

# 基于卷积神经网络的光学本征模式成像

周义深 李琪 王翔 宋嘉康 贾勘序 张轩阁 陈希浩<sup>†</sup>

辽宁大学物理学院, 沈阳 110036

## 摘要

光学本征模式成像 (Optical eigenmode imaging, OEI) 是一种基于光学系统本征模式展开与投影的计算成像方法, 通过在模式空间中对目标空间分布信息进行表征与重建, 实现目标图像的获取。然而在低采样率条件下, 传统 OEI 重建受限于模态不完备性, 导致图像细节丢失与分辨能力受限。针对此问题, 本文将深度学习引入 OEI 重建过程, 构建了一种融合物理模型约束的复数卷积神经网络框架。该方法将 OEI 物理模型嵌入非预训练复数 U-net 卷积神经网络中, 并引入注意力机制以增强多尺度特征提取能力, 从而实现了对图像细节的有效恢复。数值模拟结果表明, 该方法在低采样率下仍能有效恢复图像细节, 在峰值信噪比与结构相似性等指标上显著优于传统算法, 验证了将 OEI 物理模型与深度学习相融合的可行性, 并为相关计算成像优化提供新思路。

**关键词:** 光学本征模式成像; 图像重建; 卷积神经网络; 注意力机制

**PACS:** 42.30.Va, 42.30.Wb, 89.20.Ff, 42.30.-d

<sup>†</sup> 通讯作者: xi-haochen@163.com

## 1 引言

光学成像作为科学研究和工程应用中的核心技术之一<sup>[1]</sup>, 其分辨率长期以来受到衍射极限的制约<sup>[2]</sup>。如何在保证系统稳定性和低成本的前提下突破这一限制, 进而实现更高质量的成像, 是近年来计算光学领域的重要研究方向<sup>[3]</sup>。光学本征模式成像 (Optical eigenmode imaging, OEI) 作为一种新兴的间接成像方法<sup>[4]</sup>, 能够通过模态分解实现振幅与相位的同步获取<sup>[5]</sup>。由于目标信息可由少数主要光学本征模式高效表征, OEI 本身具有压缩采样的潜力<sup>[6]</sup>; 同时, 依托于本征模式的正交性和独立性, 它能够有效抑制像差影响, 保持图像质量与细节恢复<sup>[7]</sup>, 因此 OEI 技术逐渐受到学术界的广泛关注。其基本原理是将光场分解为一组正交的本征模式, 通过提取目标光场在本征模态下的投影系数, 从而实现复振幅的重建<sup>[8]</sup>。此外, OEI 在实验中已展现出突破瑞利判据的分辨率潜力<sup>[9]</sup>。凭借这一特性, OEI 在生物成像<sup>[10]</sup>、光谱探测<sup>[11]</sup> 以及远程感知<sup>[12]</sup> 等应用中显示出良好的发展前景。然而, OEI 也面临不可忽视的挑战。由于现有条件下可生成和测量的本征模式数量有限, 模态集合往往不完备, 从而导致高频细节难以有效恢复<sup>[13]</sup>。缺失模态所对应的信息丢失, 会引发系统点扩散函数 (Point spread function, PSF) 的旁瓣增强, 进而造成图像边缘模糊、细节丢失等问题<sup>[14]</sup>。这一问题在低采

样率条件下尤为显著，使得传统 OEI 方法在重建质量方面受到限制<sup>[15]</sup>。因此，如何在不显著增加测量复杂度的前提下提升 OEI 的重建质量，已成为亟待解决的关键问题。

近年来，深度学习技术的快速发展为计算成像领域带来了新的机遇<sup>[16]</sup>。在医学领域，深度神经网络已被广泛应用于计算机断层摄影图像重建<sup>[17]</sup>、超声图像去噪<sup>[18]</sup>等任务，大大提高了成像速度和质量。在遥感成像中，深度学习被用于处理多维度数据进行目标检测<sup>[19]</sup>和图像分类<sup>[20]</sup>；高光谱成像中<sup>[21]</sup>，深度学习已广泛用于处理高维特性<sup>[22]</sup>、应对分类挑战<sup>[23]</sup>和求解图像重建逆问题<sup>[24]</sup>，通过自主学习特征，实现了对不完整物理测量的恢复和重建质量的显著提升<sup>[25]</sup>。此外，深度学习在单像素成像 (Single-pixel imaging, SPI) 领域的应用也取得了显著进展<sup>[26]</sup>。通过结合单像素探测器和多种调制模式<sup>[27]</sup>，SPI 能够实现图像重建，而深度学习的引入能够显著提高重建质量，尤其在低采样率<sup>[28]</sup>和高噪声环境下<sup>[29]</sup>表现尤为出色。这些研究表明，深度学习不仅能够加速传统图像重建过程，还能在复杂环境下实现噪声抑制，极大地拓展了成像技术的应用范围。尤其是物理驱动型卷积神经网络 (Convolutional neural network, CNN) 方法<sup>[30]</sup>，能够将成像模型直接嵌入网络框架<sup>[31]</sup>在无监督的条件下进行优化<sup>[32]</sup>，从而避免对大规模训练数据的依赖<sup>[33]</sup>。这一特点与 OEI 的成像过程天然契合：OEI 已知的模态投影物理模型可作为约束条件引入网络优化，使得深度学习能够充分发挥特征提取与非线性映射的优势，并保证重建结果与原始目标模型的高度一致性。

因此，本文提出了一种基于卷积神经网络的 OEI 重建方法。该方法以 OEI 的物理投影模型为基础，设计了一个融合通道注意力模块与空间注意力机制模块的多尺度 U-net 卷积网络。在网络训练过程中，我们随机初始化权重，并通过与 OEI 物理测量过程的迭代交互来优化网络参数。具体而言，OEI 的物理模型用于生成预测的模态投影系数，并与仿真设定的投影数据进行对比，二者之间的误差作为损失函数，指导网络参数的更新。这样，网络的优化过程完全由物理约束驱动，无需依赖额外的训练样本或先验数据集。

## 2 原理

### 2.1 理论基础

在 OEI 成像理论中，设在探测平面上，线性偏振光场的复振幅分布为  $E(\mathbf{r})$ ，其时间演化因子约定为  $e^{i\omega t}$ ，其中  $\omega$  为光波角频率， $\mathbf{r} = (x, y)$  为成像或探测平面上的二维空间坐标矢量。该任意空间相干光场  $E(\mathbf{r})$  均可展开在一组预先选定的测试场模态集合  $\{u_j(\mathbf{r})\}_{j=1}^N$  上

$$E(\mathbf{r}) = \sum_{j=1}^N a_j u_j(\mathbf{r}), \quad (1)$$

其中  $N$  为测试场模式总数， $a_j \in \mathbb{C}$  为各模态对应的复振幅展开系数。相应的复共轭形式为

$$E^*(\mathbf{r}) = \sum_{j=1}^N a_j^* u_j^*(\mathbf{r}). \quad (2)$$

一组测试场模态  $\{u_j(\mathbf{r})\}$  若要唯一表示任意光场  $E(\mathbf{r})$ ，必须满足线性无关且完备性条件，即它们构成光场所在空间的一个基底。测试场  $u_j(\mathbf{r})$  可为复值分布（由仿真预设），用于对光场的空间结构进行表征。在感兴趣区域 (Region of interest, ROI) 上，定义光场强度泛函

$$m^{(I)}(E) = \int_{\text{ROI}} E(\mathbf{r}) E^*(\mathbf{r}) d\sigma, \quad (3)$$

其中  $d\sigma$  为面积元，积分区域 ROI 通常对应于探测器的有效感光区域。将式 (1) 与式 (2) 代入式 (3)，引入独立求和指标  $k$  以区分共轭项并定义测试场之间的相关矩阵为  $M$ ，其矩阵元为

$$M_{jk} = \int_{\text{ROI}} u_j(\mathbf{r}) u_k^*(\mathbf{r}) d\sigma, \quad (4)$$

其中  $M_{jk}$  为测试场模式  $u_j(\mathbf{r})$  与  $u_k(\mathbf{r})$  在 ROI 中的内积，用于表征各测试场模态中的空间相关性。则有

$$m^{(I)}(E) = \sum_{j=1}^N \sum_{k=1}^N a_j^* M_{jk} a_k = \mathbf{a}^\dagger M \mathbf{a}, \quad (5)$$

其中  $\mathbf{a} = [a_1, a_2, \dots, a_N]^T$  为系数列向量， $\mathbf{a}^\dagger$  其厄米共轭。对矩阵  $M$  进行本征值分解，得到的特征值  $\lambda^\ell$  和特征向量  $v_{\ell j}$  满足

$$\sum_{j=1}^N M_{jk} v_{\ell j} = \lambda^\ell v_{\ell k}. \quad (6)$$

利用特征值和特征向量，光学本征模  $\mathbb{E}_\ell(\mathbf{r})$  和其共轭形式  $\mathbb{E}_\ell^*(\mathbf{r})$  可以定义为测试场模态  $u_j(\mathbf{r})$  的线性组合

$$\mathbb{E}_\ell(\mathbf{r}) = \frac{1}{\sqrt{\lambda^\ell}} \sum_{j=1}^N v_{\ell j}^* u_j(\mathbf{r}), \quad (7)$$

$$\mathbb{E}_\ell^*(\mathbf{r}) = \frac{1}{\sqrt{\lambda^\ell}} \sum_{j=1}^N v_{\ell j} u_j^*(\mathbf{r}). \quad (8)$$

其中，特征值  $\lambda^\ell$  的大小反映了该本征模在光场强度中的贡献程度，特征向量  $v_{\ell j}$  给出了测试场模式  $u_j(\mathbf{r})$  的权重。这些光学本征模在 ROI 内满足正交归一化条件

$$\int_{\text{ROI}} \mathbb{E}_j(\mathbf{r}) \mathbb{E}_k^*(\mathbf{r}) d\sigma = \delta_{jk}. \quad (9)$$

其中  $\delta_{jk}$  为克罗内克函数，当且仅当  $j = k$  时为 1，其余为 0。对于未知的目标场  $T(\mathbf{r})$ （即目标物的复透射函数），可以将其投影到这组正交归一的本征模基底上。目标场  $T(\mathbf{r})$  在第  $\ell$  个本征模上的复投影系数定义为

$$s_\ell^* = \int_{\text{ROI}} T(\mathbf{r}) \mathbb{E}_\ell^*(\mathbf{r}) d\sigma, \quad (10)$$

其代表了该本征模式在目标光场展开中的复振幅权重，其中 ROI 表示探测器的有效感光区域。进一步，利用该组投影系数，可通过光学本征模的线性叠加重建目标场

$$T(\mathbf{r}) = \sum_{\ell} s_\ell^* \mathbb{E}_\ell(\mathbf{r}), \quad (11)$$

在理想情况下若本征模集合完备，目标场可由式 (11) 精确重建<sup>[4]</sup>。

## 2.2 图像仿真重建

在仿真过程中，光学本征模系的构造是 OEI 图像重建的基础。本文选用二维 Laguerre-Gaussian(LG) 模式作为测试场集合  $\{u_j(\mathbf{r})\}$ 。第  $j$  个测试场  $u_j$  由特定的径向阶数  $p_j$  和角向阶数  $l_j$  唯一确定，其在极坐标系  $(r, \phi)$  下的分布为

$$\begin{aligned} \text{LG}_{p,l}(r, \phi) = & \sqrt{\frac{2p!}{\pi(p+|l|)!} \frac{1}{\omega_0}} \left( \frac{\sqrt{2}r}{\omega_0} \right)^{|l|} L_p^{|l|} \\ & \times \left( 2 \frac{r^2}{\omega_0^2} \right) \exp\left(-\frac{r^2}{\omega_0^2}\right) \exp(il\phi), \end{aligned} \quad (12)$$

其中， $(r, \phi)$  为极坐标下的径向和角向坐标， $\omega_0$  为光束腰参数， $L_p^{|l|}(\cdot)$  为广义拉盖尔多项式。在后续数值计算与图像重建中，将上述解析表达式离散化映射至直角坐标系  $(x, y)$  下，记为  $u_j(x, y)$ 。基于式 (4) 定义的内积关系，我们在 ROI 区域内对矩阵  $M$  进行特征分解，得到本征值  $\lambda^\ell$  和本征向量  $v_{\ell,j}$ ，从而构建光学本征模式  $\mathbb{E}_\ell(x, y)$ ，其表达式为

$$\mathbb{E}_\ell(x, y) = \frac{1}{\sqrt{\lambda^\ell}} \sum_{j=1}^N v_{\ell,j} \cdot u_j(x, y). \quad (13)$$

为实现图像重建，设目标为  $O(x, y)$ ，计算其在光学本征模式上的投影系数  $s_\ell$

$$s_\ell = \int_{\text{ROI}} O(x, y) \cdot \mathbb{E}_\ell^*(x, y) dx dy. \quad (14)$$

最终重建的目标为：

$$O_{\text{recon}}(x, y) = \sum_{\ell} s_\ell \cdot \mathbb{E}_\ell(x, y). \quad (15)$$

通过这种方式，可以基于实际获取的复振幅进行重建。与传统的图像重建计算方法相比，基于物理驱动 DL 的重建方法被证明能够有效地避免巨大的计算负担并获得高质量的重建结果<sup>[34]</sup>。将 OEI 物理模型嵌入到随机初始化的非预训练 CNN 中，通过在网络优化过程中引入成像物理模型作为反馈约束来实现高质量的图像重建。网络以随机复值噪声图  $z$ ，输出预估的目标复场  $\tilde{T}_\theta(z)$ ，该重建过程的目标函数可表示为

$$R_{\theta^*} = \arg \min_{\theta} \mathcal{L}(\theta), \quad (16)$$

其中，损失函数  $\mathcal{L}(\theta)$  具体定义为

$$\mathcal{L}(\theta) = \sum_{\ell} |\tilde{s}_\ell - s_\ell|^2 + \zeta \cdot \text{TV}(\tilde{T}_\theta(z)) \quad (17)$$

在上述公式中， $\theta$  代表网络权重； $\mathbb{E}_\ell$  为第  $\ell$  个光学本征模式， $s_\ell$  表示仿真中基于真实物理模型得到的第  $\ell$  个本征模式的复投影系数； $\tilde{s}_\ell$  表示网络输出的预估重建复场  $\tilde{T}_\theta(z)$  与本征模  $\mathbb{E}_\ell$  之间的复内积即  $\tilde{s}_\ell = \langle \tilde{T}_\theta(z), \mathbb{E}_\ell \rangle$ ，公式 (17) 的第一项中  $|\cdot|^2$  代表复数域下的模长平方，其计算了预测系数与真实系数在复平面上的平方误差；第二项中  $\text{TV}(\cdot)$  是改善重建图像质量的正则化约束， $\zeta$  用来平衡正则化在重建过程中的权重。

重建过程的迭代步骤如下：步骤 1：将随机复值噪声图  $z$  作为网络的唯一输入，输入到随机初始化的非预训练 CNN 中，本文采用的网络架构为标准复数 U-net 与融合注意力机制的扩展版。预先计算的本征模式  $\mathbb{E}_\ell$  和仿真投影系数  $s_\ell$  在每步迭代中用于计算物理约束损失，而不作为网络的直接输入。步骤 2：网络输出重建的复场  $\tilde{T}_\theta(z)$ 。步骤 3：将网络输出  $\tilde{T}_\theta(z)$  与 OEI 的物理投影过程融合，计算网络估计的投影系数  $\tilde{s}_\ell = \langle \tilde{T}_\theta(z), \mathbb{E}_\ell \rangle$ ，并依据式 (17) 计算损失函数  $\mathcal{L}$ 。步骤 4：采用 Adam 优化器来更新神经网络权重  $\theta$ ，以最小化损失函数  $\mathcal{L}$ 。重复步骤 3 和 4，直到损失函数收敛，获得最佳权重并输出最终的重建复场。

### 2.3 网络架构

与传统的用于 SPI 图像重建的 U-net 结构不同<sup>[35]</sup>，针对以携带相位信息的 LG 模式作为测试场，构造出的光学本征模式  $\mathbb{E}_\ell$  为复数场，因此目标场在这些模式下的投影系数  $s_\ell$  亦为复数。为此，本研究设计了支持复数运算的 U-net 网络，通过将实部和虚部作为两个独立通道输入和输出，实现对复场  $T = T_r + iT_i$  的有效重建。输入的复场数据被表示为两个独立的通道：实部  $T_r$  和虚部  $T_i$ 。这一设计确保了网络在多层卷积、池化以及注意力机制处理过程中，能够同时提取和维护复场在振幅和相位上的连贯特征，这对于精确地恢复光学模场的细节信息至关重要。网络主体基于经典 U-net 的“U”字形多尺度架构，由对称的编码器和解码器组成。编码器位于左侧，负责特征提取和空间压缩：接收  $H \times W$  输入复场，首先通过连续卷积层将输入映射至高维特征空间。随后，利用如图 1(b) 所示的编码器模块执行下采样任务：用  $5 \times 5$  的卷积核通过步长为 2 的卷积操作替代传统池化层，依次经过卷积、批归一化 (Batch normalization, BN) 和 Leaky ReLU 激活，将特征图空间尺寸逐层减半，同时特征通道数逐层翻倍。最终，网络在瓶颈层特征维度达到峰值，同时空间维度被大幅压缩。解码器位于右侧，结构与编码器相互对称，其核心任务是通过转置卷积操作逐步进行上采样，将空间分辨率恢复到原始大小  $H \times W$ 。在每个上采样层之后，解码器都通过跳跃连接 (Skip connections) 将编码器中对应尺度的特征图进行通道拼接，然后进行进一步的卷积处理。这种连接机制有效融合了编码器保留的高分辨率局部细节与解码器提取的低频全局上下文特征，是 U-net 能够高效进行图像重建任务的关键设计。在网络末端，解码器通过一个卷积层将特征通道数最终恢复到 2，输出重建复场  $T$  的实部和虚部，并可选择 Sigmoid 激活函数。此外，为了将网络输出与 OEI 的物理仿真过程相连接，网络还并行计算了重建复场  $T$  在预置模场  $\mathbb{E}_\ell$  上的投影系数  $\langle T, \mathbb{E}_\ell \rangle$ 。具体而言，将网络输出的双通道分别定义为重建场的实部  $T_r$  和虚部  $T_i$ ，同时将第  $\ell$  个本征模  $\mathbb{E}_\ell$  分解为实部  $\mathbb{E}_{\ell,r}$  和虚部  $\mathbb{E}_{\ell,i}$ 。投影系数的计算通过复共轭乘积的展开形式实现

$$\langle T, \mathbb{E}_\ell \rangle = \int (T_r(x, y) + iT_i(x, y)) (\mathbb{E}_{\ell,r}(x, y) - i\mathbb{E}_{\ell,i}(x, y)) dx dy. \quad (18)$$

为了进一步提升网络对关键特征的识别和表达能力，本文引入了卷积块注意力机制模块

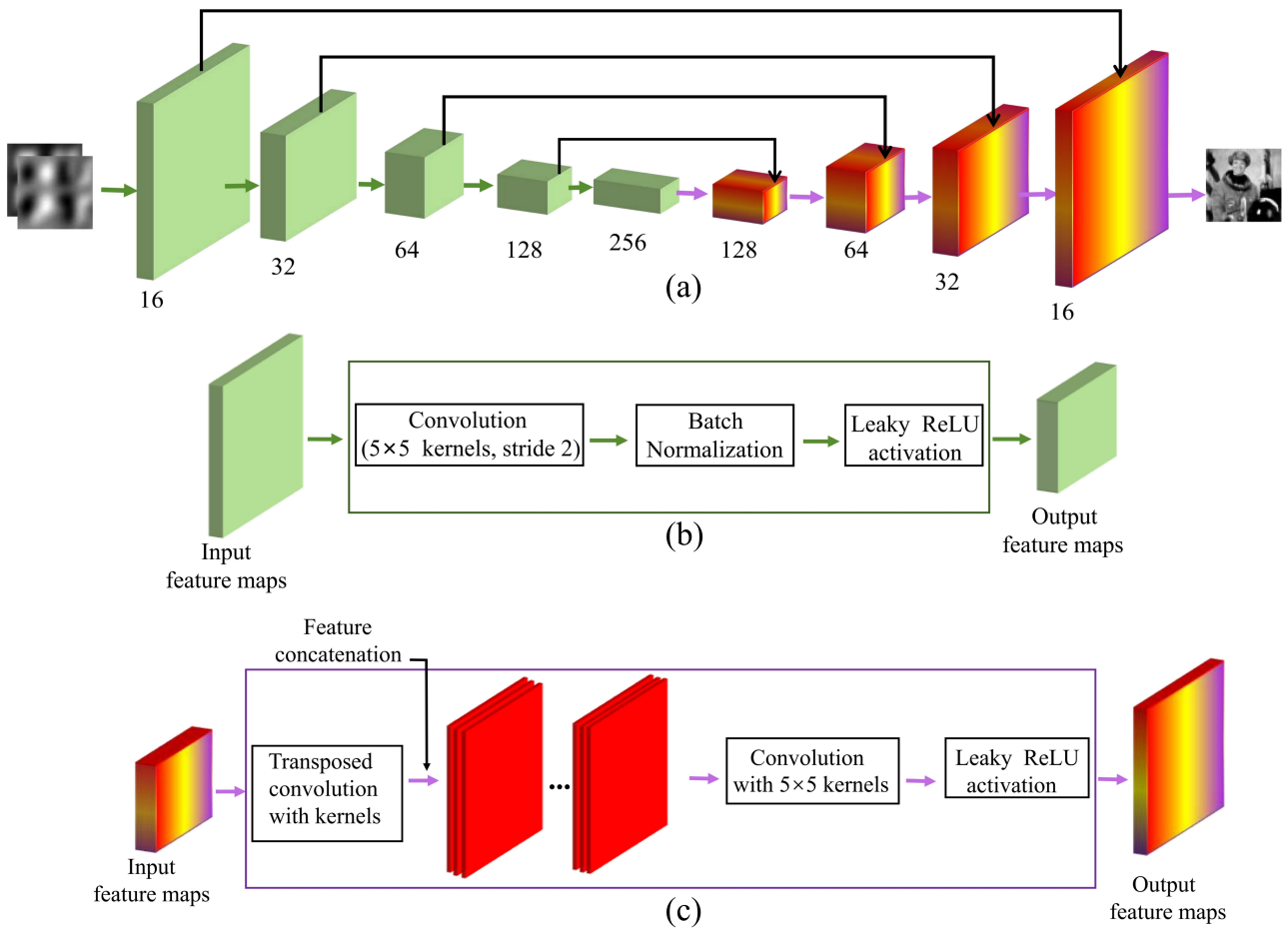


图 1 (a) 融合注意力机制的 U-net 网络架构，包含编码器层（绿色模块）和解码器层（彩色模块）；(b) 编码器架构；(c) 解码器架构。

Fig. 1 (a) The U-net architecture incorporating an attention mechanism, comprising encoder layers (green modules) and decoder layers (colored modules); (b) Encoder architecture; (c) Decoder architecture.

(Convolutional block attention module, CBAM)，并将其嵌入到解码器的每个上采样层以及瓶颈层之后。CBAM 的主要目标是通

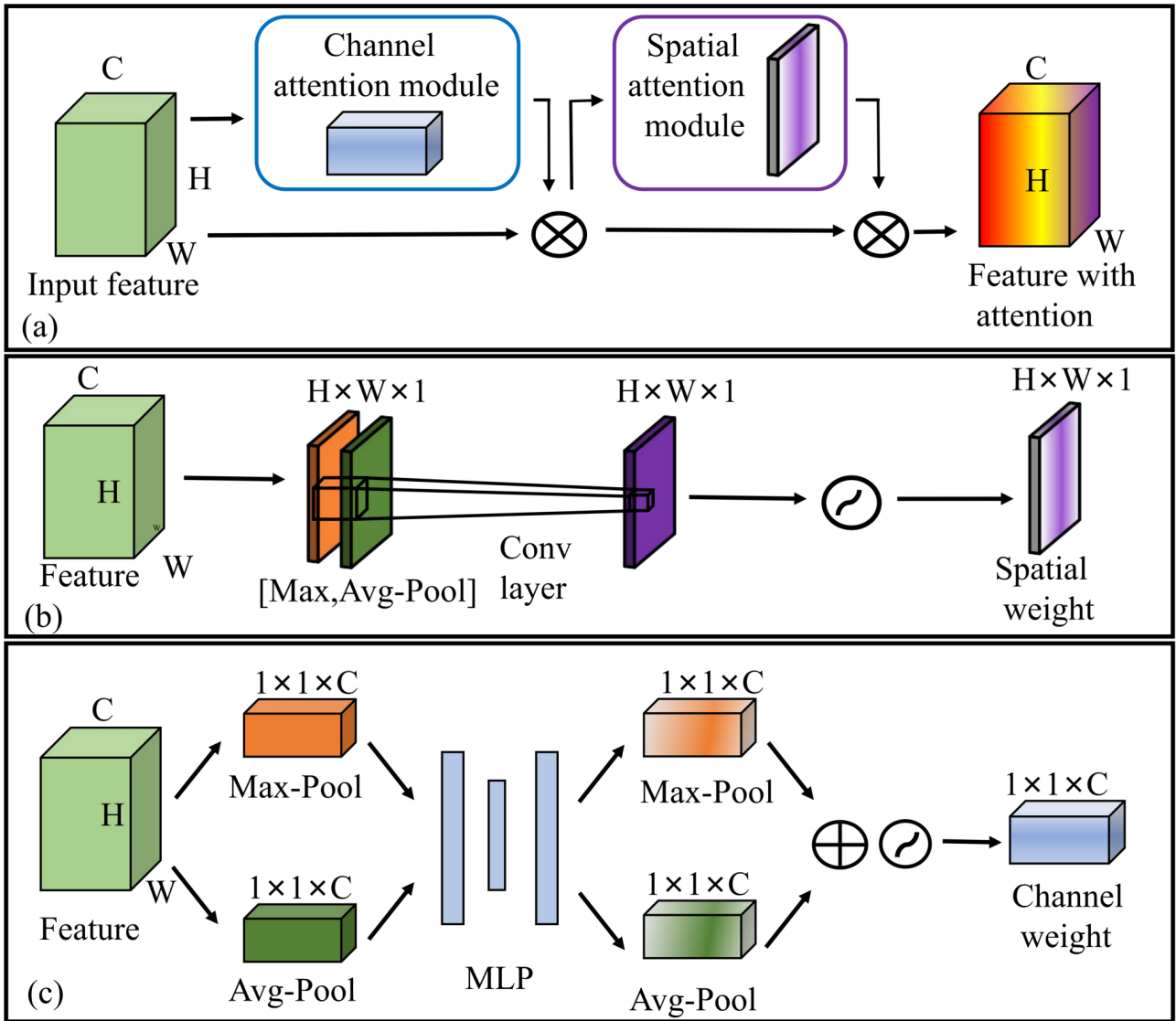


图 2 (a) 卷积块注意力机制 (CBAM) 模块总览; (b) 空间注意力模块; (c) 通道注意力模块。

Fig. 2 (a) Overview of the convolutional block attention mechanism (CBAM) module; (b) Spatial attention module; (c) Channel attention module.

过通道注意力模块 (Channel attention module, CAM) 和空间注意力模块 (Spatial attention module, SAM) 来提高模型的感知能力。CAM 用于自适应地调整各特征通道的重要性, 而 SAM 旨在赋予特征图不同空间位置以不同的权重, 两者结合后能够从通道和空间两个维度对特征进行优化加权。CAM 首先对输入特征图 (尺寸  $H \times W \times C$ ) 进行全局平均池化和全局最大池化操作, 生成两个  $1 \times 1 \times C$  的全局描述向量。这两个向量随后经过一个共享参数的两层全连接网络 (Shared MLP), 以实现通道间的交互和降维/升维操作。具体来说, 该 MLP 先将维度降至  $C/8$ , 经过 ReLU 激活后, 再恢复至原始通道数  $C$ 。将两个 MLP 输出相加后, 通过 Sigmoid 激活函数, 生成最终的  $1 \times 1 \times C$  通道注意力权重。随后, 将此权重与原始特征图进行逐通道相乘。紧接着, SAM 作用于经 CAM 加权后的特征图。SAM 首先在通道维度上进行平均池化和最大池化操作, 将原始特征图压缩为两个  $H \times W \times 1$  的单通道特征图。随

后，将两者沿通道轴拼接成  $H \times W \times 2$  的特征图，并输入到一个单卷积核的卷积层中。最终，通过 Sigmoid 激活函数将输出映射到  $(0, 1)$  区间，生成  $H \times W \times 1$  的空间注意力权重。该空间权重与经过 CAM 作用的特征图进行逐元素相乘，得到最终同时具备通道和空间注意力加权的特征图。在该过程中，我们选用 Adam 优化器，并设置初始学习率 (Learning rate, LR) 为 0.01。通过不断迭代更新网络权重，最小化损失函数，最终得到高质量重建结果。

### 3 结果与讨论

#### 3.1 光学本征模式成像超分辨的实现

在光学本征模式成像系统中，超分辨率的实现核心在于利用本征模式的线性叠加，在重构场中构建出有效宽度远小于瑞利判据的主瓣分布，从而突破传统直接成像的衍射极限。为了评估该系统的成像分辨率，我们考察系统对理想点源的响应特性，即点扩散函数。设目标场  $T(\mathbf{r})$  为位于  $\mathbf{r}_0$  处的理想点源，在数学上可由  $\delta$  函数描述，即  $T(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_0)$ 。根据公式 (10)，该点源在第  $\ell$  个本征模式上的投影系数为  $s_\ell$  为

$$s_\ell = \int \delta(\mathbf{r} - \mathbf{r}_0) \cdot \mathbb{E}_\ell^*(\mathbf{r}) d\mathbf{r} = \mathbb{E}_\ell^*(\mathbf{r}_0). \quad (19)$$

将此系数代入重建公式 (15)，即可得到系统对该点源的响应，即系统的点扩散函数  $E_{\text{PSF}}$

$$E_{\text{PSF}}(\mathbf{r}, \mathbf{r}_0) = \sum_{\ell=1}^N \mathbb{E}_\ell(\mathbf{r}) \mathbb{E}_\ell^*(\mathbf{r}_0). \quad (20)$$

$E_{\text{PSF}}$  的空间扩展程度，例如以半高宽 (Full width at half maximum, FWHM) 为指标，直接表征系统的空间分辨率。对于理想的无像差系统，该分辨率受限于光学衍射极限，其理论值通常由瑞利判据计算。对于一个焦距  $f = 400 \text{ mm}$ ，波长  $\lambda = 632 \text{ nm}$ ，透镜直径为  $d = 2 \text{ mm}$  的成像透镜系统，依据瑞利判据<sup>[36]</sup>，该系统的理论空间分辨极限  $\Delta x$  为

$$\Delta x = 1.22 \cdot \frac{f \cdot \lambda}{d}, \quad (21)$$

该系统的理论衍射极限约为  $154.4 \mu\text{m}$ 。这一分辨极限取决于衍射受限系统的 PSF，即 Airy 斑的空间展宽。作为理想点源在像平面的响应，Airy 斑的径向强度分布直接制约了系统的分辨能力，其解析表达式如下<sup>[37]</sup>

$$E(r) \propto \frac{2J_1(u)}{u}, \quad u = \frac{\pi dr}{\lambda f}, \quad (22)$$

其中  $J_1(u)$  是第一类贝塞尔函数， $u$  是衍射参数。

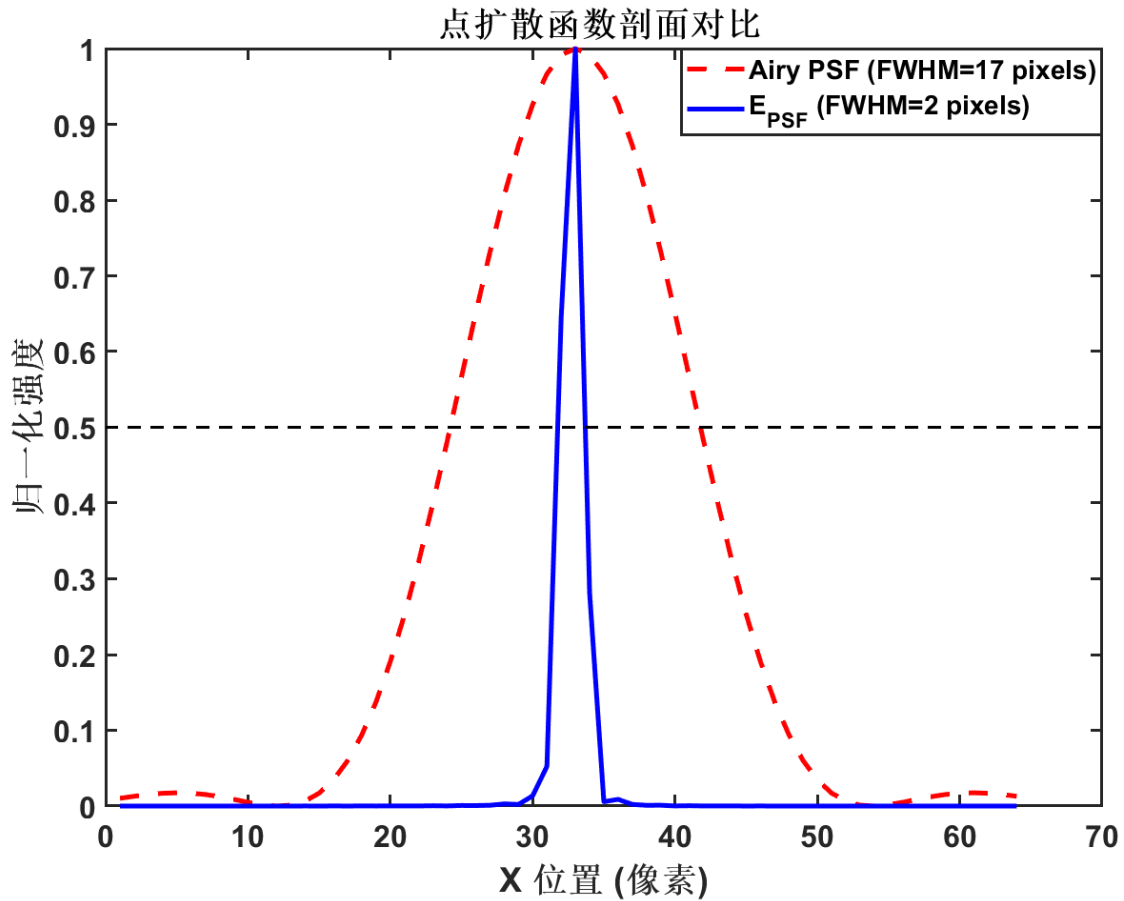


图 3 点扩散函数对比

Fig. 3 Point spread function comparison.

为验证 OEI 系统的超分辨潜力，我们以 LG 模式作为测试场，在  $64 \times 64$  像素网格构造光学本征模式，并据此计算  $E_{\text{PSF}}$  与 Airy PSF 的比较曲线。在仿真的离散化空间映射中，由 Airy PSF 空间展宽所决定的这一理论分辨率表现为约 17 个像素的特征宽度，而  $E_{\text{PSF}}$  的 FWHM 仅为 2 个像素，显著优于 Airy PSF 的 FWHM。同时  $E_{\text{PSF}}$  的主瓣更尖锐，旁瓣抑制更显著，证实 OEI 系统能够实现比传统成像方法更高的分辨率。

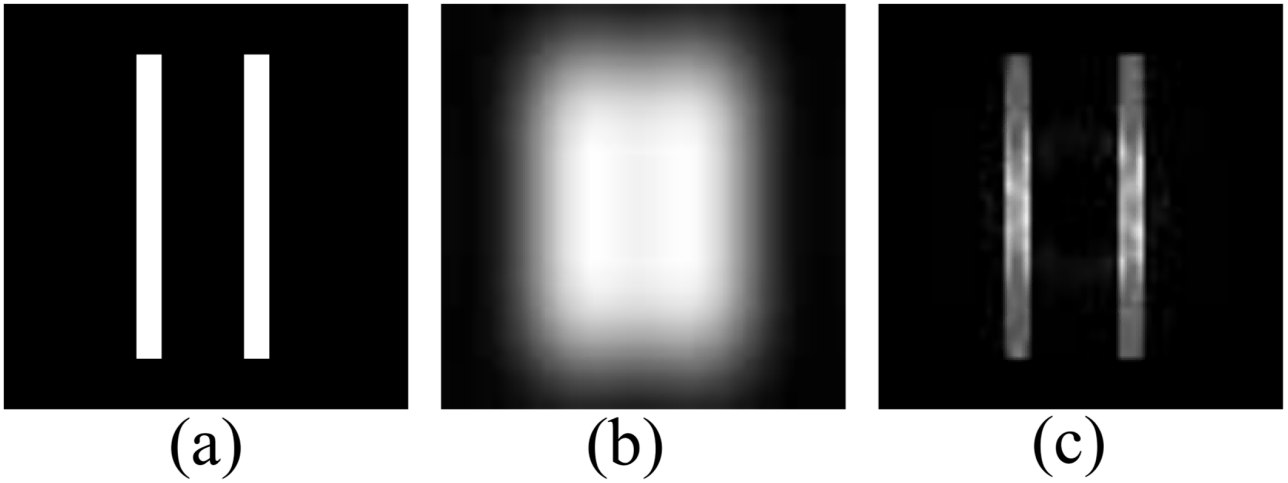


图 4 (a) 原始物体; (b) 传统成像结果; (c)OEI 重建结果。

Fig. 4 (a)Ground truth ; (b)Traditional imaging results ; (c)OEI reconstruction results .

为进一步验证 OEI 系统的超分辨潜力, 本文设计了对双缝物体的数值仿真对比。模拟参数设置如下: 采样网格为  $64 \times 64$ 。构建的双缝物体每条狭缝宽 4 个像素, 中心间距为 13 个像素。仿真中, 原始物体设定为理想的双缝矩形分布。作为对比, 衍射受限的成像效果通过原始图像与系统 Airy PSF 的卷积获得; OEI 重建则利用 1638 个光学本征模式, 通过计算模态投影系数复原目标的复场分布, 并进一步提取其强度分布以进行图像显示与评估。数值模拟结果如图 4 所示: 真实目标呈现锐利的双缝结构; 传统成像边缘显著模糊, 两缝无法有效分辨; OEI 重建则显著提升了边缘锐度和对比度, 两缝清晰可辨, 验证了其超分辨能力。

### 3.2 仿真结果对比

为验证所提出方法在 OEI 中的有效性, 本研究选取了  $64 \times 64$  像素的灰度图像“宇航员”、“房子”作为目标物体进行数值仿真。仿真过程严格依据 OEI 物理机制: 首先基于预设的采样率 (Sampling rate, SR) 生成  $N_{\text{SR}}$  个不同阶数的 LG 模式作为测试场  $\{u_j(\mathbf{r})\}$ , 其中  $N_{\text{SR}} = \text{SR} \times 64 \times 64$ 。随后, 通过计算所有  $N_{\text{SR}}$  个测试场模式之间的复共轭内积, 构造维度为  $N_{\text{SR}} \times N_{\text{SR}}$  的相关矩阵  $M$ 。接着, 矩阵  $M$  进行本征值分解, 构造一组正交归一化的光学本征模式  $\mathbb{E}_\ell(\mathbf{r})$ ; 最后计算目标在本征模基底下的投影系数。并分别利用传统 OEI、基础 U-net 以及 U-attention 网络进行图像重建。

为全面评估算法性能, 仿真设置了  $\text{SR} \in \{5\%, 10\%, 20\%, 30\%, 40\%\}$  的采样率集合。图 5 和图 6 分别展示了“宇航员”与“房子”在不同采样率下, 三种算法的重建结果对比。从视觉效果来看, 传统 OEI 算法受限于模态基底的不完备性, 成像质量随采样率降低而显著恶化。在  $\text{SR} = 10\%$  时, OEI 重建结果虽能呈现目标的全局结构, 但细节模糊; 当采样率降至  $\text{SR} = 5\%$  时, 高频信息严重丢失, 图像不仅模糊且伴有明显伪影。基础 U-net 模型在一定程度上提升了图像的对比度, 但在  $\text{SR} = 5\%$  条件下, 由于物理驱动约束在极低采样下的欠定性, 重建结果中仍伴随有斑点状伪影。相比之下, 融合了注意力机制的 U-attention 模型在视

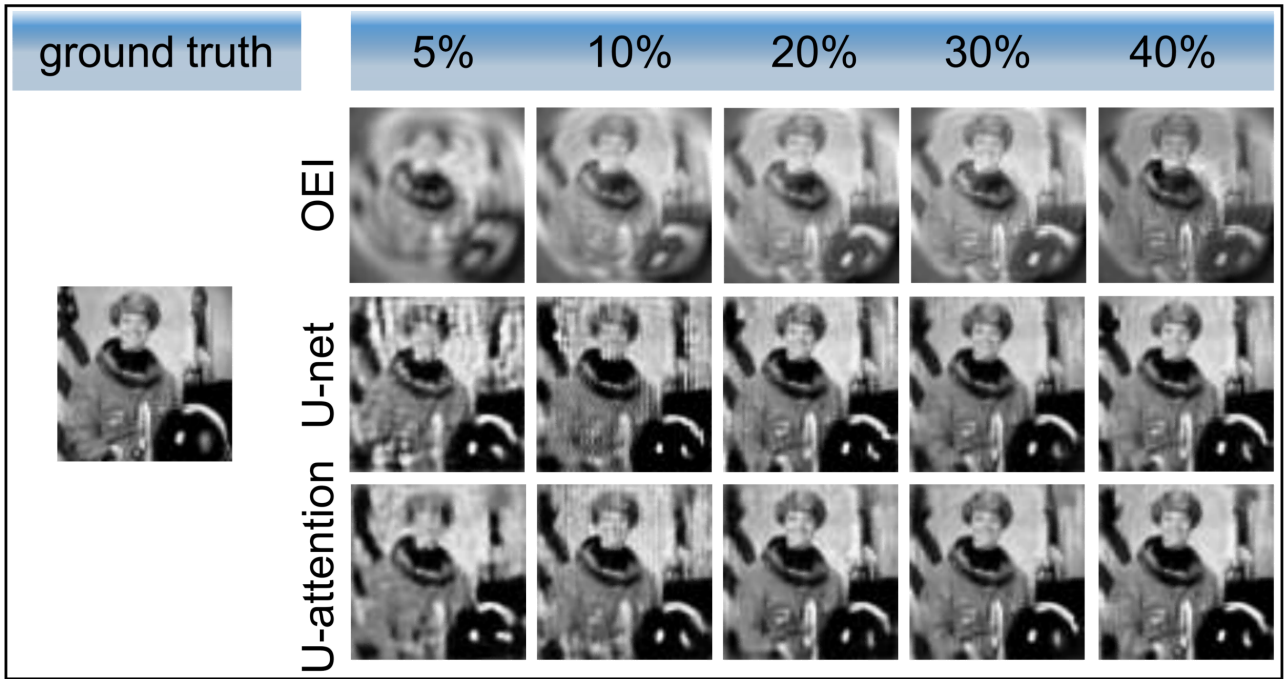


图 5 在不同采样率下对测试对象“宇航员”图像进行数值模拟，并比较不同重建算法的结果。

Fig. 5 Numerical simulations of the “astronaut” test object image were conducted at different sampling rates, and the results of various reconstruction algorithms were compared.

表 1 不同采样率与重建算法下“宇航员”图像的定量比较。

Tab. 1 Quantitative comparison of the “astronaut” image under different sampling rates and reconstruction algorithms.

Methods\Sample-rate	5%	10%	20%	30%	40%
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
OEI	15.05/0.47	14.97/0.57	15.07/0.63	15.33/0.64	17.33/0.64
U-net	17.73/0.69	19.99/0.75	21.21/0.87	23.15/0.88	23.32/0.90
U-attention	19.25/0.73	21.03/0.81	23.72/0.89	24.38/0.91	24.61/0.93

觉保真度上表现更优。得益于注意力模块对空间边缘特征的自适应强化，即使在  $SR = 5\%$  的极端条件下，虽然受限于物理极限，图像无法达到高清晰度，但相比于其他算法，该模型重构的图像表现出更好的几何规整性与结构连续性，有效地抑制了背景区域的伪影波动对主体轮廓的干扰。为进一步量化重建质量，采用峰值信噪比 (Peak signal-to-noise ratio, PSNR) 和结构相似性 (Structural similarity, SSIM)<sup>[38]</sup> 作为评价指标，两个目标物体在不同算法下定量比较重建结果分别列于表 1 和表 2。数据表明，U-attention 方法在所有采样率下的性能指标均显著优于对比方法。在极度稀疏采样条件  $SR = 5\%$  时，针对“宇航员”图像，U-attention 的 PSNR 达到 19.25 dB，SSIM 为 0.73，相比基础 U-net (PSNR 17.73 dB，SSIM 0.69) 分别提升了 1.52 dB 和 0.04。这一优势在“房子”图像的重建中同样显著，其 PSNR 达到 21.94 dB，超

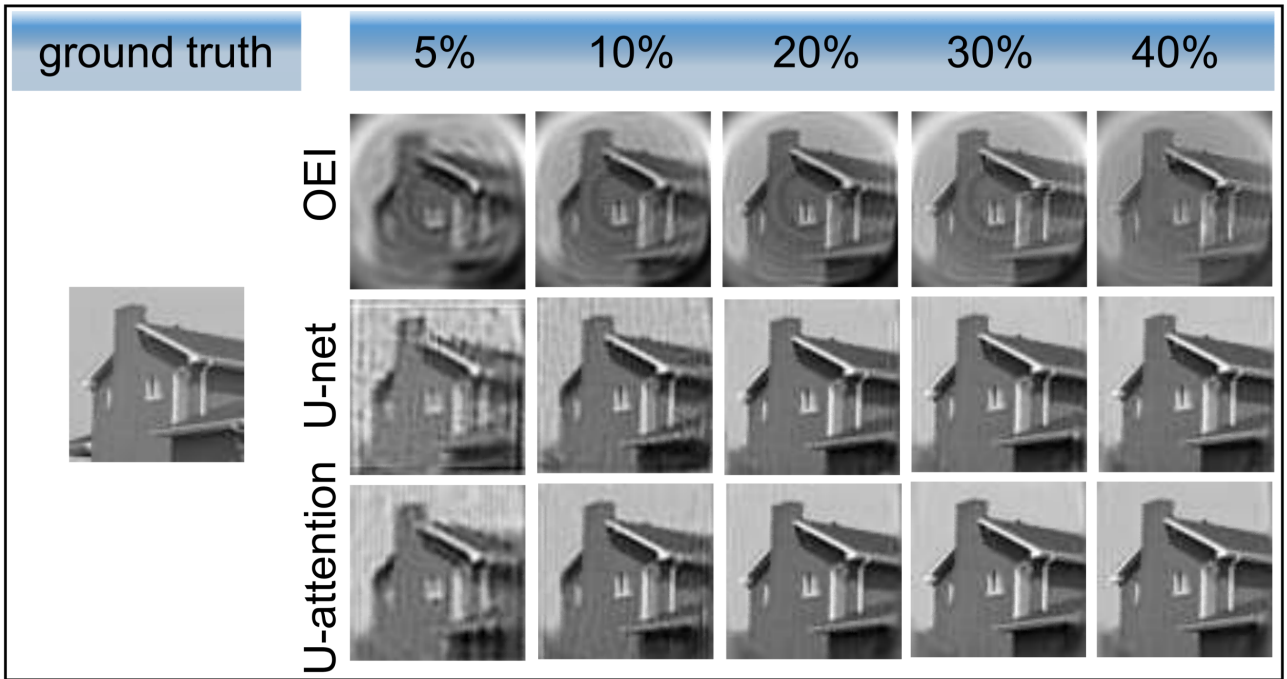


图 6 在不同采样率下对测试对象“房子”图像进行数值模拟，并比较不同重建算法的结果。

Fig. 6 Numerical simulations of the “house” test object image were conducted at different sampling rates, and the results of various reconstruction algorithms were compared.

表 2 不同采样率与重建算法下“房子”图像的定量比较。

Tab. 2 Quantitative comparison of the “house” image under different sampling rates and reconstruction algorithms.

Methods\Sample-rate	5%	10%	20%	30%	40%
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
OEI	18.38/0.45	18.97/0.55	19.36/0.66	21.05/0.69	20.63/0.71
U-net	20.62/0.48	22.37/0.61	23.41/0.71	24.02/0.76	24.37/0.79
U-attention	21.94/0.51	22.93/0.62	23.45/0.76	25.28/0.79	25.64/0.82

过传统 OEI 的 18.38 dB。这表明，即便在视觉感知较为模糊的极端低采样条件下，注意力机制依然能够通过精准捕捉核心特征权重，使重建结果接近真实值，显著缓解高频模态缺失带来的结构性损失。随着采样率增加至  $SR = 40\%$  时，该模型的优势依然稳固，“宇航员”重建图像 PSNR 提升至 24.61 dB，SSIM 达到 0.93，相比 U-net 分别实现了 1.19 dB 和 0.03 的提升。上述结果证实，即使在数据冗余度较高的情况下，注意力机制依然能够通过精细化处理不同尺度的特征图贡献，帮助网络逼近重建性能的理论上限，从而在 PSNR 和 SSIM 上均实现最优表现。

仿真结果表明，相比受限于模态不完备性的传统 OEI 以及在极低采样下细节恢复能力不足的基础 U-net，引入注意力机制的 U-attention 网络通过自适应校准特征权重，使得网络能

够优先聚焦于关键的结构特征，从而有效弥补了物理模型在欠采样下的高频信息缺失，更助推网络在全采样率范围内逼近性能理论上限，实现了对伪影抑制与结构保真度的双重优化。

## 4 结论

本文针对传统 OEI 方法在低采样率条件下重建质量受限的问题，构建了在 OEI 物理模型约束下融合注意力机制的复数卷积神经网络重建策略。研究首先从成像机理层面验证了 OEI 系统的超分辨成像潜力，仿真结果表明，OEI 系统能够有效突破传统光学系统的衍射极限，并成功实现了对瑞利判据下不可分双缝物体的清晰分辨。在此物理基础上，为解决欠采样带来的高频模态缺失难题，本文将 OEI 物理模型嵌入复数 U-net 架构作为约束，并进一步引入注意力机制以增强多尺度特征提取能力。与传统 OEI 及基础 U-net 方法相比，U-attention 方法能够有效提升重建图像的细节与信噪比，尤其在低采样率下表现出优越的性能。仿真结果表明，该方法能够在低采样率下恢复结构清晰的图像，并在 PSNR 和 SSIM 指标上显著优于对比方法。此外，该策略有效结合了物理先验与深度学习的非线性重构优势，为解决 OEI 重建中的挑战提供了一种新的思路。目前的方案主要验证了数值模拟环境下的有效性，未来研究将致力于开展实际光学实验，验证该策略在真实复杂物理环境下的鲁棒性，并进一步优化网络训练效率。

## 参考文献

- [1] Luker G D, Luker K E 2008 *J. Nucl. Med.* **49** 1
- [2] Li L Z, Yao X R, Liu X F, Yu W K, Zhai G J 2014 *Acta Phys. Sin.* **63** 224201 (in Chinese)  
[李龙珍, 姚旭日, 刘雪峰, 俞文凯, 翟光杰 2014 物理学报 **63** 224201]
- [3] Hell S W 2007 *Science* **316** 1153
- [4] De Luca A C, Kosmeier S, Dholakia K, Mazilu M 2011 *Phys. Rev. A* **84** 021803
- [5] Kosmeier S, Zolotovskaya S, De Luca A C, Riches A, Herrington C S, Dholakia K, Mazilu M 2014 *Optica* **1** 257
- [6] Mazilu M, Baumgartl J, Kosmeier S, Dholakia K 2011 *Opt. Express* **19** 933
- [7] Kosmeier S 2013 *Optical eigenmodes for illumination & imaging*. Ph.D. Dissertation, University of St Andrews
- [8] Kosmeier S, De Luca A C, Zolotovskaya S, Di Falco A, Dholakia K, Mazilu M 2013 *Sci. Rep.* **3** 1808
- [9] Piché K, Leach J, Johnson A S, Salvail J Z, Kolobov M I, Boyd R W 2012 *Opt. Express* **20** 26424

- [10] Tsampoula X, Mazilu M, Vettenburg T, Gunn-Moore F, Dholakia K 2013 *Photon. Res.* **1** 42
- [11] Johnson A S, Piché K, Salvail J Z, Leach J, Boyd R W 2013 *J. Mod. Opt.* **60** 1931
- [12] Wang W, Wang Y P, Wu Y, Yang X X, Wu Y 2014 *Opt. Lett.* **39** 2614
- [13] Kosmeier S, Mazilu M, De Luca A C, Baumgartl J, Dholakia K 2012. Complex Light and Optical Forces VI San Francisco, California, USA, Jan. p 82740I
- [14] Zhang S Z, Wang W, Yu R, Yang X X 2016 *Laser Phys. Lett.* **13** 015003
- [15] Baumgartl J, Kosmeier S, Mazilu M, Rogers E T F, Zheludev N I, Dholakia K 2011 *Appl. Phys. Lett.* **98** 181109
- [16] Lyu M, Wang W, Wang H, Wang H C, Li G W, Chen N, Situ G H 2017 *Sci. Rep.* **7** 17865
- [17] Koetzier L R, Mastrodicasa D, Szczykutowicz T P, van der Werf N R, Wang A S, Sandfort V, van der Molen A J, Fleischmann D, Willeminck M J 2023 *Radiology* **306** e221257
- [18] Cammarasana S, Nicolardi P, Patanè G 2022 *Med. Biol. Eng. Comput.* **60** 2229
- [19] Adegun A A, Fonou Dombeu J V, Viriri S, Odindi J 2023 *Sensors* **23** 5849
- [20] Thapa A, Horanont T, Neupane B, Aryal J 2023 *Remote Sens.* **15** 4804
- [21] Xu Y B, Lu L Y, Saragadam V, Kelly K F 2024 *Nat. Commun.* **15** 1456
- [22] Zhao W Z, Du S H 2016 *IEEE Trans. Geosci. Remote Sens.* **54** 4544
- [23] Liu P, Zhang H, Eom K B 2017 *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **10** 712
- [24] Tang P W, Lin C H, Liu Y R 2024 *IEEE Trans. Geosci. Remote Sens.* **62** 1
- [25] Ahmed M T, Monjur O, Kamruzzaman M 2024 *J. Food Eng.* **382** 112223
- [26] Song K, Bian Y X, Wu K, Liu H R, Han S P, Li J M, Tian J Z, Qin C B, Hu J Y, Xiao L T 2023 *arXiv: 2310.16869 [eess.IV]*
- [27] Kilcullen P, Ozaki T, Liang J 2022 *Nat. Commun.* **13** 7879
- [28] Donoho D L 2006 *IEEE Trans. Inf. Theory* **52** 1289
- [29] Duarte M F, Davenport M A, Takhar D, Laska J N, Sun T, Kelly K F, Baraniuk R G 2008 *IEEE Signal Process. Mag.* **25** 83
- [30] Lin J, Yan Q R, Lu S, Zheng Y J, Sun S D, Wei Z 2022 *Photonics* **9** 343

- [31] Wang F, Wang C L, Deng C J, Han S S, Situ G H 2022 *Photon. Res.* **10** 104
- [32] Hoshi I, Shimobaba T, Kakue T, Ito T 2020 *Opt. Express* **28** 34069
- [33] Wang F, Czarske J W, Situ G H 2025 *Adv. Photon.* **7** 054002
- [34] Wang M Y, Mao J N, Su H, Ling Y Y, Zhou C Q, Su Y K 2024 *Biomed. Opt. Express* **15** 6619
- [35] Wang C H, Li H Z, Bie S H, Lv R B, Chen X H 2023 *Photonics* **10** 224
- [36] Dang S P, Li R Z, Zhou M L, Qian J, Dan D, Yu X H, Yao B L 2022 *Infrared Laser Eng* **51** 20220735 (in Chinese) [党诗沛, 李润泽, 周美玲, 千佳, 但旦, 于湘华, 姚保利 2022 红外与激光工程 **51** 20220735]
- [37] Minin I, Minin O 2016 *Diffraction Optics and Nanophotonics*. Cham: Springer
- [38] Wang D Y, Bie S H, Chen X H, Yu W K 2024 *Photonics* **11** 174

录用稿件，非最终出版稿

# Optical eigenmode imaging based on convolutional neural networks

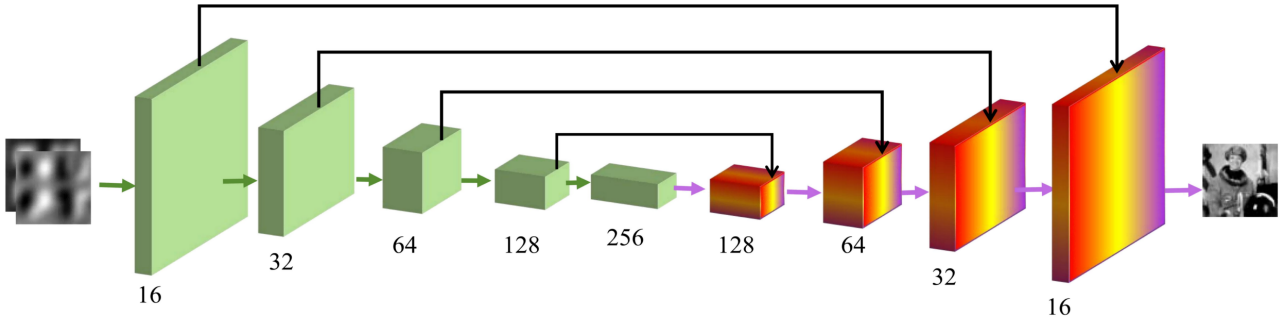
ZHOU Yishen LI Qi WANG Xiang SONG Jiakang JIA Kanxu ZHANG Xuange CHEN

Xihao <sup>†</sup>

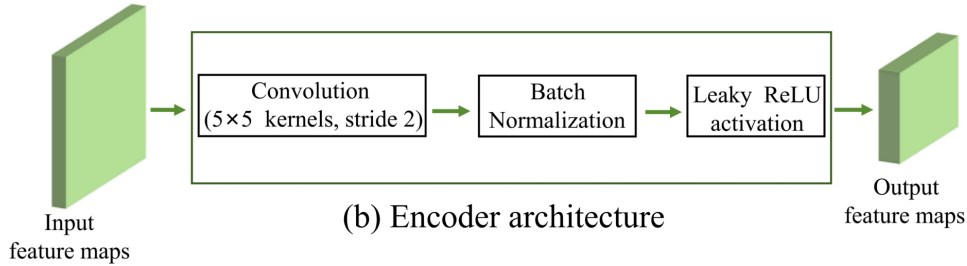
School of Physics, Liaoning University, Shenyang 110036 , China

## Abstract

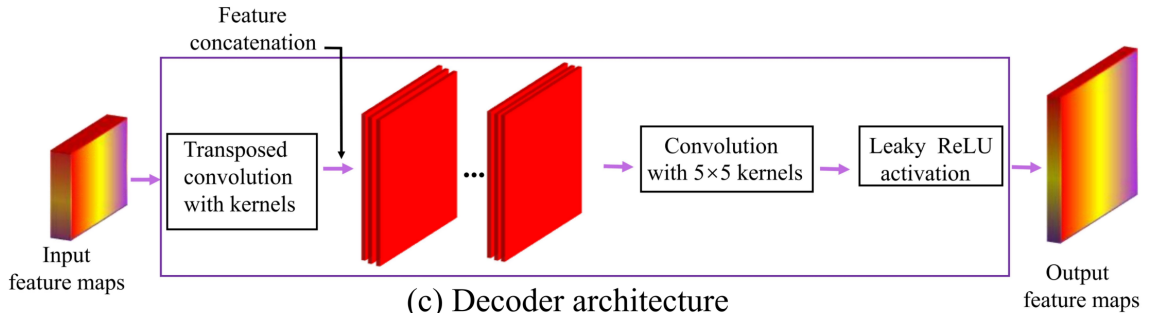
Optical eigenmode imaging (OEI) is a computational imaging method based on the expansion and projection of eigenmodes within an optical system. It acquires target images by characterizing and reconstructing spatial distribution information in the mode space. However, under low sampling rates, traditional OEI reconstruction is fundamentally limited by modal incompleteness, leading to the loss of high-frequency information. This physically manifests as enhanced



(a) Convolutional neural networks of a U-net architecture



(b) Encoder architecture



(c) Decoder architecture

side-lobes in the system's point spread function and severe artifacts in the reconstructed images, thereby compromising edge sharpness and detail fidelity. To

overcome these physical limitations, we propose a physics-driven, complex-valued convolutional neural network framework that seamlessly integrates physical laws with deep learning by incorporating the optical imaging physical model as an intrinsic constraint. Unlike conventional data-driven approaches that rely on massive paired datasets, our method utilizes an untrained complex U-net optimized through a self-supervised strategy. The network treats the real and imaginary components of the optical field as independent channels to strictly preserve phase information. Furthermore, to specifically address the issue of detail loss caused by modal incompleteness, we integrate a convolutional block attention module (CBAM) into the network's decoder. This mechanism achieves high-quality image reconstruction through adaptive feature weight adjustment.

Numerical simulations indicate that under extremely low sampling rates, where traditional OEI fails to recover meaningful structural details and the baseline U-net suffers from residual artifacts, our attention-enhanced model successfully suppresses side-lobe-induced artifacts. Comparative validations demonstrate that superior image fidelity and structural similarity are achieved compared to traditional OEI algorithms and baseline deep learning approaches, effectively mitigating structural degradation caused by modal incompleteness. This method provides a robust solution for OEI, enabling high-quality super-resolution imaging without the need for high sampling redundancy or external training datasets.

**Keywords:** optical eigenmode imaging; image reconstruction; convolutional neural networks; attention mechanisms

**PACS:** 42.30.Va, 42.30.Wb, 89.20.Ff, 42.30.-d

† Corresponding author. E-mail: xi-haochen@163.com