微波光子技术及应用专题

DOI：10.16818/j.issn1001－5868.20220102

人工智能赋能的多功能光子处理系统

邹秀婷，徐绍夫，邹卫文*

（上海交通大学 区域光纤通信网与新型光纤通信系统国家重点实验室 智能微波光波融合创新中心，上海 200240）

摘　要： 智能光子处理系统(IPS)融合了人工智能(AI)技术和光子技术，旨在实现智能、高速、大带宽、高性能的信号处理。IPS 主要包括人工智能赋能的多功能光子处理系统、光子辅助的人工智能信号处理系统和基于神经拟态的光子处理系统。文章首先简要介绍了 IPS 的概念内涵，然后重点介绍作者所在课题组在人工智能赋能的多功能光子处理系统方面的研究进展，再进一步探讨人工智能赋能研究从不可解释逐渐走向可解释的发展趋势和必要性，接着介绍该课题组已开展的具有一定可解释性的人工智能赋能的光子处理系统研究，最后对全文进行总结。

关键词： 智能光子处理系统；人工智能；深度学习；赋能；可解释

中图分类号： TN929.5；TP18　**文章编号：** 1001－5868(2022)01－0021－10

AI-powered Multifunctional Photonic Processing System

ZOU Xiuting，XU Shaofu，ZOU Weiwen

（Intelligent Microwave Lightwave Integration Innovation Center，State Key Laboratory of Advanced Optical Communication Systems and Networks，Shanghai Jiao Tong University，Shanghai 200240，CHN）

Abstract： Intelligent photonic processing system (IPS) combines artificial intelligence (AI) technology and photonic technology to achieve intelligent，high-speed，large-bandwidth and high-performance signal processing. IPS mainly includes AI-powered multifunctional photonic processing system，intelligence signal processing with photonics-facilitated AI and neuromorphic photonic processing system. In this article，the concept of IPS is briefly introduced first，then the research progresses of our resaerch group in the AI-powered multifunctional photonic processing system are described，and further the development trends and necessity of explainable AI are discussed. And the research of our group on the explainable AI-powered multifunctional photonic processing system is introduced.

Key words： intelligent photonic processing system；artificial intelligence；deep learning；powered；explainable

0　引言

未来信息处理系统，如雷达、通信、电子对抗等，面临电磁环境复杂、频谱资源占用率高、数据处理量大、场景复杂等挑战，因此对处理系统提出了高频、大带宽、低延时、智能处理等需求。光子技术充分利用了光的高速、大带宽、抗干扰等优势，能够有效克服电子技术面临的带宽限制、时钟偏移、严重串音、高衰耗等瓶颈[1-2]。人工智能(AI)旨在研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术。AI 当前存在两大研究分支：以人工神经网络为基础的深度学习[3](DL)技术和以生物神经元或生物神经网络为基础的神经拟态[4]技术。其中，深度学习利用深度神经网络，即多层人工神经网络实现智能处理。人工神经网络的基本单元是通过对生物神经元数学抽象得到的人工神经元。神经拟态利用技术手段实现生物神经元或生物神经网络行为背后的物

收稿日期：2022－01－01.

基金项目：国家重点研发计划项目(2019YFB2203700)；国家自然科学基金项目(61822508).

* **通信作者：**邹卫文　**E-mail：** wzou@sjtu.edu.cn

理机理。智能光子处理系统(IPS)将光子技术的高速、大带宽、多自由度、低功耗、抗干扰等优点和AI技术的智能处理能力有机融合,旨在实现与人脑一样聪明甚至超越人脑的、低功耗的智能处理能力。IPS包含三大层次:人工智能赋能的光子处理系统(AI-powered IPS)、光子辅助的人工智能信号处理系统(IPS with Photonics-Facilitated AI)和基于神经拟态的光子处理系统(Neuromorphic IPS)[5]。

人工智能赋能的光子处理系统利用AI算法对光子处理系统的输出进行处理,得到高质量的信号或高级信息。深度学习具有强大的特征提取能力,能够从数据中提取高级特征,实现任意复杂函数映射[6-7]。随着算力、算法和大数据的发展,近年来,深度学习成为AI的主流技术,给图像识别[8]、语音分离[9]、游戏[10]、生物医疗[11-12]、光学成像[13]、光通信[14]、光学系统设计[15-16]等领域带来重大突破。将深度学习引入光子处理系统,一方面能够克服光子处理系统的缺陷,提高系统性能,另一方面能够充分结合两者优势,激发新的处理潜能。近年来,国内外开展了一系列人工智能赋能的光子处理系统研究,包括借助深度学习技术提高布里渊光学时域分析仪对温度和应变测量的速度[17]或精度[18]、提高光子模数转换精度[19-21]、提高微波光子测频系统的频率测量精度[22-23]、提高3D微波光子雷达成像的速度和精度[24]、实现高性能光子接收系统[25-28]、实现高精度自动目标识别[29-30],以及实现宽带混合信号分选[31]等。

光子辅助的人工智能信号处理系统利用光子技术加速AI算法。深度学习作为当前计算精度最高的AI算法,需要大计算量,因而带来算力需求高、耗能严重等问题。另外,摩尔定律也逐渐逼近终点[32]。因此,亟需新的计算架构来实现高速、低功耗计算。2017年,Shen等提出基于片上马赫-曾德尔干涉仪阵列的光子神经网络计算原型[33],这一研究工作掀起了光子神经网络的研究热潮。近年来国内外报道了一系列光子神经网络相关工作,包括基于光学衍射原理的光子神经网络架构[34]、基于光学相变材料的光子神经网络架构[35]、基于波分复用和调制器阵列的光子神经网络架构[36-39]等。

基于神经拟态的光子处理系统利用光子技术从物理机理角度出发,模拟生物神经元或生物神经网络行为,实现智能处理能力。国内外近期研究成果表明光电器件具有类神经元动力学特性,能够实现多功能的IPS,包括在激光器中发现类神经元兴奋动力学响应[40]、基于半导体放大器模拟突触可塑性机制[41]、基于相变材料和环形谐振器实现具有自学习能力的全光学尖峰神经突触网络芯片[42]、基于分布式反馈激光二极管的光子神经拟态网络,以及用于时空模式识别[43]等。

以上简要介绍了IPS三大层次的内涵以及部分国内外研究成果,接下来将重点介绍本课题组在人工智能赋能的多功能光子处理系统方面的研究、AI赋能工作从不可解释逐渐走向可解释的发展趋势和必要性以及本课题组已开展的具有一定可解释性的人工智能赋能的光子处理系统工作,最后对全文进行总结。

1 人工智能赋能的多功能光子处理系统研究

光子处理系统虽然能够克服电子瓶颈,满足系统对高速、大带宽等需求,但是受限于光子链路中存在非线性会引起信号畸变以及光子系统不稳定难以调控等缺陷,系统性能不高。利用人工智能的代表性技术——深度学习对系统输出进行智能处理,能够得到高质量的系统输出。另外,光子系统的高速、大带宽优势和深度学习算法的智能处理能力结合起来不仅可以服务于光子处理系统自身,还可激发新的处理潜能,进而可以拓宽系统应用场景,更好地服务于信息系统。因此,本节将分为深度学习提升光子处理系统性能和深度学习激发新的处理潜能两部分介绍课题组已开展工作。本节介绍的人工智能赋能工作将系统和深度学习简单拼接,在设计或训练网络时,并未引入系统或领域相关知识,因此本节介绍的研究中应用的AI算法具有不可解释性,属于"黑盒"赋能。

1.1 深度学习提高光子处理系统性能

1.1.1 深度学习赋能的光子模数转换系统

图1是基于光采样和电量化的光子模数转换系统(PADC)架构。电光强度调制器将射频信号调制至光上,实现光电转换;经射频信号调制的光脉冲经过并行解复用模块完成串并转换,每路子脉冲速率降为原来脉冲的$1/N$,N为系统通道数;子脉冲经过光电转换后变成电信号,再经电模数转换器转换为数字信号。由于光脉冲速率降低,光电转换器件和电子模数转换器所需带宽仅为降速后的脉冲速率

的 1/2。理论上，该 PADC 系统能够实现高速、大带宽和高精度的模数转换功能，同时能够缓解电后端压力。但是，由于电光强度调制器的非线性电光传输函数，射频信号被调制到光上的过程会产生谐波失真。另外，串行的光脉冲被并行化到多个通道中，要实现理想的并行化，必须保证各个通道中降速后的光脉冲幅度严格相等，时间严格对齐，脉冲形状严格一致。然而，由于不可避免的工艺误差，每个通道之间总存在分光幅度不均匀、绝对光程不一致的情况，从而导致每个通道发生幅度与时间上的失配现象，引入杂散。上述缺陷恶化了光子模数转换系统的无杂散动态范围（SFDR），降低了系统模数转换的有效位数（ENOB）。

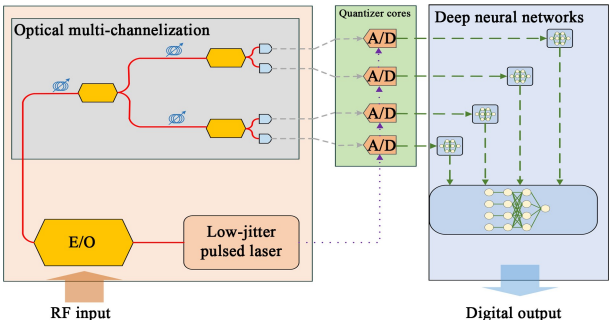


图 1　深度学习赋能的光子模数转换系统架构

本课题组提出深度学习赋能的光子模数转换系统（DL-PADC）[19]，利用深度神经网络校正了电光调制器引入的非线性，补偿了通道间的时间失配，最终实现了高速、大带宽、高精度的模数转换。具体来说，我们搭建了一个全卷积的深度残差网络（ResNet），利用梯度下降优化算法，使网络在迭代多次后能够拟合网络输入和网络参考之间的复杂函数映射，因而网络输出逼近网络参考。其中，网络输入为光子模数转换系统接收的带有杂散的射频信号的数字波形，网络参考为相应地不带有杂散的射频信号的数字波形。如图 2 所示，当输入光子模数转换系统的射频信号的频率为 3.44 GHz 时，有效位数从 4.66 bit 提高到 7.28 bit，当频率为 21.13 GHz 时，有效位数从 4.53 bit 提高到 7.07 bit。实验结果表明，系统输出经过训练好的 ResNet 网络后，网络输出其无杂散动态范围和有效位数接近系统量化噪声和抖动噪声的极限。因此，只需在光子模数转换系统后接一个 AI 算法，让网络从数据中学习数据之间的复杂映射，就能得到高性能的光子模数转换系统，避免了传统方法通过添加额外硬件电路[44-45]或提取被采信号先验信息等复杂操作[46]。

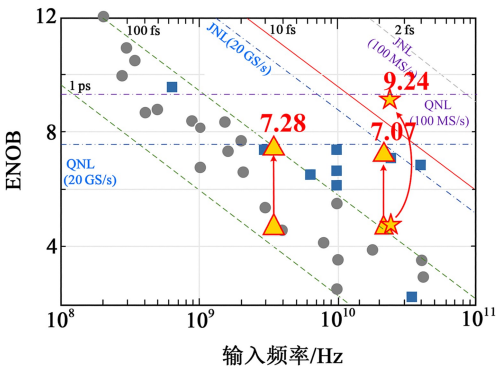


图 2　ResNet 赋能前后的光子模数转换系统的有效位数

在微波光子雷达等实际应用中，光子模数转换系统用作接收机。由于发射信号经过物体表面会发生随机散射，因此系统接收到的信号通常是宽带复杂信号。为了将上述 DL-PADC 概念推广到雷达等实际应用中，课题组进一步开展了用于宽带复杂信号接收的深度学习赋能的光子模数转换器研究[25]。具体来说，我们搭建了一个 MU-net 神经网络，遵循宽带雷达的多重散射原理模拟了数千个随机雷达回波，用于构建数据集。利用该神经网络可以在输入信号为宽带复杂信号的情况下对光子模数转换系统中的非线性失配缺陷进行校正，输出高质量的雷达回波信号。该工作以宽带回波接收为例，实验验证了 MU-net 对于宽带复杂信号恢复的可行性。结果

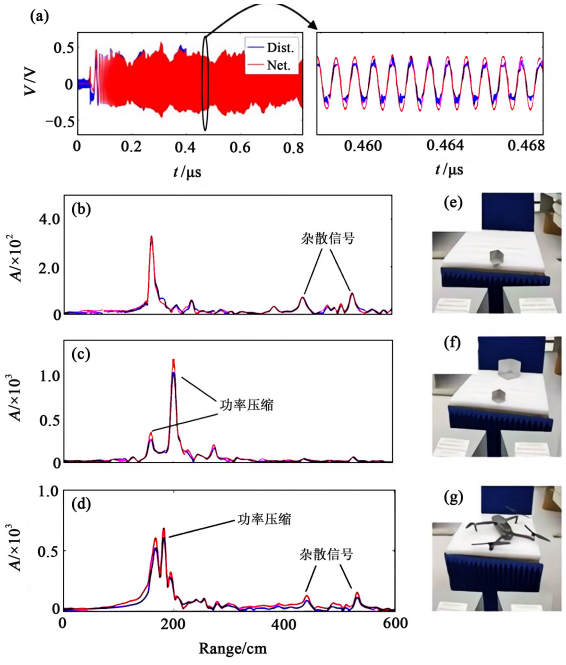


图 3　MU-net 赋能的光子模数转换系统的雷达回波目标检测结果

表明，MU-net 可以将信号的无杂散动态范围提升约 20 dB。如图 3 所示，无杂散动态范围提升后，更容易从脉冲压缩得到的一维像中探测到弱目标。另外，在输入信号功率大范围变化的情况下，MU-net 依然可以正常输出高质量信号。

1.1.2 深度学习赋能的布里渊瞬时频率测量系统

在雷达等应用中，需要探测未知微波信号的瞬时频率。前期我们开展了基于布里渊散射的宽带微波瞬时频谱测量（BIFM）技术研究[47]，利用受激布里渊散射增益谱的窄带滤波特性以及合理地设置泵浦信号的扫频形式，建立了频率与幅度之间的一一对应关系，实现了对未知宽带微波信号的瞬时频率测量。但是，由于该系统中泵浦光和探测光的偏振状态不理想、相互作用不充分、受激布里渊散射作用的媒介（如光纤）存在损耗以及电光调制器、光电探测器存在非理想等因素，光电探测器检测到的探测光的布里渊幅度增益信号存在很大噪声。因此，经频率-幅度映射后得到的未知宽带微波信号的瞬时频率误差较大（100 MHz 左右）。

基于 DL-PADC 研究思路，本课题组进行了深度学习赋能的布里渊瞬时频率测量系统（DL-BIFM）研究[22]。我们搭建了一个卷积神经网络（CNN），用来提高布里渊瞬时频率测量系统测得的微波信号的瞬时频率的精度。如图 4 所示，在 DL-BIFM 系统中，测频硬件系统对未知微波信号进行瞬时频率测量，系统输出误差比较大的瞬时频率，该瞬时频率进入训练好的卷积神经网络中，网络输出精确的瞬时频率。实验结果表明，经过网络后，瞬时频率的最大误差从 100 MHz 左右优化至 20 MHz 左右，均方根误差从 57.0 MHz 优化至 10 MHz 左右。另外，系统的灵敏度提高至 7.8 dBm，系统的动态范围更大。

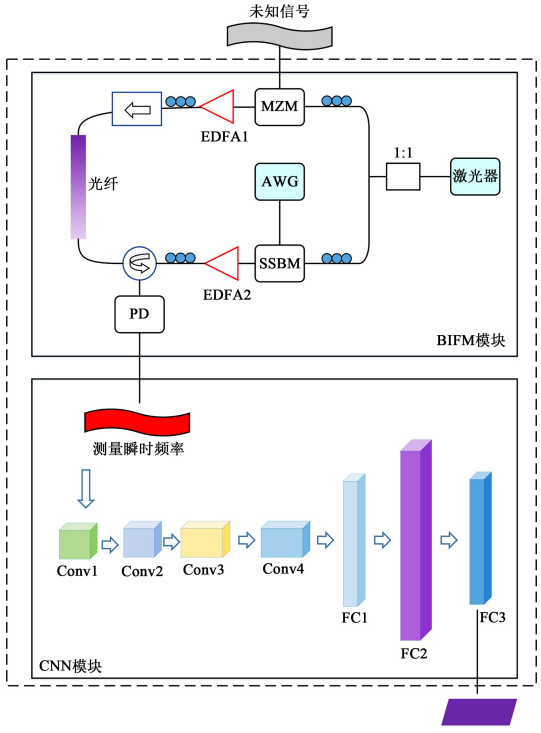


图 4 深度学习赋能的布里渊瞬时频率测量系统架构

1.1.3 深度学习赋能的光子下变频接收系统

为了接收高频、大带宽的信号，传统的电学方法需要经过一系列复杂的混频操作才能将射频信号下变频至基带。图 5 是一个由激光器、电光调制器（MZM）、光电探测器（PD）等组成的最基本的微波光子链路。该链路是一个天然的混频器，可实现将高频、大带宽信号接收下来的同时对信号完成下变频操作。但是，该链路中的光子有源元件，如激光器、电光调制器和光电探测器产生的噪声会恶化接收信号的质量。

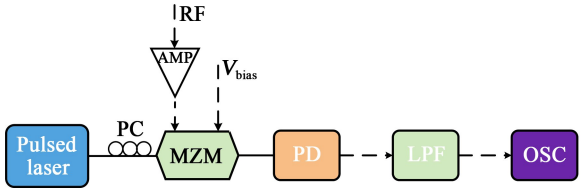


图 5 光子下变频架构

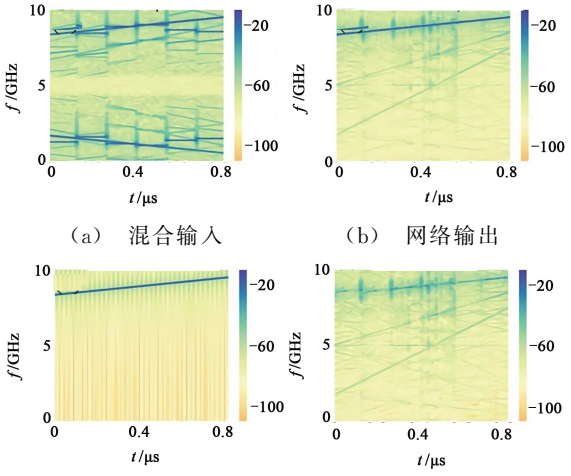
课题组提出一种基于光子下变频和深度学习的多波段低噪声信号接收系统（DL-PR）[26]。其中光子下变频系统直接接收信号，神经网络用于降低信号中的噪声。结合迁移学习，网络实现了针对不同类型或不同频带频率信号的超快速再训练能力。经神经网络降噪后，信号质量明显提升。实验结果表明，该神经网络可以对不同类型、不同信噪比、不同占空比收发信号提升 8 dB 左右信噪比。同时，利用迁移学习，网络只需要 1/3 的数据和 1/10 的时间消耗就能完成训练且达到同样降噪效果。

1.2 深度学习激发新的处理潜能

1.2.1 深度学习赋能的宽带光子接收混合信号分选

雷达、电子对抗等信息系统面临的电磁环境复杂，通常宽带接收系统接收到的信号中除了雷达自身发射的已知信号外还混杂了其他未知信号。为了从接收信号中获得想要的信息，首先要从混合信号中分离出已知信号，即实现信号分选（SS）。传统分选算法如独立成分分析[48]、变换域滤波[49]等只对一些简单的信号混合情况有效。

课题组提出基于光子模数转换系统和深度学习的宽带混合信号分选(DL-SS)研究[31]。我们搭建并训练了一个具有分选功能的扩张卷积神经网络(DFCN)。光子模数转换系统对大带宽的天线接收到的混合信号进行采样,得到数字混合信号,混合信号输入 DFCN 网络,网络从中提取想要的信号。如图 6(a)所示,光子模数转换系统接收到的混合信号中有和目标信号同频段的信号,还有系统的电光调制器引入的非线性以及通道间的时间失配引入的杂散。图 6(c)所示的线性调频信号是我们期望从混合信号中提取的信号。从图 6(b)可以看出,混合信号中的线性调频信号被提取出来。线性调频信号中存在功率较低的非线性是由 DFCN 网络的非线性带来的。进一步地,我们将 DFCN 网络应用在数字调制信号分选中,并对结果进行了定量分析。如图 7 所示,相较于分离前,二进制相移键控调制信号的平均误比特率提高了 3 个数量级。



(a)　混合输入　　(b)　网络输出

(c)　ground truth　　(d)　residual

图 6　线性调频信号分选结果

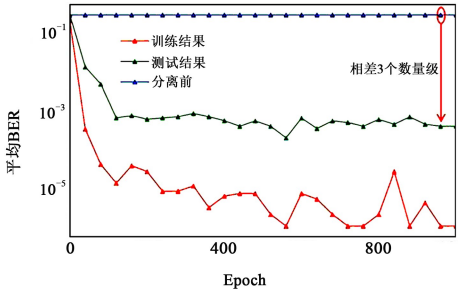
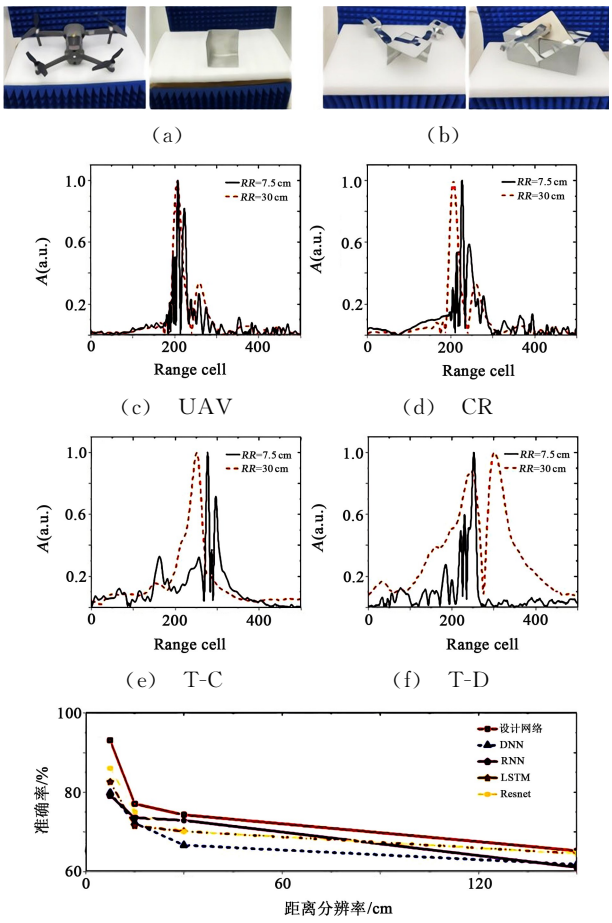


图 7　二进制相移键控调制信号分选结果

1.2.2　深度学习赋能的高精度自动目标识别

高分辨率雷达的兴起为目标识别提供了新的途径。雷达发射信号照射目标,在目标表面发生散射,接收机接收雷达回波,对雷达回波进行脉冲压缩,得到一维像。一维像反映了目标的几何结构特征,通过对其进一步分析,能够得出雷达探测到的目标类型等信息,进而指导下一步决策。雷达信号照射到不同目标,其产生的回波不同,因此可以根据一维像对目标进行识别。另外,雷达发射信号的带宽越大,距离分辨率越高,接收到的宽带回波信号经脉冲压缩后得到的一维像越细节。

通过结合光子模数转换系统大带宽的优势以及深度学习强大的自动提取特征能力,来实现高精度的自动目标识别(DL-ATR)[29]。相对于传统电子雷达,微波光子雷达系统能发射大带宽的信号去探测目标,光子模数转换系统作为微波光子雷达的接收机,能接收到大带宽的雷达回波信号[50]。我们研究了一维像距离分辨率(RR)与识别准确率的对应关系。用带宽分别为 0.1,0.5,1.0 和 2.0 GHz,相应的距离分辨率为 150,30,15 和 7.5 cm 的信号去探测四种不同的目标。然后搭建了一个卷积神经网络对一维像进行识别。实验结果如图 8 所示,当一维像距离分辨率越高时,目标识别的准确率越高。



(g)　不同分辨率下的目标识别准确率

图 8　卷积神经网络对一维像识别结果

另外，该工作设计的网络优于其他多种深度学习算法，最终自动目标识别性能比已有多数电子方案水平高十个百分点以上。

2 可解释的人工智能赋能研究

2.1 人工智能赋能的发展趋势

从数学上看，深度神经网络是一个由多层非线性函数嵌套而成的复杂函数，人们对网络输出的结果知其然而不知其所以然，因此网络具有“黑箱子”特性[51]。近年来，深度学习的表现和应用饱受质疑，网络的不可解释性引起广泛关注。另外，深度学习的“黑箱子”特性导致网络结果缺乏依据，因此在诸如医疗诊断、自动驾驶等高风险领域，盲目信任网络可能会带来灾难性的后果。2016 年，欧盟提出法规，人们有权获得 AI 决策结果的解释[52]。近年来，国内外的学术界、工业界纷纷出版了 AI 可解释相关文档。2020 年，Geirhos 等提出深度学习具有捷径学习[53]的特点，该结论从某种程度上解释了当前大多数深度学习算法应用饱受质疑的原因。当数据集不理想时，如在分类任务中，不同类别的图像背景之间明显不一样、不同类别的样本数量存在偏差等时，或当网络输入和网络参考之间存在简单映射即可使得损失函数迅速下降时，网络很容易选择最简单的问题解决方案收敛，导致虽然损失函数降得很低，但是网络并未学到物理规律，进而带来网络的泛化能力弱、鲁棒性差等问题。

为了解决当前 AI 算法“黑箱子”特性导致其在应用落地中会遇到的信任、公平、安全、伦理等问题，如图 9 所示，AI 算法赋能各个领域的相关工作，需要从“黑盒”赋能逐渐走向“灰盒”赋能，最终实现“白盒”赋能，也即 AI 赋能要逐渐具有可解释性。其中“黑盒”赋能即在训练时只让 AI 算法从数据集中学习，不加以其他知识辅助，训练好的网络也不具备任何可解释性；“灰盒”赋能可分为“前向”和“后向”两个角度。“前向”即根据领域特点设计网络结构和数据集或在训练时给网络加入领域知识，指导网络学习物理本质规律，对训练好的网络提供某种程度上的解释；“后向”是对训练好的网络进行解释，给出网络输出的依据。“白盒”赋能即所用的 AI 算法具有透明性，能够清晰解释算法是如何实现赋能的。近年来，学术界和工业界开展了大量深度学习可解释工作。当前的可解释技术大致可以分为三类：“前向”可解释、“后向”可解释性和神经网络理论。“前向”可解释即训练后得到一个具有可解释性的网络；“后向”可解释是对训练好的神经网络进行解释；神经网络理论是从深度神经网络理论出发，研究网络自身性质。“前向”可解释技术是实现“灰盒”赋能的重要技术，“后向”可解释技术以及神经网络理论是最终实现“白盒”赋能，也即网络具有“白盒”特性的关键支撑技术。实现“白盒”赋能的一条高效科学的途径是先基于“前向”可解释技术指导网络朝物理规律方向收敛，然后对训练好的网络进行“后向”可解释。深度学习可解释是当前的科研难题，上述可解释技术均处于发展的初级阶段。

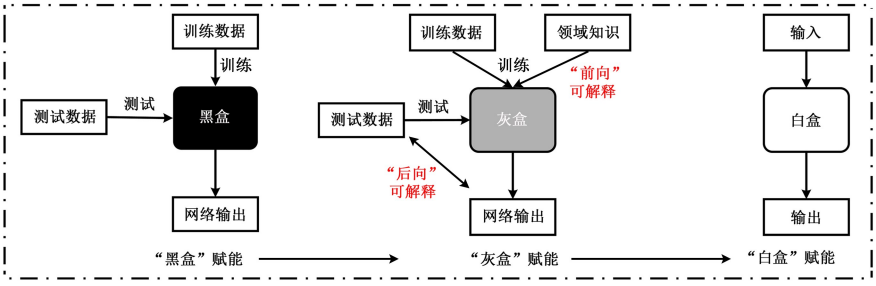


图 9 人工智能赋能的三阶段

当前常见的一种“前向”可解释技术通常是通过添加规则项等改造网络的损失函数以及从网络结构或输入层面添加领域信息，指导网络训练。Zhang 等提出基于互信息学习的图像识别网络[54]。该工作在损失函数中添加互信息项，使得网络高层的每个滤波器只学习被识别对象的某个特征，从而在测试过程中，可以根据高层滤波器学习到的特征给出网络判决依据。2017 年，Karpatne 等提出物理指导的神经网络(PGNN)[55]，PGNN 将科学知识引入网络，充分利用物理知识和数据，指导网络收敛，确保网络学到物理上一致的解。PGNN 网络在损失函数中引入科学知识，可以确保网络收敛得到物理一致的结果，同时还能降低训练集数量、提高网络的泛化能力和鲁棒性[56]。

常见的“后向”可解释技术有可视化、显著图、函数分解等操作。Zeiler 等[57]对训练好的网络做反卷积，然后可视化网络各层。结果表明网络前几层提取输入的低级特征如边缘，中间层对低级特征进行整合、提取中级特征如形状，后面几层网络提取高水平特征如物体。显著图方法[58]可以识别网络输入中与网络输出最相关的属性，因此可以从整体上检查判决依据是否合理。常见的函数分解法有深度泰勒分解[59]和逐层相关传播[60]，它们将网络输出分解至网络输入，得出网络输入的每个元素对网络输出的贡献，从而在一定程度上理解网络输出结果。

神经网络理论以研究网络自身性质为出发点，根据网络性质解释网络行为。Xu 等提出神经网络具有“频率原则”[61]和“嵌入原则”[62]等性质。“频率原则”表明在网络迭代过程中，会先拟合低频后拟合高频，此处的频率指的是神经网络的响应频率。因此当神经网络需要拟合的目标函数具有低频占优的特性，那么深度学习比较容易取得很好的效果，反之则不然[63]。“嵌入原则”指出在网络迭代过程中，宽网络的极值点“包含”所有窄网络的极值点，窄网络的局部最小值点在宽网络里会变成鞍点，可能使得宽网络更加容易训练，并且宽网络的极值点退化更高（与数据无关的退化），也即大网络像小网络却又比小网络好。“嵌入原则”为简化宽神经网络提供了解释，并揭示了训练过程中潜在的隐式低复杂度的正则化。

2.2　可解释的人工智能赋能的多功能光子处理系统研究

为了 AI 能够真正地服务于光子处理系统，也即人工智能赋能的多功能光子处理系统能够应用落地，课题组提出基于“前向”可解释和“后向”可解释充分结合的“灰盒”赋能研究。在“前向”过程中，基于对光子系统的理解设计网络架构、数据集、训练方法等，实现利用少量样本、轻巧的网络架构即可使神经网络学习到光子处理系统的系统特性或物理规律，从而使收敛后的网络具有泛化能力和鲁棒性。在“后向”过程中，结合“后向”可解释技术，使网络可理解、可操控。接下来介绍课题组已开展的“灰盒”赋能研究。

本课题组前期工作[19]初步验证了 DL-PADC 的可行性，但存在 ResNet 泛化能力不足等问题。例如，当光子模数转换系统通道间的时间失配为 35 ps 时，采集数据训练 ResNet，收敛后的网络能将失配量为 35 ps 的信号的杂散去除，但是当通道间的失配量不为 35 ps 时，网络输出仍然存在杂散分量。对于光子模数转换系统而言，当温度变化或抖动引起系统状态发生变化时，通道间的时间失配会发生改变，因此 ResNet 会失效。基于工作[19]的“黑盒”赋能思路，如果要实现网络对多种时间失配量有效，需要用多种时间失配的数据进行训练，这大大增加了数据收集压力和网络训练成本。

我们基于对光子模数转换系统的理解，提出了卷积循环自动编码器（CRAE），并重新设计了训练集[20]。对于系统来说，通道间失配量的大小反映在信号幅频域中失配成分的功率大小，因此用两组不同时间失配量的同一信号训练网络，有助于网络从输入中学习到失配特征。另外我们采用了宽带线性调频信号作为光子模数转换系统的被采样信号，对于线性调频信号而言，其频率随时间增加而增加，而由系统的通道间失配带来的杂散其频率随时间增加而减小。循环神经网络具有“记忆”功能，非常适用于处理序列数据。因此，可以借助循环神经网络提取光子模数转换系统输出（时域波形）的时序关系。如图 10 所示，卷积循环自动编码器是由卷积神经网络（CNN）、循环神经网络（RNN）、自动编码器（AE）结构组成的混合网络。当光子模数转换系统通道间的失配量为 35 和 57 ps 时，采集数据训练网络。从图 11 的箱型图可以看出，失配量从 35 ps 增加至 137 ps，系统的无杂散动态范围从 10.2 dBc 降至 −3.0 dBc，经过网络后，无杂散动态范围均高达 31.6 dBc，杂散均降至噪声水平。因此，基于对光子模数转换系统的理解设计网络和构建数据集，能够指导网络学到系统特性也即时间失配，从而实现用少量样本训练但对大量时间失配量均具有泛化能力和系统鲁棒性的网络。

为了进一步探究卷积循环自动编码器是如何实现泛化能力的以及网络各层之间的作用，对上述训练后的网络进行了“后向”可解释研究[21]。首先对网络中间层进行可视化，然后对网络各个模块进行功能分析。可视化结果表明，信号中的失配成分在经过循环神经网络层之后就已经基本去除，剩下的反卷积层用来拟合信号功率。将网络中的循环神经网络层去掉、只保留循环神经网络层、去掉反卷积层，得到三个网络，分别为 CAE，RNN，CNN + RNN。对三个网络重新训练，失配补偿的结果如图 12(b)～12(d)所示。由 CAE 网络输出结果来看，

没有 RNN 层，网络几乎输出失配，由此可知 RNN 对提取网络输入中的时序关系非常关键。若只有 RNN 层，网络不能将失配去除，由此说明卷积神经网络层起到了从网络输入中提取抽象、高级信息的作用。图 12(d)表明如果将反卷积层完全去除，将丢失部分信息，网络不能完全去除失配。结合可视化结果，将卷积循环自动编码器中的两层反卷积层去除，再将卷积层和反卷积层的滤波器数据降低一半，得到简化后的卷积循环自动编码器。对其重新训练，发现简化后的卷积循环自动编码器的失配补偿能力基本不变，帧率提升了一倍。

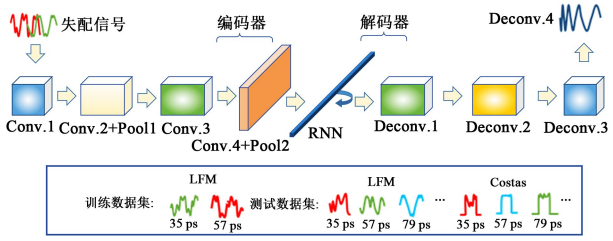


图 10 卷积循环自动编码器(CRAE)

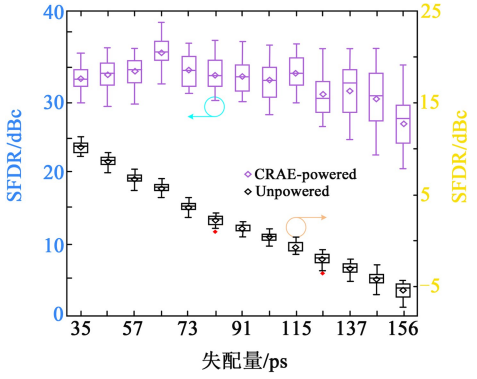


图 11 CRAE 赋能的光子模数转换系统的无杂散动态范围的箱型图

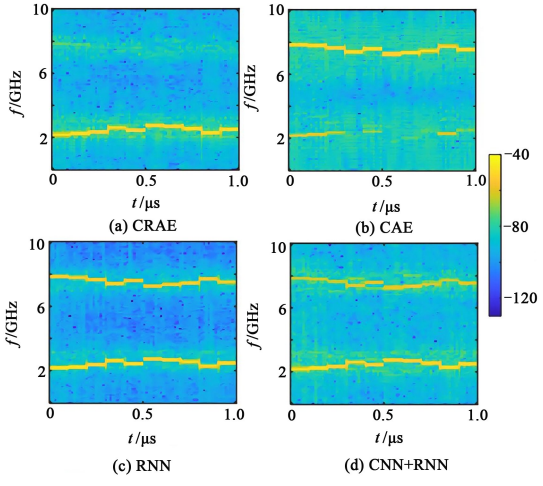


图 12 CRAE 网络模块功能分析

进一步，我们将简化后的卷积循环自动编码器应用至通道数更多的光子模数转换系统中，网络训练方式不变。实验结果如图 13 所示，当通道间的最大失配量为 156 ps 时，简化后的卷积循环自动编码器将系统的无杂散动态范围从 5.2 dBc 提升至 24.6 dBc，简化后的网络也具有很强的鲁棒性和泛化能力。

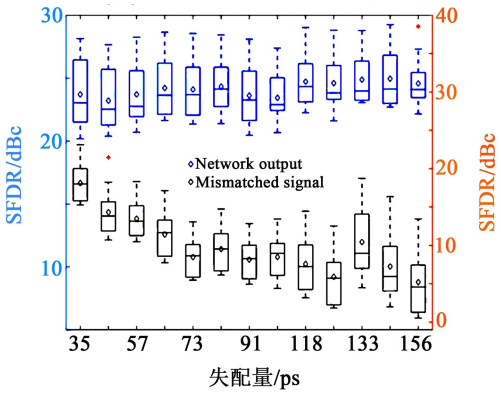


图 13 简化后的 CRAE 赋能的光子模数转换系统的无杂散动态范围的箱型图

3 总结

综上所述，本文在引言部分简要介绍了 IPS 的三大层次的内涵和部分国内外研究工作。在正文部分首先介绍了课题组在人工智能赋能的多功能光子处理系统研究，然后阐述了 AI 赋能工作从不可解释走向可解释的发展趋势和必要性，最后介绍了课题组已开展的“灰盒”赋能研究。下一步，课题组将继续开展具有可解释性的人工智能赋能的光子处理系统研究，与此同时继续开展光子辅助的人工智能信号处理系统研究和基于神经拟态的光子处理系统研究。期望 IPS 的三大层次协同发展，不断成熟，最终能够实现与人脑一样聪明甚至超越人脑的、低功耗的智能处理能力，服务于未来信息处理系统。

参考文献：

[1] Capmany J, Novak D. Microwave photonics combines two worlds[J]. Nature Photonics, 2007, 1(6): 319-330.

[2] Yao J P. Microwave photonics[J]. J. of Lightwave Technol., 2009, 27(3): 314-335.

[3] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

[4] Smith L S. Neuromorphic Systems: Past, Present and Future [M]. New York: Springer, 2010.

[5] Zou W, Ma B, Xu S, et al. Towards an intelligent photonic system[J]. Science China: Information Sciences, 2020, 63

(6): 1-17.

[6] Ismailov V E. Approximation by neural networks with weights varying on a finite set of directions[J]. J. of Mathematical Analysis and Appl., 2012, 389(1): 72-83.

[7] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators[J]. Neural Networks, 1989, 2(5): 359-366.

[8] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25: 1097-1105.

[9] Ephrat A, Mosseri I, Lang O, et al. Looking to listen at the cocktail party: A speaker-independent audio-visual model for speech separation[J]. ACM Trans. on Graphics, 2018, 37(4): 03619.

[10] Silver D, Schrittwieser J, Simonyan K, et al. Mastering the game of go without human knowledge[J]. Nature, 2017, 550(7676): 354-359.

[11] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold[J]. Nature, 2021, 596(7873): 583-589.

[12] Anthimopoulos M, Christodoulidis S, Ebner L, et al. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network[J]. IEEE Trans. on Medical Imaging, 2016, 35(5): 1207-1216.

[13] Rivenson Y, Göröcs Z, Günaydin H, et al. Deep learning microscopy[J]. Optica, 2017, 4(11): 1437-1443.

[14] Khan F N, Fan Q, Lu C, et al. An optical communication's perspective on machine learning and its applications[J]. J. of Lightwave Technol., 2019, 37(2): 493-516.

[15] Qian C, Zheng B, Shen Y, et al. Deep-learning-enabled self-adaptive microwave cloak without human intervention[J]. Nature Photonics, 2020, 14(6): 383-390.

[16] Peurifoy J, Shen Y, Jing L, et al. Nanophotonic particle simulation and inverse design using artificial neural networks[J]. Science Advances, 2018, 4(6): eaar4206.

[17] Azad A K, Wang L, Guo N, et al. Signal processing using artificial neural network for BOTDA sensor system[J]. Opt. Express, 2016, 24(6): 6769-6782.

[18] Wang B, Wang L, Guo N, et al. Deep neural networks assisted BOTDA for simultaneous temperature and strain measurement with enhanced accuracy[J]. Opt. Express, 2019, 27(3): 2530-2543.

[19] Xu S F, Zou X T, Ma B, et al. Deep-learning-powered photonic analog-to-digital conversion[J]. Light: Science & Applications, 2019, 8(1): 1-11.

[20] Zou X, Xu S, Deng A, et al. Photonic analog-to-digital converter powered by a generalized and robust convolutional recurrent autoencoder[J]. Opt. Express, 2020, 28(26): 39618-39628.

[21] Zou X, Xu S, Zou W. Visualizing and simplifying convolutional recurrent autoencoder for mismatch compensation of channel-interleaved photonic analog-to-digital converter[J]. Opt. Lett., 2021, 46(13): 3167-3170.

[22] Zou X, Xu S, Li S, et al. Optimization of the Brillouin instantaneous frequency measurement using convolutional neural networks[J]. Opt. Lett., 2019, 44(23): 5723-5726.

[23] Zhou Y, Zhang F, Shi J, et al. Deep neural network-assisted high-accuracy microwave instantaneous frequency measurement with a photonic scanning receiver[J]. Opt. Lett., 2020, 45(11): 3038-3041.

[24] Sun G, Zhang F, Gao B, et al. Photonics-based 3D radar imaging with CNN-assisted fast and noise-resistant image construction[J]. Opt. Express, 2021, 29(13): 19352-19361.

[25] Xu S, Wan J, Wang R, et al. Modified deep-learning-powered photonic analog-to-digital converter for wideband complicated signal receiving[J]. Opt. Lett., 2020, 45(19): 5303-5306.

[26] Yi S, Xu S, Zou W. Multi-band low-noise microwave-signal-receiving system with a photonic frequency down-conversion and transfer-learning network[J]. Opt. Lett., 2021, 46(23): 5982-5985.

[27] Zhou Y, Zhang F, Sun G, et al. Machine learning based LFM signal recovery for fiber-connected radar networks[C]// Optoelectronics and Communications Conf. Optical Society of America, 2021: M4B. 2.

[28] Liu E, Yu Z, Wan Z, et al. Linearized wideband and multi-carrier link based on TL-ANN[J]. Chinese Opt. Lett., 2021, 19(11): 113901.

[29] Wan J, Xu S, Zou W. High-accuracy automatic target recognition scheme based on a photonic analog-to-digital converter and a convolutional neural network[J]. Opt. Lett., 2020, 45(24): 6855-6858.

[30] Wan J, Xu S, Zou W. High-accuracy automatic target recognition scheme based on a photonic analog-to-digital converter and a convolutional neural network[J]. Opt. Lett., 2020, 45(24): 6855-6858.

[31] Wang R, Xu S, Chen J, et al. Ultra-wideband signal acquisition by use of channel-interleaved photonic analog-to-digital converter under the assistance of dilated fully convolutional network[J]. Chinese Opt. Lett., 2020, 18(12): 123901.

[32] Waldrop M M. The chips are down for Moore's law[J]. Nature News, 2016, 530(7589): 144.

[33] Shen Y, Harris N C, Skirlo S, et al. Deep learning with coherent nanophotonic circuits[J]. Nature Photonics, 2017, 11(7): 441-446.

[34] Lin X, Rivenson Y, Yardimci N T, et al. All-optical machine learning using diffractive deep neural networks[J]. Science, 2018, 361(6406): 1004-1008.

[35] Feldmann J, Youngblood N, Wright C D, et al. All-optical

spiking neurosynaptic networks with self-learning capabilities [J]. Nature, 2019, 569(7755): 208-214.

[36] Xu S, Wang J, Wang R, et al. High-accuracy optical convolution unit architecture for convolutional neural networks by cascaded acousto-optical modulator arrays[J]. Opt. Express, 2019, 27(14): 19778-19787.

[37] Xu S, Wang J, Zou W. Optical patching scheme for optical convolutional neural networks based on wavelength-division multiplexing and optical delay lines[J]. Opt. Lett., 2020, 45(13): 3689-3692.

[38] Xu S, Wang J, Zou W. Optical convolutional neural network with WDM-based optical patching and microring weighting banks[J]. IEEE Photon. Technol. Lett., 2020, 33(2): 89-92.

[39] Xu S, Wang J, Shu H, et al. Optical coherent dot-product chip for sophisticated deep learning regression[J]. Light: Science & Applications, 2020, 10: 221.

[40] Prucnal P R, Shastri B J, Thomas F D L, et al. Recent progress in semiconductor excitable lasers for photonic spike processing[J]. Adv. Opt. Photonics, 2016, 8(2): 228-299.

[41] Goi E, Zhang Q, Chen X, et al. Perspective on photonic memristive neuromorphic computing[J]. PhotoniX, 2020, 1(1): 1-26.

[42] Feldmann J, Youngblood N, Wright C D, et al. All-optical spiking neurosynaptic networks with self-learning capabilities [J]. Nature, 2019, 569(7755): 208-214.

[43] Ma B, Chen J, Zou W. A DFB-LD-based photonic neuromorphic network for spatiotemporal pattern recognition [C]// Optical Fiber Communication Conf. Optical Society of America, 2020: M2K. 2.

[44] Juodawlkis P W, Twichell J C, Betts G E, et al. Optically sampled analog-to-digital converters[J]. IEEE Trans. on Microwave Theory and Techniques, 2001, 49(10): 1840-1853.

[45] Zhang H, Zou W, Yang G, et al. Dual-output modulation in time-wavelength interleaved photonic analog-to-digital converter based on actively mode-locked laser[J]. Chinese Opt. Lett., 2016, 14(3): 030602.

[46] Yang G, Zou W, Yu L, et al. Compensation of multi-channel mismatches in high-speed high-resolution photonic analog-to-digital converter[J]. Opt. Express, 2016, 24(21): 24061-24074.

[47] Zou W, Long X, Li X, et al. Brillouin instantaneous frequency measurement with an arbitrary response for potential real-time implementation[J]. Opt. Lett., 2019, 44(8): 2045-2048.

[48] Davies M E, James C J. Source separation using single channel ICA[J]. Signal Proc., 2007, 87(8): 1819-1832.

[49] Smaragdis P. Blind separation of convolved mixtures in the frequency domain[J]. Neurocom., 1998, 22(1/3): 21-34.

[50] Ghelfi P, Laghezza F, Scotti F, et al. A fully photonics-based coherent radar system[J]. Nature, 2014, 507(7492): 341-345.

[51] Castelvecchi D. Can we open the black box of AI? [J]. Nature News, 2016, 538(7623): 20.

[52] Goodman B, Flaxman S. European union regulations on algorithmic decision-making and a "right to explanation"[J]. AI Magazine, 2017, 38(3): 50-57.

[53] Geirhos R, Jacobsen J H, Michaelis C, et al. Shortcut learning in deep neural networks [J]. Nature Machine Intelligence, 2020, 2(11): 665-673.

[54] Zhang Q, Wu Y N, Zhu S C. Interpretable convolutional neural networks[C]// Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 2018: 8827-8836.

[55] Karpatne A, Watkins W, Read J, et al. Physics-guided neural networks (PGNN): An application in lake temperature modeling [J]. arXiv preprint, 2017, 1710.11431.

[56] Borzyszkowski B, Damaszke K, Romankiewicz J, et al. Physics-guided neural networks (PGNNs) to solve differential equations for spatial analysis[J]. Bulletin of the Polish Academy of Sciences: Technical Sciences, 2021: e139391-e139391.

[57] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks[C]// European Conf. on Computer Vision, Springer, Cham, 2014: 818-833.

[58] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254-1259.

[59] Montavon G, Lapuschkin S, Binder A, et al. Explaining nonlinear classification decisions with deep taylor decomposition[J]. Pattern Recognition, 2017, 65: 211-222.

[60] Bach S, Binder A, Montavon G, et al. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation [J]. PloS one, 2015, 10(7): e0130140.

[61] Xu Z Q J, Zhang Y, Luo T, et al. Frequency principle: Fourier analysis sheds light on deep neural networks[J]. arXiv preprint, 2019, 1901.06523.

[62] Zhang Y, Zhang Z, Luo T, et al. Embedding principle of loss landscape of deep neural networks[J]. arXiv preprint, 2021, 2105.14573.

[63] Xu Z Q J, Zhang Y, Xiao Y. Training behavior of deep neural network in frequency domain[C]// Inter. Conf. on Neural Information Processing. Springer, 2019: 264-274.

作者简介：

邹秀婷(1995—)，女，电子信息专业博士研究生，主要从事可解释的智能光子信号处理研究；

邹卫文(1980—)，男，博士，教授，博士生导师，主要从事智能微波光波融合技术研究。