

基于深度学习的环芯光纤精确模式分解方法

田梓辰¹, 裴丽^{1*}, 王建帅¹, 白冰¹, 胡恺华¹, 郑晶晶¹, 沈磊^{2,3,4}, 徐文轩¹¹北京交通大学全光网络与现代通信网教育部重点实验室, 北京 100044;²光纤光缆制备技术国家重点实验室, 湖北 武汉 430073;³长飞光纤光缆股份有限公司, 湖北 武汉 430073;⁴湖北光谷实验室, 湖北 武汉 430073

摘要 提出了一种无需预训练的卷积神经网络(CNN)模式分解(MD)算法,该算法利用具有不同感受野的分支结构提升神经网络的学习能力,用于对环芯少模光纤(FM-RCF)复杂耦合进行高精度表征。在仿真测试中,所提方法的模态相对相位和模态权重误差分别减小为传统CNN-MD的1/8与1/5。实验中利用CCD采集RCF输出端实际光斑图像作为算法输入,对3个模式叠加状态下的光斑进行模态权重和相对相位分析,得到重建光斑与实际光斑之间的相似度高于90%。所提算法不易陷入局部最优且无需迭代,处理时间为ms量级,可为RCF的模式表征及性能预测提供理论与技术支持。

关键词 光通信; 少模光纤; 模式分解; 深度学习; 环芯光纤

中图分类号 TN929.11

文献标志码 A

DOI: 10.3788/AOS221559

1 引言

近年来,针对少模光纤(FMF)的研究在基础理论和实际应用方面都取得了很大的进展,推动了光纤通信多个方向的进步。在传输方面,基于FMF的模分复用(MDM)技术被认为是一种极具潜力的克服当前单模光纤(SMF)光通信系统容量危机的方法^[1-2];在光纤激光器制造方面,具有大模场面积的FMF用于抑制非线性效应以及提高损伤阈值,以实现高功率光纤激光器^[3];在非线性现象方面,FMF可以作为时空锁模机制^[4]及克尔非线性光束清理^[5]的理想研究平台。其中,少模环芯光纤(RCF)的芯子由折射率(RI)凹陷的中心区与高折射率的外层环芯组成,使得传导模式均被限制在环芯区之中。这种独特的纤芯结构使得RCF在光纤放大器的模式增益均衡^[6]、MDM系统中多输入多输出(MIMO)简化^[7-8]、涡旋光束生成^[9]等方面具有广泛的应用前景。但是与SMF相比,RCF在实际使用中受到外界环境等因素干扰或用于长距离传输时,容易引起模式间的耦合。RCF中各个模式的模场分布重叠度高、识别难度大,使得针对RCF的耦合解析相比于其他种类的FMF难度更高。

模式分解(MD)是一种用于光纤模式耦合特性表征的关键技术。通过获取光纤中各模式模态权重(ρ^2)与模态相对相位(θ)信息,可进一步分析出FMF中的

模式耦合变化情况并可从完整的光纤光场中推算出相关联的光束属性。当前,一系列基于不同原理的MD方法如空间和光谱域分辨成像(F²)法^[10]、光学相关分析(OCA)法^[11]、逆矩阵求解法^[12]、随机梯度下降(SPGD)法^[13]、卷积神经网络(CNN)法^[14-17]等相继被提出。其中,CNN-MD作为一种非迭代的数值算法,所构建的神经网络在训练完成后只需经历一次前向传播过程便可完成MD,具有分解速度快的优势。而且,在实际使用中,CNN-MD仅使用由CCD采集的待测光纤的输出端光斑作为算法输入,对实验设备及环境的要求低,数据处理操作简单。此外,CNN对初始值的选取并不敏感,不易陷入局部最优解中,可以保持MD的高精度与稳定性。

近年来,CNN逐渐成为MD领域主流的研究方法。2019年,An等^[14]利用VGG-CNN在5个模式叠加情况下完成了高精度且实时的MD(VGG-MD),但该方法需要先使用包含 10^7 量级图片的ImageNet数据集对神经网络进行预训练,耗时很长。2020年,Fan等^[17]通过添加与近场(NF)和远场(FF)光斑图相关的额外损失函数更新CNN,实现了6个模式叠加情况下的高精度MD(NFFL-CNN-MD)。但是用于CNN训练的光斑图数量高达 10^6 张,这对计算机硬件要求过高且导致训练时间较长。2021年,Zhu等^[16]采用光斑图与相位分布图(PD)的组合作为CNN输入,基于ResNet-18

收稿日期: 2022-08-02; 修回日期: 2022-08-26; 录用日期: 2022-09-19; 网络首发日期: 2022-09-29

基金项目: 国家重点研发计划(2018YFB1801003)、国家自然科学基金(61827817)

通信作者: *lpei@bjtu.edu.cn

型 CNN 成功实施了 6 个模式的高精度分解 (DH-ResNet18-MD)。但是拍摄相位分布图依赖高精度全息装置,大大增加了成本与操作复杂度。同年,Rothe 等^[15]利用多达 121 层的 DenseNet-CNN 成功在 8 模式叠加的情况下实现了高精度 MD (DenseNet121-MD)。但是由于 DenseNet-121 网络中涉及大量的通道拼接操作以及底层特征提取上的冗余问题,需要两个 GPU 同时加速才能完成训练,这使得计算机硬件和算力资源消耗过大。在测试光纤种类方面,当前的 CNN-MD 工作大多聚焦于表征传统的阶跃 FMF。但是 RCF 所特有的高 RI 环芯区结构将带来比阶跃 FMF 更高的模式分布相似程度,引起更复杂的模式重叠和更明显的模式耦合现象,这使得目前的 CNN-MD 算法在针对 RCF 的高精度 MD 实现上显得不足,限制了其分解效果与应用范围。

本文提出并设计了一种无需预训练的 CNN-MD (PFCNN-MD)。该方案在针对 RCF 耦合模式的模态权重和模态相对相位识别中具有优良的表现。在仿真测试中,8 个模式叠加情况下的模态权重和模态相对相位误差分别低于 0.95% 和 1.92%,分别减小为传统 CNN-MD 方案的 1/5 和 1/8,完成一次 MD 的时间为 9 ms。在实验验证中,利用 CCD 拍摄的真实光斑源图

像作为 CNN 输入,在 3 个模式叠加情况下重建光斑图像与源图像之间的相似度高于 90%。PFCNN-MD 的精度高、速度快,可实现实时 MD,这将为环芯少模光纤 (FM-RCF) 中传输模式特性的定量分析以及进一步应用提供了有力的支撑。

2 环芯光纤模式分解原理

2.1 环芯光纤模式特性

如图 1(a) 所示,RCF 中心区域与包层的 RI 相同,高 RI 的环芯层位于中心区域与包层之间,其内径 r_{in} 与外径 r_{out} 分别为 $2.85 \mu\text{m}$ 和 $8.40 \mu\text{m}$,环形纤芯折射率 n_{co} 与包层折射率 n_{cl} 之间的 RI 差 Δn 为 0.014。在 1550 nm 工作波长处,该光纤支持 12 个模式,其中一些典型的模场分布如图 1(b) 所示,可以发现各传导模式被限制在了高 RI 环芯层中。图 1(c) 展示了 RCF 中支持传输的前 8 个模式以随机模态权重和模态相对相位叠加得到的 6 幅纤芯光斑图。虽然光斑图之间的模态权重和模态相对相位数值差异非常大,但从图中可以看出,大部分的光束功率被限制在了相同的区域内。不同光斑之间的强度分布重叠度高,这将为针对 RCF 耦合情况的解析带来严峻的挑战。

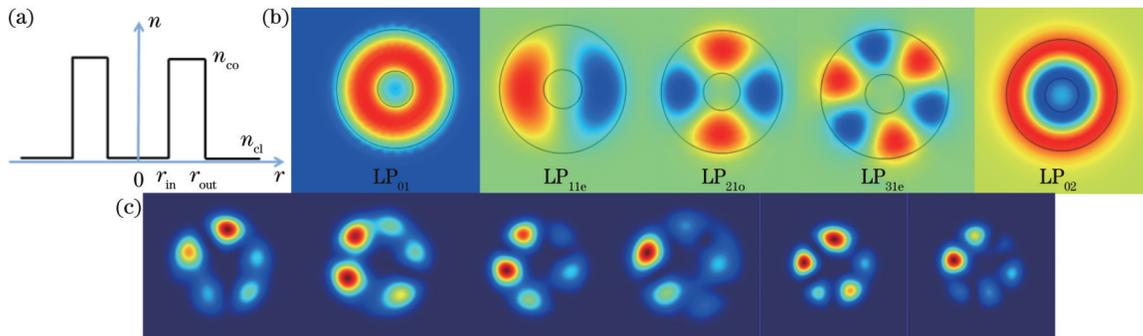


图 1 环芯少模光纤 (FM-RCF) 特性示意图。(a) 折射率分布;(b) 光纤模式典型模场分布;(c) 8 模式 (LP_{01} 、 LP_{11o} 、 LP_{11e} 、 LP_{21o} 、 LP_{21e} 、 LP_{31e} 、 LP_{31o} 、 LP_{02}) 叠加情况下的纤芯光斑图

Fig. 1 Diagram of FM-RCF characteristics. (a) Refractive index profile; (b) typical mode field profiles of fiber modes; (c) 8-mode (LP_{01} , LP_{11o} , LP_{11e} , LP_{21o} , LP_{21e} , LP_{31e} , LP_{31o} , and LP_{02}) superposition beam patterns in fiber core

在 RCF 中,耦合模场分布 $U(r, \varphi)$ 可以看作是各本征模模场的线性组合^[16],即

$$U(r, \varphi) = \sum_{n=1}^N \rho_n e^{i\theta_n} \psi_n(r, \varphi), \quad (1)$$

$$\sum_{n=1}^N \rho_n^2 = 1, \quad (2)$$

式中: $\psi_n(r, \varphi)$ 表示光纤中第 n 个本征模所具有的归一化电场分布; N 为光纤中可支持传输的总模式数量; $\theta_n \in [-\pi, \pi]$ 。本文利用线偏振模 (LP) 来统一描述各个本征模式。第 n 个模式对应的模态权重和模态相对相位分别采用 ρ_n^2 和 θ_n ^[14] 表示。

光束的强度分布 $I(x, y)$ 可描述为

$$I(x, y) = |U(x, y)|^2. \quad (3)$$

光强分布可由 CCD 或 CMOS 相机在光纤出射端口拍摄成为近场光斑图。CNN-MD 方法的目标在于通过分析光斑图获取图中包含的 $2N-1$ 个模态系数值 (包括 N 个模态权重和 $N-1$ 个模态相对相位)。一旦所有的模态系数被确定,便可以根据式 (1)、(3) 重建出对应的光斑图像。

2.2 无需预训练的卷积神经网络模式分解

在传统 CNN-MD 方法中,首先需要使用大型计算机视觉数据库对 CNN 进行预训练,以提升神经网络对图片复杂特征的提取能力。然后,利用待测光纤的结构参数计算出各模式归一化的场分布,基于为各模式

随机赋予的模式系数生成大量仿真灰度光斑图作为数据集;相应的模式系数值则被设置为光斑图的标签。生成的光斑图数据集分成训练集、验证集和测试集三个部分。使用仿真光斑图构成的训练集对 CNN 进行再次训练,帮助神经网络从光斑图中学习所蕴含的模式特征。在训练过程中,将训练集多次迭代后输入神经网络中,使得 CNN 的权重与偏置参数朝着使网络输出与标签之间差异最小化的方向更新,直至网络收敛。使用验证集在不同的迭代轮次中实现网络拟合情况监测和网络超参数调整。当再次训练完成的 CNN 经过测试集的泛化性最终检验后,便可对实际拍摄的光斑图实施 MD,整个模式系数预测过程只需 CNN 进行一次前向传播计算便可完成。

但是,预训练过程所使用的大型计算机视觉数据库如 ImageNet 数据集等,通常具有多达 10^7 量级的带标签的图像数据,因此,需要耗费长达数十天甚至更久的时间才能完成 CNN 的预训练,这极大地增加了 CNN-MD 在训练阶段的时间成本与算力资源需求,严重限制了其实际的应用范围。

本文所设计的无需预训练的 PFCNN-MD 架构如图 2 所示,它只由 8 个模块组成,在仅使用一个

GPU 加速的情况下就可以实现对神经网络的快速训练。在模块 3 到模块 7 中,设置了 InceptionNet 型^[18]分支结构来拓宽网络的宽度与深度。通过在各分支上组合多样化的小尺寸卷积核和池化层来并行提取特征,不仅使神经网络对不同尺度的特征都具有强大的学习能力,还有效地避免了网络层数增多带来的容易过拟合以及计算资源使用效率低的问题。因此,所设计的 CNN 结构并不需要进行长达数十天甚至更久的预训练,增强了网络对图片特征的提取能力。所设计结构可以直接在光斑图训练集上实现复杂模式特征的学习与提取。作为 CNN 输入的光斑图由于是单通道灰度图,所以将模块 1 的 7×7 卷积层的输入通道数设置为 1。除了处于模块 8 末端的 1×1 卷积层以外,其余模块的卷积层后面都伴随着一个批归一化(BN)层以及一个 Leaky_ReLU 激活函数。在模块 8 的末端添加了 sigmoid 激活函数,用于保证网络所预测的 $2N-1$ 个模式系数在 MD 任务中的有效性。为了消除 MD 中模式相对相位复共轭因子带来的相位模糊问题,真实的模式相对相位值使用 cos 函数进行处理并线性缩放至 $[0, 1]$ 后作为用于神经网络训练的标签。

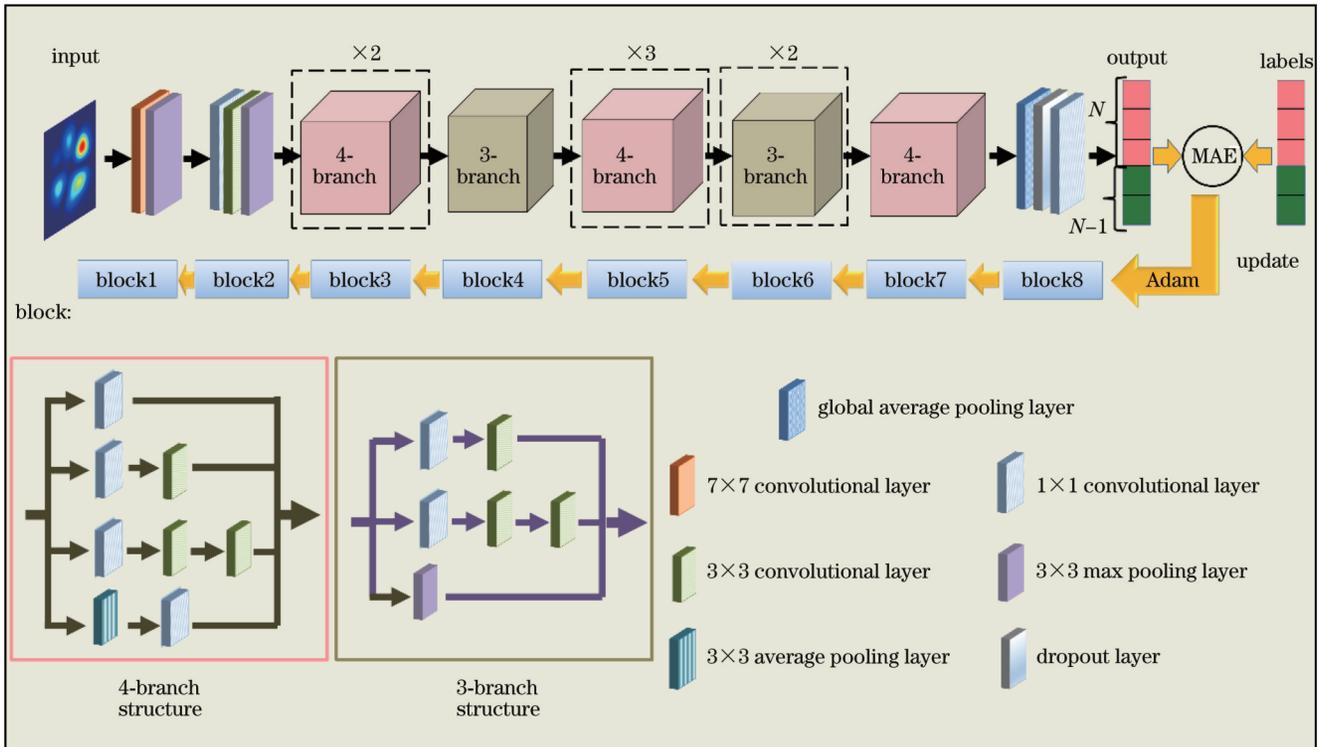


图 2 本文提出的卷积神经网络结构示意图
Fig. 2 Illustration of proposed CNN architecture

采用平均绝对误差 (MAE) 作为损失函数,即

$$L = \frac{1}{T} \sum_{j=1}^T \sum_{i=1}^{2N-1} |y_o^{(j)}[i] - y_i^{(j)}[i]|, \quad (4)$$

式中: T 代表训练集的样本数量; N 表示光纤中所支持的模式数量; y_i 和 y_o 分别为标签向量和 CNN 预测数

值。由于网络预测值的范围为 $[0, 1]$, 相较于传统数值 MD 方法使用的均方根误差 (MSE), 使用 MAE 作为损失函数将会使得网络收敛速度变得更快, 更适用于模式系数分析任务的误差计算。在优化器选择方面, 利用 Adam 优化器来更新网络参数。

3 结果与讨论

本文提出的 CNN 架构部署在配置了一个 AMD R9-5900X CPU 和一个 NVIDIA RTX 3060 GPU 的台式计算机中,其中神经网络每一层的权重和偏置参数都使用随机初始化。

3.1 数值仿真测试与分析

在仿真测试中,选择了图 1 所示的 RCF 来验证 PFCNN-MD 算法的性能。在 1550 nm 的工作波长下,将光纤中支持传输的前 8 个模式 (LP₀₁、LP_{11e}、LP_{11o}、LP_{21e}、LP_{21o}、LP_{31e}、LP_{31o}、LP₀₂ 模式,下标 e 和 o 用于区分偶模与奇模) 作为示例。根据模式简并性,本文分析 4 种模式组合 (前 3、5、7、8 个模式叠加组成的模式组) 情况下的分解性能。针对这 4 种模式叠加情况,在测试过程中分别生成了 100000、150000、200000、250000 幅 128×128 分辨率下的仿真光斑图作为训练集。在不进行预训练的情况下,直接利用由光斑图构造的训练集对初始化的 CNN 进行训练。针对每一种模式叠加情况都生成另外 2000 幅相同分辨率下的仿真光斑图作为测试集。模块 8 中的丢弃层比率在测试中设置为 0.8, Adam 优化器的动量参数则设置为 $\beta_1=0.9$, $\beta_2=0.999$ 。随着训练轮次 (epoch) 的增加,设置的学习率逐渐从 10^{-4} 下降到 10^{-6} 。

将模态权重误差 $\Delta\rho^2=|\rho_p^2-\rho_t^2|$ 和模态相对相位误差 $\Delta\theta=|\theta_p-\theta_t|/2\pi$ 作为指标分析 MD 的精度^[14],其中下标 p 和 t 分别表示网络预测和实际的模态系数。利用相关性系数来表征基于预测的模态系数重建的光斑图像与源光斑图像之间的相似程度,相关性系数^[8]可表示为

$$C = \left| \frac{\iint \Delta I_r(r, \varphi) \Delta I_m(r, \varphi) r dr d\varphi}{\sqrt{\iint \Delta I_r^2(r, \varphi) r dr d\varphi \iint \Delta I_m^2(r, \varphi) r dr d\varphi}} \right|, \quad (5)$$

式中: $\Delta I_j(r, \varphi) = I_j(r, \varphi) - \bar{I}_j$ ($j=r, m$), \bar{I}_j 表示重建图像光束强度 I_r 和实际图像光束强度 I_m 之间的平均值。相关性系数的取值范围为 0~100%。

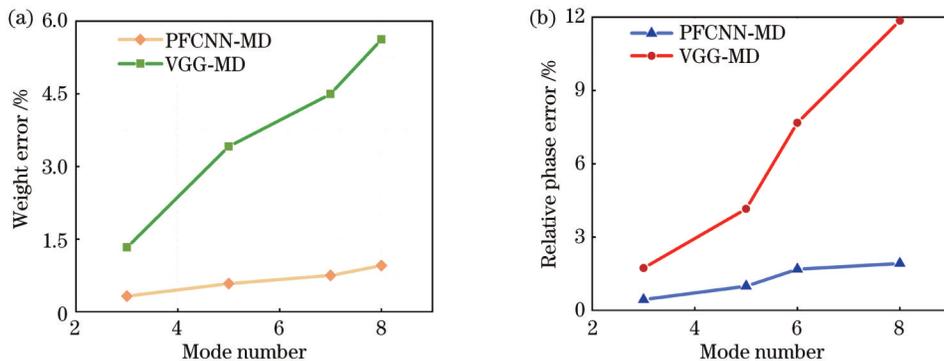


图 4 模态系数误差对比示意图。(a)模态权重误差;(b)模态相对相位误差

Fig. 4 Comparison of modal coefficient error. (a) Modal weight error; (b) modal relative phase error

图 3 展示了 PFCNN-MD 与传统的 VGG-MD 方案^[14]在不使用预训练的条件下,直接使用光斑图数据集训练时损失函数在每个 epoch 中的变化情况。可以发现:即使在 3 模式叠加情况下,VGG-MD 的损失函数也一直在较高的损失值上处于振荡状态而无法收敛,这意味着 CNN 无法从光斑图训练集中学习到有效的特征,使得 MD 的误差非常大。而 PFCNN-MD 中的损失函数则快速收敛至 0.007。本文所设计的 CNN 架构在 4 种不同模式组叠加情况测试中都可以在约 55 个 epoch 的训练之内达到收敛状态,极大地减少神经网络训练的时间成本和计算资源的消耗。

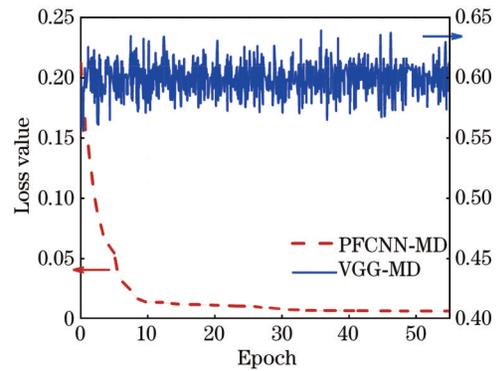


图 3 3 模式叠加情况下两种不同 MD 方案下作为训练轮次函数的卷积神经网络损失

Fig. 3 Loss value of CNN as a function of epochs for two MD methods in 3-mode superposition case

在 3 个模式到 8 个模式叠加的情况下,未经过预训练的 PFCNN-MD 与经过 ImageNet 数据集预训练的 VGG-MD 的模态权重和模态相对相位误差结果如图 4 所示。从图 4(a)中可知:在 3 模式叠加情况下,PFCNN-MD 的模态权重误差仅为 0.32%;当模式数增长至 8 个,相比于 VGG-MD,模态权重误差从 5.62% 减小至 0.95%,精度提升了 5 倍以上。而在图 4(b)展示的模态相对相位误差对比中可以看出:3 模式叠加情况下,模态相位误差从 1.73% 下降到了 0.43%;而在 8 个模式情况下,相对相位误差则从

11.86% 下降到了 1.92%，精度提升了约 8 倍。结果表明：PFCNN-MD 算法拥有强大的特征学习能力，即使在不使用预训练的情况下仍可获取高精度的 MD 结果。与经过预训练的传统 CNN-MD 相比，PFCNN-MD 算法仍然具有大幅度误差减小与训练时间减少的优势。

图 5 则展示了不同模式叠加情况下一些典型的重建光斑图像示例，其中包括基于 PFCNN-MD 重建的光斑与源光斑图像对，以及二者之间对应的残差图像

和相关性系数值。在 3 个模式叠加情况下，重建光斑图像与源图像之间的平均相关性系数达到 99.99%，所提出的算法执行一次 MD 的时间为 5 ms。当模式数上升至 8 个时，平均相关性系数仍然保持在 99.60%，神经网络完成一次 MD 也仅需 9 ms。这证明重建的图像与源光斑图像之间具有高度相似性，即使在多模式叠加的情况下也可以实现高精度的实时 MD。图 5 中 ORI 为源光斑图像，REC 为重建光斑图像，RES 为残差图，COR 为相关性系数。

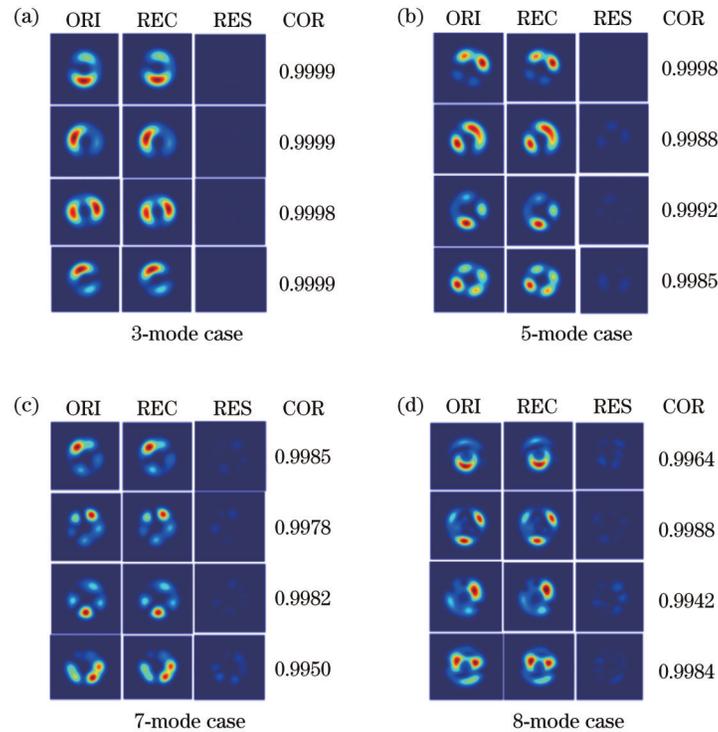


图 5 不同模式叠加情况下的光斑图经典重建案例。(a) 3 模式案例；(b) 5 模式案例；(c) 7 模式案例；(d) 8 模式案例

Fig. 5 Typical reconstructed spot examples for different mode superposition. (a) 3-mode case; (b) 5-mode case; (c) 7-mode case; (d) 8-mode case

为了进一步表征 PFCNN-MD 的性能，如表 1 所示，将所提出的方案与近年来出现的 4 种 CNN-MD 方法（VGG^[14]、NFFL-CNN^[17]、DH-ResNet18^[16] 和 DenseNet121^[15]）在多种指标上进行了对比。从表 1 中可知，在 8 模式叠加情况下，即使 RCF 具有更复杂的模式重叠和耦合特性，PFCNN 分解精度也仍高于其他 CNN 方案在阶跃型 FMF 时的分解精度。PFCNN 执行一次 MD 消耗的时间处于 ms 量级。并且，PFCNN-MD 无需预训练过程，只需要近场 (NF) 光斑图用于神经网络的训练和算法输入，在具有更低实验操作复杂度与设备要求的同时，对于训练的硬件设备要求也降低，还节省了大量用于预训练的时间成本。此外，PFCNN 只需要使用单个 GPU 便可实现更高精度的 MD，节省了大量的计算机算力资源和能耗。但是，在 8 模式叠加情况下，相比于 VGG、NFFL-CNN 和 DH-ResNet18 算法只需要少于 40 轮 (epoch) 的训练

便可实现神经网络收敛，PFCNN 需要经历约 55 轮训练才能使网络收敛，耗费的训练时间更长。不过，PFCNN 方案可以提前完成网络训练，在 MD 过程中利用训练好的网络进行一次前向传播计算便可获取高精度结果，消耗时间仅处于 ms 量级。因此，通过多种 CNN-MD 方案之间的性能对比可以证明，PFCNN 算法具有更强的学习能力，在无需预训练的情况下便可实现高精度且实时的 MD，对实验和计算设备要求较低，操作简便，有利于 CNN-MD 应用范围的扩大，在更复杂的 RCF 模式耦合特性表征上可获取高精度的结果。

3.2 实验测试与分析

基于全光纤器件的光斑采集实验装置如图 6 所示，根据实验条件，利用光斑采集装置获取真实的光斑图像，再利用完成训练的 PFCNN 算法对真实光斑图实施 MD，对 3 个模式叠加情况下的 PFCNN-MD 方案

表 1 8 模式叠加情况下多种 CNN-MD 方法性能比较
Table 1 Comparison of multiple CNN-MD methods in 8-mode superposition case

Method	FMF RI profile	$\Delta\rho^2 / \%$	$\Delta\theta / \%$	One MD time / ms	Pre-training dataset	Input request	Equipment requirement
VGG	Step-index	2.20	3.90	~21	ImageNet	One NF	Single GPU
NFFL-CNN	Step-index	1.10	2.25	~52	None	One NF	Single GPU
DH-ResNet18	Step-index	1.25	2.01	~10	None	One NF One PD	Single GPU
DenseNet121	Step-index	1.80	2.60	~6	None	One NF	Two GPUs
PFCNN	Ring-core	0.95	1.92	9	None	One NF	Single GPU

性能进行了检验。在真实光斑采集实验装置中,可调制激光器(Santec TSL-510)作为光源,发射的光束经偏振控制器后被光分束器分为三束。三通道的光子灯笼(LP₀₁、LP_{11e}和 LP_{11o})作为模式转换器,其输出端直接与 RCF 相连接,实现三个模式的同时激发。在 1550 nm 工作波长下,RCF 的环形纤芯的内径与外径分别为 5.80 μm 和 9.40 μm ,且具有标准的包层尺寸。

环形纤芯与包层的折射率差 Δn 为 0.012,中心区域的 RI 与包层相等。经 RCF 传输后的光束在输出端经过物镜扩束后由 CCD 相机(Bobcat-320-GigE, 像素尺寸: 20 μm)拍摄并成像。通过调节偏振控制器及弯曲光纤,总共采集了 100 幅真实的 3 个模式叠加情况下具有不同分布的光斑图像。图 7 中 real 为真实光斑图,REC 为重建光斑图,COR 为相关性系数。

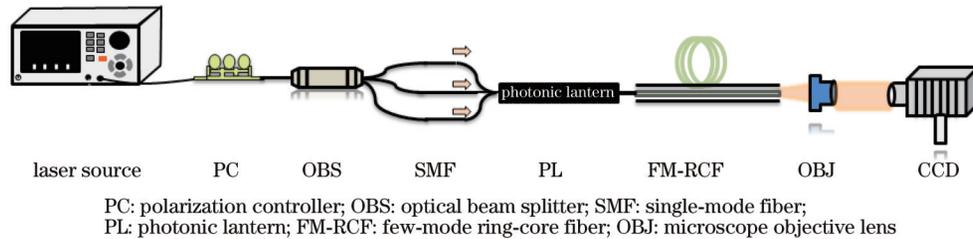


图 6 实际光斑采集装置示意图

Fig. 6 Experimental setup for recording real beam patterns

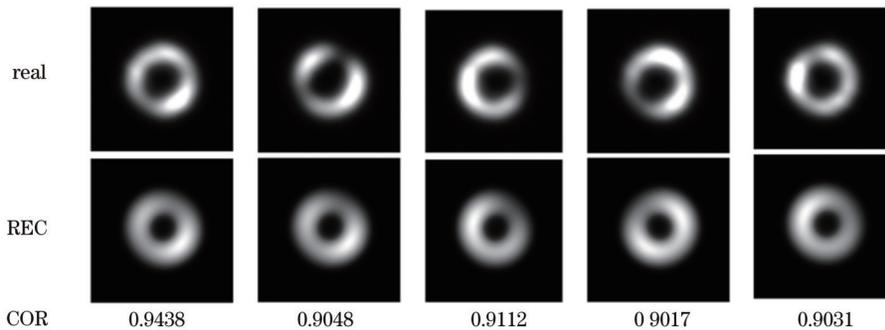


图 7 3 模式叠加实验光斑经典重建示例。

Fig. 7 Typical reconstructed examples based on real spot patterns for 3-mode superposition case

将实际拍摄的光斑图处理成 128×128 分辨率的图片后作为算法输入,对其模态系数进行预测,进而重建光斑图像,并计算其与实际的光斑图像之间的相关性系数。选取了其中的 5 对原图像及重建图像作为示例,如图 7 所示。测试数据集中的真实图像和重建图像之间的相关性系数平均值为 90.01%。实验 MD 的精度与数值仿真结果相比存在一定的差距,这是由于:实验光斑样本由普通 CCD 拍摄,存在器件噪声以及中心偏移问题,给 MD 精度带来负面影响。后续实验优化中将引入去噪设备以及图像中心自适应调控算法,提升 PFCNN-MD 的实际抗干扰能力。

4 结 论

针对 RCF 中模场重叠与模式耦合表征较为复杂的现象,提出了一种高精度 PFCNN-MD 算法。所提出的算法在不使用预训练的情况下,可快速完成训练并获取高精度的 RCF 模态系数。从仿真与实验两个方面对 PFCNN-MD 的性能进行了检验。在数值仿真中,8 模式叠加情况下的模态权重和模态相对相位误差低于 0.95% 和 1.92%,分别减小为传统 CNN-MD 的 1/5 和 1/8,完成一次 MD 的时间为 9 ms。在实验中,实际拍摄的光斑图像与重建光斑图像的相关性高

于 90%。所提出的 PFCNN 算法在 RCF 实时 MD 和有关模式耦合特性的表征中显示出巨大潜力。

参 考 文 献

- [1] 郑晶晶, 宋豫婧, 裴丽, 等. 面向空分复用的矩形辅助环芯光纤设计及特性[J]. 光学学报, 2022, 42(16): 1606006.
Zheng J J, Song Y J, Pei L, et al. Design and characteristics of rectangular-assistant ring-core fiber for space division multiplexing[J]. Acta Optica Sinica, 2022, 42(16): 1606006.
- [2] 雷晓, 任芳, 张艺赢, 等. 面向模分复用的沟槽-纳米孔辅助双包层弱耦合少模光纤[J]. 光学学报, 2021, 41(23): 2306003.
Lei X, Ren F, Zhang Y Y, et al. Trench-nanopore assisted double-clad weakly coupled few-mode fiber for mode division multiplexing[J]. Acta Optica Sinica, 2021, 41(23): 2306003.
- [3] 马绍朔. 抗弯曲大模场面积少模光纤及少模光纤激光器的研究[D]. 北京: 北京交通大学, 2019: 14-20.
Ma S S. Research on bending resistant large mode area few-mode fiber and few-mode fiber laser[D]. Beijing: Beijing Jiaotong University, 2019: 14-20.
- [4] Wright L G, Christodoulides D N, Wise F W. Spatiotemporal mode-locking in multimode fiber lasers[J]. Science, 2017, 358(6359): 94-97.
- [5] Krupa K, Tonello A, Shalaby B M, et al. Spatial beam self-cleaning in multimode fibres[J]. Nature Photonics, 2017, 11(4): 237-241.
- [6] 阮江冉, 裴丽, 郑晶晶, 等. 基于包层泵浦的 4 模掺铒光纤放大器的增益均衡[J]. 光学学报, 2022, 42(4): 0406001.
Ruan J R, Pei L, Zheng J J, et al. Gain equalization of 4-mode erbium-doped fiber amplifier based on cladding pump[J]. Acta Optica Sinica, 2022, 42(4): 0406001.
- [7] Tan H Y, Zhang J W, Liu J, et al. Low-loss ring-core fiber supporting 4 mode groups[C]//2019 Conference on Lasers and Electro-Optics (CLEO), May 5-10, 2019, San Jose, CA, USA. New York: IEEE Press, 2019.
- [8] Zuo M Q, Ge D W, Liu J X, et al. Long-haul intermodal-MIMO-free MDM transmission based on a weakly coupled multiple-ring-core few-mode fiber[J]. Optics Express, 2022, 30(4): 5868-5878.
- [9] Zhu L, Li J X, Zhu G X, et al. First demonstration of orbital angular momentum (OAM) distributed Raman amplifier over 18-km OAM fiber with data-carrying OAM multiplexing and wavelength-division multiplexing[C]//2018 Optical Fiber Communications Conference and Exposition (OFC), March 11-15, 2018, San Diego, CA, USA. New York: IEEE Press, 2018.
- [10] 张澍霖, 冯国英, 周寿桓. 基于空间域和频率域傅里叶变换 F^2 的光纤模式成分分析[J]. 物理学报, 2016, 65(15): 154202.
Zhang S L, Feng G Y, Zhou S H. Fiber modal content analysis based on spatial and spectral Fourier transform[J]. Acta Physica Sinica, 2016, 65(15): 154202.
- [11] Zhao Y, Huang S J, Yan C. Parallel measurement of multiple linear polarization modes in few-mode optical fibers using spatial light modulators[J]. Optical Engineering, 2019, 58(9): 094103.
- [12] Manuylovich E, Donodin A, Turitsyn S. Intensity-only-measurement mode decomposition in few-mode fibers[J]. Optics Express, 2021, 29(22): 36769-36783.
- [13] Li L, Leng J Y, Zhou P, et al. Multimode fiber modal decomposition based on hybrid genetic global optimization algorithm[J]. Optics Express, 2017, 25(17): 19680-19690.
- [14] An Y, Huang L J, Li J, et al. Learning to decompose the modes in few-mode fibers with deep convolutional neural network[J]. Optics Express, 2019, 27(7): 10127-10137.
- [15] Rothe S, Zhang Q, Koukourakis N, et al. Intensity-only mode decomposition on multimode fibers using a densely connected convolutional network[J]. Journal of Lightwave Technology, 2021, 39(6): 1672-1679.
- [16] Zhu Z H, Xiao Y Y, Yao R M. CNN-based few-mode fiber modal decomposition method using digital holography[J]. Applied Optics, 2021, 60(24): 7400-7405.
- [17] Fan X J, Ren F, Xie Y L, et al. Mitigating ambiguity by deep-learning-based modal decomposition method[J]. Optics Communications, 2020, 471: 125845.
- [18] Sergey I, Christian S. Batch normalization: accelerating deep network training by reducing internal covariate shift[EB/OL]. [2015-06-01]. <https://proceedings.mlr.press/v37/ioffe15.html>.

Accurate Mode Decomposition for Ring Core Fibers Based on Deep Learning

Tian Zichen¹, Pei Li^{1*}, Wang Jianshuai¹, Bai Bing¹, Hu Kaihua¹, Zheng Jingjing¹,
Shen Lei^{2,3,4}, Xu Wenxuan¹

¹Key Laboratory of All Optical Network and Advanced Telecommunication Network, Ministry of Education, Beijing Jiaotong University, Beijing 100044, China;

²State Key Laboratory of Optical Fiber and Cable Manufacture Technology, Wuhan 430073, Hubei, China;

³Yangtze Optical Fiber and Cable Joint Stock Limited Company (YOFC), Wuhan 430073, Hubei, China;

⁴Hubei Optics Valley Laboratory, Wuhan 430073, Hubei, China

Abstract

Objective The core of a ring-core few-mode fiber (RCF) is composed of a central refractive index (RI) depression region and an outer high RI ring. The RCF plays an important role in modal gain equalization, mode division multiplexing transmission, and vortex beam generation. In order to figure out the mode coupling and reveal the associated beam properties in few-mode fibers (FMFs), the mode decomposition (MD) techniques are required, which can obtain the modal weight (ρ^2) and modal relative phase difference (θ) from modal superposition images. However, the RCF has complex modal overlap and

mode coupling because the power of each mode is confined to the same high RI ring region. As a result, the mode coupling analysis of RCFs faces severe challenges. In this paper, we propose a pretraining-free CNN-MD algorithm (PFCNN-MD) based on a convolutional neural network (CNN) for high-accuracy characterization of complex couplings in RCFs, and the algorithm uses branch structures with different receptive fields to improve the learning ability of neural networks.

Methods In the PFCNN-MD, the normalized field distribution of each supported mode is calculated based on the tested RCF's structural parameters first. After that, massive simulated grayscale beam patterns can be generated numerically with random modal coefficients as the dataset. The corresponding modal coefficient values are set as the label. The generated beam pattern dataset is divided into three parts: training set, validation set, and test set. The CNN is trained by using the training set, which helps the neural network learn the modal features from the beam patterns. During the training process, the training set is iteratively input into the network. The weight and bias parameters of the CNN are updated by minimizing the difference between the network output and the label until the CNN converges. The validation set is used to monitor network fitting and tune network hyperparameters. After the trained CNN has been examined for generalization based on the test set, MD can be implemented on the real beam pattern. The entire process can be completed with only one forward propagation calculation by the trained CNN. The designed PFCNN-MD architecture consists of eight blocks (Fig. 2). In blocks 3 to 7, the InceptionNet-type branch structures are set to increase the width and depth of the network. A variety of small-sized convolution kernels and pooling layers are combined on each branch to extract features in parallel. This not only enables the neural network to have a strong learning ability for features of different scales but also effectively avoids the defects of overfitting and inefficient use of computing resources. Therefore, the designed CNN structure does not require pre-training to enhance the network's ability. Instead, the extraction of complex modal features can be realized directly on the beam pattern training set.

Results and Discussions In the simulation, the first eight modes (LP_{01} , LP_{11e} , LP_{11o} , LP_{21e} , LP_{21o} , LP_{31e} , LP_{31o} , and LP_{02}) that support the selected RCF are used as examples with a wavelength of 1550 nm. Four mode-superposition cases (superposition of the first three, five, seven, and eight modes) are analyzed to test the performance of the PFCNN-MD. By comparing the PFCNN-MD with traditional VGG-MD schemes without pre-training, it can be found that even in the three-mode case, the loss function in the VGG-MD always oscillates at higher loss values and cannot converge, while that in the PFCNN-MD quickly converges to 0.007 (Fig. 3). Moreover, the designed PFCNN architecture can achieve convergence within about 55 epochs of training in the four mode superposition cases. The error results of ρ^2 and θ in the superposition case with three to eight modes are compared between the non-pre-trained PFCNN-MD and the VGG-MD pre-trained on the ImageNet dataset. It can be found that in the eight-mode case, compared with that by VGG-MD, the error of ρ^2 is reduced from 5.62% to 0.95%, and the accuracy is improved by more than five times. The error of θ drops from 11.86% to 1.92%, and the accuracy is improved by about eight times (Fig. 4). In the three-mode case, the average correlation coefficient between the reconstructed beam pattern and the original beam pattern reaches 99.99%, and one MD consumes 5 ms. When the mode number becomes eight, the average correlation remains at 99.60%, and the trained CNN completes one MD in only 9 ms (Fig. 5). In order to further characterize the performance of PFCNN-MD, the proposed method is compared with four CNN-MD methods (VGG, NFFL-CNN, DH-ResNet18, and DenseNet121) emerging in recent years on various indexes (Table 1). In the eight-mode case, even if the RCF has more complex modal coupling characteristics, the accuracy of the PFCNN is still higher than that of other CNN-MD schemes in the step-index FMF. The time consumed by PFCNN to perform one MD is on the order of ms. In addition, the PFCNN-MD doesn't require the pre-training process and only needs a near-field (NF) beam pattern for training and algorithm input. It has lower experimental operation complexity and equipment requirements. Much time for pre-training is also saved. Moreover, the PFCNN only uses a single GPU to achieve more accurate MD, saving many computer computing resources and reducing energy consumption. Although PFCNN needs to go through about 55 epochs of training to converge the network, it can complete in advance. In the experiment, an experimental setup based on an all-fiber device is used to collect real beam pattern to examine the practical performance of the PFCNN. The average correlation coefficient between the real and reconstructed patterns in the test dataset is 90.01% (Fig. 7).

Conclusions In this paper, a high-precision PFCNN-MD algorithm is proposed to solve the problem of complex mode coupling characterizing in the RCF. The proposed algorithm can fast complete the training and obtain high-precision MD results without pre-training. The performance of PFCNN-MD is tested from both simulation and experiment. In the simulation, compared with that by the traditional CNN-MD, errors of ρ^2 and θ in the eight-mode case are lower than 0.95% and 1.92%, which are decreased by 80% and 87.5%, respectively. One MD consumes 9 ms. In the experiment, the correlation between the real and the reconstructed beam patterns is higher than 90%. The PFCNN-MD algorithm shows great potential in real-time MD and the characterization of the RCF's mode coupling properties.

Key words optical communications; few-mode fiber; mode decomposition; deep learning; ring-core fiber