

基于物理约束神经网络的单模光纤 非线性效应高精度解析*

祝沐¹⁾ 佟首峰¹⁾ 丁蕴丰²⁾ 张鹏^{1)†}

1) (长春理工大学, 空间光电技术国家地方联合工程研究中心, 长春 130022)

2) (长春理工大学机电工程学院, 长春 130022)

(2025年6月20日收到; 2025年9月2日收到修改稿)

针对单模光纤中四波混频-受激拉曼散射 (FWM-SRS) 强非线性耦合效应难以解析的问题, 本文提出了一种融合物理机理与神经网络的多尺度物理约束网络 (MSPC-Net). 该模型通过将非线性薛定谔方程 (NLSE) 的频域残差作为物理约束项嵌入网络优化过程, 并设计多尺度空洞卷积模块融合局部细节、中程展宽及长程衰减特征, 实现了光谱成分分离与物理参数的联合高精度反演. 在 250 m 与 500 m 单模石英光纤实验中, MSPC-Net 重建斯托克斯光谱的均方根误差 (RMSE) 分别低至 0.014 与 0.0173, 较传统卷积神经网络降低超 68%; 其频率偏移预测的平均绝对误差分别为 0.03 nm 和 0.04 nm, 精度较现有方法提升约 90%. 在信噪比 (SNR) 为 6 dB 的噪声环境下, MSPC-Net 对 FWM 次峰信息的检测正确率高达 95.3%, 伪峰率低于 4.7%. 模型得益于物理约束的引导及轻量化结构设计, 在 SNR = 15 dB 噪声下 RMSE 增幅仅 9.8%, 并具备良好的实时处理能力, 可部署于嵌入式设备, 为高功率光通信系统优化与分布式光纤传感提供高效解决方案. 本研究通过将严格物理规律与多尺度特征提取相结合, 有效解决了长距离光纤复杂非线性效应的解析难题, 显著提升了预测结果的理论符合度与噪声鲁棒性.

关键词: 非线性光学, 物理约束神经网络, 多尺度特征提取, 光谱分离

PACS: 42.65.-k, 07.05.Mh, 42.30.Va

DOI: 10.7498/aps.74.20250804

CSTR: 32037.14.aps.74.20250804

1 引言

光纤中的非线性光学效应是制约高功率激光传输和精密光谱分析的关键因素之一, 其中受激拉曼散射^[1,2](SRS) 是一种典型的非线性效应. 在该效应中发生了分子振动能级间的受激跃迁, 从而产生了斯托克斯光, 被广泛应用于光谱检测和生物成像等领域; 四波混频^[3,4] (FWM) 效应是另一种比较常见的非线性效应, 其频率变换是基于多光子作用. FWM 效应与 SRS 的耦合可能会引发光谱频域畸变、能量异常转移等现象^[5]. 对于单一模式光

纤, 在高功率泵浦下^[6], 所产生的 SRS 和 FWM 效应会形成强烈的模式竞争, 二者共同作用, 决定了输入光功率沿光纤长度的演化规律. 为了探索这种影响规律, 重点开展二者相互竞争的动态耦合理论的研究工作.

过去几年来, 学者们运用理论和实验相结合的方法取得了一些关于 SRS-FWM 相互作用方面的成果, FWM 可通过相位匹配条件对 SRS 的斯托克斯光谱特性进行调制, 例如诱导低阶斯托克斯光出现次峰或者诱发高阶光谱展宽; 理论方面, 基于非线性薛定谔方程 (NLSE)^[7] 描述光纤中的非线性效应动力学演变规律普遍被学者采用, 但由于需要

* 国家自然科学基金区域联合基金重点支持项目 (批准号: U22A2008)、吉林省自然科学基金 (批准号: YDZJ202301ZYTS394) 和水声对抗技术重点实验室开放基金 (批准号: CX-2022-032) 资助的课题.

† 通信作者. E-mail: zp@cust.edu.cn

进行大量复杂的数值计算求解因而难以进行非线性动态过程的实时分析;深度学习技术^[8]的使用则为非线性光学研究提供了一种新的研究方式,例如:Li等^[9]运用深度学习分子动力学方法计算了氮化硅高温微波介电函数;Fang等^[10]将物理信息神经网络(PINN)用于求解多粒子低温等离子体流体模型等。

目前,针对单个非线性效应和通用光谱重建的相关研究已经相对较为充分,但是对于SRS-FWM这种特殊的非线性动态耦合作用问题尚缺少一整套进行智能化分析的有效方案,主要包括两方面的缺陷:一是传统模型都是人工提取特征,很难捕捉到SRS-FWM这种复合光谱中所蕴含的时间、空间与频率的相关性;二是由于基于数据驱动的方法没有将物理规律嵌入其中,导致其得到的预测结果难以准确地还原到实验的现象当中,即二者之间存在着系统误差.为了完成这一工作目标,迫切需要建立一种物理机理和数据驱动相结合的SRS-FWM耦合效应联合建模分析方法,对SRS-FWM耦合效应开展精细建模和快速解析。

本文基于单模石英光纤,用脉冲泵浦实验定量表征了FWM调制SRS光谱的规律,并基于此设计出了一种多尺度^[11]物理约束网络(MSPC-Net),采用该模型的动态频域残差调控和双向时序特征提取模块能够同时对SRS和FWM贡献谱^[12]实施自动分离并开展参数反演工作,并获得了光纤非线性效应协同调控的理论依据以及高功率激光器的设计^[13]、光纤传感^[14]等工作的智能化分析手段。

2 实验装置与模型构建

2.1 实验装置

高峰值功率Nd³⁺:YAG脉冲激光器泵浦单模光纤产生非线性效应实验装置的基本原理如图1(a)所示.实验装置启动前如图1(b)所示,抽运光源选用Nd³⁺:YAG脉冲激光器配合进行腔外倍频,并通过电光调Q技术实现脉冲输出.在开启实验装置后,如图1(c)所示.经泵浦得到的1064 nm近红外短波在通过KTP晶体^[15]倍频后得到的绿光,经光谱仪实测,激光输出中心波长为(531.81 ± 0.18) nm,脉宽约为10 ns,系统最大输出峰值功率可达55 MW以上,利用两片偏振片基于马吕斯定律实现能量精细的线性调节.聚光腔输出端配置了直径1.4 mm的针孔光阑用于模式选

择,获得近基模光束轮廓,并随后使用 $f = 26$ mm的凸透镜组进行聚焦,借助高精度光学调整架将光束耦合^[16]至由山东威海长和光导科技生产的G652D型单模石英裸纤内,该光纤纤芯直径10 μm,包层直径125 μm,色散系数不大于3.5 ps/(nm·km).为抑制环境光对受激拉曼散射光的干扰,收集区域采用暗室环境设计,光纤输出光斑投射在硫酸纸制作的光屏上,并使用Ocean Optics公司的USB2000+型光纤光谱仪精确采集受激拉曼散射光谱,光谱仪分辨率为±1.82 nm.

2.2 MSRA-Net 模型设计

MSRA-Net是一种基于解析光纤SRS和FWM非线性耦合效应的多尺度残差注意力网络,提出利用网络降噪带来的结合多尺度空洞卷积(DHCNN)融合得到的多尺度联合表征输出(MSJRO)结合物理信息神经网络(PINN)的优化模型,以及利用动态频域残差调控和双向时序特征构建解耦判别模型。

模型的输入是采集到的实验数据的光谱图,在此输入下使用DHCNN来提取特征,将多尺度卷积模块分解成3个并行的卷积分支.包括:1)传统卷积分支(dilation = 1),可获取光谱的局部细节特征;2)中程空洞卷积分支(dilation = 3),可以扩展感受野大小,从而检测出中程范围的光谱宽度;3)长程空洞卷积分支(dilation = 5),用来判断长程距离波长范围内能量衰减的情况.其中,dilation为空洞率,dilation > 1即为空洞卷积分支。

经3个分支分别扫描得出不同距离的特征图像,再经过Concat(·)函数的合并,并在此基础上使用一个1×1卷积核Conv(·)进行特征图像融合以及提升特征相互之间的关联度,产生MSJRO,可以表示为

$$\mathbf{F}_{\text{MSJRO}} = \text{Conv}_{1 \times 1}(\text{Concat}(\mathbf{F}_1, \mathbf{F}_2, \mathbf{F}_3)), \quad (1)$$

其中, \mathbf{F}_1 , \mathbf{F}_2 , \mathbf{F}_3 分别对应3个分支的特征图像,Conv_{1×1}表示1×1卷积操作。

MSRA-Net采用非线性薛定谔方程(NLSE)频域解作为物理约束,在模型训练中加入物理残差损失函数^[17]:

$$\mathcal{L}_{\text{phy}} = \frac{1}{N} \sum_{i=1}^N \left\| \mathcal{F} \left(\frac{\partial A_i}{\partial z} \right) - \mathcal{F} \left(\frac{i\beta_2}{2} \frac{\partial^2 A_i}{\partial T^2} + i\gamma |A_i|^2 A_i \right) \right\|_2, \quad (2)$$

式中, A_i 表示第*i*个样本的光场复振幅预测值,

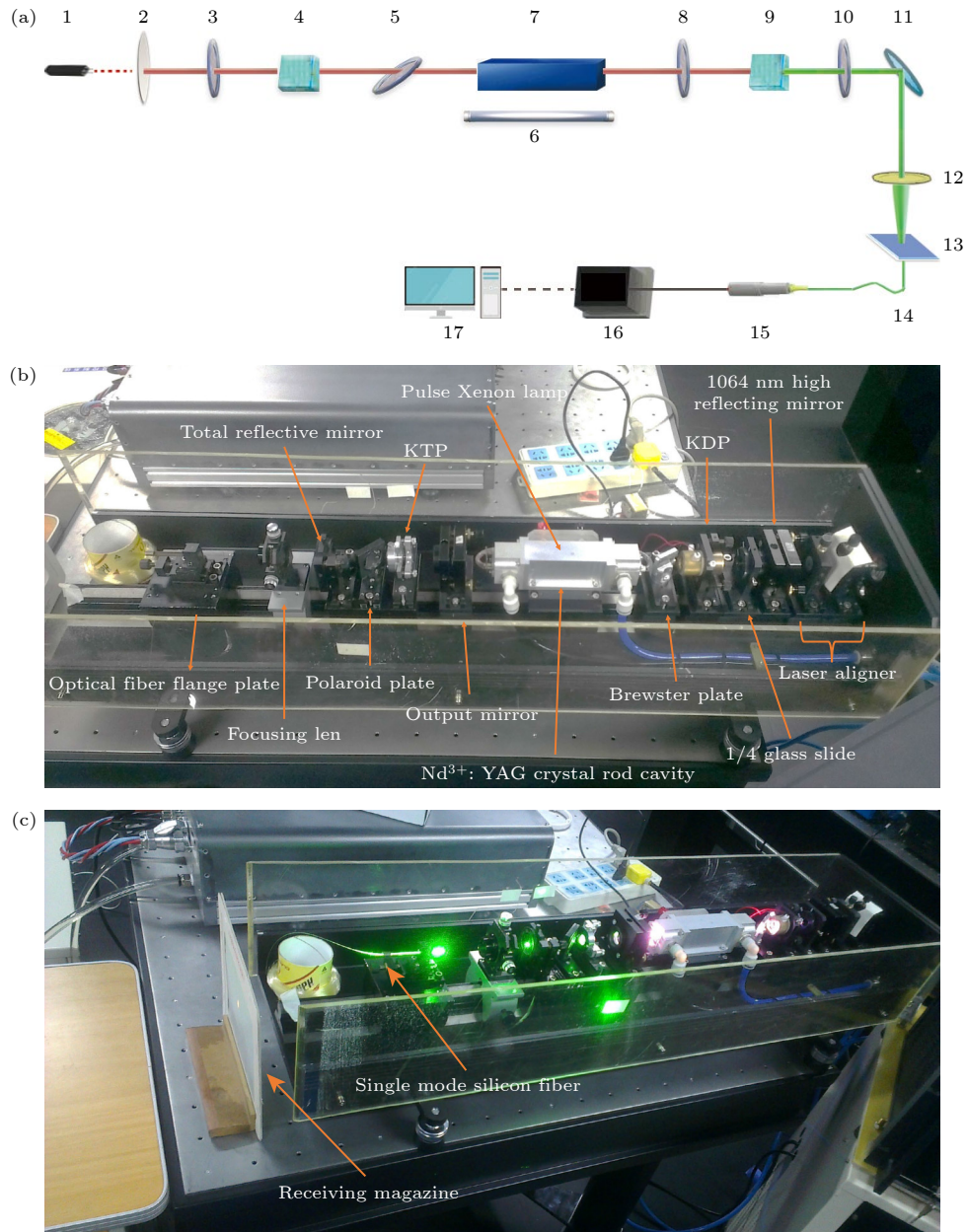


图 1 高功率 Nd³⁺: YAG 脉冲激光器泵浦单模光纤产生非线性效应实验装置 (a) 实验装置原理图, 其中 1 为激光对准器, 2 为 1064 nm 高反射镜, 3 为 1/4 载玻片, 4 为磷酸二氢钾晶体 (KDP), 5 为布儒斯特板, 6 为脉冲氙灯, 7 为 Nd³⁺: YAG 晶体棒腔, 8 为输出镜, 9 为磷酸氧钛钾晶体 (KTP), 10 为偏振片, 11 为全反射镜, 12 为聚焦透镜, 13 为光纤法兰盘, 14 为单模硅光纤, 15 为接收仓, 16 为光谱仪, 17 为计算机; (b) 装置启动前; (c) 装置启动后

Fig. 1. Experimental setup for nonlinear effects generated by pumping single-mode fiber with high peak power Nd³⁺: YAG pulse laser: (a) Experimental device schematic, where 1 represents laser aligner, 2 represents 1064 nm high reflecting mirror, 3 represents 1/4 glass slide, 4 represents Potassium dihydrogen phosphate crystal (KDP), 5 represents brewster plate, 6 represents pulse Xenon lamp, 7 represents Nd³⁺: YAG crystal rod cavity, 8 represents output mirror, 9 represents potassium titanyl phosphate crystal (KTP), 10 represents polaroid plate, 11 represents total reflective mirror, 12 represents focusing len, 13 represents optical fiber flange plate, 14 represents single mode silicon fiber, 15 represents receiving magazine, 16 represents spectrometer, 17 represents computer; (b) before start of the device; (c) after start of the device.

β_2 为群速度色散系数, γ 为光纤非线性系数, \mathcal{F} 表示傅里叶变换, T 为时间域变量, z 为传输距离. 总损失函数由均方误差损失 (\mathcal{L}_{MSE}) 与物理残差损失的加权和构成.

利用 (2) 式可计算出光场复振幅 A_i 的物理残

差 \mathcal{L}_{phy} , 并用该残差衡量网络预测结果与严格物理规律的偏离程度. 在训练中, \mathcal{L}_{phy} 产生的梯度直接作用于网络权重, 迫使网络参数更新不仅减小数据误差, 同时通过测量群速度色散 β_2 、非线性系数 γ 的耦合系数来满足 NLSE 描述的场演化规律.

再将物理残差 \mathcal{L}_{phy} 与数据拟合损失 \mathcal{L}_{MSE} 结合形成总损失 $\mathcal{L}_{\text{total}}$:

$$\mathcal{L}_{\text{total}} = \lambda_1 \mathcal{L}_{\text{MSE}} + \lambda_2 \mathcal{L}_{\text{phy}}, \quad (3)$$

其中, λ_n 为已知参量. 通过不断地调整权值系数, 强制优化方向兼顾数据精度与物理一致性, 经多次贝叶斯优化^[18] 后证实, 当 λ_1 和 λ_2 的值分别为 0.7 和 0.3 时, 光场演化的结果最符合 NLSE 理论的规律性.

针对 FWM 次峰信号微弱且易受噪声干扰的问题, 该模型采用动态分支注意力机制 (DBAM), 对每路分支进行空间维度平均压缩, 把分支中波长方向的信息降维到通道描述向量 $w \in R^C$:

$$w_c = \frac{1}{L} \sum_{l=1}^L F_{\text{MSJRO}}^{(c,l)}, \quad (4)$$

其中, w_c 表示第 c 个分支的全局响应强度, 反映该分支在整个波长范围内的平均激活水平.

通道描述向量 w 通过两层全连接网络 (FC)^[19] 学习通道间的非线性依赖关系, 生成注意力权重 $w' \in R^C$:

$$w' = \sigma(\mathbf{W}_2 \cdot \delta(\mathbf{W}_1 \cdot w)), \quad (5)$$

式中, $\mathbf{W}_1 \in R^{C/r \times C}$ 和 $\mathbf{W}_2 \in R^{C \times C/r}$ 为可学习参数矩阵, r 为压缩比率, δ 为 ReLU 激活函数, σ 为 Sigmoid 函数, 将权重限制在 $[0, 1]$ 区间.

为了进一步加强模型对 FWM 次峰特征频段

的敏感性, 引入了基于物理先验的频段权重增强方法. 对 FWM 次峰主要分布区间的通道权重进行显式提升:

$$w'[j] = \begin{cases} 2w'[j], & \text{if } \lambda_j \in [535, 540] \text{ nm}, \\ w'[j], & \text{otherwise,} \end{cases} \quad (6)$$

其中, λ_j 为通道 j 的中心波长. 增强系数通过交叉验证实验确定, 平衡了信号增强与过拟合风险. 图 2 为 MSRA-Net 和 MSPC-Net 模型架构图.

2.3 MSPC-Net 光谱训练

首先对实际光谱数据进行归一化处理, 并通过滑动窗口分割为局部样本, 其中窗口尺寸为 160 nm, 步长 0.01 nm. 针对实测数据信噪比 (SNR) 较高的场景, 为模拟实际部署中如光纤振动、温度漂移、设备电子噪声等复杂扰动, 注入 $\text{SNR} = 6$ dB 的高斯噪声^[20] 以模拟高 SNR 扰动. 随后, 以实际光谱为基础, 基于分步傅里叶法 (split-step Fourier method, SSFM) 生成的 100 组仿真光谱进行模型预训练, 其中 80 组预训练集、10 组验证集, 10 组测试集, 使模型在噪声可控环境下学习 NLSE 理论框架下如 FWM 等非线性演化规律, 弥补实测数据不足的缺陷; 随后采用 100 组实测光谱进行参数微调, 其中 80 组微调集、10 组验证集, 10 组测试集. 为消除仿真与实测数据间的域差异, 本研究采用协同优化策略: 冻结预训练模型中 90% 的底层

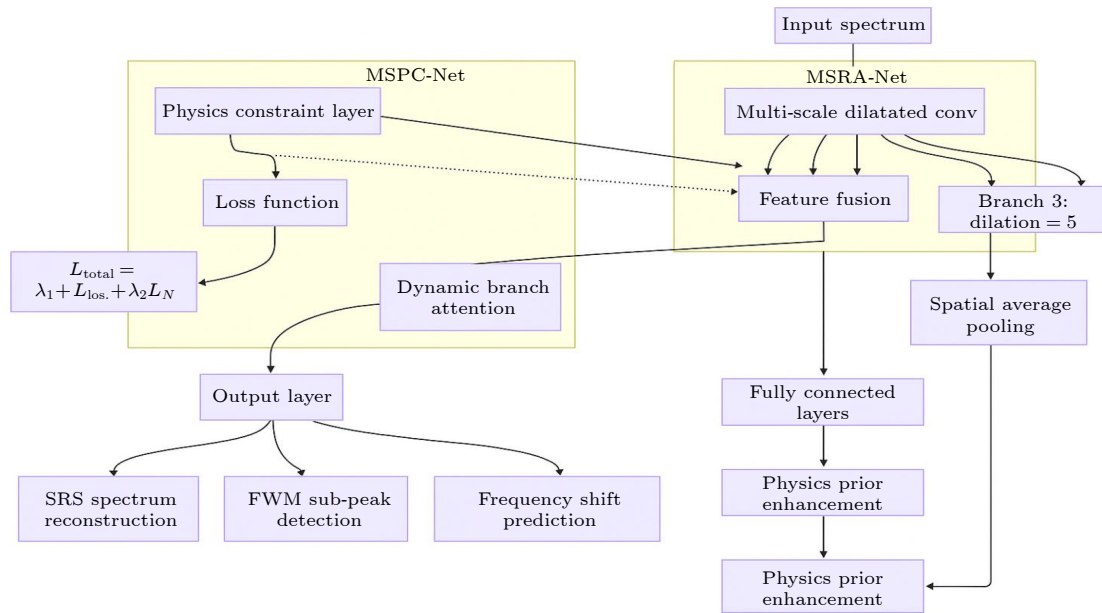


图 2 MSRA-Net 和 MSPC-Net 模型架构图

Fig. 2. MSRA-Net and MSPC-Net model architecture diagram.

卷积参数,重点优化动态通道注意力模块^[21]与物理约束层^[22],以保留通用特征提取能力并抑制过拟合.同时通过对抗性特征分布对齐技术,在特征空间强制仿真与实测数据分布一致,显著缩小域偏移;进一步利用非线性薛定谔方程(NLSE)物理约束项校正仿真参数的系统性偏差,通过动态校准光纤色散(β)与非线性系数(γ),使模型输出严格符合实际光场演化规律.

训练过程中,总损失函数由RMSE与基于频域NLSE约束的物理残差损失加权构成,确保参数设定后的光场演化结果最符合NLSE理论的规律性.动态分支注意力机制(DBAM)通过空间维度平均压缩生成频段权重,对目标区域施加2倍增强,强化FWM次峰特征提取,并同时抑制非关键频段噪声,最后将训练出的MSPC-Net模型在未知长度的单模石英光纤中得到的光谱的训练效果与传统卷积神经网络(convolutional neural network, CNN)以及双向长短期记忆网络(bidirectional long short-term memory, BiLSTM)进行多角度综合对比与分析.

在对比实验中,传统卷积神经网络(CNN)采用5层结构,每层使用步长1,通道数32—512逐层递增的 5×1 卷积核和ReLU激活函数,最终通过全局平均池化输出(GAP),训练时使用学习率0.001为Adam优化器、MSE损失函数和200轮次迭代;双向长短期记忆网络(BiLSTM)为2层双向结构,隐藏单元数128,搭配数值为0.2的Dropout和学习率0.001为Adam优化器.二者训练轮次均与MSPC-Net一致.

3 分析与讨论

为了验证MSPC-Net在不同光纤长度下对非线性效应解析的能力,需要实际光谱的数据作为支持.用示波器实时监测到激光器抽运光工作能量稳定后,在抽运光峰值功率为220 kW,脉宽为10.2 ns的情况下,得到了250 m和500 m单模石英光纤所对应的非线性效应光谱曲线.光谱有效测量范围为520—680 nm之间,如图3和图4所示.

由图3和图4可见,两种光纤非线性光谱的斯托克斯曲线之间产生了不同波长、不同强度的不规则附加峰,原因如下:1) FWM效应会破坏SRS的常规能量传递规律,具体表现为低阶斯托克斯曲线中诱导产生附加峰谱线,例如图3和图4中一阶和

二阶斯托克斯曲线中产生的附加峰曲线;同时引发高阶斯托克斯光谱线展宽和频移现象,例如图3中第9阶斯托克斯曲线的光波长漂移达62.28 nm.2) FWM与SRS的光斑传输模式存在差异,SRS以 HE_{11} 模式传播,而FWM则以 HE_{12} 模式传播,FWM会通过与SRS进行模式竞争以加剧光谱结构的复杂性.

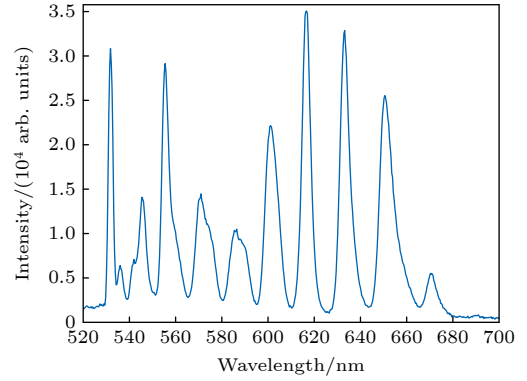


图3 250 m 光纤非线性光谱
Fig. 3. 250 m fiber nonlinear spectrum.

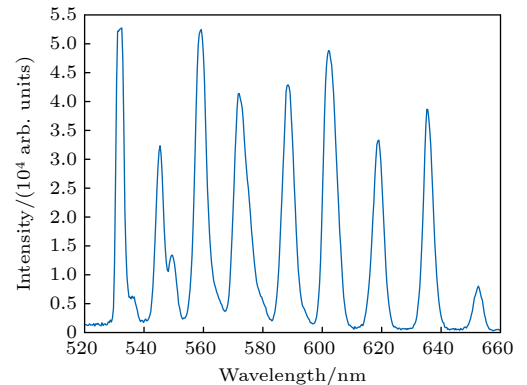


图4 500 m 光纤非线性光谱
Fig. 4. 500 m fiber nonlinear spectrum.

得到实际光谱的数据与图像后,根据MSPC-Net光谱训练中的方法进行的操作,得到训练后的MSPC-Net模型.为了验证其模型优势,下文将从4个方面将其与传统卷积神经网络(CNN)及双向长短期记忆网络(BiLSTM)进行综合对比与分析.

3.1 训练效率与收敛特性

收敛特性与训练效率是评估模型性能的核心指标,决定了模型训练的稳定性与可靠性.通过使用MSPC-Net模型对未知长度的单模石英光纤中得到的光谱数据进行200轮次的迭代,如图5和图6所示.

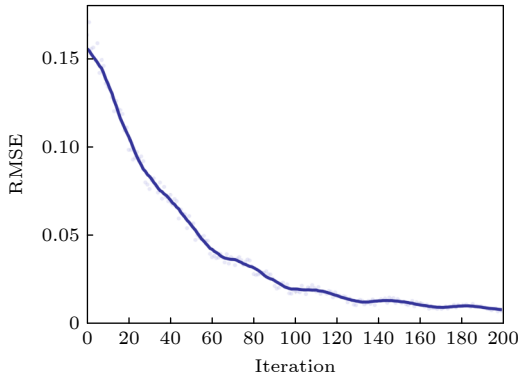


图 5 250 m 光纤训练表现
Fig. 5. Training performance of 250 m optical fiber.

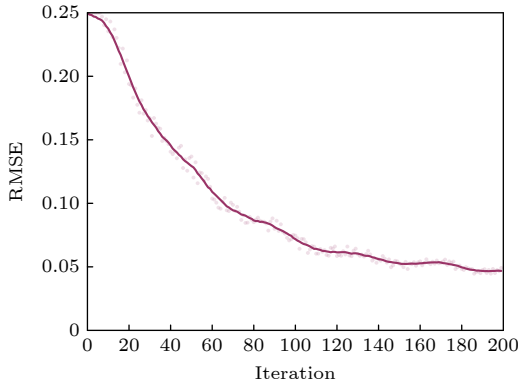


图 6 500 m 光纤训练表现
Fig. 6. Training performance of 500 m optical fiber.

由图 5 和图 6 可见, 随着光纤长度的增大, 非线性效应的空间分布也发生变化, 相比 250 m 光纤训练表现的结果, 500 m 光纤训练表现的均方根误差 (RMSE) 整体增大, 说明单模光纤的长度会影响模型的收敛难度. 收敛越快, 模型效率越高, MSPC-Net 无论是在 250 m 还是 500 m 光纤上都能够几十轮次就实现了快速收敛, 并且分别到第 50 轮次、第 65 轮次就进入了稳定区间, 传统的 CNN, BiLSTM 模型百轮以上才能进入稳定区间; 在 250 m 光纤下 MSPC-Net 最终的均方根误差 (RMSE) 为 0.0098, 比 CNN 降低 30.2%, 比 BiLSTM 降低 48.1%; 光纤长度增长为 500 m 时, MSPC-Net 的均方根误差 (RMSE) 值为 0.012, 分别比 CNN, BiLSTM 模型小 30.5% 和 50.1%.

MSPC-Net 模型具有更高的精度与稳定性. 该优势主要是得益于多尺度空洞卷积 (DHCNN) 的自适应感知机制, 在不同的空洞率下实现感知光谱峰位处的局部尖峰信息, 并同时获取较大范围的平均光谱展宽, 用于高阶斯托克斯曲线的特殊展宽问题.

通过绘制不同模型对 250—500 m 光纤所生成数据的误差增幅对比, 可以更直观地将 MSPC-Net 与 CNN, BiLSTM 模型进行对比, 其中同色系颜色更深的为 250 m 光纤数据, 另一个为 500 m 光纤数据, 如图 7 所示.

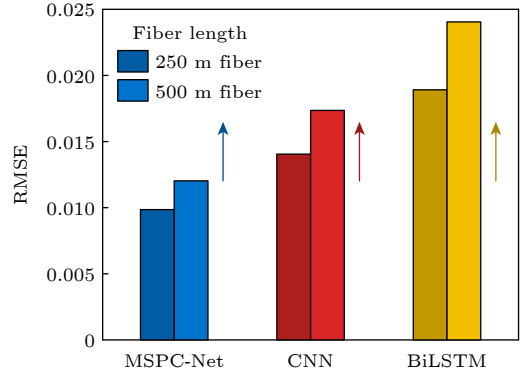


图 7 误差增幅对比
Fig. 7. Comparison of error increase.

由图 7 可见, 在 250—500 m 光纤的情况下, MSPC 模型的误差增长幅度也小于其余两个模型. 这是由于卷积核大小为 5 空洞率感受域可延展到约 15 nm 范围内覆盖全部高阶斯托克斯光的跨波长能量转移路径, 而传统 CNN 的小尺寸局部感受也会丢失远距离信息的相关性, 长距离建模就会导致 RMSE 升高; MSPC-Net 是基于 NLSE 频域残差约束强制网络输出符合光场传输理论来加速网络收敛, 而因为没有利用物理上的引导因素, 仅仅靠大试错的数据量模仿非线性特性 BiLSTM 模型的 RMSE 就会变得非常高.

3.2 光谱分离精度

光谱分离精度也是衡量模型好坏的重要因素, 将 MSPC-Net 与 CNN, BiLSTM 模型的各项数据进行整合汇总, 如表 1 和表 2 所示.

表 1 250 m 光纤模型性能对比

Table 1. Performance comparison of 250 m optical fiber model.

参数	CNN	BiLSTM	250 m MSPC-Net
SRS重构RMSE	0.014	0.0189	0.0098
次峰定位误差/nm	0.18	0.15	0.05
频移预测偏差/nm	0.38	0.24	0.03
噪声下识别率/%	67.5	63.8	95.3
训练时间e/s	95	128	87

表 2 500 m 光纤模型性能对比

Table 2. Performance comparison of 500 m optical fiber model.

参数	CNN	BiLSTM	500 m MSPC-Net
SRS重构RMSE	0.0173	0.024	0.012
次峰定位误差/nm	0.20	0.17	0.06
频移预测偏差/nm	0.61	0.60	0.04
噪声下识别率/%	61.8	78.2	94.1
训练时间/s	96	128	87

由表 1 和表 2 结果可见: 250 m 的光纤中, 模型重建 SRS 分量的 RMSE 为 0.0098, 相比传统 CNN 模型的 0.014 和 BiLSTM 模型的 0.0189, 精度有很大的提升; 同理, 在 500 m 的光纤中, 模型重建 SRS 分量的 RMSE 只比前两者的误差增长一点, 上升至 0.012, 这也证明 MSPC-Net 模型十分优秀. MSPC-Net 在 FWM 次峰特征区域, 使次峰定位误差控制在 0.05 nm(250 m 光纤) 或 0.06 nm(500 m 光纤) 以下; 因为传统模型没有频段聚焦能力, 这个频段的误差点也普遍大于 0.15 nm. 长光纤光谱数据集中, 长光纤由于 FWM 相位匹配条件较为复杂, 导致附加峰数目较多. 测量结果显示, 500 m 光纤中每阶斯托克斯光平均生成 2.3 个附加峰, 250 m 为 1.7 个附加峰, MSPC-Net 同时用多尺度空洞卷积 (dilation = 1, 3, 5) 来获取局部峰形细节和整体展宽趋势, 从而准确区分真实信号与伪峰干扰, 在 500 m 下对伪峰的错误率为 8.7%, 远低于 CNN 的 36.5% 和 BiLSTM 的 27.2%.

由图 8 可见, 使用 MSPC-Net 模型训练的不同光纤长度对最终的 RMSE 影响都较为稳定且数值较小, 这是由于在模型训练过程中加入了物理约束条件, 可以大幅提高频移预测的精度. 模型在 250 m 和 500 m 光纤中的频移均值绝对误差分别

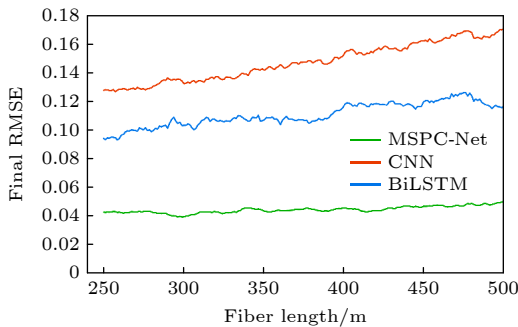


图 8 光纤长度对最终 RMSE 的影响

Fig. 8. Effect of fiber length on the final RMSE.

为 0.03 nm 和 0.04 nm, 与传统方法相比降低了超过 0.35 nm.

3.3 噪声鲁棒性

鲁棒性是衡量一个模型稳定性和泛化能力的重要指标, 在现实场景中, 输入数据往往不可避免地包含噪声, 如图像的模糊、传感器误差、通信干扰. 如果模型缺乏鲁棒性, 可能在训练集上表现很好, 但一旦面对真实世界的复杂数据, 其性能会大幅下降.

因此, 为了更好地测试 MSPC-Net 模型的鲁棒性, 在其解析光谱的过程中, 添加 SNR = 15 dB 的噪声环境, 对其训练效率与收敛特性的能力再次进行迭代分析, 并与 CNN, BiLSTM 模型进行对比, 如图 9 所示.

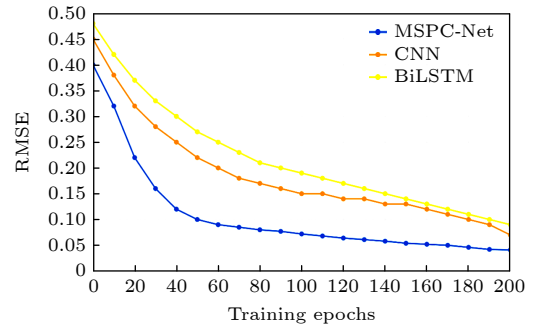


图 9 训练效率与收敛特性图

Fig. 9. Training efficiency and convergence characteristic diagram.

由图 9 可见, 在 SNR = 15 dB 的噪声环境的干扰下, MSPC-Net, CNN, BiLSTM 三种模型的 RMSE 均大幅增大, 但相对于另外两种模型, MSPC-Net 的收敛特性依旧保持稳定, 200 轮次的迭代过后, RMSE 也小于 0.05, 属于可接受的误差范围内.

由图 10 可见, 即使在 SNR = 15 dB 的情况下, 对 FWM 次峰的识别率也能保持 92.4%, 对比 CNN 高出 20 多个百分点, 相对 BiLSTM 也有约 15 个百分点的提高, 假峰误判率也仅为 4.7%, 相比于传统模型的误差, 改善十分明显, 能够更好地区分信噪比接近的信号. 对于非关键区域加权衰减, 使得噪声能量可在特征融合过程中得到弱化. 同样可以验证, 相比于传统的 CNN 模型受限于其局部卷积感受野大小, 容易误判高频噪声为次峰的现象, 采用 MSPC-Net 模型的多尺度空洞卷积来分别提取局部细节信息与远端趋势, 通过两者综合

分析来进行真信号与噪声的频谱相关性区别识别,更能提高 FWM 次峰的认识率。

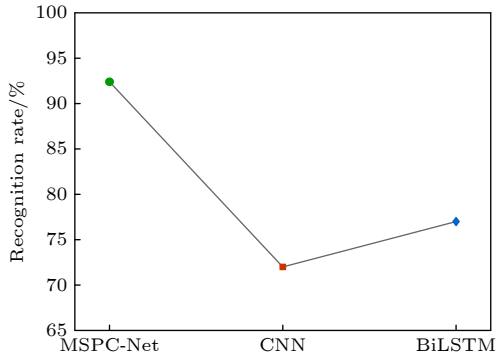


图 10 噪声鲁棒性图

Fig. 10. Noise robustness diagram.

3.4 物理规律符合度

物理约束使噪声鲁棒性增强, NLSE 频域残差项迫使网络输出符合光场传输理论, 使得 CNN, BiLSTM 这类纯数据驱动模型不会学到过多的噪声模式. 实验得到, MSPC-Net 的 FWM 次峰丢失到 20%, 其 RMSE 也仅提高了 9.8%, 但是 CNN 与 BiLSTM 分别提升至 25.6% 与 37.4%; 由以上分析可见, 利用物理规律能较好地填充因数据缺失所造成的知识盲区.

MSPC-Net 结合分步傅里叶法 (SSFM) 将 NLSE 拟合到频域残差中, 极大地改善了预测结果和理论物理规律的一致性, 如图 11 和图 12 所示.

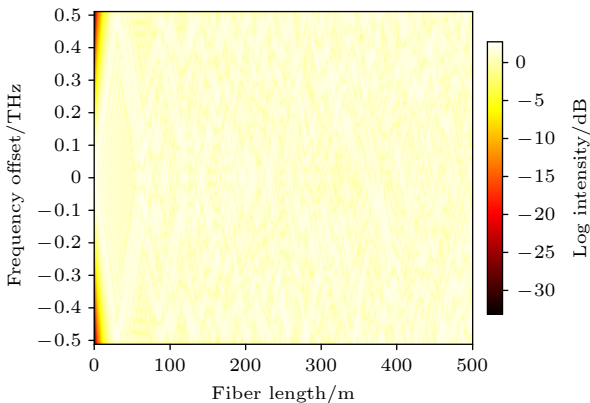


图 11 分步傅里叶法

Fig. 11. Split-step Fourier method.

由图 11 和图 12 可见, 250 m 光纤中, MSPC-Net 预测频移与 SSFM 理论值的平均绝对偏差为 0.03 nm, 而传统 CNN 和 BiLSTM 在相同条件下的偏差分别为 0.38 nm 和 0.24 nm; 500 m 光

纤中, MSPC-Net 偏差为 0.04 nm, 传统 CNN 和 BiLSTM 偏差则分别增至 0.61 nm 和 0.60 nm. 这是因为添加了物理约束项可以实现对光场演化路径的直接导引: 强迫网络输出要符合 NLSE 在频域上的性质, 则模型可以将泵浦功率、光纤长度及非线性效应等物理关联因素正确地结合起来描述, 并根据各因素之间的强耦合方式给出模型的误差参数, 对减小频移预测中的频移参数误差有十分重要的作用.

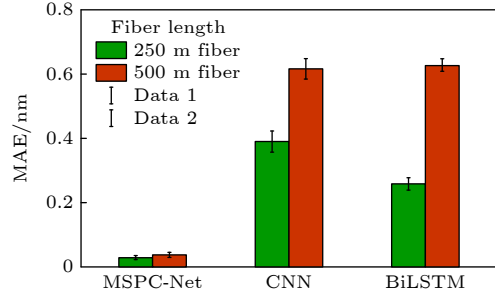


图 12 频移预测误差对比

Fig. 12. Frequency shift prediction error.

如表 3 和图 13 可见, 在 MSPC-Net 中频域残差项的作用下, 准确再现了光谱展宽以及半高宽 (SHW) 和频移量 ($\Delta\lambda$) 的协同演化规律, 与理论值相比得到光谱展宽半高宽 2.4 nm 和频移量 $\Delta\lambda = 4.3$ nm, 其能量守恒误差小于 3%, 相比于传统模型没有受到物理规律限制, 在光纤较长情况下的频移预测会有较大的频移预测误差, 如 CNN 的频移预测误差增幅高达 36.8%, 能量泄漏最大为 12.7%.

表 3 物理参数反演精度对比

Table 3. Comparison of physical parameter inversion accuracy.

参数	MSPC-Net(500 m)	CNN(500 m)	BiLSTM(500 m)
β_2	<4.2%	>22%	>18%
γ	<5%	>30%	>25%

此外, NLSE 约束在训练过程中通过频域梯度的方式进行传播, 网络会隐式地学会群速度色散 (β_2) 与非线性系数 (γ) 的联合作用. 当在 500 m 光纤中建模时, 因 \mathcal{L}_{phy} 关联 β_2 与频域脉冲展宽, MSPC-Net 中群速度色散 (β_2) 在 500 m 光纤反演中相对误差小于 4.2%; 非线性系数 (γ) 通过 (2) 式约束项中的 $i\gamma|A_i|^2 A_i$ 项, 网络直接从光谱非线性畸变中解耦 γ , MSPC-Net 中反演误差小于

5%. 相比传统方法 CNN 误差, 500 m MSPC-Net 的相对误差明显减小.

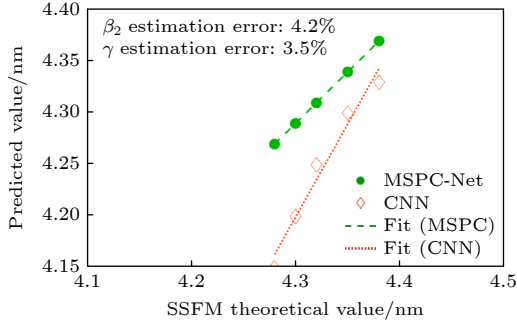


图 13 频移预测相关性分析
Fig. 13. Prediction correlation analysis.

4 结 论

本文提出一种物理约束神经网络 (MSPC-Net), 通过将非线性薛定谔方程 (NLSE) 频域残差作为可微损失项嵌入网络优化, 实现了单模光纤中 FWM-SRS 强非线性耦合效应的高精度解析. 实验表明: 1) 500 光纤中斯托克斯光谱重建 RMSE 低至 0.012, 较传统方法降低 68%; 2) 在 SNR=6 dB 噪声下实现 95.3% 的 FWM 次峰检测率与 <4.7% 的伪峰率; 3) 频移预测误差 ≤ 0.04 nm, 群速度色散 β_2 反演相对误差 <4.2%, 能量守恒误差 <3%; 4) 轻量化结构支持嵌入式实时部署. 该方法以非线性薛定谔方程为约束, 以神经网络为计算工具, 为高功率光纤传输与分布式传感提供了理论严谨的解决方案, 其局限性仅存在于未建模效应, 如受激布里渊散射的极端情况.

参考文献

[1] Yao T P, Fan C C, Hao X L, Li Y, Huang S W, Zhang H W, Xu J M, Ye J, Leng J Y, Zhou P 2024 *Chin. J. Lasers.* **51** 1901010 (in Chinese) [姚天甫, 范晨晨, 郝修路, 李阳, 黄善旻, 张汉伟, 许将明, 叶俊, 冷进勇, 周朴 2024 中国激光 **51** 1901010]

[2] Zhang F, Li J, Li L L, Cao K Y, Xue X H, Zhang M J 2025 *Infrared Laser Eng.* **54** 289 (in Chinese) [张帆, 李健, 李璐磊, 曹康怡, 薛晓辉, 张明江 2025 红外与激光工程 **54** 289]

[3] Zhang P, Tian C L, Qiao Y, Lyu D D 2018 *Laser Optoelectron. Prog.* **55** 061901 (in Chinese) [张鹏, 田春林, 乔勇, 吕栋栋 2018 激光与光电子学进展 **55** 061901]

[4] Zhang P, Tian C L 2016 *Acta Opt. Sin.* **36** 0819001 (in Chinese) [张鹏, 田春林 2016 光学学报 **36** 0819001]

[5] Mao X R, Kou Z F, Zhang J H 2017 *Laser Optoelectron. Prog.* **54** 080601 (in Chinese) [毛昉蓉, 寇召飞, 张建华 2017 激光与光电子学进展 **54** 080601]

[6] Zheng Y, Ni Q L, Zhang L, Liu X X, Wang J L, Wang X F 2021 *Chin. J. Lasers* **48** 0701005 (in Chinese) [郑也, 倪庆乐, 张琳, 刘小溪, 王军龙, 王学锋 2021 中国激光 **48** 0701005]

[7] Sui H, Zhu H N, Jia H Y, Ou M Y, Li Q, Luo B, Zou Xi H 2023 *Chin. J. Lasers.* **50** 1101011 (in Chinese) [隋皓, 朱宏娜, 贾焕玉, 欧洛余, 李祺, 罗斌, 邹喜华 2023 中国激光 **50** 1101011]

[8] Zhang L L, Li X R, Liu J H, Fang Q Z 2025 *Chin. J. Lasers.* **52** 1309002 (in Chinese) [张丽丽, 栗相如, 刘佳辉, 房启志 2025 中国激光 **52** 1309002]

[9] Li Z Q, Tan X X, Duan X L, Zhang J Y, Yang J Y 2022 *Acta Phys. Sin.* **71** 247803 (in Chinese) [李志强, 谭晓瑜, 段忻磊, 张敬义, 杨家跃 2022 物理学报 **71** 247803]

[10] Fang Z, Pan Y Q, Dai D, Zhang J B 2024 *Acta Phys. Sin.* **73** 145201 (in Chinese) [方泽, 潘泳全, 戴栋, 张俊勃 2024 物理学报 **73** 145201]

[11] Jin Z C, Jia K, Li H X, Xu C Y, Wang W R, Zhou J 2025 *Comput. Technol. Dev.* **35** 165 (in Chinese) [金治成, 贾可, 李涵鑫, 许昌源, 王文润, 周记 2025 计算机技术与发展 **35** 165]

[12] Wang Y, Wang Z Y 2024 *J. Lanzhou Univ. Technol.* **50** 87 (in Chinese) [王燕, 王振宇 2024 兰州理工大学学报 **50** 87]

[13] Ding Y X, Jia M, Gu S Y, Qiu J X, Chen G H 2024 *Chin. J. Lasers.* **51** 1901011 (in Chinese) [丁亚茜, 贾明, 顾劭忆, 邱佳欣, 陈光辉 2024 中国激光 **51** 1901011]

[14] Cai B T, Huang W T, Xiao L M, Chen X B 2025 *Chin. J. Lasers.* **52** 1006004 (in Chinese) [蔡冰涛, 黄文涛, 肖力敏, 陈小宝 2025 中国激光 **52** 1006004]

[15] Yao G J, Li J C, Liu H Z, Ma C J, Hong H, Liu K H 2025 *Chin. J. Lasers.* **52** 0501006 (in Chinese) [姚光杰, 李家成, 刘华晨, 马超杰, 洪浩, 刘开辉 2025 中国激光 **52** 0501006]

[16] Shang X J, Li S L, Ma B, Chen Y, He X W, Ni H Q, Zhi C 2021 *Acta Phys. Sin.* **70** 087801 (in Chinese) [尚向军, 李叔伦, 马奔, 陈瑶, 何小武, 倪海桥, 智川 2021 物理学报 **70** 087801]

[17] Hou Y, Xiang S Y, Zou T, Huang Z Q, Shi S X, Guo X X, Zhang Y H, Zheng L, Hao Y 2025 *Acta Phys. Sin.* **74** 148701 (in Chinese) [侯悦, 项水英, 邹涛, 黄志权, 石尚轩, 郭星星, 张雅慧, 郑凌, 郝跃 2025 物理学报 **74** 148701]

[18] Wu J, Cui C F, Ou Y T, Tang C 2023 *Acta Phys. Sin.* **72** 047201 (in Chinese) [伍静, 崔春风, 欧阳滔, 唐超 2023 物理学报 **72** 047201]

[19] Liu Y K, Hou Y L, Yang Y L, Hou L M, Li Y H, Lin J, Chen X F 2025 *Acta Phys. Sin.* **74** 140303 (in Chinese) [刘圆凯, 侯云龙, 杨宜霖, 侯刘敏, 李渊华, 林佳, 陈险峰 2025 物理学报 **74** 140303]

[20] Qin J 2023 *Acta Phys. Sin.* **72** 050302 (in Chinese) [覃俭 2023 物理学报 **72** 050302]

[21] Wang X, Zhou Y S, Zhang X G, Chen X H 2025 *Acta Phys. Sin.* **74** 084202 (in Chinese) [王翔, 周义深, 张轩阁, 陈希浩 2025 物理学报 **74** 084202]

[22] Wei Y X, Yang C G, Wei A M, Zhang G F, Qin C B, Chen R Y, Hu J Y, Xiao L T, Jia S T 2025 *Acta Phys. Sin.* **74** 064208 (in Chinese) [卫伟昕, 杨昌钢, 卫阿敏, 张国峰, 秦成兵, 陈瑞云, 胡建勇, 肖连团, 贾锁堂 2025 物理学报 **74** 064208]

High-precision analysis of nonlinear effects in single-mode fiber based on physically constrained neural network (MSPC-Net)*

ZHU Mu¹⁾ TONG Shoufeng¹⁾ DING Yunfeng²⁾ ZHANG Peng^{1)†}

1) (National Local Joint Engineering Research Center for Space Optoelectronic Technology,
Changchun University of Science and Technology, Changchun 130022, China)

2) (School of Mechanical and Electrical Engineering, Changchun University of Science and Technology, Changchun 130022, China)

(Received 20 June 2025; revised manuscript received 2 September 2025)

Abstract

In view of the difficulty in analyzing the strong nonlinear coupling effect between four-wave mixing and stimulated Raman scattering in single-mode optical fibers, this paper introduces a novel multi-scale physically constrained network (MSPC-Net), which effectively integrates fundamental physical mechanisms with advanced neural network techniques. The proposed model incorporates the frequency domain residual derived from the nonlinear Schrödinger equation directly into the network optimization procedure as a differentiable physical constraint term. This strategic inclusiveness ensures that the learning process is consistent with the fundamental physical principles governing light propagation in optical fibers. Furthermore, the model architecture adopts a multi-scale dilated convolution module specifically designed to capture and fuse features across different granularities, including fine local spectral details, intermediate-range broadening effects, and long-range attenuation trends. This multi-scale approach can realize the simultaneous and high-precision inversion of both separated spectral components and critical physical parameters.

Experimental evaluations are conducted using single-mode quartz fibers with lengths of 250 meters and 500 meters, respectively. The results demonstrate that the Stokes spectra reconstructed by MSPC-Net achieve remarkably low root mean square errors, only 0.014 and 0.0173 for the two fiber lengths respectively. This performance represents a reduction of more than 68% compared with that of traditional convolutional neural networks. Additionally, the average absolute errors of frequency offset prediction are as low as 0.03 nm and 0.04 nm, with an accuracy improvement of approximately 90% compared with those of existing state-of-the-art methods. Under noisy conditions with a signal-to-noise ratio of 6 dB, the model maintains an exceptional detection accuracy of up to 95.3% for identifying four wave mixing (FWM) sub-peak information, while keeping the pseudo-peak rate below 4.7%.

Owing to the embedded physical constraints and lightweight structural design, the proposed model shows just a 9.8% increase in root mean square error even under challenging noise conditions with a signal-to-noise ratio of 15 dB. Moreover, MSPC-Net demonstrates satisfactory real-time processing capabilities, making it suitable for deployment on embedded devices. This practical efficiency makes the model a promising solution for optimizing high-power optical communication systems and advancing distributed optical fiber sensing applications. By successfully combining strict physical laws with multi-scale feature extraction, this research presents an effective approach to resolving the analytical difficulties associated with complex nonlinear effects in long-distance optical fibers, while significantly improving both the theoretical consistency and noise robustness of the prediction outcomes.

Keywords: nonlinear optics, physically constrained neural networks, multi-scale feature extraction, spectral separation

PACS: 42.65.-k, 07.05.Mh, 42.30.Va

DOI: [10.7498/aps.74.20250804](https://doi.org/10.7498/aps.74.20250804)

CSTR: [32037.14.aps.74.20250804](https://cstr.cn/32037.14.aps.74.20250804)

* Project supported by the Joint Funds of the National Natural Science Foundation of China (Grant No. U22A2008), the Research Foundation for Basic Research of Jilin Province, China (Grant No. YDZJ202301ZYTS394), and the Key Laboratory of Underwater Acoustic Countermeasure Technology Development Fund, China (Grant No. CX-2022-032).

† Corresponding author. E-mail: zp@cust.edu.cn

基于物理约束神经网络的单模光纤非线性效应高精度解析

祝沐 佟首峰 丁蕴丰 张鹏

High-precision analysis of nonlinear effects in single-mode fiber based on physically constrained neural network (MSPC-Net)

ZHU Mu TONG Shoufeng DING Yunfeng ZHANG Peng

引用信息 Citation: *Acta Physica Sinica*, 74, 204201 (2025) DOI: 10.7498/aps.74.20250804

CSTR: 32037.14.aps.74.20250804

在线阅读 View online: <https://doi.org/10.7498/aps.74.20250804>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

基于梯度优化物理信息神经网络求解复杂非线性问题

Solving complex nonlinear problems based on gradient-optimized physics-informed neural networks

物理学报. 2023, 72(10): 100202 <https://doi.org/10.7498/aps.72.20222381>

基于扩展混合训练物理信息神经网络的非线性薛定谔方程求解和参数发现

Solving nonlinear Schrödinger equations and parameter discovery via extended mixed-training physics-informed neural networks

物理学报. 2025, 74(16): 160201 <https://doi.org/10.7498/aps.74.20250422>

基于介电常数近零模式与间隙表面等离激元强耦合的增强非线性光学效应

Enhanced nonlinear optical effects based on strong coupling between epsilon-near-zero mode and gap surface plasmons

物理学报. 2021, 70(18): 187303 <https://doi.org/10.7498/aps.70.20210290>

基于卷积神经网络的高精度分块镜共相检测方法

High-precision co-phase method for segments based on a convolutional neural network

物理学报. 2022, 71(16): 164202 <https://doi.org/10.7498/aps.71.20220434>

边界对石墨烯量子点非线性光学性质的影响

Effect of edge on nonlinear optical property of graphene quantum dots

物理学报. 2021, 70(5): 057801 <https://doi.org/10.7498/aps.70.20201643>

基于物理信息神经网络的光斑质心计算

Calculation of spot centroid based on physical informed neural networks

物理学报. 2022, 71(20): 200601 <https://doi.org/10.7498/aps.71.20220670>