

引用格式: LIU Xiaoqing, LI Bolin, LUAN Guangrui, et al. Research on Adaptive Error Compensation in Optical Diffraction Neural Networks[J]. Acta Photonica Sinica, 2026, 55(7):0705001

刘晓庆,李柏霖,栾广瑞,等. 光学衍射神经网络的自适应误差补偿研究[J]. 光子学报, 2026, 55(7):0705001

光学衍射神经网络的自适应误差补偿研究

刘晓庆,李柏霖,栾广瑞,林剑

(上海理工大学 智能科技学院, 上海 200093)

摘 要: 光学衍射神经网络在高效计算中极具潜力,但在实际部署中不可避免的器件制造与装配对准误差会导致严重的波前畸变,限制了其推理性能。针对这个问题,设计了一种用于误差补偿的光学衍射神经网络,其中由固定相位层和位于光路末端的可编程调制层组成。为实现动态校正,在训练过程中引入了自适应策略,通过反向精准更新调制层的振幅掩模来重构畸变波前,从而减弱物理失准对系统性能的影响。以手写数字二分类任务为例,该架构实现了对光场偏离的有效校正,使实验推理准确率从初始的 81% 提升至 92% 左右,且在 3 个和 6 个像素的二维位移扰动下网络仍能保持性能稳定。在更复杂的时尚物品四分类任务中,准确率由 76% 提升至 88% 左右。证明了该光学架构可以有效补偿硬件物理误差,并为实现低成本、高鲁棒性的光学计算奠定了基础。

关键词: 深度学习; 光学衍射神经网络; 自适应训练; 误差补偿; 系统鲁棒性

中图分类号: O436

文献标识码: A

doi: 10.3788/gzxb20265507.0705001

0 引言

人工智能(Artificial Intelligence, AI)和深度学习技术的发展,在计算机科学应用(例如图像分类^[1]、语音识别^[2]、游戏^[3])到科学研究(例如医学诊断^[4]、智能成像^[5]、行为科学^[6])都取得了巨大的成功。在这些主流应用中,深度学习方法也被用来解决逆成像问题^[7, 8]。但随着对计算算力和能效的需求呈指数级增长,传统的电子计算平台正面临着摩尔定律^[9]失效和冯·诺依曼^[10, 11]架构瓶颈的严峻挑战。光学神经网络(Optical Neural Network, ONN)凭借光子的高传播速度,高并行性,低串扰以及低能耗特性^[12, 13],被认为是突破电子计算限制的下一代计算范式^[14]。人们提出了各种 ONN 架构,例如光学干涉神经网络^[15]、光学衍射深度神经网络(Diffractive Deep Neural Network, D²NN)^[16, 17]及光子储存计算^[18]。其中,D²NN作为一种基于自由空间光衍射的物理计算架构,利用多层无源衍射元件对光场进行调制,能够以光速完成复杂的图像分类^[19]和特征提取任务。

当前的 D²NN 架构通常在电子计算机上进行计算机模拟训练,以获得其物理实现的设计。通过对光与物质相互作用进行建模,学习网络参数,并确定将结构部署在光子器件上——或是通过 3D 打印^[20]、光刻等工艺制造为固定衍射层,或是加载到空间光调制器(Spatial Light Modulator, SLM)和数字微镜器件(Digital Micromirror Device, DMD)等可编程器件上^[21]。然而,由于物理实现的非理想性,固定元件不可避免地存在加工误差,且多层系统在搭建时面临微米级的对准偏差及层间距的误差。离线训练的数字模型无法感知这些实际物理环境中的随机误差,导致构建的物理系统在光波前调制上与设计预期产生严重偏离,极大地限制了 D²NN 实际应用中的准确率与鲁棒性。

为缓解上述问题,使光学计算系统能感知并修正自身物理误差,研究人员提出了原位训练(In-situ training)策略,即直接在物理光学系统中对可调参数进行优化。然而,现有的原位训练方法仍面临多严峻挑战。主流的基于梯度的训练通常依赖于反向传播(Backpropagation, BP)算法^[19],这需要构建复杂的可逆光路或

第一作者: 刘晓庆(2001—),女,学术,硕士,主要研究方向为光学衍射神经网络。Email: 18001411001@163.com

通讯作者: 林剑(1979—),男,副教授,博士,主要研究方向为光学衍射神经网络。Email: jianlin@usst.edu.cn

收稿日期: 2026-03-09; 录用日期: 2026-04-09

干涉系统来测量梯度^[21],对实验环境要求极高且难以扩展。因此,当前具有更深配置的D²NN^[20-25]在物理系统中实现复杂的网络架构面临着挑战。

针对多层D²NN在物理实现过程中不可避免的制造误差与装配偏差问题,本文提出了一种结构化的误差补偿策略。不同于依赖全可调衍射层或构建复杂可逆光路实现原位反向传播的方法,本工作采用“高分辨固定衍射层 + 末端低分辨可编程调制层”的混合架构。固定相位层负责完成主要特征提取任务,而可编程层则作为波前校正单元,通过自适应更新实现对系统误差的动态补偿。该结构在保证高空间分辨率的同时,显著降低了对多层精密机械对准与全动态调制硬件的依赖。在训练策略上,本文构建了物理感知的光电闭环反馈框架。系统首先通过共轭探测分支获取实际传播光场的强度分布,并结合预训练相位恢复网络重建复振幅信息。在此基础上,利用角谱传播模型进行梯度计算与参数更新,实现对可编程层振幅掩模的迭代优化。该方法不依赖严格的光学可逆传播结构,而是在物理可测强度约束下实现近似梯度估计与误差补偿,从而提高系统在非理想物理环境中的推理稳定性与鲁棒性。

1 原理与方法

1.1 光学衍射神经网络的物理架构

本文构建的光学衍射神经网络,如图1所示,主要由光学输入、固定相位层、可编程层以及输出面组成。固定相位层通过高精度3D打印实现,其利用1.8 μm的像素尺寸在微米级尺度内集成了10000个衍射神经元。这种高度并行性使得系统能够在光波穿过的瞬间完成大规模特征加权运算,在并行计算能力与单位体积计算密度方面具有潜在优势。尽管通过其极高的像素密度实现了高效的光学特征提取,但一旦物理形态制造完成便无法更改,因此对系统误差十分敏感。为了克服这一物理局限,系统在固定相位层后方引入了由DMD构成的可编程调制层。该层作为波前校正器,通过动态更新权重参数来调制并修正固定相位层造成的光场畸变。其中“动态”补偿是指系统通过可编程层的主动波前重构,实现对非理想物理环境的自适应适配。此外,DMD作为补偿器件自身也存在量化误差、响应非均匀性及配准偏差等非理想因素。本文采用闭环反馈策略,将DMD与前序固定相位层作为整体光路系统统一迭代校正,因此在补偿固定相位层误差的同时,也可在一定程度上减弱DMD自身非理想性带来的不利影响。此方案在保持系统高分辨率的同时,能提高系统的抗误差鲁棒性。

通过采用角谱法构建正向传播模型,来模拟该架构中的光波传播行为。该方法基于标量衍射理论,能够精确描述光波在自由空间中的演化过程。假设第 k 层输出平面的复振幅分布为 $U_k(x, y)$,传播距离 z 后到第 $k+1$ 层输入面的光场 $U_{k+1}(x, y)$ 可表示为:

$$U_{k+1}(x, y) = \mathcal{F}^{-1} \left\{ \mathcal{F} \left\{ U_k(x, y) \right\} \cdot H(f_x, f_y) \right\} \# (1)$$

其中, $\mathcal{F}$ 和 $\mathcal{F}^{-1}$ 分别表示二维傅里叶变换及其逆变换^[23], (f_x, f_y) 为空间频率。 $H(f_x, f_y)$ 为自由空间的传递函数,描述了光波在自由空间中传播的相位变化,具体表达式为:

$$H(f_x, f_y) = \exp \left(j \frac{2\pi}{\lambda} z \sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2} \right) \# (2)$$

其中, λ 为工作波长,在硬件系统具备有效物理响应的前提下,仅需调整模型参数即可实现向不同工作波长的理论迁移, j 为虚数单位, f_x 和 f_y 为空间频率。

为了实现图像分类任务,首先在输出探测器平面定义了 K 个互不重叠的信号响应区域,每个区域对应一个类别。系统的最终优化目标是调节DMD层的振幅分布参数,使得当输入特别类别的样本时,光场能量能够最大程度地汇聚在对应的信号区域内,同时抑制其他区域的背景噪声。为了量化这一目标,本文采用均方差作为损失函数来衡量物理输出与理想目标之间的差异。损失函数定义如下:

$$L = \frac{1}{K} \sum_{i=1}^K \left(\tau_{\text{out}}^{(i)} - \mathcal{T}^{(i)} \right)^2 \# (3)$$

式中, K 为分类类别总数。 $\tau_{\text{out}}^{(i)}$ 表示第 i 个探测区域内的积分光强,该值由角谱模型计算得到的二维光强分布 $I_{\text{out}}(x, y)$ 经区域积分获得。 $\mathcal{T}^{(i)}$ 为预设的目标值向量。

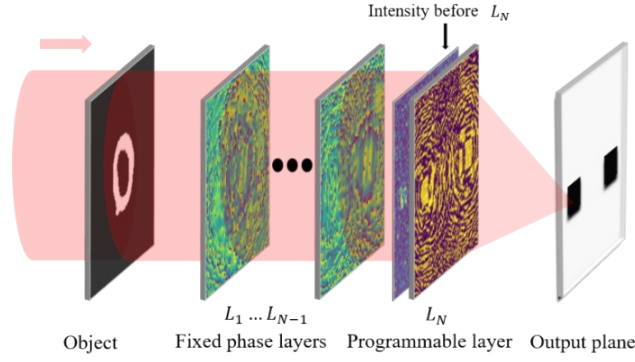


图1 本文设计的D²NN模型
Fig.1 The D²NN model designed in this paper

1.2 基于相位重建的自适应训练策略

基于上述物理架构,为了有效弥补固定相位层的制造与装配过程引入的系统误差,本文采用了一种物理感知的自适应训练策略^[24],完整的闭环流程如图2所示。这里所谓“物理感知”,是指参数更新过程中所使用的梯度建立在实测物理场约束之上,而非仅依赖理想传播模型。定义物理前向传播过程为 $U_{\text{out}} = f(U_{\text{in}}; \theta, \omega, \Delta) \#(4)$

其中 U_{in} 为输入光场, θ 为固定层参数, ω 为可编程层权重, Δ 为系统误差。系统实际测得的光强为 $I_{\text{meas}} = |U_{\text{out}}|^2 = |f(U_{\text{in}}; \theta, \omega, \Delta)|^2 \#(5)$

该流程的核心在于构建一个光电闭环反馈系统,使可编程DMD层主动适应并校正物理光路中的波前畸变。

整个误差补偿过程分为物理前向传播与反向更新两个阶段。首先,光束穿过带有物理误差的固定相位层($L_1 \cdots L_{N-1}$),由系统获取可编程层(L_N)前方实际畸变的光强分布(如图1所示)。在实际物理部署中,该步骤可通过引入与 L_N 平面光学共轭的探测分支来实现,从而确保记录的光强能精确反映入射至调制层前的波前状态。随后,为了获取计算梯度所需的复振幅信息,引入预训练的U-Net^[25]神经网络作为数字相位恢复模块,对该光强分布进行相位重建^[26]。该U-Net模块预先基于大量散斑数据进行了全监督学习训练,即便面对复杂的畸变光场,其预测相位在整体轮廓及局部纹理上均能与真值保持高度一致。

最后,系统结合实际捕获的光强与U-Net重建的相位,合成出用于反向传播的复振幅。电子网络基于此复振幅计算当前的系统分类误差损失,并经反向传播获得物理梯度,以此迭代更新可编程层的振幅调制参数。在该补偿过程中,DMD通过振幅调制对误差光场中各局域波分量进行加权,并借助后续自由空间传播中的衍射与干涉效应,在输出面重构目标能量分布。因此,尽管振幅调制不直接作用于相位项,但仍可通过改变各波分量的相干叠加关系,补偿由相位误差引起的能量扩散与焦点偏移。通过这种迭代优化,可编程层能够逐渐形成与物理误差互补的输出调制分布,从而实现系统性能的恢复。

由于系统中仅能直接测量光强分布而无法直接获取精确相位信息,本文采用基于深度学习的相位恢复

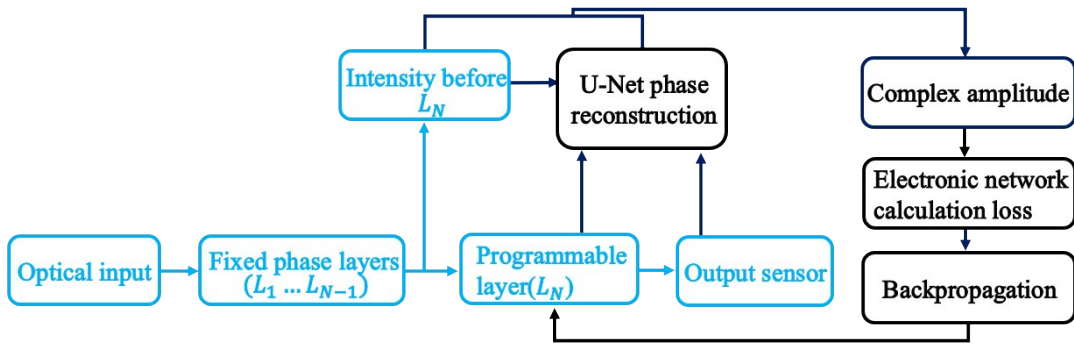


图2 自适应训练过程示意图
Fig.2 Schematic Diagram of the Adaptive Training Process

模块对复振幅进行近似重建,因此计算得到的梯度为基于重建复振幅的近似物理梯度,而非严格真实物理梯度。若记真实梯度为 g_i ,本文实际使用的近似物理梯度为 $\hat{g}_i$,则可表示为 $\hat{g}_i = g_i + e_i$,其中 e_i 表示由相位恢复误差及复振幅近似重建引入的梯度估计误差。当该误差受到控制且步长选取适当时,近似梯度更新仍可有效引导损失函数下降,并使系统收敛至稳定补偿状态。实验结果表明,在实际误差尺度范围内,该近似梯度能够有效引导可编程层收敛至稳定补偿状态。进一步地,U-Net相位恢复模块的有效性依赖于训练数据分布与真实物理误差分布之间的匹配程度。当实际误差处于训练数据覆盖范围内时,网络能够较稳定地恢复相位整体结构;而当实际误差超出训练分布时,相位重建误差将增大,并进一步导致梯度估计偏差,从而影响DMD掩模更新质量、补偿收敛速度及最终识别精度。

2 仿真分析

为了验证所提出架构的误差补偿能力,本文基于角谱传播模型构建了仿真环境(波长638 nm,相位分辨率 100×100 ,单像素 $1.8 \mu\text{m}$)。在手写数字数据集(Modified National Institute of Standards and Technology, MNIST)^[19]二分类任务上的基础评估表明,在理想无误差条件下,三种不同层数的混合架构均展现出了优异的特征提取能力,其分类准确率均达到95%以上(图3(b))。

然而,物理对准误差会严重影响这种理想性能。如图3(a)与(b)所示,当所有相位层在 x 和 y 方向同时引入5个像素(约 $9 \mu\text{m}$)的二维位移误差时,各架构的混淆矩阵对角线特征变得模糊,准确率大幅衰退,这种现象证实了仅依赖离线训练无法应对物理制造与装配过程中的随机偏差。

针对该问题,系统引入了基于DMD层的自适应训练策略来动态补偿波前畸变。图3(c)详细展示了这一动态修复过程。图中Epoch0的数据点代表了系统在未进行补偿,仅受误差影响时的初始准确率状态。在引入自适应训练后,三种架构均在10个Epoch内实现了快速收敛。更进一步的极端位移测试(图3(d))证实了该策略的卓越鲁棒性:当误差增至8像素时,未补偿系统的性能急剧恶化(单层架构的准确率跌破20%),而自适应补偿后的多层架构即使在严重错位下,仍能保持约85%的高精度,展现出较强的物理误差容忍能力。

最后,为进一步量化该策略对性能的还原程度,我们将补偿后的光学架构与“理想全可变衍射层架构”进行了对比验证。如图4所示,对比结果表明,尽管本文提出的架构主要由不可调的固定相位层构成,但系统通过单层DMD的动态调制成功突破了这些固定层的物理限制。最终补偿精度(黄色条形图)与理想架构相比,性能偏差较小。这有力地证明了,通过末端的主动波前重构,能够有效抵消前端固定相位层的物理缺陷,以低成本的硬件配置实现了等效于全动态架构的高精度计算能力。

3 实验与结果

3.1 实验装置

搭建光学系统来验证所提出的自适应训练来弥补系统误差的有效性,如图5(a)所示。638 nm的线偏振光经DMD1(西安中科微星,DMD-2K090,2560×1600像素)调制后,穿过由透镜L1和L2组成的4f缩束系统到达固定相位层Layer 1。如图5(b)所示,该层通过高精度3D打印技术^[20]制备(单像素尺寸 $1.8 \mu\text{m}$),负责特征提取与相位调制。随后,光波经过由透镜L3和L4组成的光束扩展器进行扩束,并传输至分束器(Beam Splitter, BS)。为了精准监控入射波前,我们严格校准了光路,使CMOS Camera1相机(Hamamatsu, C14440-20UP,分辨率2304×2304,像素尺寸 $6.5 \mu\text{m} \times 6.5 \mu\text{m}$)的靶面与透射光路中的DMD2(西安中科微星,DMD-2K095,1920×1080像素)关于BS呈光学共轭关系。反射光束被CMOS Camera1相机捕捉作为下一层衍射层的输入,透射光束经过DMD2调制,它通过加载自适应的振幅掩模来动态补偿前序光路中的制造与装配误差。DMD2的像素翻转频率为10 kHz以上,为系统的快速重构提供了硬件基础。最后,调制后的光场通过由透镜L5和L6组成的光束扩展器成像到CMOS Camera2相机(Andor, ZYLA-5.5-CL10,分辨率2560×2160,像素尺寸 $6.5 \mu\text{m} \times 6.5 \mu\text{m}$),用于记录系统的最终推理结果及分类光斑的能量分布,完成整个计算过程。

从系统实现层面出发,本文将所构建的实验平台与代表性的原位训练^[21]方案进行系统级比较,如表1所

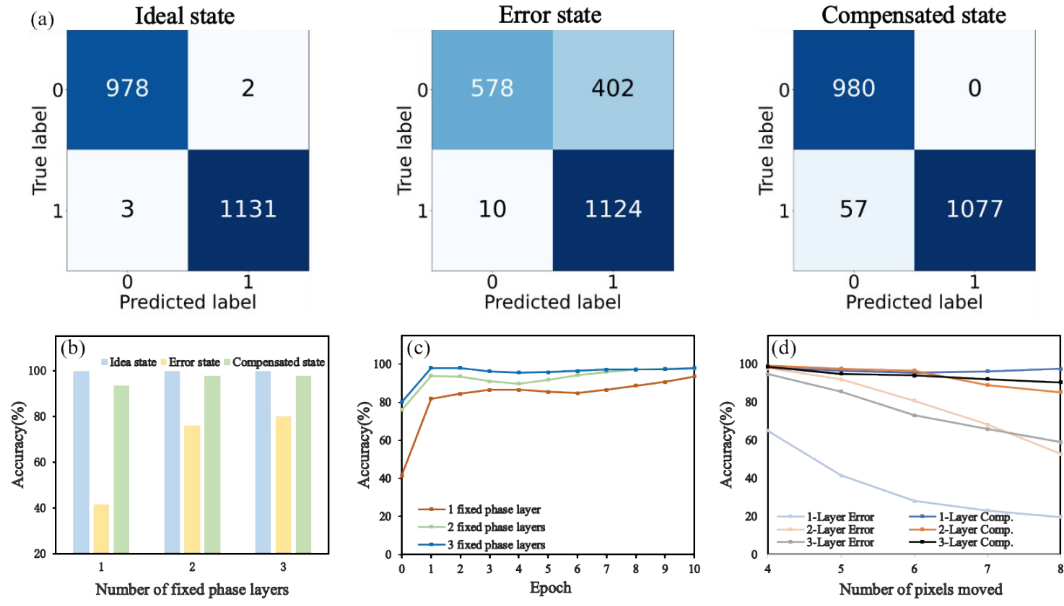


图3 二分类任务下自适应误差补偿策略的综合性能验证。(a)以三层固定相位层为例,在理想条件、误差条件(5像素位移)及补偿后的混淆矩阵对比;(b)不同数量的固定相位层在三种状态下的柱状图对比;(c)不同数量的固定相位层在弥补误差过程中的准确率曲线;(d)极端条件下系统的鲁棒性分析

Fig.3 Comprehensive performance validation of adaptive error compensation strategies for binary classification tasks. (a) Comparing confusion matrices under ideal conditions, error conditions (5-pixel displacement), and after compensation, using a three-layer fixed-phase layer as an example; (b) Bar chart comparison of different numbers of fixed-phase layers across three states; (c) Accuracy curves for different numbers of fixed-phase layers during error compensation; (d) Robustness analysis of systems under extreme conditions

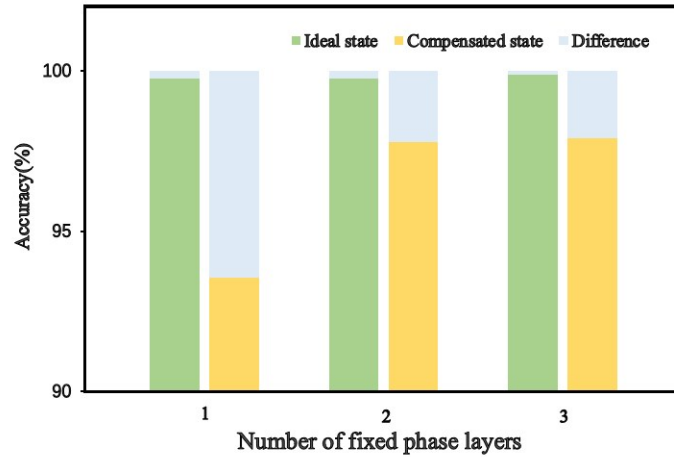


图4 不同配置下理想状态与补偿状态的定量对比

Fig.4 Quantitative comparison of ideal state and compensated state under different configurations

示。该原位训练方案采用多层可编程SLM级联,并结合相移数字全息和误差光场生成模块实现原位梯度测量,具有较高的可调自由度,但系统硬件链路相对复杂。相比之下,本文仅对单个末端DMD层进行在线更新,从而降低了可调制层数量与系统实现复杂度。原位训练方案在训练自由度和误差补偿能力方面具有优势,而本文方案则更强调在较低硬件复杂度下实现对固定衍射层宏观物理误差的有效补偿。

3.2 实验结果

我们构建了一个由一层固定相位层和一层末端可编程层组成的衍射神经网络,来验证自适应训练在弥补误差方面的有效性。该模型使用MNIST训练集进行计算机模拟训练,并在测试集中的2000张数字图像上实现了97.3%的盲测准确率。训练目标设置为将输入手写数字单独映射到输出层上的2个预定义区域,

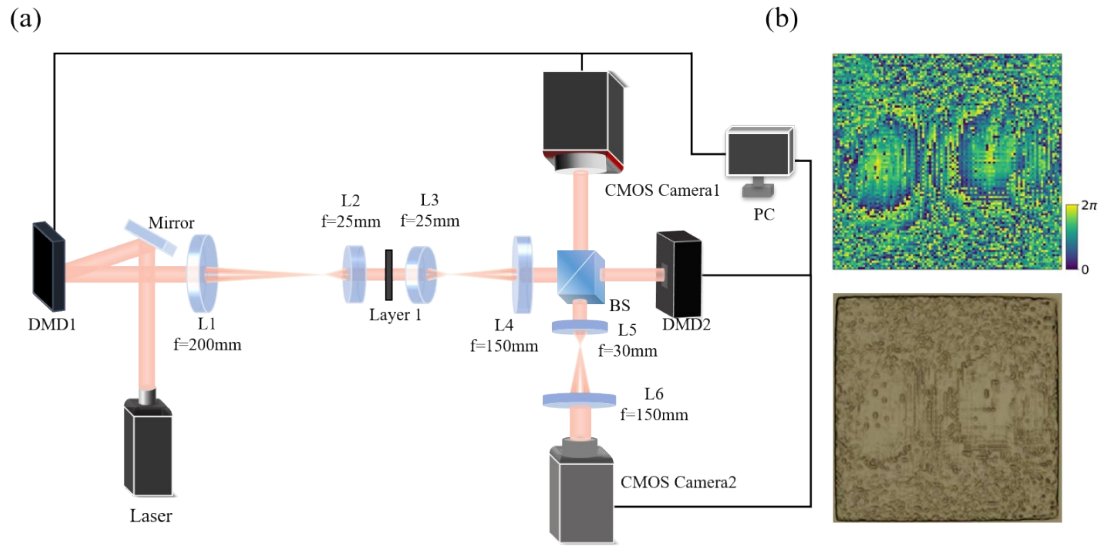


图5 实验装置与器件表征。(a)实验光学系统装置图;(b)3D打印固定相位层的设计相位与实物表征
Fig.5 Experimental Setup and Device Characterization. (a) Schematic diagram of an experimental optical system; (b) Design phase and physical characterization of 3D-printed fixed-phase layers

表1 本文方案与代表性原位训练方案的系统级比较

Table 1 System-level comparison between the proposed scheme and a representative in situ training method

Method category	Number of adjustable layers	Reverse-measured optical path	Main hardware components	Training time	performance
in situ training method	Multiple	Yes	Multiple SLMs, phase-shifting digital holography, error field generation module	Approximately 0.08 s per update	Comparable to in silico training; suitable for full in situ optimization
Proposed scheme	One terminal DMD	No	Fixed phase layers, DMD, CMOS, and control unit	Approximately < 0.1 s per update	Effective compensation under physical errors

通过寻找光信号最大的目标区域来确定分类结果。根据收敛后的相位参数制造物理固定衍射层,并将预训练得到的初始振幅权重加载至DMD2,以此作为实验的初始基准状态。图6(a)展示了实验测试集上的分类统计结果:在未进行自适应优化时,由于光场聚焦偏差,系统存在明显的误判现象,整体准确率仅为81%左右。相比之下,完成自适应训练后,混淆矩阵对角线上的正确分类样本数显著增加,整体实验分类准确率提升至约92%。物理实验准确率与理想仿真结果高度贴合,表明DMD自身的非理想因素在当前误差范围内并未主导系统性能变化,其影响已通过自适应策略得到有效抑制,验证了该方法在实际硬件上修正推理偏差的有效性。

上述宏观分类性能的恢复,本质上源于系统对畸变光场的微观重塑。如图6(c)和6(d)所示,在无自适应训练阶段,到达探测面的光斑能量发散,DMD2上的权重分布也未能针对当前物理环境进行有效调控。通过自适应训练,DMD2逐步演化出具有特定干涉条纹特征的权重掩模,成功对畸变波前进行了二次整形,使得输出光场在目标探测区域形成了高对比度的清晰聚焦点,证明了系统光场调控能力的显著恢复。

此外,为了探究系统在应对不同程度物理失准时的鲁棒性,实验中对固定衍射层施加了不同的二维平移误差,设置了0像素、3像素和6像素三个误差等级。上述误差量级一方面用于模拟实际装调过程中常见

的位置失准,另一方面也与实验中位移台可稳定实现的位移范围相匹配,从而便于在可控条件下评估系统的补偿能力。如图6(b)所示,在所有测试条件下,系统均表现出较快的收敛速度。具体而言,系统在5次迭代内即可完成对误差的有效补偿并趋于稳定。其次,在系统闭环补偿速度方面,我们实验测得单次闭环更新时间在0.1 s以内。实验结果表明,在无额外人为位移的基准状态下,系统保持了最高的分类精度;而当引入3个像素和6个像素的显著对准误差时,尽管系统的初始推理能力受到抑制,但通过更新DMD2的振幅掩模,系统仍能准确率恢复至90%以上,说明所提出的闭环自适应补偿机制对装调误差具有较好的容忍能力和稳定补偿效果。

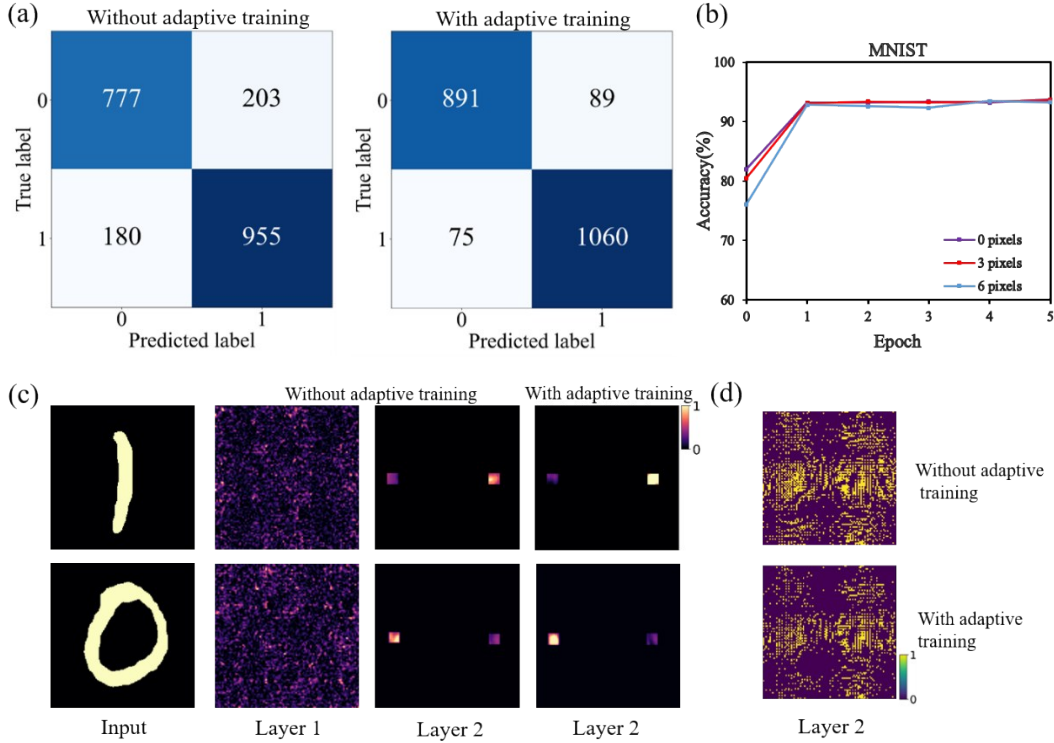


图6 基于MNIST二分类任务下自适应训练的误差补偿实验结果。(a)补偿前后的分类性能对比;(b)不同物理误差条件下的鲁棒性验证;(c)系统中光场传播的可视化;(d)DMD振幅调制掩模的变化

Fig.6 Experimental results on error compensation using adaptive training for the MNIST binary classification task. (a) Comparison of classification performance before and after compensation; (b) Robustness verification under different physical error conditions; (c) Visualization of light field propagation in the system; (d) Variations in DMD amplitude modulation masks

在上述基于单层固定相位层的MNIST实验基础上,为进一步验证所提方法在更复杂分类任务和更深层光学结构中的有效性,本文补充构建了一个由两层固定相位层和一层末端可编程层组成的衍射神经网络,并在Fashion-MNIST四分类任务上进行了实验验证。图7(a)给出了典型输入样本在补偿前后的输出光场分布。可以看出,在无自适应训练条件下,部分样本的能量未能有效聚焦至对应目标探测区域,输出面存在较明显的能量扩散与类别串扰;经过自适应训练后,DMD重塑了针对复杂纹理的补偿波前,使输出平面在预定义的目标区域形成更高对比度的聚焦光斑。对应的混淆矩阵如图7(b)所示,自适应训练后对角元素显著增强,非对角元素整体减小,平均识别准确率由初始约76%提升至约88%。上述结果进一步表明,所提出的自适应补偿策略不仅适用于较简化的单固定层结构,在两层固定相位层的更复杂光学网络中同样具有良好的误差补偿能力与泛化效果。同时也需要指出,随着任务复杂度进一步提高,系统对波前整形精度、相位恢复精度及补偿结构自由度的要求也会相应增加,从而对最终识别性能构成更高挑战。

4 结论

本文提出了一种结合固定衍射层与可编程调制层的混合型光学衍射神经网络架构,并构建了基于物理

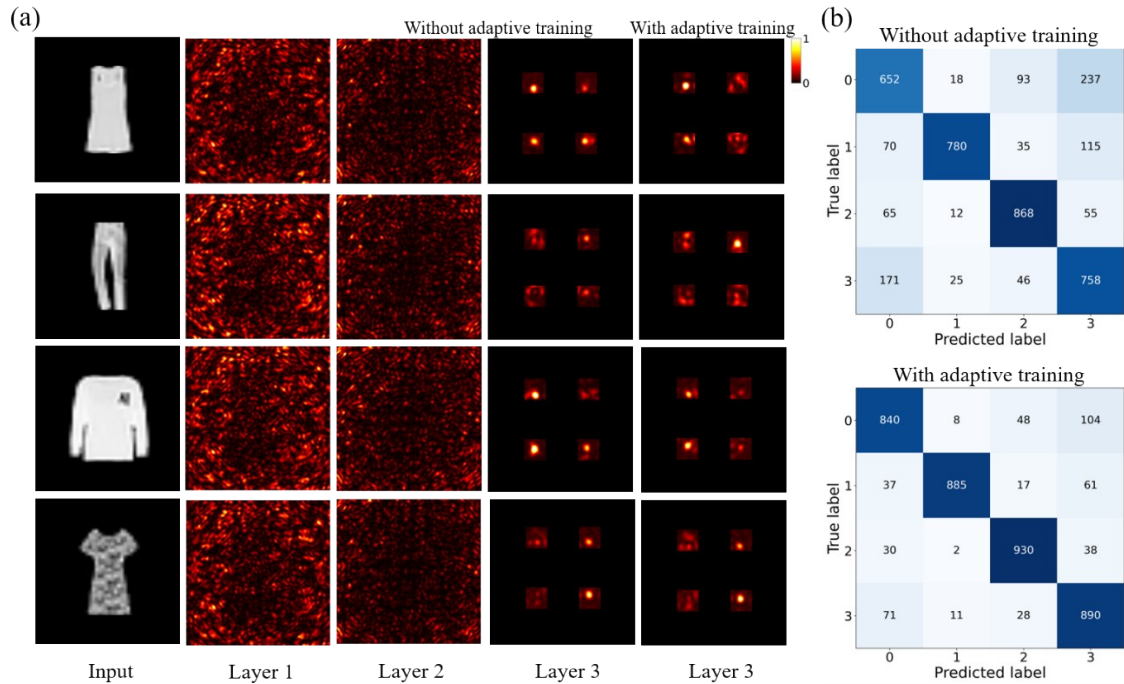


图7 Fashion-MNIST 四分类任务自适应补偿的实验结果。(a)系统中输出光场分布;(b)补偿前后的分类性能对比
Fig.7 Experimental results of adaptive compensation for the Fashion-MNIST four-class task (a) Output light field distribution in the system; (b) Comparison of classification performance before and after compensation

感知的自适应训练策略,用于补偿固定衍射层系统在实际制造与装配过程中产生的物理误差。通过数值仿真与物理实验验证,在MNIST二分类任务中,该方法将存在系统误差条件下的实验推理准确率由约81%提升至92%,同时在引入3像素与6像素的二维平移误差时仍能保持90%以上的分类精度,体现出良好的误差容忍能力。在两层固定相位层结构下的Fashion-MNIST四分类实验中,自适应补偿后系统识别准确率由约76%提升至约88%,表明所提方法在更复杂分类任务和更深层光学结构中同样具有良好的有效性与泛化能力。实验结果表明,单层可编程调制单元能够通过波前重构有效缓解前端固定衍射层的物理失准影响,从而缩小数字模型与物理实现之间的性能差距。该策略在不依赖多层全可调结构或严格光学可逆路径的条件下,实现了对宏观装配误差的自适应补偿,为构建低成本、高鲁棒性的实用化光学计算系统提供了一种可行路径。系统的补偿上限受限于调制器的空间自由度,且在处理极端畸变时表现出能量利用率与补偿精度间的物理权衡。未来工作可进一步结合更高精度的相位恢复模型与多层补偿结构,以拓展其在复杂任务与大规模衍射网络中的应用潜力。

参考文献

- [1] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]//Advances in Neural Information Processing Systems 25. 2012: 1097-1105.
- [2] HINTON G, DENG L, YU D, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups [J]. IEEE Signal Processing Magazine, 2012, 29(6): 82 - 97.
- [3] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of Go with deep neural networks and tree search [J]. Nature, 2016, 529(7587): 484 - 489.
- [4] ESTEVA A, ROBICQUET A, RAMSUNDAR B, et al. A guide to deep learning in healthcare [J]. Nature Medicine, 2019, 25(1): 24 - 29.
- [5] BARBASTATHIS G, OZCAN A, SITU G. On the use of deep learning for computational imaging [J]. Optica, 2019, 6(8): 921 - 943.
- [6] MATHIS A, MAMIDANNA P, CURY K M, et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning [J]. Nature Neuroscience, 2018, 21(9): 1281 - 1289.
- [7] JIN K H, MCCANN M T, FROUSTEY E, et al. Deep convolutional neural network for inverse problems in imaging [J]. IEEE Transactions on Image Processing, 2017, 26(9): 4509 - 4522.
- [8] SINHA A, LEE J, LI S, et al. Lensless computational imaging through deep learning [J]. Optica, 2017, 4(9): 1117 -

- 1125.
- [9] PATTERSON D, GONZALEZ J, LE Q, et al. Carbon emissions and large neural network training [EB/OL]. 2021 [2026-03-29]. <https://arxiv.org/abs/2104.10350>
 - [10] MEROLLA P A, ARTHUR J V, ALVAREZ-ICAZA R, et al. A million spiking-neuron integrated circuit with a scalable communication network and interface [J]. *Science*, 2014, 345(6197): 668 - 673.
 - [11] PEI J, DENG L, SONG S, et al. Towards artificial general intelligence with hybrid Tianjic chip architecture [J]. *Nature*, 2019, 572(7767): 106 - 111.
 - [12] ZHANG Q, YU H, BARBIERO M, et al. Artificial neural networks enabled by nanophotonics [J]. *Light: Science & Applications*, 2019, 8(1): 42.
 - [13] MINZIONI P, LACAVA C, TANABE T, et al. Roadmap on all-optical processing [J]. *Journal of Optics*, 2019, 21(6): 063001.
 - [14] XIANG Shuiying, WANG Yizhi, NIU Xinran, et al. Integrated photonic computing: current status, challenges, and prospects (Invited)[J]. *Acta Photonica Sinica*, 2025, 54(9): 0954205.
 - [15] SHEN Y, HARRIS N C, SKIRLO S, et al. Deep learning with coherent nanophotonic circuits [J]. *Nature Photonics*, 2017, 11(7): 441 - 446.
 - [16] 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.
 - [17] YAN T, WU J, ZHOU T, et al. Fourier-space diffractive deep neural network [J]. *Physical Review Letters*, 2019, 123(2): 023901.
 - [18] LARGER L, BAYL6N-FUENTES A, MARTINENGHI R, et al. High-speed photonic reservoir computing using a time-delay-based architecture: Million words per second classification [J]. *Physical Review X*, 2017, 7(1): 011015.
 - [19] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
 - [20] LUO Y, MENGU D, YARDIMCI N T, et al. Design of task-specific optical systems using broadband diffractive neural networks [J]. *Light: Science & Applications*, 2019, 8(1): 112.
 - [21] ZHOU T, FANG L, YAN T, et al. In situ optical backpropagation training of diffractive optical neural networks [J]. *Photonics Research*, 2020, 8(6): 940 - 953.
 - [22] DOU H, DENG Y, YAN T, et al. Residual D2NN: training diffractive deep neural networks via learnable light shortcuts [J]. *Optics Letters*, 2020, 45(10): 2688 - 2691.
 - [23] CAO Liangcai, HE Zehao, ZHAO Yan, et al. Angle-spectrum-propagation holographic algorithm for layered structures (invited)[J]. *Infrared and Laser Engineering*, 2018, 47(6): 0603001.
 - [24] ZHOU T, LIN X, WU J, et al. Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit [J]. *Nature Photonics*, 2021, 15(5): 367 - 373.
 - [25] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]// *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015*. Cham: Springer, 2015: 234-241.
 - [26] YANG Huizhe, ZHANG Haoran, LIU Jin, et al. Application of deep neural network in TIE wavefront sensing[J]. *Acta Photonica Sinica*, 2024, 53(12): 1201001.

Research on Adaptive Error Compensation in Optical Diffraction Neural Networks

LIU Xiaoqing, LI Bolin, LUAN Guangrui, LIN Jian

(University of Shanghai for Science and Technology, School of Artificial Intelligence Science and Technology, Shanghai 200093, China)

Abstract: To address the systematic errors inevitably introduced during the physical fabrication and assembly of multilayer diffractive deep neural networks (D²NN), this paper aims to validate and implement a physically aware adaptive training strategy. By incorporating programmable layers into the diffractive architecture, an optoelectronic closed-loop feedback system is constructed to enable online correction and compensation for wavefront aberrations in the actual physical optical path. The objective of this study is to quantitatively evaluate the effectiveness of this strategy in mitigating physical environmental non-idealities, ensuring high reliability and inference fidelity during practical deployment. The proposed method constructs

an optoelectronic closed-loop feedback framework with a hybrid “fixed layer + programmable layer” architecture based on physically aware adaptive training. During the physical forward propagation phase, detectors conjugate to the modulation surface are introduced within the diffraction layer to capture the intensity distribution of actual distortions. During backward updating, a pre-trained U-Net is employed as a digital phase-recovery module to reconstruct the complex amplitude, which is then used to estimate physical gradients and iteratively update the amplitude modulation parameters of the digital micromirror device (DMD) layer, thereby enabling reconstruction of the degraded wavefront. To comprehensively validate this approach, this paper employs a combined simulation and experimental methodology: First, an end-to-end simulation environment is established using the angular spectrum propagation model with parameters matched to the physical system. Subsequently, a corresponding physical optoelectronic experimental platform is constructed, integrating a high-precision 3D-printed fixed phase layer, a DMD serving as a programmable diffractive layer, a complementary metal-oxide-semiconductor (CMOS) detector, and a computer control unit to implement a complete closed-loop adaptive compensation strategy. Experiments are performed on the Modified National Institute of Standards and Technology (MNIST) binary classification task and the Fashion-MNIST four-class classification task at a wavelength of 638 nm to systematically evaluate the dynamic compensation capability and fault tolerance of the proposed architecture under mechanical displacement errors. The results indicate that, under ideal error-free conditions, simulation analysis shows that the system can achieve a theoretical classification accuracy of over 95% in the MNIST binary classification task. When a two-dimensional displacement error of 5 pixels (9 μm) is introduced in both the x and y directions across all fixed phase layers, the classification performance of the uncompensated system degrades significantly. By adopting an adaptive training strategy, the model converged rapidly within 10 iteration cycles, restoring the accuracy to a level close to the ideal. Furthermore, under extreme displacement errors of 8 pixels, this strategy successfully improved the initial accuracy from below 20% to approximately 90% while maintaining system stability. In the MNIST physical experiment, faced with real-world alignment deviations and manufacturing errors, the system improved the initial recognition accuracy from 81% to 92% by iteratively optimizing the DMD modulation parameters. Even when 3-pixel and 6-pixel two-dimensional displacement disturbances were artificially introduced, the network maintained stable classification performance, effectively mitigating the negative effects. To further validate the universality of this strategy under complex decision boundaries, this study conducted additional validation on the Fashion-MNIST four-class classification task. Compared to simple MNIST digits, clothing images possess more complex spatial texture features, placing higher demands on the accuracy of wavefront reconstruction. Experimental results show that the adaptive compensation strategy successfully increased the experimental inference accuracy for this task from an initial 76% to 88%. The spot, which had previously diverged due to error accumulation, refocused within the target detection area following adaptive optimization. This successful validation across datasets provides physical evidence of the architecture's robust adaptability when handling high-dimensional feature extraction tasks. In summary, the physically aware adaptive training strategy effectively alleviates the performance degradation of multilayer D2NNs caused by physical misalignment in practical applications. By introducing a DMD-based compensation mechanism, the system effectively neutralizes multi-scale macro-level misalignment errors. Consistency between simulations and experiments fully demonstrates that this optoelectronic closed-loop feedback architecture significantly narrows the performance gap between digital models and physical entities. This approach does not rely on stringent hardware alignment accuracy or small perturbation assumptions, providing an efficient and flexible technical pathway for robust computation of diffractive neural networks in non-ideal physical environments.

Key words: Deep learning; Optical diffraction neural network; Adaptive training; Error compensation; System robustness

OCIS Codes: 200.4700; 050.1970; 230.6120; 100.5070

CSTR: 32255.14.gzxb20265507.0705001