

深度学习策略下光纤中超短脉冲非线性传输过程表征及控制研究进展

隋皓¹, 朱宏娜^{1*}, 贾焕玉¹, 欧洛余², 李祺¹, 罗斌², 邹喜华²

¹西南交通大学物理科学与技术学院, 成都 四川 610031;

²西南交通大学信息科学与技术学院, 成都 四川 610031

摘要 常规数值求解方法在表征光纤中超短脉冲的非线性传输过程时存在计算量大、效率低等局限。随着人工智能的快速发展,深度学习技术展现出了强大的计算能力、广泛的适用范围、良好的硬件移植性,在光纤中超短脉冲非线性传输过程表征和控制研究中具有巨大潜力。本文概述了深度学习技术及其在预测光纤中超短脉冲传输、超短脉冲重构及参数估计方面的研究进展,同时展望了深度学习与光纤中超短脉冲非线性传输这一新兴交叉技术的发展方向和挑战。

关键词 光纤光学; 超短激光脉冲传输; 非线性薛定谔方程; 光纤非线性效应; 深度学习

中图分类号 O437 文献标志码 A

DOI: 10.3788/CJL230508

1 引言

作为超短光脉冲传输及控制的重要应用方向,光纤中超短激光脉冲的非线性传输过程研究是光纤光学、超快光学和光信息处理的重要基础,促进了光纤激光器、光纤放大器以及光纤通信、传感技术的快速发展^[1-6]。光纤中激光脉冲的非线性传输可通过非线性薛定谔方程(NLSE)描述,当色散和非线性效应同时存在时,此复合偏微分方程通常没有解析解,需要采用分步傅里叶法(SSF)、有限差分法等数值方法进行求解。为了达到较高的精度,SSF等方法需要设置合适的步长和迭代次数,从而导致计算复杂性提升。此外,NLSE系统对初始脉冲参数和光纤参数非常敏感,较小的参数变化可导致超短激光脉冲非线性传输过程显著变化^[1,4]。因此,采用常规SSF等数值方法对光纤中激光脉冲非线性传输进行研究或设计优化需要开展大量参数化模拟,存在超大参数空间扫描计算复杂度高、效率低、运算量大等问题^[4,7-9]。

随着人工智能技术的快速发展,机器学习算法已在超短脉冲激光器中实现了广泛应用,包括智能锁模控制、脉冲压缩等^[10-13]。其中,深度学习技术展现出了强大的计算能力、广泛的适用策略、良好的硬件移植性,使其成为控制和解决大量高维数据问题的有效途径^[14-15]。不同于模型驱动、数值求解方式,基于数据驱动的深度学习方法从一系列数据中学习内在的输入-

输出对应关系,使大量光脉冲数据集的访问更加快速便捷,同时也开拓了处理、表征和控制光纤中超短脉冲非线性传输的另一种途径,使超短激光脉冲设计和操作更加准确,获取和处理所需数据更加高效。此外,基于数据驱动的深度学习方法还展现出了更高的计算效率以及更低的能耗等潜在优势。目前,深度学习技术已被广泛应用于超快光学^[4,15]、高速光通信^[16-18]、光纤激光^[3,19-20]、光纤传感^[21]、光网络^[22]、光学成像^[23]、计算光学^[14,24]、激光与物质相互作用^[25]、光学器件^[26]、新型光学神经网络^[27]等领域。

本文首先概述了深度学习技术和常用的深度神经网络,接着介绍了深度学习方法在预测自相位调制、光孤子传输及超连续谱产生等光纤中超短脉冲非线性传输方面的研究成果,以及深度学习方法在解决光纤中超短脉冲非线性传输过程逆问题(如超短脉冲重构、激光脉冲参数估计及优化)方面的研究成果。这些成果展示了深度学习技术在光纤中超短脉冲非线性传输过程表征和控制中的应用和潜力。最后本文简要总结了这一新兴跨学科研究方法面临的挑战以及未来的发展方向。

2 深度学习概述

深度学习技术通过一系列不同函数的复合形成神经网络,实现输入向量到输出向量的映射^[28]。在给定足够多的网络层数、层内单元数以及标定数据后,深度

收稿日期: 2023-02-10; 修回日期: 2023-03-29; 录用日期: 2023-04-10; 网络首发日期: 2023-04-18

基金项目: 国家重点研发计划项目(2019YFB1803500)、四川省科技计划项目(2020YJ0016)

通信作者: *hznzhu@swjtu.edu.cn

神经网络(DNN)可以通过“学习”来表示输入-输出数据间高度复杂的函数关系。“学习”是指一个计算机算法被经验 E 改进后,在某类任务 T 上性能度量 P 有所提升^[3]。常见的任务 T 包含回归、分类等,相应的性能度量 P 可能是预测误差、准确率等。绝大多数深度学习算法可解释为在整个数据集上获取经验。数据集是样本集合的统称,例如 $D = \{(y_i, x_i)\}_{i=1, \dots, n}$, 其中: x_i 被称作样本,可由向量、数组或矩阵等任意数据形式组成,可代表原始数据或已量化的特征; y_i 为标签,是期望的网络输出,如正确的预测结果、正确的分类结果等; n 为样本的数量。一般情况下,数据集可由数值仿真或实验方式得到。对于光纤中的超短脉冲传输,如图 1(a)所示,可将深度学习任务分为预测和反演两类。输入的超短光脉冲在光纤中传输,最终输出受色散和各类非线性效应影响的光脉冲。将根据输入脉冲分布获取输出脉冲分布的学习任务定义为预测,即寻求 $f: x \rightarrow y, y^* = f(x)$ 。例如:由初始脉冲预测输出脉冲以及由初始脉冲建模整个脉冲传输过程等任务。同时,将从输出脉冲分布映射至输入脉冲分布的学习任

务定义为反演,即 $f: y \rightarrow x, x^* = f(y)$,包括从输出脉冲预测初始脉冲、从输出脉冲估计和优化初始脉冲或系统参数、设计初始脉冲以实现特定的脉冲传输过程等。

深度学习算法的工作流程如图 1(b)所示,除了准备数据集,网络结构设计也是关键的步骤。面对不同的任务目标和应用场景,往往需要不同的网络构架。常用的神经网络包括前馈神经网络(FNN)、循环神经网络(RNN)、卷积神经网络(CNN)、生成对抗网络(GAN)、自编码器(VAE)、深度置信网络(DBN)及图神经网络(GNN)等。目前主流的DNN构架是图 1(c)所示的包含多个隐藏层的FNN、RNN和CNN,其中FNN由输入层、多个隐藏层和输出层组成,多用于分类任务和基于模式分类的学习控制中^[29]。RNN由循环单元连接组成,当前输出由输入及前过程产生的“记忆特征”决定,其擅长处理时序和序列数据,已被广泛应用于自然语言处理等领域^[30]。CNN中的卷积层和池化层提取输入数据的各维度特征,之后经过特征映射用于回归或分类任务中。CNN已被广泛应用于处理图像数据,包括特征提取、图像增强及实现超分辨等^[31]。

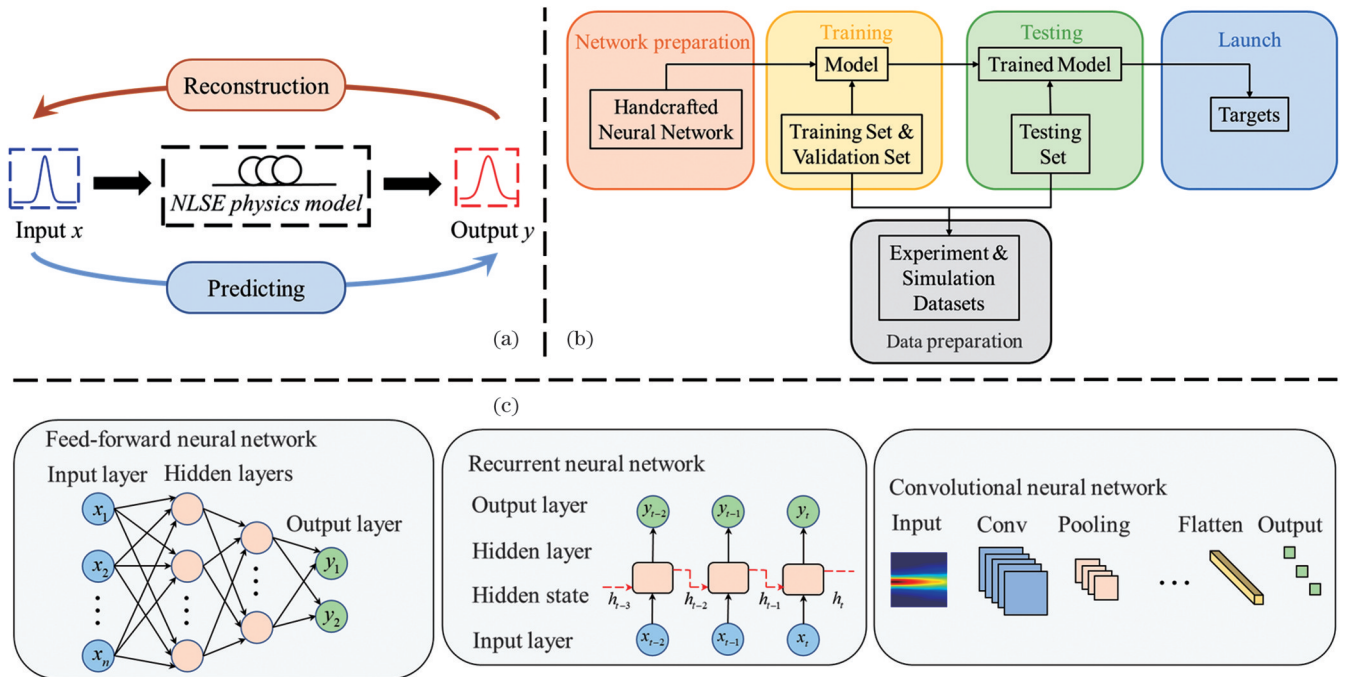


图 1 深度学习方法的概述。(a)深度学习在非线性脉冲传输建模中的应用;(b)深度学习算法的工作流程;(c)常见的神经网络结构
Fig. 1 Overview of deep learning methods. (a) Application of deep learning methods in modeling nonlinear pulse propagation; (b) workflow of deep learning methods; (c) common neural network structures

DNN 训练是通过搜索算子确定网络权重的过程。网络权重优化是训练的核心,大多数深度学习问题可以归结为优化问题^[32]。在实践中,各类梯度下降算法(如随机梯度下降算法、Adam 等)被用于 DNN 训练^[33],以优化由一系列权重构成的网络映射函数,直至网络的输出误差小于设定的阈值。该过程可以看作是网络映射函数在给定数据集和模型参数空间上的优化

表示。为显示网络模型在训练过程中的性能,有时可以从训练数据集中适当分离出验证数据集。训练结束后,DNN 需要进行测试,以验证网络的泛化性等关键性能,训练好的 DNN 便可应用于既定任务中。泛化性用于表征网络在之前未观测(训练)到的新输入上的表现。一般来说,测试集是有相同或相似统计分布的独立于训练集的数据集^[28]。

3 光纤中超短脉冲非线性传输过程的表征

本节介绍深度学习方法在表征光纤中超短脉冲非线性传输方面的应用,包括对超短脉冲调制不稳定性、光孤子传输和超连续谱产生等非线性传输过程的预测。采用 DNN 对特定脉冲参数空间内光脉冲非线性传输过程进行近似表示,即用训练好的网络映射函数在特定数据空间近似替代 NLSE 物理系统,可以达到“隐藏物理模型”的目的,同时可以兼顾预测精度和计算效率^[7]。

3.1 色散和自相位调制

Boscolo 等^[34-35]采用 FNN 预测了考虑增益和损耗、群速度色散(GVD)和自相位调制(SPM)相互作用的光纤中的自相似抛物脉冲传输,实现了不同传输距离下输出脉冲强度和频谱的精确预测。Jiang 等^[36]采用物理信息神经网络(PINN)求解 NLSE,预测了超短脉冲在 GVD 和 SPM 共同影响下的传输过程,并讨论了 PINN 在不同初始脉冲功率和形状(对称双峰脉冲、不对称双峰脉冲等)下的预测精度和计算复杂度。

不同于完全数据驱动的深度学习方法,PINN 通过改造损失函数,增加了对神经网络的物理约束。通常,这些约束可以看作是网络模型的先验条件,为网络训练提供正则化项。相较于传统的 DNN, PINN 在收敛速度、小样本泛化性等方面具有潜在优势^[37-38]。

Wang 等^[39]将 PINN 与传统谱方法相结合预测了 NLSE 系统中由调制不稳定性产生的光学怪波和呼吸子。呼吸子泛指一类具有周期性分布结构的非线性波,与非线性系统中普遍存在的调制不稳定性紧密相关。该方法修正了常规方法在调制不稳定性边界条件微小扰动下引起的较大模拟误差,实现了适用于周期和非周期边界下光学怪波和呼吸子的高精度数值求解。如图 2(a)所示,Jiang 等^[2]采用 PINN 预测了 GVD 和 SPM 共存、三阶色散和 SPM 共存、自变陡效应(SS)和脉冲内拉曼散射(IRS)等情况下光纤中超短脉冲的非线性传输过程,图中展示了 PINN 在时域和频域上的高精度预测结果。此外,Jiang 等还讨论了 PINN 在多种情况下的预测精度和计算复杂度。Jiang 等采用了在网络输入中耦合物理参数的训练策略,提升了 PINN 在不同应用场景下的泛化能力。

Wang 等^[40]采用双边长短时记忆网络(BiLSTM)对 GVD 和 SPM 共存的光纤信道进行建模,并在开关键控和脉冲幅度调制四路信号两种应用场景下进行了测试,其结果在时域和频域等多个特征维度上与理论值相差较小,且相比常规数值模拟方法的运算量显著降低。同时,Wang 等对网络泛化性、硬件移植性进行了讨论。Yang 等^[41]采用生成对抗网络(GAN)对包

含 GVD、SPM 和掺铒光纤放大器自发辐射噪声的光纤信道进行建模,并将模型计算量降低至常规 SSF 的 2%。Yang 等^[42]采用特征解耦分布模型对波分复用光纤通信系统进行建模,其中:信道线性效应,如衰减、色散等,采用物理模型计算;信道非线性效应,如 SPM、交叉相位调制(XPM)和四波混频(FWM),采用 BiLSTM 网络进行建模和预测。对比常规数值模拟方法,特征解耦分布模型的计算效率提升了 2 个数量级。

3.2 光孤子

如图 2(b)所示,Salmela 等^[4]采用长短时记忆网络(LSTM)预测了皮秒脉冲在高非线性光纤反常色散区传输产生的高阶光孤子压缩过程。基阶光孤子是沿光纤长度方向形状和频谱都不发生变化的特殊脉冲,而高阶光孤子会在 GVD 和 SPM 的共同作用下发生周期性演化,并在孤子周期整数倍处恢复到初始形状。Salmela 等的研究表明,LSTM 能在时域和频域上精确预测高阶光孤子的传输过程,预测结果与同参数下的实验结果相吻合。Teġin 等^[43]通过相似 LSTM 结构研究了渐变型多模光纤中高功率飞秒脉冲非线性传输过程,实现了不同传输距离下光脉冲时域和频域强度的精确预测。同时,Teġin 等还对渐变型多模光纤中 1 GW 飞秒脉冲的空间强度演化进行了预测。Jiang 等^[2]采用 PINN 对 GVD 和 SPM 共同作用下基阶光孤子、高阶光孤子和 PAM4 格式随机信号的传输过程进行了预测。

同时,研究人员也对 PINN 的网络模型进行了改进,以增强网络在特定任务上的预测精度及泛化性。Wu 等^[44]采用两个 PINN 子网络分别处理实部和虚部数据,提升了模型对复函数偏微分方程的求解能力,并通过改进网络实现了光纤中单孤子、双孤子和明暗孤子碰撞等非线性脉冲传输过程的预测和参数估计。Fang 等^[45]提出了基于能量守恒定律的 PINN 并将其用于包括基阶光孤子传输、孤子相互作用等非线性动力学过程的建模。相比于常规 PINN,改进网络的预测准确性和泛化性都有所提升。Fang 等^[46]提出了一种基于导数信息约束 PINN 的方法,该方法在网络损失中引入非线性偏微分方程解的导数信息,并采用自适应权值和可变学习率,以加速网络收敛速度。优化后的 PINN 被用于预测基于高阶 NLSE 的各类孤子的非线性传输过程。对比常规 PINN 网络,优化后的 PINN 网络的预测精度提升了数倍。Zhu 等^[9]提出了包含加权损失函数的双子网络 PINN,研究了超短脉冲在光纤传输过程中的高阶非线性效应,并利用该网络预测了基于高阶 NLSE 的混合型单/双光孤子和光孤子分子的非线性传输过程。

3.3 光脉冲的其他非线性传输过程

DNN 也被用于预测超连续谱产生、FWM 等光

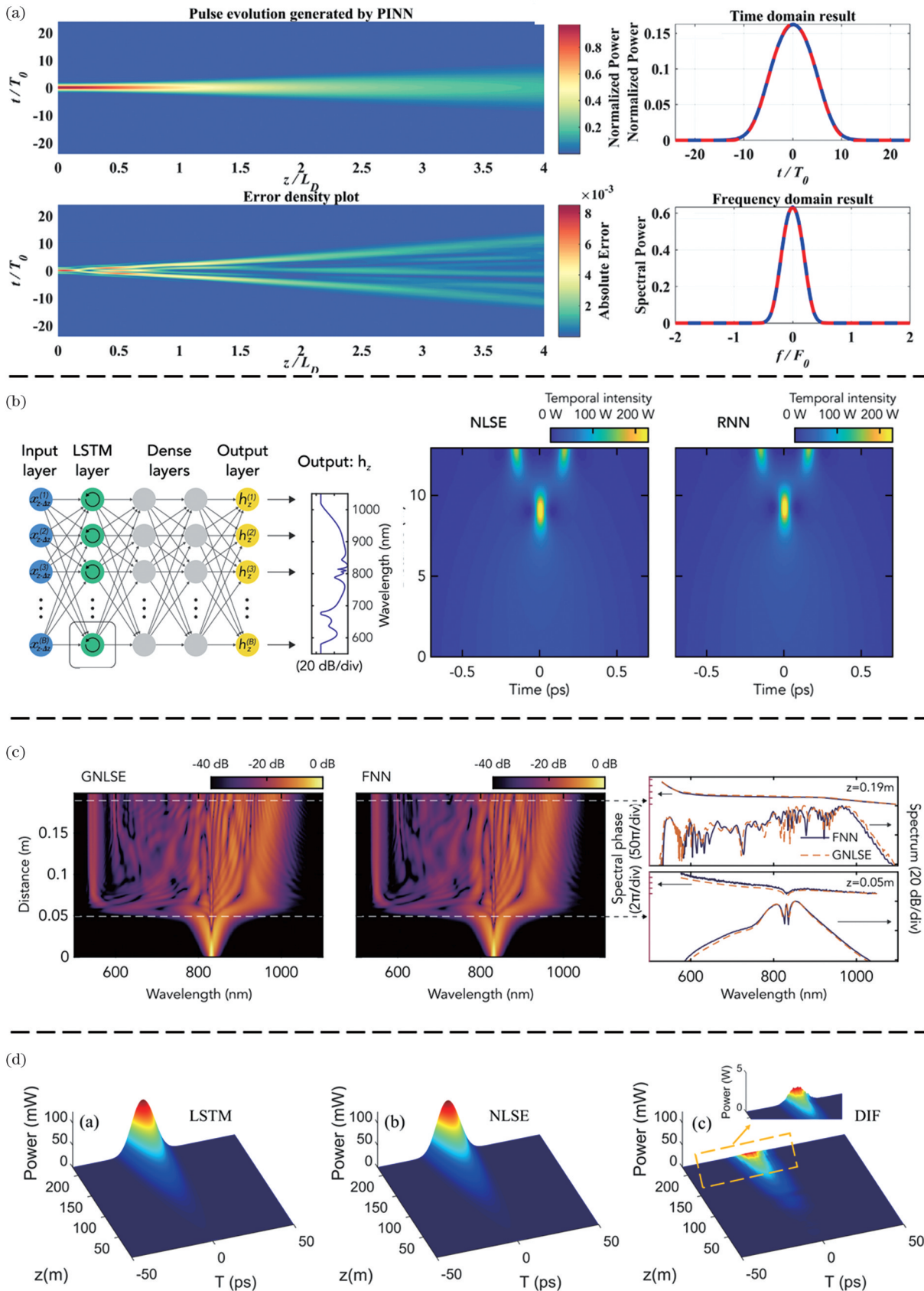


图 2 深度学习应用在预测光纤超短脉冲非线性传输过程中的应用。(a) 预测群速度色散和自相位调制共同作用下的超短脉冲传输^[2]; (b) 预测高阶光孤子传输^[4]; (c) 预测超连续谱产生^[47]; (d) 预测光纤参量放大系统中信号光脉冲的传输^[48]

Fig. 2 Deep learning applications in predicting the nonlinear ultrashort pulse propagation in optical fiber. (a) Predicting the ultrashort pulse propagation under GVD and SPM^[2]; (b) predicting the higher-order soliton propagation^[4]; (c) predicting the supercontinuum generation^[47]; (d) predicting the signal pulse propagation in optical fiber parametric amplification system^[48]

纤中的其他超短脉冲非线性传输过程。Salmela 等^[4]采用 LSTM 预测了不同功率下飞秒脉冲在高非线性光纤反常色散区的传输和超连续谱产生。LSTM 网络在时域和频域上的输出均与实验结果吻合。如图 2(c) 所示, Salmela 等^[47]采用 FNN 预测了光纤中超短脉冲的非线性传输和超连续谱产生, 图中所示结果显示深度学习预测方法与理论值差异较小。Salmela 等还对比了 FNN 和 RNN 在训练时间和内存需求上的差异, 结果表明: FNN 的计算复杂度更低, 在相同的算力和数据规模下更有优势。2021 年, Sui 等^[48]采用 LSTM 预测了光纤参量放大系统中不同初始脉冲功率和脉宽条件下信号光脉冲的放大传输过程, 预测结果如图 2(d) 所示。可见, 与理论值相比, 网络的预测误差较小。隋皓等^[49]采用 CNN 预测了啁啾信号光脉冲在全光纤超短脉冲啁啾放大中的非线性放大过程, 结果显示: 在保持较高预测精度的基础上, CNN 的预测时间为常规 SSF 的 1/10。

DNN 结构也是影响网络收敛速度、预测精度和泛化性的重要因素。Gautam 等^[50]基于 5 种不同初始脉冲和光纤参数下生成的数据集, 研究了 FNN、CNN、LSTM 等 6 种不同类型神经网络对高非线性光纤中超短脉冲非线性传输过程预测的精度和计算复杂度, 并讨论了网络结构对预测性能和泛化性的影响。Martins 等^[8]采用 LSTM 和卷积长短时记忆网络对 13 m 内高非线性光纤中的超短脉冲非线性传输过程进行了预测, 并研究了两种网络在不同初始脉冲功率、脉宽及脉冲形状下的泛化性, 最后得出了 LSTM 预测精度更高、泛化性更好的结论。

4 光纤中超短脉冲非线性传输的逆问题

深度学习方法已被广泛应用于解决相干成像^[51-52]、散射介质成像^[53-54]等领域中的逆问题, CNN 也被成功应用于超短脉冲重构^[7, 55-56]。本节主要介绍深度学习方法在解决光纤中超短脉冲非线性传输逆问题上的研究进展, 包括光纤中超短光脉冲重构和激光脉冲参数估计及优化两方面。

4.1 超短光脉冲重构

2022 年, Stanfield 等^[57]采用 DNN 基于 SPM 引起的非线性频谱变化重建初始飞秒脉冲, DNN 在随机初始相位的数据集上开展训练。基于受 SPM 影响的频谱和初始频谱, DNN 能够精确重建初始脉冲的相位, 重建结果如图 3(a) 所示。Stanfield 等在实验数据上的验证结果显示该方法的重建精度良好, 且对噪声有较高的鲁棒性。Sui 等^[58]采用 CNN 研究了光纤参量放大系统内信号光脉冲传输和光孤子对相互作用这两类过程中初始光脉冲的重建, 并研究了

不同输入脉冲参数及多种数据集下 CNN 对初始脉冲强度的重建精度, 讨论了 CNN 在不同结构和不同噪声水平下的泛化性。2022 年, Sui 等^[59]将 NLSE 与 DNN 耦合, 通过两者的耦合产生物理损失, 实现了无标签网络训练方法, 该训练方法在受 GVD 和 SPM 共同影响的超短脉冲重构任务中实现了与监督学习相近的收敛性和预测精度。两种训练方法下的脉冲重建结果如图 3(b) 所示。Gautam 等^[50]采用不同的神经网络重构受色散和非线性效应共同影响的超短光脉冲, 并研究和对比了多种神经网络在不同光纤参数和不同初始脉冲参数数据集上的重建精度和计算时间。Sidelnikov 等^[60]采用 CNN 替代常规数字反向传播算法 (DBP), 实现了长距离超高容量光纤传输系统中的 SPM 和 XPM 补偿。Sidelnikov 等认为 CNN 相较于 DBP 算法具有更优的性能和更低的计算复杂度。

4.2 激光脉冲参数估计及优化

光纤中超短脉冲的非线性传输过程参数估计和系统优化是激光控制及特性表征的重要研究方向。Tay 等^[61]采用 FNN 实现了双泵浦光纤参量放大系统的泵浦光功率和波长优化设计。通过 FNN 学习光纤参量放大器增益与泵浦光功率、波长的映射关系, 网络能够针对特定增益设计最优泵浦光功率及波长。2022 年, He 等^[62]采用结合 LSTM 和 CNN 的 LSTM-DenseNets 预测了被动锁模光纤激光器中光孤子的特征参数, 实现了光孤子时间间隔和相位差的精确提取, 预测结果如图 3(c) 所示。同时, He 等根据提取的特征参数开展了光孤子非线性传输过程的预测, 显著提升了计算效率。Herrera 等^[63]采用 FNN 和 CNN 从飞秒光孤子非线性频谱演化中估计了光孤子峰值功率和脉宽, 并研究了两种神经网络在不同量子噪声水平下对光孤子参数预测的准确性, 结果显示 CNN 对噪声的鲁棒性更好。如图 3(d) 所示, Jiang 等^[64]采用轻量级 FNN 实现了对光束质量因子的高精度快速估计, 并在 50 个少模光纤样本实验测试集上实现了较低的平均估计误差。Salmela 等^[47]采用 FNN 从超连续谱强度中预测了光孤子的时间特征。FNN 将超连续谱强度矢量作为输入, 经过两个隐藏层后输出光孤子的峰值功率、脉宽和时间延迟。FNN 网络能在广泛的频谱范围内保持较高的精度, 具有优异的性能。Zhu 等^[9]提出了带加权损失函数的双子网络 PINN, 并基于 SS 和 IRS 共同影响下的光孤子非线性传输过程逆向推导出了高阶 NLSE 系统中的物理系数。Li 等^[65]采用不同的网络结构重构光孤子分子, 研究了三种不同结构的 CNN 在多种情况下对孤子分子相位差的重构精度, 并讨论了最优的 CNN 结构。

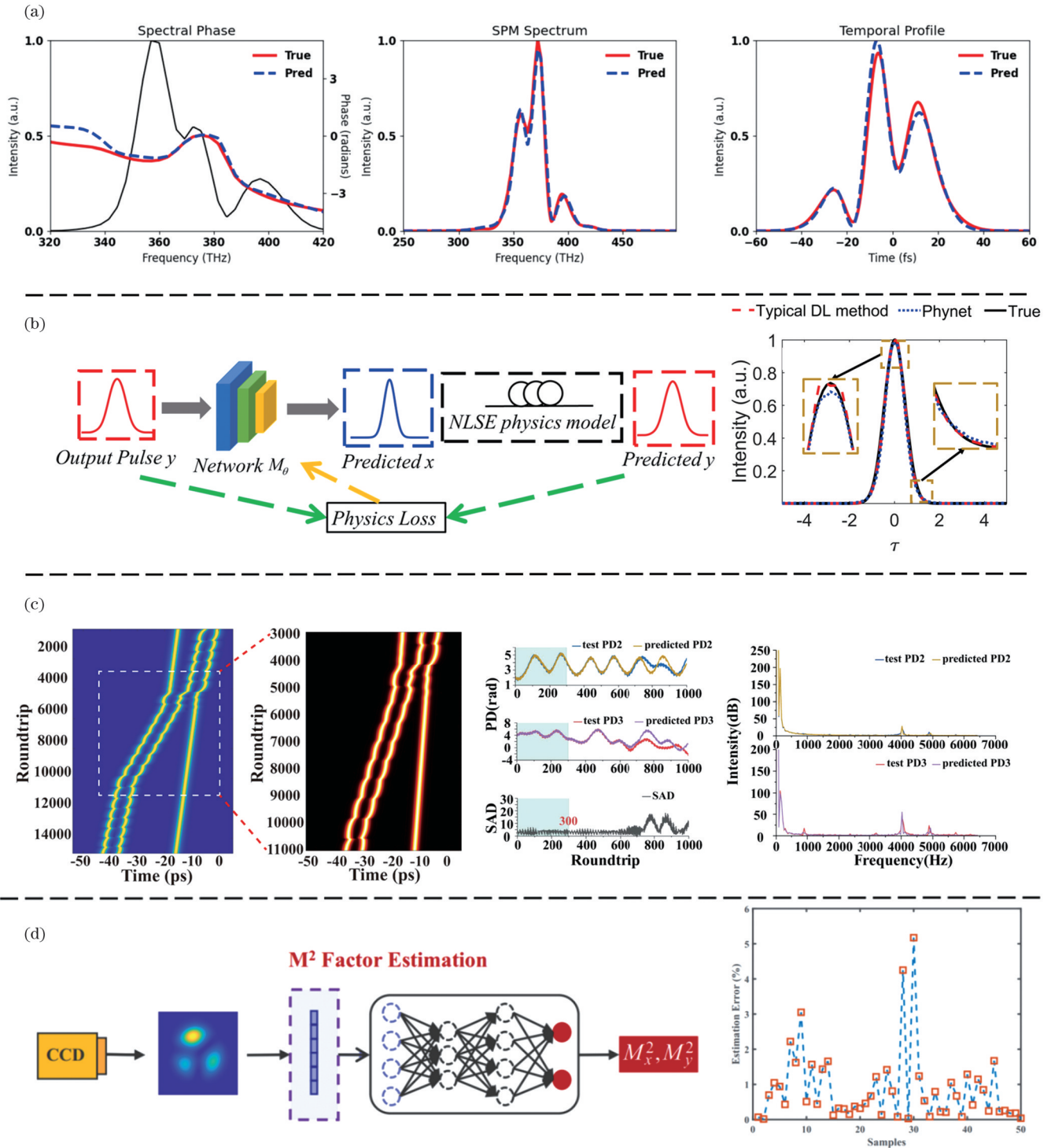


图3 深度学习方法在解决光纤中超短脉冲非线性传输逆问题上的应用。(a)基于自相位调制引发的非线性频谱变化重建超短脉冲^[57];(b)基于物理模型的深度学习方法重构初始脉冲^[59];(c)预测光孤子及光孤子分子的特征参数^[62];(d)少模光纤中的光束质量因子估计^[64]

Fig. 3 Deep learning applications in solving the inverse problems of the ultrashort laser pulse propagation in optical fiber. (a) Reconstruction of ultrashort pulse based on the nonlinear spectral changes induced by SPM^[57]; (b) physics-based deep learning method for the reconstruction of the initial pulse^[59]; (c) predicting typical parameters of single soliton and soliton molecule^[62]; (d) estimation of M^2 factor for few-mode fibers^[64]

5 结束语

本文综述了深度学习方法在光纤中超短脉冲非线性传输过程表征和控制方面的最新进展,包括采用

DNN 预测超短脉冲非线性传输、超短光脉冲重建、激光脉冲参数估计和优化等。相较于模型驱动的传统算法, DNN 从数据中拟合输入-输出间的映射关系。轻量化网络可显著提升计算效率,降低算力需求。目前,

深度学习方法已经实现了较高的预测精度,且相较于传统方法大幅度提升了计算效率^[40,47]。除了加速数值计算之外,深度学习方法也展现出下列潜在优势^[4,40]:

1) 可对缺少精确数学理论或物理模型的复杂的传输场景进行数据驱动建模,弥补常规模型驱动方法的不足;2) 可实现光脉冲传输系统端对端建模,避免常规模型驱动方法中函数结构嵌套、重复迭代等,从而有效降低了仿真系统的复杂度。然而,将深度学习算法应用于光纤中超短脉冲的传输仍然面临一些挑战。一个重要的研究方向是将深度学习构架与物理模型相结合,通过在损失函数中加入物理约束^[37]、将物理模型与神经网络耦合^[59]、在神经网络输入中内嵌物理参数^[2]等方法来提升网络的泛化性。迁移学习、强化学习等更科学的训练方式也有利于增强网络在实际系统中的可扩展性,减少网络训练时间和数据成本,拓展深度学习方法的应用场景。

参 考 文 献

- [1] Agrawal G P. Nonlinear fiber optics[M]. 4th ed. Amsterdam: Academic Press, 2007.
- [2] Jiang X T, Wang D S, Fan Q R, et al. Physics-informed neural network for nonlinear dynamics in fiber optics[J]. *Laser & Photonics Reviews*, 2022, 16(9): 2100483.
- [3] Jiang M, Wu H S, An Y, et al. Fiber laser development enabled by machine learning: review and prospect[J]. *Photonix*, 2022, 3(1): 1-27.
- [4] Salmela L, Tsipinakis N, Foi A, et al. Predicting ultrafast nonlinear dynamics in fibre optics with a recurrent neural network [J]. *Nature Machine Intelligence*, 2021, 3(4): 344-354.
- [5] Zou X H, Lu B, Pan W, et al. Photonics for microwave measurements[J]. *Laser & Photonics Reviews*, 2016, 10(5): 711-734.
- [6] 胡明列, 蔡宇. 中红外波段超快光纤激光器研究进展[J]. *中国激光*, 2020, 47(5): 0500009.
Hu M L, Cai Y. Research progress on mid-infrared ultrafast fiber laser[J]. *Chinese Journal of Lasers*, 2020, 47(5): 0500009.
- [7] Genty G, Salmela L, Dudley J M, et al. Machine learning and applications in ultrafast photonics[J]. *Nature Photonics*, 2021, 15(2): 91-101.
- [8] Martins G R, Silva L C B, Segatto M E V, et al. Design and analysis of recurrent neural networks for ultrafast optical pulse nonlinear propagation[J]. *Optics Letters*, 2022, 47(21): 5489-5492.
- [9] Zhu B W, Fang Y, Liu W, et al. Predicting the dynamic process and model parameters of vector optical solitons under coupled higher-order effects via WL-tsPINN[J]. *Chaos, Solitons and Fractals*, 2022, 162: 112441.
- [10] Wu X Q, Peng J S, Boscolo S, et al. Intelligent breathing soliton generation in ultrafast fibre lasers[J]. *Laser Photonics Reviews*, 2022, 16(2): 2100191.
- [11] Wu X Q, Zhang Y, Peng J S, et al. Farey tree and devil's staircase of frequency-locked breathers in ultrafast lasers[J]. *Nature Communications*, 2022, 13(1): 1-10.
- [12] Andral U, Fodil R S, Amrani F, et al. Fiber laser mode locked through an evolutionary algorithm[J]. *Optica*, 2015, 2(4): 275-278.
- [13] Pu G Q, Yi L L, Zhang L, et al. Intelligent programmable mode-locked fiber laser with a human-like algorithm[J]. *Optica*, 2019, 6(3): 362-369.
- [14] 王一同, 周宏强, 闫景道, 等. 基于深度学习算法的计算光学研究进展[J]. *中国激光*, 2021, 48(19): 1918004.
Wang Y T, Zhou H Q, Yan J X, et al. Advances in computational optics based on deep learning[J]. *Chinese Journal of Lasers*, 2021, 48(19): 1918004.
- [15] 彭家俊, 李晓辉, 郝孙凡, 等. 机器学习策略下的超快光子学设计: 回顾与展望(特邀)[J]. *光子学报*, 2022, 51(8): 0851518.
Peng J J, Li X H, Xi S F, et al. Intelligent ultrafast photonics based on machine learning: review and prospect(invited)[J]. *Acta Photonica Sinica*, 2022, 51(8): 0851518.
- [16] 蒙建宇, 张洪波, 张敏, 等. 基于 IPCA-DNN 算法的光纤非线性损伤补偿[J]. *光学学报*, 2021, 41(24): 2406002.
Meng J Y, Zhang H B, Zhang M, et al. Nonlinear damage compensation of optical fiber based on IPCA-DNN algorithm[J]. *Acta Optica Sinica*, 2021, 41(24): 2406002.
- [17] 鞠京泽, 刘庆添, 李宏钊, 等. 基于卷积神经网络的多任务光性能监测[J]. *光学学报*, 2022, 42(22): 2206002.
Ju J Z, Liu Q T, Li H Z, et al. Multi-task optical performance monitoring based on convolutional neural network[J]. *Acta Optica Sinica*, 2022, 42(22): 2206002.
- [18] 胡进坤, 郭晓洁, 李建平, 等. 基于深度学习的多模光纤通信系统的模式与模群识别[J]. *光学学报*, 2022, 42(4): 0406004.
Hu J K, Guo X J, Li J P, et al. Deep learning-based recognition of modes and mode groups in multimode optical fiber communication system[J]. *Acta Optica Sinica*, 2022, 42(4): 0406004.
- [19] Yan Q Q, Deng Q H, Zhang J, et al. Low-latency deep-reinforcement learning algorithm for ultrafast fiber lasers[J]. *Photonics Research*, 2021, 9(8): 1493-1501.
- [20] 张逸文, 蔡宇, 苑莉薪, 等. 基于循环神经网络的超短脉冲光纤放大器模型(特邀)[J]. *红外与激光工程*, 2022, 51(1): 20210857.
Zhang Y W, Cai Y, Yuan L X, et al. Ultra-short pulse fiber amplifier model based on recurrent neural network(invited) [J]. *Infrared and Laser Engineering*, 2022, 51(1): 20210857.
- [21] 李治龙, 张卫华, 王一民, 等. 基于机器学习的布里渊光时域分析传感系统温度提取研究进展[J]. *激光与光电子学进展*, 2021, 58(13): 1306022.
Li Z L, Zhang W H, Wang Y M, et al. Advances of machine learning in Brillouin optical time domain analysis sensing systems for temperature extraction[J]. *Laser & Optoelectronics Progress*, 2021, 58(13): 1306022.
- [22] Musumeci F, Rottondi C, Nag A, et al. An overview on application of machine learning techniques in optical networks[J]. *IEEE Communications Surveys and Tutorials*, 2019, 21(2): 1383-1408.
- [23] 孟球, 胡海峰, 胡金洲, 等. 基于深度学习的多模光纤散射介质成像重建[J]. *中国激光*, 2020, 47(12): 1206005.
Meng L, Hu H F, Hu J Z, et al. Image reconstruction of multimode fiber scattering media based on deep learning[J]. *Chinese Journal of Lasers*, 2020, 47(12): 1206005.
- [24] Yu T, Zhang S J, Chen W, et al. Phase dual-resolution networks for a computer-generated hologram[J]. *Optics Express*, 2022, 30(2): 2378-2389.
- [25] Midtvedt D, Mylnikov V, Stilgoe A, et al. Deep learning in light-matter interactions[J]. *Nanophotonics*, 2022, 11(14): 3189-3214.
- [26] Pipek J. Simulation-based machine learning for optoelectronic device design: perspectives, problems, and prospects[J]. *Optical and Quantum Electronics*, 2021, 53(4): 175.
- [27] Qu Y R, Zhu H Z, Shen Y C, et al. Inverse design of an integrated-nanophotonics optical neural network[J]. *Science Bulletin*, 2020, 65(14): 1177-1183.
- [28] Goodfellow I, Bengio Y, Courville A. Deep learning[M]. Cambridge: The MIT Press, 2016.
- [29] Smirnov E A, Timoshenko D M, Andrianov S N. Comparison of regularization methods for ImageNet classification with deep convolutional neural networks[J]. *AASRI Procedia*, 2014, 6: 89-94.
- [30] Yu Y, Si X S, Hu C H, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. *Neural Computation*, 2019, 31(7): 1235-1270.
- [31] 张顺, 龚怡宏, 王进军. 深度卷积神经网络的发展及其在计算机

- 视觉领域的应用[J]. 计算机学报, 2019, 42(3): 453-482.
- Zhang S, Gong Y H, Wang J J. The development of deep convolution neural network and its applications on computer vision [J]. Chinese Journal of Computers, 2019, 42(3): 453-482.
- [32] Bennett K P, Parrado-Hernández E. The interplay of optimization and machine learning research[J]. Journal of Machine Learning Research, 2006, 7: 1265-1281.
- [33] Ruder S. An overview of gradient descent optimization algorithms [EB/OL]. (2016-09-15)[2022-11-09]. <https://arxiv.org/abs/1609.04747>.
- [34] Boscolo S, Dudley J M, Finot C. Modelling self-similar parabolic pulses in optical fibres with a neural network[J]. Results in Optics, 2021, 3: 100066.
- [35] Boscolo S, Finot C. Artificial neural networks for nonlinear pulse shaping in optical fibers[J]. Optics & Laser Technology, 2020, 131: 106439.
- [36] Jiang X T, Wang D S, Fan Q R, et al. Solving the nonlinear Schrödinger equation in optical fibers using physics-informed neural network[C]//2021 Optical Fiber Communications Conference and Exhibition (OFC), June 6-10, 2021, San Francisco, CA, USA. New York: IEEE Press, 2021.
- [37] Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations [J]. Journal of Computational Physics, 2019, 378: 686-707.
- [38] Raissi M. Deep hidden physics models: deep learning of nonlinear partial differential equations[EB/OL]. (2018-01-20)[2022-11-09]. <https://arxiv.org/abs/1801.06637>.
- [39] Wang R Q, Ling L M, Zeng D L, et al. A deep learning improved numerical method for the simulation of rogue waves of nonlinear Schrödinger equation[J]. Communications in Nonlinear Science and Numerical Simulation, 2021, 101: 105896.
- [40] Wang D S, Song Y C, Li J, et al. Data-driven optical fiber channel modeling: a deep learning approach[J]. Journal of Lightwave Technology, 2020, 38(17): 4730-4743.
- [41] Yang H, Niu Z K, Xiao S L, et al. Fast and accurate optical fiber channel modeling using generative adversarial network[J]. Journal of Lightwave Technology, 2021, 39(5): 1322-1333.
- [42] Yang H, Niu Z K, Zhao H C, et al. Fast and accurate waveform modeling of long-haul multi-channel optical fiber transmission using a hybrid model-data driven scheme[J]. Journal of Lightwave Technology, 2022, 40(14): 4571-4580.
- [43] Tegin U, Dinç N U, Moser C, et al. Reusability report: predicting spatiotemporal nonlinear dynamics in multimode fibre optics with a recurrent neural network[J]. Nature Machine Intelligence, 2021, 3(5): 387-391.
- [44] Wu G Z, Fang Y, Wang Y Y, et al. Predicting the dynamic process and model parameters of the vector optical solitons in birefringent fibers via the modified PINN[J]. Chaos, Solitons & Fractals, 2021, 152: 111393.
- [45] Fang Y, Wu G Z, Wen X K, et al. Predicting certain vector optical solitons via the conservation-law deep-learning method[J]. Optics & Laser Technology, 2022, 155: 108428.
- [46] Fang Y, Wang Y Y, Liu W, et al. Data-driven prediction of soliton solutions of the higher-order NLSE via the strongly-constrained PINN method[J]. Computers & Mathematics With Applications, 2022, 127: 144-153.
- [47] Salmela L, Hary M, Mäbed M, et al. Feed-forward neural network as nonlinear dynamics integrator for supercontinuum generation[J]. Optics Letters, 2022, 47(4): 802-805.
- [48] Sui H, Zhu H N, Wu J, et al. Modeling pulse propagation in fiber optical parametric amplifier by a long short-term memory network [J]. Optik, 2022, 260: 169125.
- [49] 隋皓, 朱宏娜, 张妍, 等. 基于深度学习的光纤超短脉冲啁啾放大研究[J]. 光学学报, 2023, 43(3): 0319001.
- Sui H, Zhu H N, Zhang Y, et al. Ultrashort chirped pulse amplification in fiber based on deep learning[J]. Acta Optica Sinica, 2023, 43(3): 0319001.
- [50] Gautam N, Choudhary A, Lall B. Comparative study of neural network architectures for modelling nonlinear optical pulse propagation[J]. Optical Fiber Technology, 2021, 64: 102540.
- [51] Shoreh M H, Kamilov U S, Papadopoulos I N, et al. A learning approach to optical tomography[J]. Optica, 2015, 2(6): 517-522.
- [52] Rivenson Y, Wu Y C, Ozcan A. Deep learning in holography and coherent imaging[J]. Light: Science & Applications, 2019, 8(1): 1-8.
- [53] Borhani N, Kakkava E, Moser C, et al. Learning to see through multimode fibers[J]. Optica, 2018, 5(8): 960-966.
- [54] Li Y Z, Xue Y J, Tian L. Deep speckle correlation: a deep learning approach towards scalable imaging through scattering media[J]. Optica, 2018, 5(10): 1181-1190.
- [55] Zahavy T, Dikopoltsev A, Moss D, et al. Deep learning reconstruction of ultrashort pulses[J]. Optica, 2018, 5(5): 666-673.
- [56] Kleinert S, Tajalli A, Nagy T, et al. Rapid phase retrieval of ultrashort pulses from dispersion scan traces using deep neural networks[J]. Optics Letters, 2019, 44(4): 979-982.
- [57] Stanfield M, Ott J, Gardner C, et al. Real-time reconstruction of high energy, ultrafast laser pulses using deep learning[J]. Scientific Reports, 2022, 12(1): 1-11.
- [58] Sui H, Zhu H N, Cheng L, et al. Deep learning based pulse prediction of nonlinear dynamics in fiber optics[J]. Optics Express, 2021, 29(26): 44080-44092.
- [59] Sui H, Zhu H N, Luo B, et al. Physics-based deep learning for modeling nonlinear pulse propagation in optical fibers[J]. Optics Letters, 2022, 47(15): 3912-3915.
- [60] Sidelnikov O, Redyuk A, Sygletos S, et al. Advanced convolutional neural networks for nonlinearity mitigation in long-haul WDM transmission systems[J]. Journal of Lightwave Technology, 2021, 39(8): 2397-2406.
- [61] Tay K G, Pakarzadeh H, Huong A, et al. Gain prediction of dual-pump fiber optic parametric amplifier based on artificial neural network[J]. Optik, 2022, 253: 168579.
- [62] He J Y, Li C Y, Wang P, et al. Soliton molecule dynamics evolution prediction based on LSTM neural networks[J]. IEEE Photonics Technology Letters, 2022, 34(3): 193-196.
- [63] Herrera R A. Evaluating a neural network and a convolutional neural network for predicting soliton properties in a quantum noise environment[J]. Journal of the Optical Society of America B, 2020, 37(10): 3094-3098.
- [64] Jiang M, An Y, Huang L J, et al. M^2 factor estimation in few-mode fibers based on a shallow neural network[J]. Optics Express, 2022, 30(15): 27304-27313.
- [65] Li C Y, He J Y, Liu Y G, et al. Comparing performance of deep convolution networks in reconstructing soliton molecules dynamics from real-time spectral interference[J]. Photonics, 2021, 8(2): 51.

Nonlinear Propagation Representation and Control for Ultrashort Pulse in Optical Fibers Based on Deep Learning

Sui Hao¹, Zhu Hongna^{1*}, Jia Huanyu¹, Ou Mingyu², Li Qi¹, Luo Bin², Zou Xihua²

¹*School of Physical Science and Technology, Southwest Jiaotong University, Chengdu 610031, Sichuan, China;*

²*School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, Sichuan, China*

Abstract

Significance Nonlinear ultrashort laser pulse propagation in optical fibers, which is the physical principle of fiber-based optical devices, optical signal transmission, and processing, comprises a series of complex nonlinear dynamics. It finds extensive application in the fields of fiber lasers, fiber amplifiers, and fiber communications. Generally, nonlinear ultrashort pulse propagation is governed by the nonlinear Schrödinger equation (NLSE) and can be solved using model-driven methods such as the split-step Fourier (SSF) and finite-difference methods. However, NLSE-based systems are sensitive to both the initial pulse and fiber parameters, making it difficult for traditional numerical methods to control the complex nonlinear pulse evolution in a time-efficient manner.

As a powerful tool for system parameter optimization and the construction of models of complex dynamics from observed data, deep learning (DL) algorithms have recently been applied to ultrafast photonics, optical communications, optical networks, optical imaging, and the modeling and control of nonlinear pulse propagation to reap the benefits of purely data-driven methods without any underlying governing equations. In this paper, the current key technologies and applications of the DL method for predicting nonlinear pulse dynamics in fiber optics, reconstructing ultrashort pulses, and evaluating critical pulse characteristics are summarized, and the development trends are predicted.

Progress First, a brief introduction to the DL method and practical DNN is presented. Second, the applications of DL for predicting nonlinear ultrashort pulse propagation are listed. Several types of neural networks, *i. e.*, LSTM, CNN, and FNN, have been applied to predict nonlinear pulse evolution, *i. e.*, predicting the effects of GVD and SPM on ultrashort pulse propagation, higher-order soliton compression, and supercontinuum generation, in both the temporal and spectral domains with high prediction precision. Moreover, DL methods are used for modeling optical fiber channels, resulting in a significant reduction in computation demand. Further, the PINN is verified in multiple nonlinear pulse propagations governed by the NLSE, which considers the physical and boundary constraints of the physics model. Subsequently, the optimized PINN, *i. e.*, subnet structure and adapted loss function, is applied to solve the NLSE and predict the nonlinear soliton dynamics with higher prediction accuracy and generalizability.

Third, the DL applications for solving the inverse problems of nonlinear propagation of ultrashort pulses are discussed. Therefore, FNNs and CNNs are utilized to reconstruct the ultrashort pulse and counteract the effects of nonlinearity without prior knowledge. The ultrashort pulse profiles are precisely recovered using SPM and four-wave mixing effects. Furthermore, CNNs have been applied as alternatives to the DBP algorithm to compensate for the nonlinear distortions in the fiber-optic transmission systems. In addition, DNNs are used extensively in parameter estimation and optimization of optical fiber systems, including the optimal design of the pump power and pump wavelength in the FOPA (fiber optical parametric amplifier) system, predicting the collision between a single soliton and soliton molecule, realizing the extraction of important soliton characteristics, evaluating soliton properties in a quantum noise environment, and estimating the M^2 factor in few-mode fibers.

Conclusions and prospects The DL methods have become a development frontier and research hotspot in the field of predicting, modeling, controlling, and designing nonlinear pulse propagation in optical fibers. Compared to the conventional SSF method, lightweight neural networks can significantly improve computing efficiency and reduce computing demand, making it simple and convenient to study nonlinear pulse dynamics and optimize fiber-based optical systems. In addition, DL methods have the following potential advantages: (1) They can conduct pure data-driven modeling for complex propagation scenarios that lack accurate mathematical theories or physical models. (2) They can achieve flexible end-to-end modeling for typical nonlinear dynamics or transmission systems, avoiding a nested function structure and repeated iterations, which effectively reduces the complexity of the simulation system.

However, the generalization of neural networks is a critical issue that restricts the prediction of precision and accuracy. Several methods have been developed to improve the generalization of neural networks, such as adding physical constraints to the loss function, coupling a physical model to a neural network, and embedding physical parameters into the neural network input. Further, scientific training methods such as transfer learning and reinforcement learning are conducive to enhancing the scalability of the network in an actual system and reducing the time and data cost of network training. As alternatives to the traditional numerical method, the application of DL methods could aid in the understanding of nonlinear ultrashort pulse propagations as well as the design and optimization of ultrashort pulse-based optical systems.

Key words fiber optics; ultrashort laser pulse propagation; nonlinear Schrödinger equation; nonlinear optical fiber effect; deep learning method