; 背景区域噪声标准差 (0.000 4~0.003 0 rad) 较最优对比方法降低 82.6%~95.8%; 实测台阶高度系统偏差 $\leq 3.74\%$, 测量局部波动达到亚纳米级 ($s = 0.42$ nm), 该方法为微纳尺度的三维形貌测量提供了有效解决方案。

关键词: 数字全息显微; 相位解包裹; 像差补偿; 多光谱通道注意力机制

中图分类号: TH744; O436 **文献标识码:** A

doi: 10.37188/OPE.20253315.2441

CSTR: 32169.14.OPE.20253315.2441

MScU-Net for simultaneous phase unwrapping and aberration correction in digital holography

YAN Liping, MIAO Xiaozhe, HUANG Liu*

(School of Information Science and Engineering (School of Cyber Science and Technology),
Zhejiang Sci-Tech University, Hangzhou 310018, China)

* Corresponding author, E-mail: huangliu@zstu.edu.cn

Abstract: Phase retrieval in digital holographic microscopy contends with unwrapping errors induced by noise and phase distortions arising from aberrations. Traditional sequential approaches are constrained by limited computational efficiency and error propagation, motivating the development of synchronous solution paradigms. A multi-spectral channel U-Net (MScU-Net) is proposed for phase retrieval in digital holographic microscopy. The network embeds a Transformer-enhanced multi-spectral channel attention mechanism within a ResU-Net framework, enabling end-to-end nonlinear mapping that simultaneously performs phase unwrapping and aberration compensation directly from noisy, distorted wrapped phase maps. Experimental results demonstrate substantial improvements: peak signal-to-noise ratio (PSNR) attains 29.82 dB, structural similarity index measure (SSIM) is maintained at 0.96, and background noise standard deviation (0.000 4-0.003 0 rad) is reduced by 82.6%-95.8% relative to a state-of-the-art method.

收稿日期: 2025-06-26; 修订日期: 2025-07-24.

基金项目: 浙江省自然科学基金资助项目 (No. LQ23E050020); 浙江省领雁计划资助项目 (No. 2024C01174); 浙江省科技创新领军人才资助项目 (No. 2022R52052)

Step-height measurements show a systematic deviation of $\leq 3.74\%$ from white-light interferometry references, with local measurement fluctuations reaching sub-nanometer precision ($s = 0.42 \text{ nm}$). These findings indicate that MScU-Net offers an effective synchronous solution for accurate micro and nanoscale 3D topography measurement.

Key words: digital holographic microscopy; phase unwrapping; aberration correction; multi-spectral channel attention mechanism

1 引 言

微纳米尺度表面形貌测量在精密制造、生物医学和材料科学等领域具有重要应用。针对微结构的三维复杂形貌特征与亚微米尺度特性,数字全息显微术作为一种非接触、高分辨率光学测量方法,提供了有效的解决方案^[1]。数字全息显微术通过计算全息图的相位分布重构样本形貌。然而,由于相位包裹效应,相位数据被截断在 $[-\pi, \pi]$ 内,需经相位解包裹才能获得连续的三维形貌分布^[2]。噪声干扰显著增加了相位解包裹算法的复杂度,要求算法兼具高精度、强稳定性和抗噪能力;同时,系统像差需在相位解包裹后通过额外校正步骤补偿,以实现准确的样本相位重建。因此,在噪声和像差耦合干扰下实现从包裹相位到真实表面形貌的高精度重构,是数字全息显微技术的核心挑战。

相位解包裹方法主要分为传统算法和深度学习算法两大类。传统算法包含路径跟踪法和最小范数法两个分支切法^[3]和质量引导法^[4]等路径跟踪法,通过局部路径规划策略展开相位数据,能有效规避噪声区域,计算效率较高,但对复杂形貌适应性较差,容易产生误差传播。尽管卡尔曼滤波和遗传模拟退火^[5]等改进策略优化路径规划,但噪声敏感性问题仍未完全解决。加权最小二乘法^[6]和基于压缩感知的相位校正模型^[7]等最小范数法,采用全局优化框架求解相位分布,在噪声环境下具有更好的稳定性,但计算复杂度高,且依赖准确的权重矩阵设计,在处理低质量相位数据时精度受限。这些传统算法局限性促使研究者转向端到端的深度学习解决方案。

基于深度学习的解包裹算法主要分为端到端回归方法和深度学习-传统算法融合方法两大类。端到端回归方法通过深度学习架构如 PHU-Net^[8]和 PU-GAN^[9]等,直接建立从包裹相位到真

实相位的非线性映射关系。这类方法的优势在于其能够隐式学习并补偿系统像差,通过训练数据中包含的像差特征^[10],而无需显式的像差建模步骤。近期研究通过引入 Transformer 自注意力机制^[11-12]等结构,显著提升网络对噪声干扰和 2π 相位跳跃的识别能力,在保持端到端处理优势的同时进一步提高了重建精度。

相位解包裹的深度学习方法-传统算法融合方法通过结合数据驱动模型与物理约束,显著提升传统算法性能,包括:(1)结合深度学习无效点去除和传统路径跟踪的混合空间相位解包裹方法^[13-14];(2)融合深度学习相位调制提取与传统排序算法的单帧相位解包裹方法^[15];(3)基于深度学习特征提取与传统图切割优化的可微分 NeuralPUMA 方法^[16];(4)包裹相位计数方法^[17]。其中,包裹相位计数方法将相位解包裹问题转化为像素级包裹次数的分类任务,通过深度学习模型(如 PhaseNet)预测各像素点的 2π 跳变次数,再结合传统相位积分算法完成最终相位重建。这类方法通过构建物理先验数据集和设计自监督损失函数,有效提升了在低信噪比和复杂背景条件下的分类准确性。深度学习方法与传统算法的融合策略既保留了传统算法的物理合理性,又通过数据驱动优化提升了计算效率。然而,这类融合方法的架构设计复杂,需解决深度学习模块与传统算法间的接口兼容问题,且相位解包裹性能受限于传统算法的理论约束,如路径依赖导致误差传播,病态优化引发数值不稳定,拓扑限制阻碍复杂相位重建等问题。

相位像差补偿方法主要分为物理补偿、数值补偿和深度学习补偿三大类。物理补偿方法通过光路优化直接校正像差,如搭建远心显微结构^[18]、引入液体透镜^[19]或相同的显微物镜、记录无物体全息图^[20],需要精密光学调整以及后续的数值补偿残余像差^[21]。数值补偿基于数学模型

拟合像差相位,主要包括 Zernike 多项式拟合^[22]、主成分分析^[23]和交替方向法^[24]等。数值方法计算量大,且预设像差模型可能无法适应复杂像差建模,如对非对称、高阶像差的精确补偿。深度学习补偿方法包括基于 U-Net 的背景分割与 Zernike 多项式拟合^[25-26]、基于两阶段生成对抗网络(GAN)实现无物体全息图直接映射^[27],以及通过深度残差网络 ResNet-50^[28]预测 Zernike 多项式系数等方法。Li 等提出了一种基于 PACU-net3+网络^[29]的数字全息显微镜相位像差的同步补偿与相干噪声去除方法,建立包裹相位图作为输入,无噪声样本相位图作为标签的数据集,简化了传统两步相位处理流程(先相位补偿后去噪),其重建结果与传统方法相当。深度学习像差补偿方法自适应性强、计算效率高,但依赖大量训练数据,泛化能力受限。

目前,分步式相位解包裹和相位像差补偿方法存在计算效率低与误差传递的固有缺陷。分阶段深度网络架构(由独立的相位解包裹网络和相位像差补偿网络串联构成),虽然能通过模块化设计和分步训练实现处理流程的智能化和计算并行性,但仍存在显著局限性。首先,需要分别构建两个专用网络架构并独立训练,不仅增加了模型开发复杂度,还需要准备两套大规模训练数据集,其中包裹相位-连续相位数据集用于解包裹网络,像差相位-补偿相位数据集用于像差补偿网络;其次,级联架构导致相位误差在解包裹和像差补偿两个阶段之间传递累积,且整体计算效率较低。这些局限性的根源在于解包裹与像差补偿的分离处理模式,开发能够同步实现相位解包裹和像差补偿的一体化网络架构,已成为提升相位成像性能的必然技术路径。本文提出一种基于 MScU-Net 的数字全息显微相位同步求解方法,通过在 ResU-Net 架构中集成 Transformer 增强的多光谱通道注意力机制,实现了相位解包裹与像差补偿的联合优化。该方法利用离散余弦变换(Discrete Cosine Transform, DCT)提取包裹相位的全局频谱特征,有效捕捉 2π 相位跳变,并通过空间域注意力机制建立局部相位梯度与全局像差的映射关系。这一同步求解方法不仅简化了系统复杂度,还避免了传统分步处理中的误差传递问题,为复杂光学环境下的

高精度相位测量提供了可靠的解决方案。

2 方法和原理

设 $\varphi_o(x, y)$ 表示样本的真实相位分布, $\psi(x, y)$ 为通过数字全息显微系统采集并经数值重建得到的包裹相位数据。在实际测量过程中,由于系统噪声和光学系统像差干扰, $\psi(x, y)$ 与 $\varphi_o(x, y)$ 之间存在如下数学关系:

$$\psi(x, y) = \text{angle} \left\{ \exp \left[j\varphi_o(x, y) + j\varphi_a(x, y) + j\varphi_n(x, y) \right] \right\}, \quad (1)$$

其中: $\text{angle}(\cdot)$ 为反正切变换,将相位数据截断在 $[-\pi, \pi]$ 内; $\varphi_a(x, y)$ 和 $\varphi_n(x, y)$ 分别表示像差相位和噪声相位。

这里提出一种端到端深度学习架构 MScU-Net,通过构建 $\psi(x, y) \rightarrow \varphi_o(x, y)$ 的非线性映射模型,直接从受噪声和像差干扰的包裹相位中恢复出真实连续的样本相位。图 1 为 MScU-Net 实现相位解包裹与像差补偿的一体化相位恢复流程。

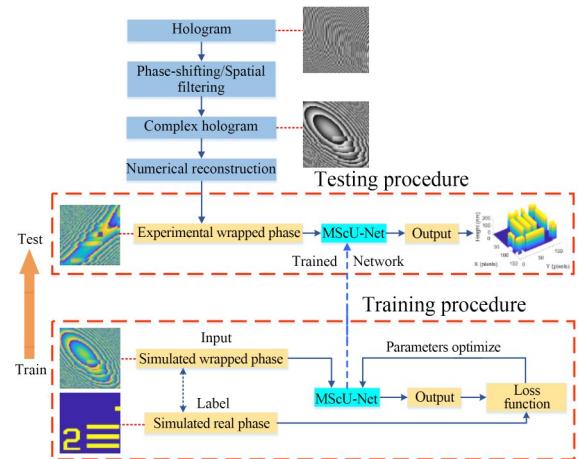


图 1 MScU-Net 网络相位恢复流程

Fig. 1 Flowchart of phase retrieval using MScU-Net network

MScU-Net 网络的训练流程中,采用仿真数据集进行监督学习,其中输入(Input)为含噪声和像差的仿真包裹相位 $\psi(x, y)$,标签(Label)为对应的样本相位 $\varphi_o(x, y)$ 。通过前向传播获得预测相位后,计算其与真实相位间的损失函数,并采用反向传播算法优化 MScU-Net 网络参数。

MScU-Net网络的测试流程包含以下步骤:

(1)搭建同轴或离轴数字全息显微系统,采集样本全息图(Hologram);

(2)采用相移运算或空间滤波提取仅含正一项的复全息图(Complex hologram);

(3)通过角谱衍射理论数值再现(Numerical reconstruction)物光复振幅 $U(x,y)$;

(4)经反正切变换 $\text{angle}[U(x,y)]$ 提取实验包裹相位 $\phi(x,y)$ (Experimental wrapped phase);

(5)将 $\phi(x,y)$ 输入经过训练收敛的 MScU-Net 网络,输出样本连续相位 $\phi_o(x,y)$;

(6)根据相位与高度转换公式计算样本表面形貌分布。

2.1 深度学习网络架构

受 ResU-Net^[10] 和 FcaNet^[30] 网络启发,本文设计的 MScU-Net 网络架构采用 U 形结构,主要由收缩路径、桥梁路径和扩张路径组成,如图 2 所示,收缩路径逐层下采样提取高层次特征,每层包含两个 3×3 卷积,配合最大池化实现降维,并通过多频残差块实现频率信息提取。桥梁路径连接收缩路径与扩张路径,通过跳跃连接增强特征表达能力,去除池化操作,仅使用卷积和残差连接,以保持高层特征一致性,提高训练稳定性。扩张路径逐步上采样恢复特征尺寸,每层包含反卷积和两个 3×3 卷积,通道数逐层减半,与收缩

路径特征通过跳跃连接融合,保留高分辨率信息。多频残差模块采用 DCT 进行多光谱通道的频率提取,使特征包含更多高频信息,提升网络对纹理和边缘特征的提取能力。表 1 详细列出了 MScU-Net 层次架构。

MScU-Net 中多频残差模块采用多分支残差结构和多光谱通道注意力机制相结合。如图 3(a) 所示,该模块首先采用五分支并行设计(含 1×1 卷积与多层 3×3 卷积组)实现多尺度特征提取。每个分支负责提取不同感受野范围的特征,以平衡局部细节和全局上下文信息。其中,Branch0 采用 1×1 卷积,用于提取低级特征并降低计算量,而 Branch1~Branch4 通过逐步增加 3×3 卷积堆叠的层数,以提取小尺度局部特征和大尺度全局信息。最终,所有分支的输出在通道维度拼接,并与 Branch0 的输出通过残差连接相加,从而增强梯度流动并防止信息丢失。

在传统多光谱通道注意层(Multi Spectral Attention Layer, MSAL)中,通常采用全连接层进行通道维度建模,但其无法有效捕捉不同空间位置间的相互依赖关系,缺乏空间维度上的全局建模能力。为了克服该问题,本研究将 Transformer 机制引入多光谱通道注意层,构建具有全局空间-通道交互能力的增强型注意力机制,结构如图 3(b) 所示。

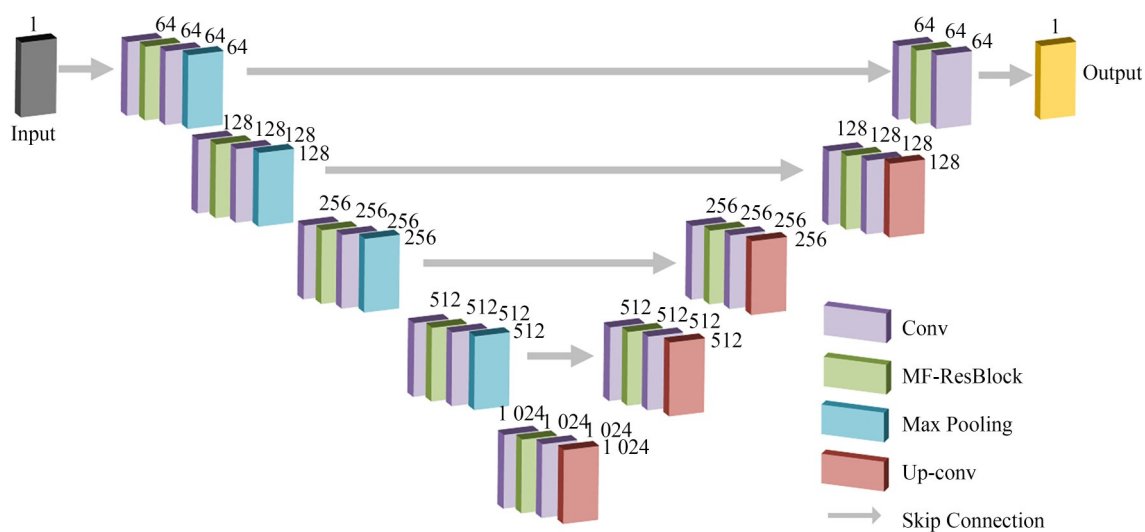


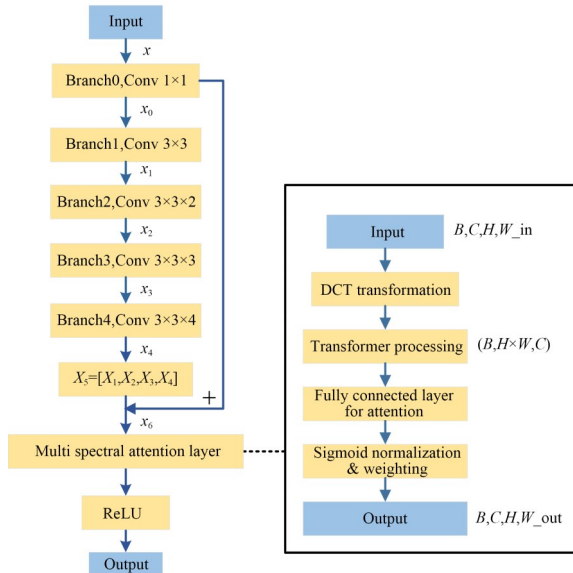
图 2 MScU-Net 网络架构

Fig. 2 MScU-Net network architecture

表 1 MScU-Net 层次架构

Tab. 1 Hierarchical architecture of MScU-Net

Layer	Input channels	Operator module	Output channels
1	1	DownB	64
2	64	DownB	128
3	128	DownB	256
4	256	DownB	512
5	512	MF-ResB	1024
6	1024	UpB	512
7	512	UpB	256
8	256	UpB	128
9	128	UpB	64
10	64	Outconv	1



(a) 多频残差模块(MF-ResBlock)架构
(a) Architecture of multi frequency residual block (MF-ResBlock)

(b) 融合Transformer的多光谱通道注意力层(Multi Spectral Attention Layer)结构
(b) Structure of Transformer-enhanced multi spectral attention layer

图 3 融合后的 MF-ResBlock 架构

Fig. 3 Integrated MF-ResBlock architecture

融合 Transformer 的多光谱通道注意力层的核心步骤如下:

(1) 多频段压缩

假设 $X \in \mathbf{R}^{C \times H \times W}$ 是网络中的图像特征图, C 是通道数, H 和 W 分别是特征的高度和宽度。首先将输入特征图 X 沿通道维度分割为 n 个部分 $[X^0, X^1, \dots, X^{n-1}]$, 其中 $X^i \in \mathbf{R}^{C' \times H \times W}$, $i \in \{0, 1, \dots,$

$n-1\}$, $C' = C/n$ 。然后进行多频率分量压缩, 对每个部分 X^i 进行二维离散余弦变换(2DDCT)^[28], 可用公式表示为:

$$Freq^i = 2DDCT^{u_i, v_i}(X^i) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} X_{:,h,w}^i \times B_{h,w}^{u_i, v_i} \quad s. t. \quad i \in \{0, 1, \dots, n-1\}, \quad (2)$$

其中: $Freq^i \in \mathbf{R}^{C'}$ 是 DCT 压缩后的 C' 尺度向量; (u_i, v_i) 表示与 X^i 对应的频率分量 2D 索引; $X_{:,h,w}^i$ 表示输入特征图中第 i 部分; $B_{h,w}$ 是 DCT 基函数。

(2) 多光谱特征向量构建

将所有 DCT 压缩向量进行拼接, 构建多光谱特征向量, 得到:

$$Freq = compress(X) = cat([Freq^0, Freq^1, \dots, Freq^{n-1}]), \quad (3)$$

其中: $Freq \in \mathbf{R}^C$ 为得到的多光谱向量, $cat(\cdot)$ 表示拼接操作。

(3) Transformer 空间-通道建模

将 $Freq$ 的空间维度展平, 形成 Transformer 输入格式: $V = Flatten(Freq) \in \mathbf{R}^{B \times (H \times W) \times C}$, 其中 B 为批次大小, $H \times W$ 为空间 token 数。此时每个空间位置 (h, w) 对应一个 C -维 token。

Transformer 编码器通过多头自注意力 (Multi-Head Attention, MHA) 和前馈网络 (Feed-Forward Network, FFN) 实现空间-通道全局建模。该过程可表示为:

$$\begin{aligned} Z &= LayerNorm(V + MHA(V)) \in \mathbf{R}^{B \times (H \times W) \times C} \\ Z' &= LayerNorm(Z + FFN(Z)) \end{aligned} \quad (4)$$

其中: $LayerNorm$ 是层归一化函数, 对特征维度进行归一化; $MHA(\cdot)$ 通过缩放点积注意力计算空间-通道交互权重, 用于多头拼接与投影; $FFN(\cdot)$ 用于对每个 token 独立进行非线性变换。

(4) 注意力生成

将 Transformer 输出 $Z' \in \mathbf{R}^{B \times (H \times W) \times C}$ 通过全连接层 (FC) 映射到原始空间尺寸, 通过 Sigmoid 激活函数生成归一化的空间-通道注意力权重向量 $ms_att \in \mathbf{R}^{B \times C \times H \times W}$, 最终与输入特征图逐通道相乘, 构成完整的增强型多光谱通道注意力机制。

$$ms_att = sigmoid(fc[Z']), \quad (5)$$

其中: ms_att 用于对特征图进行空间-通道加权; $sigmoid(\cdot)$ 表示 Sigmoid 激活函数; $fc[\cdot]$ 表示全连接层。

MF-ResBlock 模块融合了基于 Transformer 的多光谱通道注意力机制,通过 DCT 频谱分解和空间-通道全局建模,协同提取局部细节与全局上下文特征,显著提升网络表达能力与模型收敛效率。

2.2 仿真数据集生成

本研究采用样本绘制方法制作仿真数据集,制作流程如图 4 所示,具体步骤如下:

(1) 选取具有代表性的微纳结构样本,包括 USAF 1951 分辨率板、微纳器件和 MEMS 芯片等,通过仿真建模生成 128×128 像素的相位图,相位值随机分布在 $1.3 \sim 2.3$ rad 内。为提升数据真实性,对这些相位图进行随机变换操作,包括翻转、剪切等。

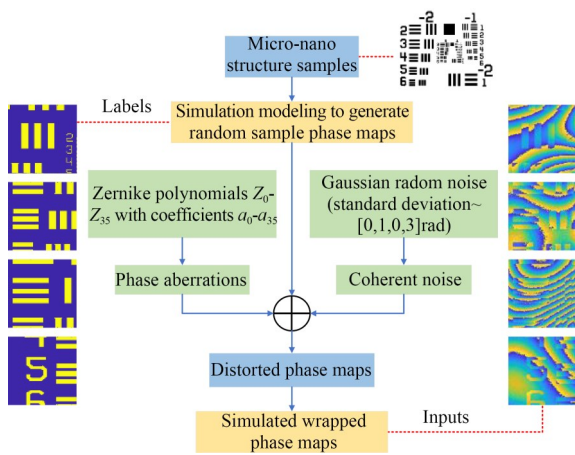


图 4 仿真数据集生成流程

Fig. 4 Flowchart of simulation dataset generation process

(2) 采用 1~36 项 Zernike 多项式 ($Z_0 \sim Z_{35}$) 构建相位像差模型,覆盖至 7 阶像差模式。随机生成各阶 Zernike 系数 ($a_0 \sim a_{35}$) 实现多样化的像差分布,有效模拟实际测量中的复杂光学像差。

(3) 为模拟实际测量中的相干噪声,在样本相位中添加高斯随机噪声,其均值 $\mu=0$,标准差 $\sigma \in [0.1, 0.3]$ rad,并与像差叠加生成失真相位图。

(4) 将含噪声的失真相位图 Φ 压缩至 $[-\pi, \pi]$ 区间 ($\Phi_{\text{wrapped}} = \text{angle}[\exp(j\Phi)]$),生成由包裹相位图(输入)和样本相位图(标签)组成的仿真数据集。

该仿真数据集完整保留了 Zernike 像差和高斯噪声的复合畸变特征,为相位恢复提供了高保真训练样本。

2.3 网络实现细节

这里采用基于 Python 3.10.13 的 PyTorch version: 2.2.2+cu121 框架实现了 MScU-Net 网络架构,在配备 NVIDIA GeForce RTX 3090i GPU 和 Intel Xeon Gold 5118 CPU (2.3 GHz, 382 GB RAM) 的工作站上进行训练与测试。网络训练采用 128×128 像素图像 (35 500 对样本, batch size=64),通过基于自适应动量估计 (Adaptive Moment Estimation, ADAM) 的优化器 (学习率设置为 10^{-4}) 优化均方误差 (Mean Squared Error, MSE) 损失函数,完成 100 个训练周期耗时约 12 h。测试表明,网络对单帧 128×128 像素图像的相位恢复时间约为 8 ms。

图 5 对比了 MScU-Net 与 ResU-Net 的训练和验证损失曲线。如图 5(a) 所示,两种网络均表现出稳定的收敛趋势,但 MScU-Net 在优化过程中展现出更高性能。在训练初期,MScU-Net 的损失下降速度显著快于 ResU-Net,表明它具有更高效的收敛特性。最终,MScU-Net 的训练损失稳定在约 0.0060 rad^2 ,最低达到 0.0058 rad^2 ,而 ResU-Net 的最低损失为 0.0123 rad^2 ,始终高于前者。结果表明,MScU-Net 在收敛速度和最终优化精度上均优于 ResU-Net。

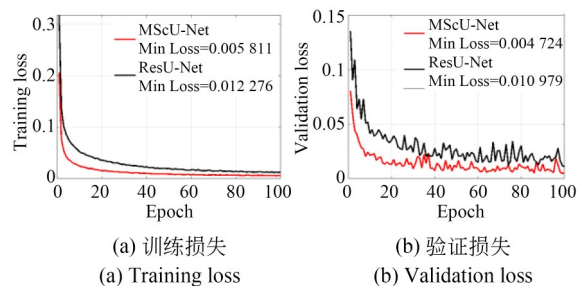


图 5 MScU-Net 与 ResU-Net 训练和验证损失曲线

Fig. 5 Training and validation loss curves of MScU-Net and ResU-Net

图 5(b) 表明,相较于训练损失曲线,验证损失呈现出更显著的波动特征。MScU-Net 在相位恢复任务中展现出更优异的泛化性能,其验证损失从初始值 0.0805 rad^2 稳定收敛至 0.0047 rad^2 ,明显小于 ResU-Net (最终收敛值 0.0110 rad^2)。这一对比结果充分证明 MScU-Net 对含噪声和畸变的数据具有更强的稳定表征能力,能够实现更精确的相位恢复。

2.4 消融实验

为评估 MScU-Net 网络架构中各模块的贡献,设计了系统性消融实验方案。在保持仿真数据集和训练超参数一致的前提下,分别移除 Transformer 与 MF-ResBlock 关键模块,以评估其对网络性能的影响。实验设置 4 个对比组: Baseline 为原始 ResU-Net 架构; MScU-Net without Transformer 为移除 Transformer 模块的 MScU-Net 架构; MScU-Net without MF-ResBlock 为移除多频残差模块的 MScU-Net 架构; MScU-

Net 为完整架构。定量评估指标中, PSNR 用于评估相位重建质量, SSIM 反映结构保持能力, RMSE 用于定量分析相位误差。所有对比组仅存在单一结构差异, 确保实验结果的可比性和模块贡献度分析的准确性。消融实验的定量评估结果见表 2。结果表明, 各结构模块均对 MScU-Net 的整体性能产生显著影响, 完整 MScU-Net 模型在 PSNR (29.82 dB)、SSIM (0.960 2) 和 RMSE (0.066 2 rad) 三项指标上均达到最优。

表 2 仿真数据集上的消融实验结果

Tab. 2 Ablation experiment results on simulated dataset

模型配置	PSNR/dB	SSIM	RMSE/rad	训练时间/h
Baseline	26.04	0.953 1	0.073 2	8.032
MScU-Net without Transformer	28.62	0.956 3	0.082 8	9.185
MScU-Net without MF-ResBlock	23.19	0.952 6	0.121 4	7.458
MScU-Net	29.82	0.960 2	0.066 2	11.663

由实验结果可知, MF-ResBlock 模块对性能影响最为关键, 其移除导致 PSNR 下降 22.3% (6.63 dB) 且 RMSE 增加 83.3%; Transformer 模块的移除则使 RMSE 增加 25.1%。这验证了 MF-ResBlock 模块通过多频残差机制增强局部相位细节的建模能力, 而 Transformer 模块则通过空间-通道协同注意力机制提升全局建模效果。MScU-Net 模型在保持 SSIM > 0.96 高结构保真度的同时, 将 RMSE 降低 9.6%, 虽增加 45.2% 的训练成本, 却显著提升了相位恢复的整体性能。

3 实验结果及分析

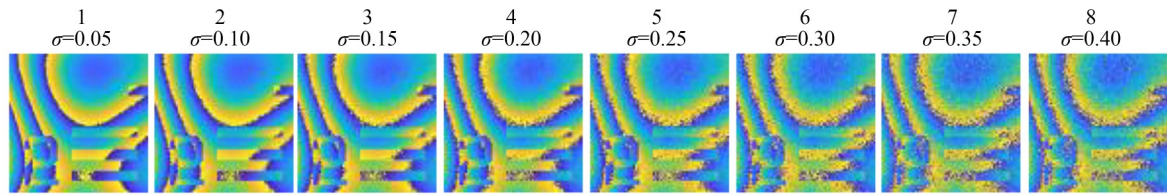
3.1 仿真结果

为定量评估所提方法的抗噪性能, 采用 USAF 1951 分辨率板包裹相位图作为基准相位图, 如图 6(b) 所示。在基准相位中引入相同成分的系统像差, 随后叠加递增的高斯噪声 ($\sigma = 0.05 \sim 0.40$ rad, $\Delta\sigma = 0.05$ rad), 对应的噪声等级编号为 1~8, 如图 6(a) 所示。这种复合退化模型既包含确定性像差又涵盖随机噪声, 可模拟实际光学系统中遇到的混合型相位失真。

实验采用 3 种方法进行系统性对比分析: (1)

ResU-Net 网络^[10]; (2) DeepLabV3+ 网络^[31]; (3) 分步式相位恢复方法 (基于可靠度掩膜的加权最小二相位解包裹算法^[6] + 基于曲线拟合预处理和背景提取的相位畸变补偿算法^[22], 简称为 RM-PU-PEAC)。所有深度学习方法均在完全相同的条件 (数据集、超参数和硬件环境) 下训练。4 种方法的相位恢复结果如图 6(c1)~(c4) 所示。

由图 6 可以看出, MscU-Net 在低中噪声水平 ($\sigma = 0.05 \sim 0.30$ rad) 下能完整恢复 USAF 1951 相位结构, 细节分辨率与基准相位基本一致。当噪声增强至 $\sigma \geq 0.35$ rad, 总体相位恢复效果尚可, 但结构边缘出现轻微模糊与畸变, 背景相位也出现局部异常, 如图 6(c1) 所示。ResU-Net 在低噪声 ($\sigma \leq 0.15$ rad) 时相位恢复效果良好, 但随着噪声增强, 相位异常区域显著扩展, $\sigma \geq 0.30$ rad 时出现全局性失真, 如图 6(c2) 所示。DeepLabV3+ 即使在低噪声下已显现结构边缘模糊和非物理性伪影, 如图 6(c3) 所示。其相位恢复结果与基准相位存在显著差异, 噪声适应性较差。分步式相位恢复方法 (RM-PU-PEAC) 在 $\sigma \leq 0.20$ rad 时能维持主体结构轮廓, 但背景噪声明显; 当噪声增强至 $\sigma \geq 0.25$ rad, 相位图被噪声严重污染, 样本结构几乎被完全破坏, 如图 6(c4) 所示。这是由于噪声过大导致可靠性掩膜失效, 解



(a) 噪声递增的失真包裹相位图

(a) Distorted wrapped phase maps with increasing noise levels



(b) 基准包裹相位图

(b) Ground-truth wrapped phase maps



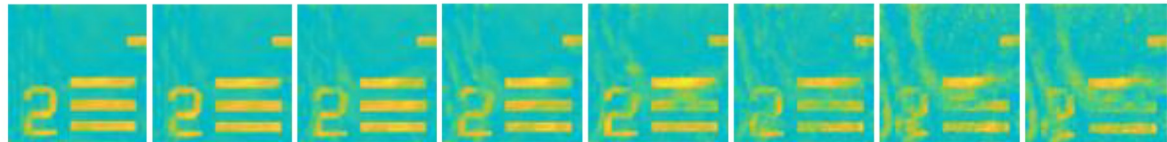
(c1) MScU-Net相位恢复结果

(c1) Chase retrieval results of MScU-Net



(c2) ResU-Net相位恢复结果

(c2) Phase retrieval results of ResU-Net



(c3) DeepLabV3+相位恢复结果

(c3) Phase retrieval results of DeepLabV3+



(c4) RMPU-PEAC相位恢复结果

(c4) Phase retrieval results of RMPU-PEAC

图 6 抗噪性能测试

Fig. 6 Noise robustness evaluation

包裹误差通过最小二乘优化过程全局传播。该现象揭示了传统解包裹算法在自适应噪声抑制方面的固有局限。

为了系统评估 4 种方法的相位恢复精度,采用 PSNR, SSIM 和 RMSE 3 项指标对相位恢复结果进行定量分析,评估结果如图 7 所示。结果

表明:MScU-Net 在抗噪性能方面展现出显著优势,其 RMSE 稳定在 $0.047\ 2 \sim 0.085\ 1$ rad (波动为 ± 0.019 rad), PSNR ($21.44 \sim 26.51$ dB) 和 SSIM ($0.86 \sim 0.93$) 均保持最高水平。

相比之下,分步式相位恢复方法的噪声稳定性最差,其性能指标随噪声增强而系统性劣化:

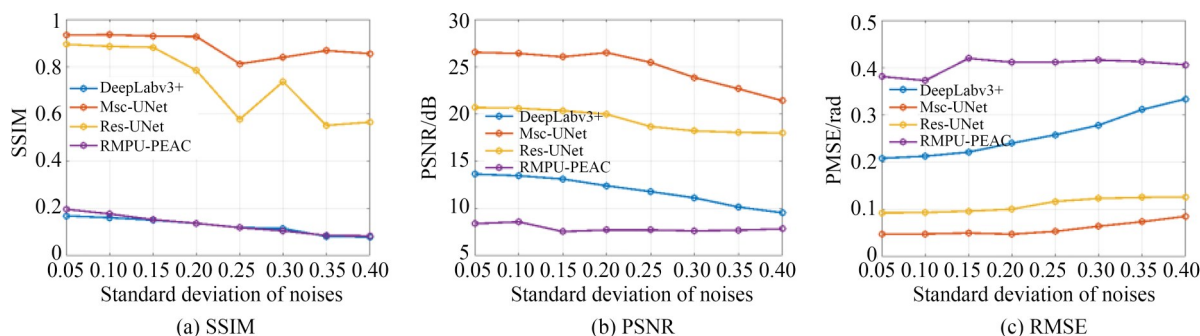


图 7 抗噪性能定量评估曲线

Fig. 7 Quantitative evaluation curves of noise robustness

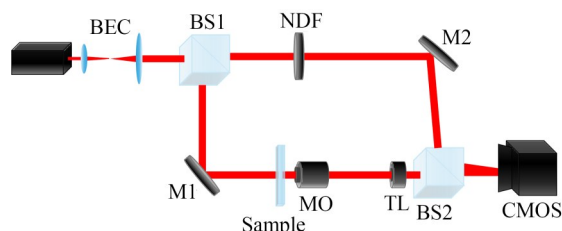
RMSE 上升至 0.419 4 rad, 同时 PSNR 和 SSIM 分别被限制在 8.576 3 dB 和 0.196 3 以下。虽然 ResU-Net 和 DeepLabV3+ 的性能优于分步式相位恢复方法,但在高噪声下仍存在明显的性能衰减。综合三项指标,MScU-Net 展现出最优的相位恢复精度和噪声稳定性。

3.2 实验结果

3.2.1 透射式数字全息显微实验

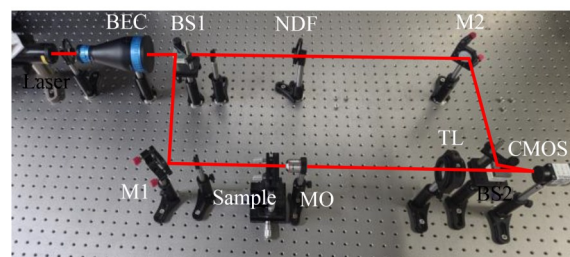
为了测试所提出的 MScU-Net 模型在复杂真实场景中的相位恢复能力和稳定性,搭建了透射式离轴数字全息显微系统,如图 8 所示。采用波长为 594.1 nm 激光器作为照明光源,激光束经过准直扩束器和分光镜(BS1)后分为物光束和参考光束。物光路经 10X/0.25 NA 显微物镜与 180 mm 套管透镜组成远心显微模块,将样品清晰成像于 CMOS 靶面(5 120×5 120 像素,2.5 μm 像元)。参考光束通过中性密度滤光片衰减光强后,与物光束在 CMOS 靶面上倾斜干涉形成离轴全息图。实验样本为通过激光无掩膜直写技术在石英玻璃(折射率 $n=1.46$)表面直接烧蚀制备的 USAF 1951 分辨率板。

对实验采集到的 USAF 1951 分辨率板全息图进行数值再现和相位恢复,图 9(a)为 USAF 1951 分辨率板的包裹相位图(128×128 像素)。图 9(b1)~(b4)分别为 MScU-Net, ResU-Net, DeepLabV3+ 和 RMPU-PEAC 4 种方法分别对图 9(a)进行相位恢复后得到的 USAF 1951 分辨率板相位分布。相位恢复结果表明,MScU-Net 恢复的 USAF 1951 分辨率板相位分布兼具结构保真度与平滑性,如图 9(b1)所示。ResU-



(a) 光路示意图

(a) Optical path schematic



(b) 实验装置

(b) Experimental setup

BEC: beam expander and collimator;
NDF: neutral density filter;
MO: microscope objective (Olympus, 10 $\times$, NA=0.25);
TL: tube lens ($f=180$ mm);
M1/M2: mirrors;
BS1/BS2: beam splitters

图 8 透射式离轴数字全息显微系统

Fig. 8 Transmission off-axis digital holographic microscopy system

Net, DeepLabV3+ 和 RMPU-PEAC 的相位恢复结果均呈现出结构表面粗糙度的异常增加,和背景区域的非均匀相位波动,如图 9(b2)~(b4)所示。在 MScU-Net 结构上表面的相位标准差($\sigma=0.018$ 0 rad)较 ResU-Net, DeepLabV3+ 和 RMPU-PEAC 分别降低 53.7%, 75.5% 和 79.4%。特别地,其在背景区域 σ 仅为 0.000 4

rad, 相较 3 种对比方法降幅达 82.6%~99.3%, 验证了 MScU-Net 在实际噪声抑制方面的优势。

图 9(c)展示了 MScU-Net 的相位分布(图 9(b1))经像素标定和透射式相位高度换算($h = \varphi_0 \lambda / 2\pi \Delta n$, 波长 $\lambda = 594.1 \text{ nm}$, 折射率之差 $\Delta n = 0.46$)后, 得到的分辨率板三维轮廓。以白光干涉仪(WLI, Zygo, NewView 9000)的测量结果为基准, 对 4 种方法的相位保真度和测量精度进行系统性评估。图 9(d1)和 9(d2)分别为沿着图 9(a)

中水平和垂直方向红色虚线提取四种方法的分辨率板轮廓曲线。可以看出, MScU-Net 展现出最优的相位保真度, 恢复的轮廓曲线在背景区域与基准线高度重合, 在结构区域存在纳米级表面扰动和局部形变。相比之下, 3 种对比方法则表现出不同程度的轮廓偏离和结构失真。在分辨率板台阶高度测量中, MScU-Net 的测量值为 $(284.88 \pm 1.98) \text{ nm}$, 与 WLI 测量值 $(0.284 \text{ } \mu\text{m})$ 的相对偏差为 +0.31%, 验证了该方法在微纳尺度形貌测量中的准确度。

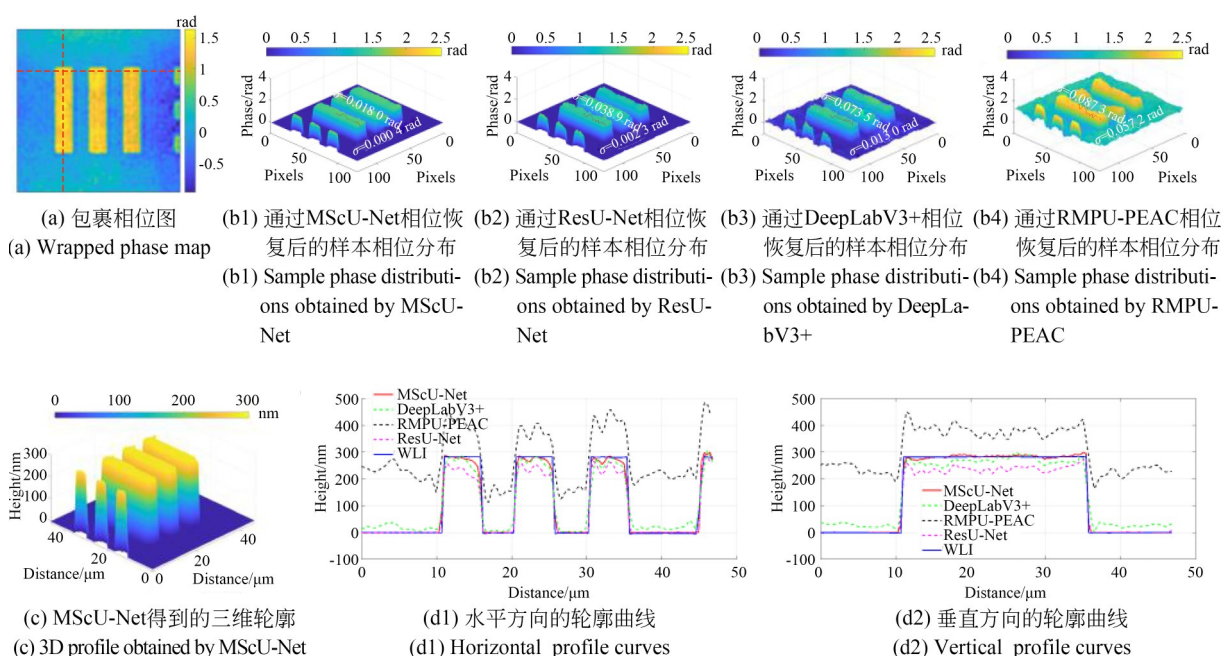


图 9 USAF 1951 分辨率板的实验结果

Fig. 9 Experimental results of USAF 1951 resolution target

3.2.2 反射式数字全息显微实验

实验搭建了一个反射式离轴数字全息显微系统(无套管透镜校正二次像差, 且非像面成像), 对两种典型结构的微纳标准器(折线型与三角型)进行了定量表征。图 10(a)为折线型微纳标准器的包裹相位分布。该相位图包含离轴数字全息显微系统典型的像差特征: 离轴干涉引入的一次像差表现为线性载频条纹, 而未经套管透镜校正的二次像差则呈现为同心环状条纹。

图 10(b1)~(b4)分别为 MScU-Net, ResU-Net, DeepLabV3+ 和 RMPU-PEAC 4 种方法分

别对图 10(a)进行相位恢复后得到的折线型微纳标准器相位分布。如图所示, 在 MScU-Net 结构上表面的相位标准差($\sigma = 0.0257 \text{ rad}$)较 ResU-Net, DeepLabV3+ 和 RMPU-PEAC 分别降低 53.7%, 69.1% 和 93.0%。MScU-Net 在背景区域 σ 为 0.0011 rad, 相较 3 种对比方法降幅达 85.3%~99.0%, 再次验证 MScU-Net 在实际噪声抑制方面的显著优势。

图 10(c)展示了 MScU-Net 的相位分布(图 10(b1))经像素标定和反射式相位高度换算($h = \varphi_0 \lambda / 4\pi$, $\lambda = 632.8 \text{ nm}$)得到的折线型微纳标准器三维轮廓图。图 10(d1)和 10(d2)分别为沿着图

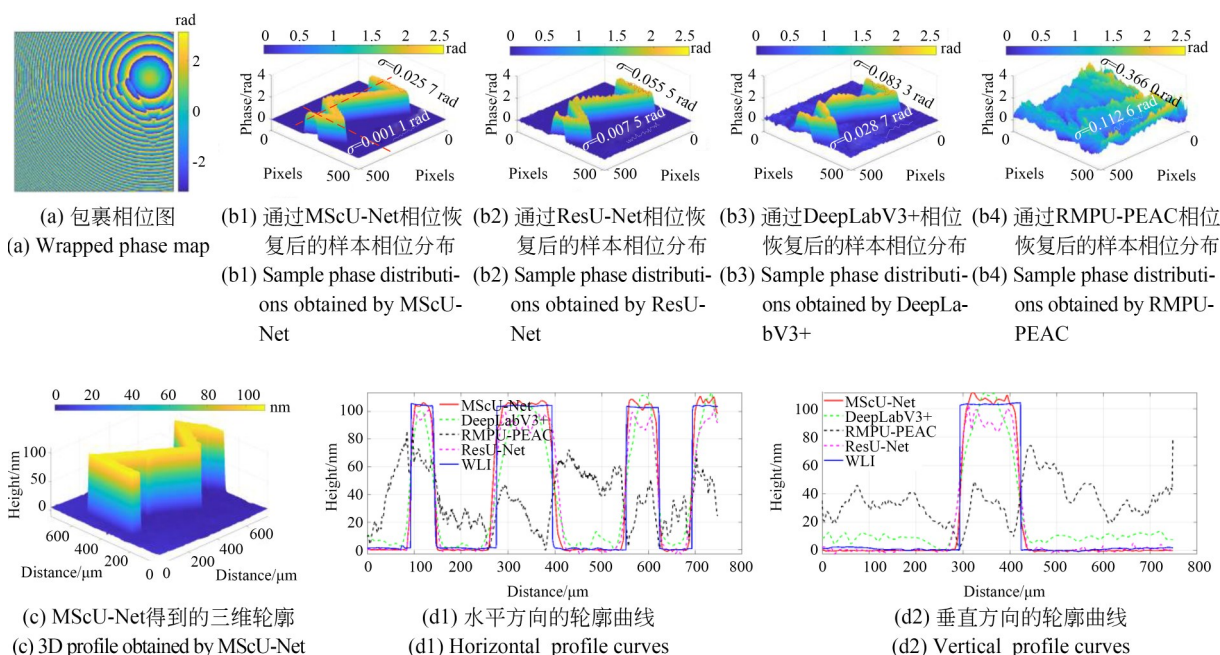


图 10 折线型微纳标准器的实验结果

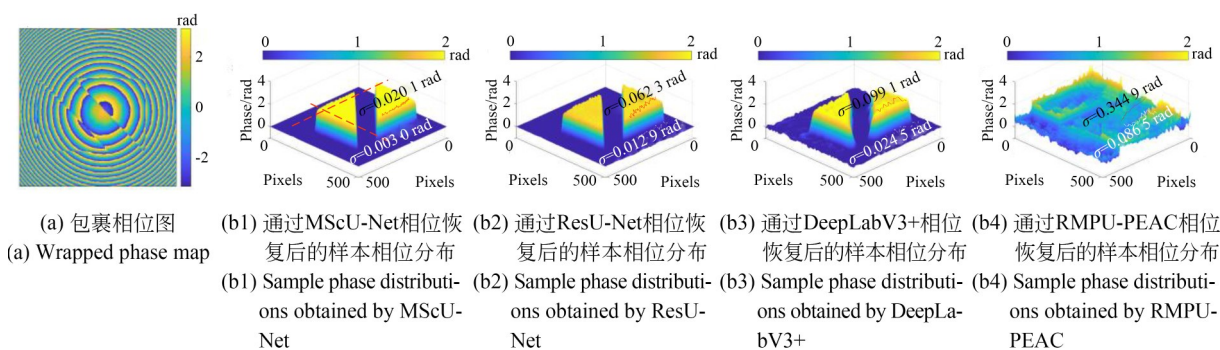
Fig. 10 Experimental results of polyline-type micro-nano standard artifact

10(b1)中水平和垂直方向红色虚线提取四种方法的轮廓曲线。在折线型微纳标准器台阶高度测量中, MScU-Net 测量值为 (105.81 ± 0.44) nm, 与 WLI 测量值 $(0.102 \mu\text{m})$ 的相对偏差为 $+3.74\%$ 。

图 11(a)为三角型微纳标准器的包裹相位图,由其同心环状条纹可知主要像差是由物镜引入的二次像差。图 11(b1)~(b4)分别为 MScU-Net, ResU-Net, DeepLabV3+ 和 RMPU-PEAC, 对图 11(a)相位恢复后得到的三角型微纳标准器相位分布。如图所示,在 MScU-Net 结构上表面

的相位标准差 ($\sigma=0.020 \text{ rad}$) 较 ResU-Net, DeepLabV3+ 和 RMPU-PEAC 分别降低 67.7% , 79.7% 和 94.2% 。MScU-Net 在背景区域 σ 为 0.003 rad , 相较 3 种对比方法降幅达 $76.7\% \sim 96.5\%$ 。

图 11(c)展示了三角型微纳标准器三维轮廓图。图 11(d1)和 11(d2)分别为沿着图 11(b1)中水平和垂直方向红色虚线提取 4 种方法的三角型微纳标准器轮廓曲线。MScU-Net 的台阶高度测量值为 $(102.34 \pm 0.42) \text{ nm}$, 与 WLI 测量值 $(0.104 \mu\text{m})$ 的相对偏差分别为 -1.60% 。



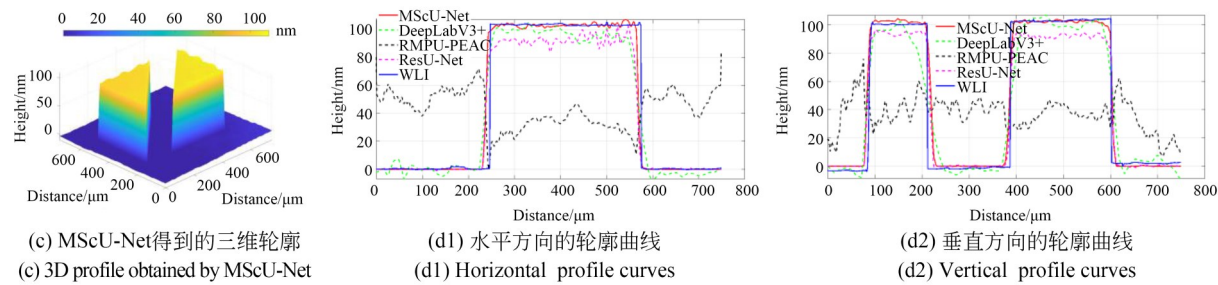


图 11 三角型微纳标准器的实验结果

Fig. 11 Experimental results of triangular micro-nano standard artifact

表 3 统计了 4 种方法对 3 种微纳器件包裹相位图进行相位恢复后的相位标准差,表 4 则系统展示了 MScU-Net 方法对 3 种微纳器件的台阶高度测量数据,包括与白光干涉仪(WLI)基准值的对比分析。定量数据表明,MScU-Net 在 3 种微纳器件上实现了最低相位噪声水平,背景区域标准差(0.000 4~0.003 0 rad)比最优对比方法

(ResU-Net)降低 82.6%~95.8%,样品区域(0.018 0~0.025 7 rad)降低 33.8%~53.7%,均优于现有方法。

表 4 表明,MScU-Net 台阶高度测量与 WLI 基准的系统偏差 $\leq 3.74\%$,测量结果的局部波动性达到亚纳米量级(空间分布标准差 $s = 0.42$ nm,折线型微纳标准器)。

表 3 四种方法对三种微纳器件相位恢复的噪声水平对比

Tab. 3 Comparison of noise levels in phase retrieval for three micro-nano devices by four methods

Method	Phase standard deviation σ/rad					
	Background regions			Sample regions		
	USAF 1951	Polyline	Triangular	USAF 1951	Polyline	Triangular
MScU-Net	0.000 4	0.001 1	0.003 0	0.018 0	0.025 7	0.020 1
ResU-Net	0.002 3	0.007 5	0.012 9	0.038 9	0.055 5	0.062 3
DeepLabV3+	0.013 0	0.028 7	0.024 5	0.073 5	0.083 3	0.099 1
RMPU-PEAC	0.057 2	0.112 6	0.086 5	0.087 3	0.366 0	0.344 9

表 4 MScU-Net 和白光干涉仪对三种微纳器件的测量结果对比

Tab. 4 Comparative of measurement results for three micro-nano devices between MScU-Net and WLI

Measured sample	Step height measurement data		Relative deviations/%
	WLI/ μm	MScU-Net/nm	
USAF 1951 resolution target	0.284	284.88 ± 1.98	+0.31
Polyline-type micro-nano standard artifact	0.102	105.81 ± 0.44	+3.74
Triangular micro-nano standard artifact	0.104	102.34 ± 0.42	-1.60

4 结 论

本文提出了一种基于多光谱通道注意力机制与 Transformer 融合的深度神经网络 MScU-Net,用于数字全息显微术中相位解包裹与像差补偿的同步求解。通过在 ResU-Net 框架中引

入多光谱通道注意力层,利用 DCT 提取全局频谱信息并构建空间-通道联合注意力机制,增强高频噪声抑制与细节保留能力,有效解决了传统分步处理中的误差传递问题。实验结果表明,MScU-Net 在 PSNR (29.82 dB) 和 SSIM (0.960 2) 等关键指标上均达到最优性能,实际

背景区域噪声较最优对比方法降低 82.6%~95.8%,实测台阶高度系统偏差 $\leq 3.74\%$,测量局部波动达到亚纳米级($s=0.42\text{ nm}$)。虽然 MScU-Net 训练成本增加 45.2%,计算效率存在优化空间,但显著的性能提升使其在工业检测和科学研究中具有重要应用价值,适用于半导体检测、MEMS 芯片测量和精密光学等微纳计量场景。

参考文献:

- [1] 黄琳琳,刘丙才,王红军,等. 共路剪切数字全息自参考聚焦显微成像[J]. 光学精密工程, 2024, 32(23): 3416-3423.
HUANG L L, LIU B C, WANG H J, *et al.* Self-reference focusing imaging for common-path shearing digital holographic microscopy [J]. *Optics and Precision Engineering*, 2024, 32(23): 3416-3423. (in Chinese)
- [2] 康乐谦,杨鹏程,吕秋娟,等. Tsallis 相对熵引导的包裹相位修补[J]. 光学精密工程, 2024, 32(8): 1130-1139.
KANG L Q, YANG P C H, LÜ Q J, *et al.* Wrapping phase repair method based on Tsallis Relative Entropy evaluation [J]. *Optics and Precision Engineering*, 2024, 32(8): 1130-1139. (in Chinese)
- [3] GOLDSTEIN R M, ZEBKER H A, WERNER C L. Satellite radar interferometry: Two-dimensional phase unwrapping [J]. *Radio Science*, 1988, 23(4): 713-720.
- [4] AREVALILLO-HERRAEZ M, VILLATORO F R, GDEISAT M A. A robust and simple measure for quality-guided 2D phase unwrapping algorithms [J]. *IEEE Transactions on Image Processing*, 2016, 25(6): 2601-2609.
- [5] 孙煜雅. 光栅投影法的包裹相位展开方法研究[D]. 太原: 中北大学, 2022.
SUN Y Y. *Research on Wrapping Phase Unwrapping Method of Grating Projection Method* [D]. Taiyuan: North University of China, 2022. (in Chinese)
- [6] YAN L P, ZHANG H Y, ZHANG R, *et al.* A robust phase unwrapping algorithm based on reliability mask and weighted minimum least-squares method [J]. *Optics and Lasers in Engineering*, 2019, 112: 39-45.
- [7] LI S J, ZHANG S B, GAO Y D, *et al.* Time series phase unwrapping algorithm using L P-norm optimization compressive sensing [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2023, 117: 103182.
- [8] ZHOU H Y, CHENG C L, PENG H, *et al.* The PHU-NET: a robust phase unwrapping method for MRI based on deep learning [J]. *Magnetic Resonance in Medicine*, 2021, 86(6): 3321-3333.
- [9] ZHOU L F, YU H W, PASCAZIO V, *et al.* PU-GAN: a one-step 2-D InSAR phase unwrapping based on conditional generative adversarial network [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5221510.
- [10] WANG K Q, LI Y, QIAN K M, *et al.* One-step robust deep learning phase unwrapping [J]. *Optics Express*, 2019, 27(10): 15100-15115.
- [11] SUN G W, LI B Y, LI Z, *et al.* Phase unwrapping based on channel transformer U-Net for single-shot fringe projection profilometry [J]. *Journal of Optics*, 2024, 53(4): 3331-3341.
- [12] 张家骏, 康敬, 刘冀钊, 等. 利用图像平滑结构信息指导图像修复[J]. 光学精密工程, 2024, 32(4): 549-564.
ZHANG J J, LIAN J, LIU J Z H, *et al.* Using image smoothing structure information to guide image inpainting [J]. *Opt. Precision Eng.*, 2024, 32(4): 549-564. (in Chinese)
- [13] LUO X L, SONG W Z, BAI S L, *et al.* Deep learning-enabled invalid-point removal for spatial phase unwrapping of 3D measurement [J]. *Optics & Laser Technology*, 2023, 163: 109340.
- [14] JIANG X, WANG J, FANG Y Q, *et al.* Spatial phase unwrapping approach for single-frame 3D shape measurement based on deep learning [J]. *Applied Optics*, 2024, 63(20): 5269.
- [15] LI D X, XIE X M. Deep learning-based phase unwrapping method [J]. *IEEE Access*, 2023, 11: 85836-85851.
- [16] LATTARI F, CANNICI M, MURDACA G, *et al.* NeuralPUMA: learning to phase unwrap

作者贡献声明:

严利平: 测量方法, 论文构思, 论文审核与编辑写作, 提供资源, 项目管理;

苗小喆: 实验设计与验证, 实验结果整理和分析, 软件, 初稿写作;

黄柳: 实验结果分析, 获取资助, 论文审核与编辑写作。

- through differentiable graph cuts[J]. *IEEE Transactions on Signal Processing*, 2023, 72: 321-336.
- [17] ZHANG Z H, WANG X X, LIU C X, *et al.* Efficient and robust phase unwrapping method based on SFNet [J]. *Optics Express*, 2024, 32 (9) : 15410-15432.
- [18] DOBLAS A, SÁNCHEZ-ORTIGA E, MARTÍNEZ-CORRAL M, *et al.* Accurate single-shot quantitative phase imaging of biological specimens with telecentric digital holographic microscopy [J]. *Journal of Biomedical Optics*, 2014, 19(4): 046022.
- [19] DENG D N, PENG J Z, QU W J, *et al.* Simple and flexible phase compensation for digital holographic microscopy with electrically tunable lens [J]. *Applied Optics*, 2017, 56(21): 6007.
- [20] COLOMB T, KÜHN J, CHARRIÈRE F, *et al.* Total aberrations compensation in digital holographic microscopy with a reference conjugated hologram [J]. *Optics Express*, 2006, 14(10): 4300-4306.
- [21] CHOI I, LEE K, PARK Y. Compensation of aberration in quantitative phase imaging using lateral shifting and spiral phase integration[J]. *Optics Express*, 2017, 25(24): 30771-30779.
- [22] HUANG L, YAN L P, CHEN B Y, *et al.* Phase aberration compensation of digital holographic microscopy with curve fitting preprocessing and automatic background segmentation for microstructure testing [J]. *Optics Communications*, 2020, 462: 125311.
- [23] LAI X M, XIAO S, XU C, *et al.* Aberration-free digital holographic phase imaging using the derivative-based principal component analysis [J]. *Journal of Biomedical Optics*, 2021, 26(4): 046501.
- [24] HUANG Z Z, CAO L C. Phase aberration separation for holographic microscopy by alternating direction sparse optimization[J]. *Optics Express*, 2023, 31(8): 12520-12533.
- [25] NGUYEN T, BUI V, LAM V, *et al.* Automatic phase aberration compensation for digital holographic microscopy based on deep learning background detection [J]. *Optics Express*, 2017, 25 (13): 15043-15057.
- [26] MA S J, FANG R, LUO Y, *et al.* Phase-aberration compensation *via* deep learning in digital holographic microscopy [J]. *Measurement Science and Technology*, 2021, 32(10): 105203.
- [27] MA S J, LIU Q, YU Y T, *et al.* Quantitative phase imaging in digital holographic microscopy based on image inpainting using a two-stage generative adversarial network [J]. *Optics Express*, 2021, 29(16): 24928-24946.
- [28] HUANG L, TANG J J, YAN L P, *et al.* Wrapped phase aberration compensation using deep learning in digital holographic microscopy [J]. *Applied Physics Letters*, 2023, 123(14): 141109.
- [29] LI D J, LI Z X, DING W X, *et al.* Simultaneous phase aberration compensation and denoising for quantitative phase imaging in digital holographic microscopy with deep learning [J]. *Applied Optics*, 2024, 63(26): 6931.
- [30] QIN Z Q, ZHANG P Y, WU F, *et al.* FcaNet: frequency channel attention networks [C]. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 10-17, 2021, Montreal, QC, Canada. IEEE, 2021: 763-772.
- [31] ZHANG T, JIANG S W, ZHAO Z X, *et al.* Rapid and robust two-dimensional phase unwrapping *via* deep learning [J]. *Optics Express*, 2019, 27 (16): 23173-23185.

作者简介:



严利平(1977—),女,四川内江人,教授,博士生导师,2000、2003年于重庆大学分别获得学士和硕士学位,2014年于浙江理工大学获得博士学位,主要从事数字全息微纳结构测量、激光干涉精密位移测量和激光跟踪大尺寸空间坐标测量方面的研究。E-mail: yanliping@zstu.edu.cn

通讯作者:



黄柳(1987—),女,湖北汉川人,博士,讲师,2009年、2012年于中国计量大学分别获得学士和硕士学位,2021年于浙江理工大学获得博士学位,主要从事数字全息测量、微纳器件与芯片测试方面的研究。E-mail: huangliu@zstu.edu.cn