

光学学报

基于非线性超表面的多任务光电混合神经网络(特邀)

罗栩豪^{1,2,3,4},董思禹^{1,2,3,4*},魏泽勇^{1,2,3,4},王占山^{1,2,3,4},程鑫彬^{1,2,3,4**}¹同济大学物理科学与工程学院精密光学工程技术研究所,上海 200092;²先进微结构材料教育部重点实验室,上海 200092;³上海市数字光学前沿科学研究基地,上海 200092;⁴上海市全光谱高性能光学薄膜器件及应用专业技术服务平台,上海 200092

摘要 提出一种新型的光电混合神经网络计算模型,通过非线性超表面对线性和非线性光波进行多维度光场调控,并行处理多个通道的信息。非线性超表面采用U型纳米天线,基于自旋和几何相位对基频和倍频光波的相位进行编码和复用,结合后端的浅层电子神经网络分别实现对手写数字图像输入的高精度分类(盲测准确率为95.53%)和高质量重建(平均峰值信噪比大于26 dB)。所提方法为发展光学、光电、电子模拟和数字计算技术提供了新思路,从而实现大规模并行、高通量和更广泛的人工智能系统。

关键词 物理光学; 光电计算; 人工智能; 非线性超表面; 并行处理; 数字分类; 图像重建

中图分类号 O436

文献标志码 A

DOI: 10.3788/AOS240437

1 引言

光学、电子计算系统正在试图理解大量的视觉数据,它们来源于自动驾驶、机器视觉、医疗诊断、遥感、国防和万物互联等领域^[1]。这些数据是由人工智能(AI)算法解释的,其中电子深度神经网络正迅速成为视觉数据处理的标准算法^[2-4]。尽管电子神经网络灵活且高效,但是受到隧道效应和寄生电容的影响,电子芯片的摩尔定律正在逼近其极限^[5-6]。利用光子作为计算载体的光学神经网络具有高速、低功耗、高通量和大规模并行的固有优点^[7-13]。随着光学、光电子和电子学的深入集成,光电混合网络模型为计算系统的密度、速度和能量提供了数量级的提升,逐渐成为模拟处理器的重要组成部分。

超表面由一层超薄的人工原子组成^[14-17]。精心设计超表面的形状、对称性和空间定向,几乎可以任意控制光的相位、振幅和偏振状态^[18-22]。它在线性和非线性光学系统中都有强大的光场调控能力^[23-25]。近年来,人们利用非线性超表面实现了许多优异的光学性能,并将其应用于波束调控、全息成像、光学图像加密等领域^[26-29]。基于几何相位效应,非线性超表面的一个巨大优势就是,它能够同时实现对各种高阶频次谐波的相位进行连续调控,因此可以对不同频率光波携

带的信息进行并行处理^[30-31]。非线性超表面的这些优势和性能在光电计算系统中有着良好的应用前景,并将极大地扩展计算的维度和数据存储容量,但是目前还缺乏相关的探索和研究工作。

鉴于此,本文提出一种基于非线性超表面的光电混合神经网络模型,它可以通过编码线性与非线性光波相位,同时执行多重的机器学习任务。该模型利用联合优化法,保证了每个任务的并行处理和高性能。以手写数字为网络的输入,概念性地验证了基频和倍频分别执行分类和重建图像的任务,获得95.53%的盲测准确率和26.13的平均峰值信噪比(PSNR)。非线性超表面采用单层的U型纳米天线阵列作为光学前端,它输出4倍的压缩光学信息,紧随其后的电子后端则采用浅层的人工神经网络。这种联合设计的架构呈现出几个重要优势:1)光学前端压缩了电子网络所需的信息,显著降低了系统的功耗和复杂性;2)通过添加非线性倍频光通道,可以实现大规模、高容量的并行计算,将为各种新型机器视觉系统的发展提供支撑。这种多任务的方案有望在光纤通信中同时传输多个独立的光信号、解析多个光学频率的信息、实现多线程光计算等,将推动光通信的发展和实现更通用的类脑智能架构。

收稿日期: 2024-01-02; 修回日期: 2024-03-14; 录用日期: 2024-03-28; 网络首发日期: 2024-04-12

基金项目: 国家自然科学基金(61925504, 6201101335, 62020106009, 61621001)、上海市科委项目(17JC1400800, 20JC1414600, 21JC1406100)、中国博士后科学基金(2020TQ0227, 2021M702471)

通信作者: *dongsy@tongji.edu.cn; **chengxb@tongji.edu.cn

2 光电网络模型与超表面结构设计

2.1 非线性光电混合网络模型

图 1(a)所示为基于非线性超表面的光电混合网络架构,通过基频 ω 和倍频 2ω 的相位复用实现物品分类和编码重建(输入的物品被前端光学网络生成无法直接解读的乱码,并被后端的电子网络解码重建)的并行任务。它由光学神经网络前端和电子神经网络后端两个主要部分组成。本文选择 MNIST 数据集^[32]进行概念展示,以手写数字“6”为例,它被编码为圆偏振(σ)光波的振幅或相位分布,并作为模型的输入,通过一层非线性超表面后衍射到其后 $5\ \mu\text{m}$ 位置,通过 ω 和 2ω 的两套相位进行衍射模拟计算,获得各自的光学输出。分别取 ω 和 2ω 光学输出的部分像素,通过计算机中的两个全连接浅层神经网络来执行各自的任务。需要注意的是,本文只选取基频和倍频两个通道实现两种功能作为概念展示,实际上可以对更多的高频通道引入所提模型来实现更多的功能。数学上,为了匹配具有 $28\ \text{pixel} \times 28\ \text{pixel}$ 的手写数字真实图像,输入是一个

28×28 矩阵,非线性超表面同样包含 28×28 个单元来分别调制 ω 和 2ω 的两套几何相位[比如 $\exp(i\theta\sigma)$ 和 $\exp(i\theta\sigma)$]。通过超表面后的光学输出也包含 28×28 个像素点,选取其中心部分的 14×14 个像素点(这里仅为个例,实际上可以选取输出的 784 个像素点中的任意多个像素点)的振幅 A 并进行平方律运算后得到 A^2 。 A^2 又作为电子网络的输入,对于分类任务,采用两个隐含层的浅层神经网络:第一层具有 196×49 个神经元,激活函数为 $\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}}$;第二层具有 49×10 个神经元,激活函数为 $\text{Softmax}(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}} \in (0, 1)$ 。输出为具有 10 个元素的向量,对应于 10 个类别(数字 0~9)的概率。对于编码重建任务,同样采用隐含层为两层但构型不同的浅层神经网络,其中第一层具有 196×441 个神经元,第二层具有 441×784 个神经元,激活函数均为 Sigmoid,输出为具有 28×28 个元素的矩阵,即重建后的图像。在兼顾模型

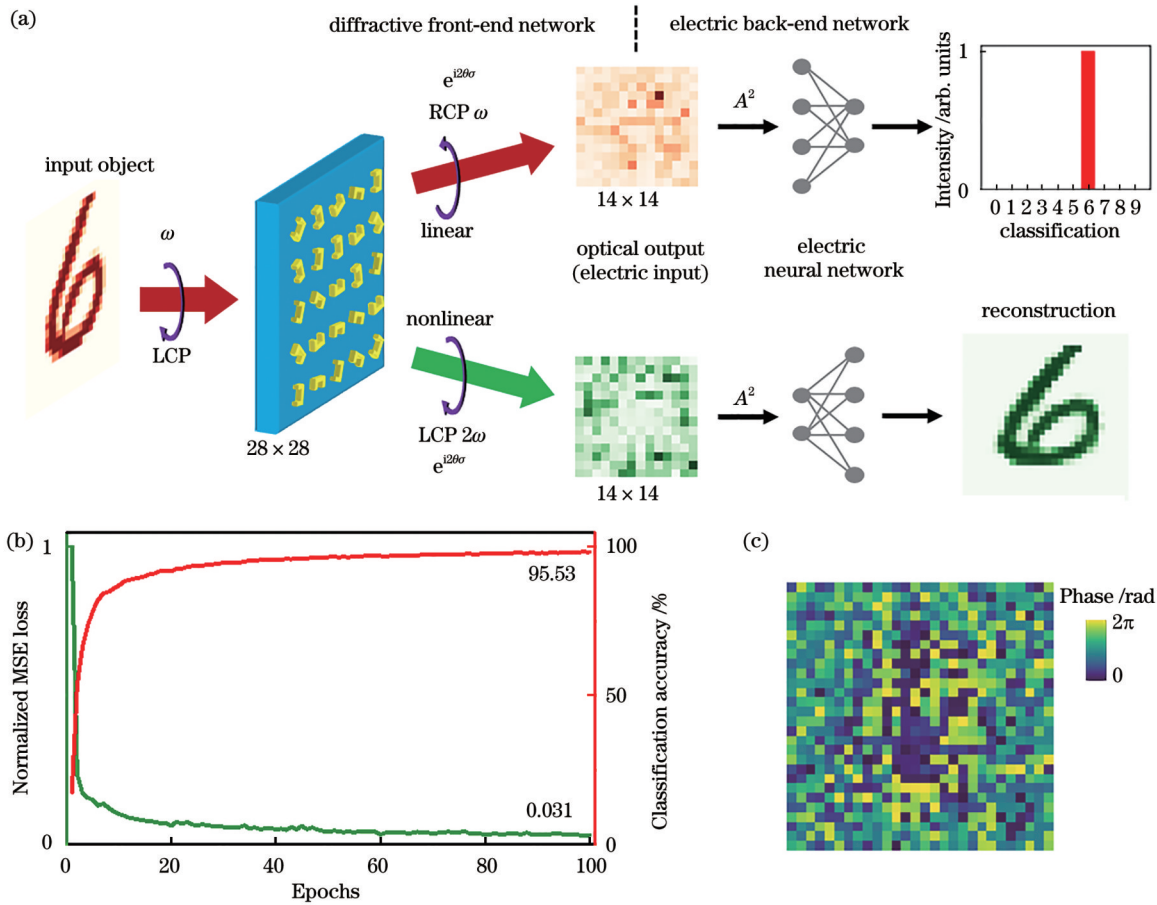


图 1 基于非线性超表面的多任务光电混合神经网络模型和训练。(a)网络模型简图,对手写数字输入实现并行分类和图像重建;(b)分类和重建网络的训练收敛结果;(c)训练网络后得到的非线性超表面的基频相位分布

Fig. 1 Model and training of a multitasking optoelectronic hybrid neural network based on a nonlinear metasurface. (a) Sketch of the network model with handwritten digital inputs for parallel classification and image reconstruction; (b) training convergence results of the classification and reconstruction network; (c) fundamental frequency phase distribution of the nonlinear metasurface obtained after training the network

的简洁性和网络性能的情况下,所提电子网络均选择具有两个隐含层的浅层神经网络,这是因为一个隐含层性能较差且多个隐含层将增加计算资源和内存。

为了提高分类和编码重建的整体性能,将分类准确率和重建质量作为一个合目标,并分别分配权重系数 γ 和 $1 - \gamma$,总的损失函数可以表示为

$$L = \gamma L_c + (1 - \gamma) L_m(O_{\text{recon}}, O_{\text{input}}), \quad (1)$$

式中: L_c 为分类部分的损失项,采用交叉熵损失函数 $L_c = -\sum_{i=1}^n \bar{x}_i \log(x_i)$,其中 x_i 为数字分类电子网络通过最后一层 Softmax 激活函数后求解得到的预测标签, \bar{x}_i 为正确标签; L_m 为编码重建部分的损失项,采用均方误差函数 $L_m = \frac{1}{n} |y_i - \bar{y}_i|$,其中 y_i 为图像重建电子网络通过最后一层 Sigmoid 激活函数后求解得到的预测值, \bar{y}_i 为正确值,通过对重建图像 (O_{recon}) 与输入的真实图像 (O_{input}) 进行像素化的比较计算得到。

每个超表面单元的相位 $\exp(i2\theta\sigma)$ 和 $\exp(i\theta\sigma)$ 是网络需要学习的参数,将前端的光学网络和后端的两个电子网络作为一个整体进行训练,利用梯度下降算法和误差反向传播算法对识别和重建的合目标损失函数进行优化,网络训练完成后通过提取复用的相位 $\exp(i2\theta\sigma)$ 和 $\exp(i\theta\sigma)$ 确定了唯一的超表面结构。所提非线性光电混合网络模型在台式计算机 (CPU: 英特尔 (R) Core(TM) i9-13900K; RAM: 128 GB; GPU: NVIDIA GeForce RTX 4090; Windows 10 操作系统) 上运行,收敛时间为 125.49 s,其收敛速度比同等条件下的纯电网络收敛速度提升了约 12%。训练后结果如图 1(b) 所示,分类部分在 MNIST 测试数据集上达到了 95.53% 的盲测精度,并且编码重建的归一化损失函数也收敛到 0.031 左右。由于要兼顾复用的两种功能,所提网络的分类准确率和重建质量相对于单个同类型的网络略微降低。优化学习过程中,超表面的

相位分布是一个 28×28 的变值矩阵,即具有 784 个可调制相位单元。以基频为例,利用梯度下降算法对初始解为 $0 \sim 2\pi$ 的随机值进行优化,随着迭代次数的增加,每个相位单元都会根据梯度作出相应的调整,模型最终收敛时获得理想的相位分布。优化后的超表面基频相位轮廓如图 1(c) 所示,可以发现,相位分布具有一些显著特性,较大值 (约为 2π) 和较小值 (约为 0) 主要分布在中心部分,而周围的相位较为均衡。该复用的光电混合网络模型除了具有并行运算的优势外,还通过一个超薄的光学前端显著减少了后端输入像素的数量,使得输入到电子网络的数据量被压缩到 1/4 (从 28×28 个像素点到 14×14 个像素点),降低了全连接电子网络的复杂性。此外,其他高频光波如三倍频、四倍频等也可以加载到网络中,以实现更多功能的复用。

2.2 非线性超表面单元设计

为了实现网络模型中非线性复用功能以及获得高的二次谐波产生 (SHG) 效率,超表面采用图 2(a) 所示的一重对称性 (C_1) U 型谐振单元^[25,31,33]。基于几何相位 [也称为 Pancharatnam-Berry (PB) 相位],通过旋转它的空间角度可以实现连续的相位调控。当自旋态为 σ 的圆偏振光束入射时,其中 $\sigma = \pm 1$ 分别表示左旋和右旋圆偏振,取向角为 θ (相对于实验室空间坐标) 的 U 型谐振单元具有各向异性的响应,自旋态为 σ 的光束会发生偏振转换,导致发射光束中出现混合自旋态。对于 n 次谐波,它与入射光束具有相同和相反自旋态的谐波所携带的相位分别为 $(n - 1)\sigma\theta$ 和 $(n + 1)\sigma\theta$ 。理论上, C_1 型超表面单元能产生任意倍频的几何相位,因此它具有很多相位复用通道。本文选择基频下相反自旋态的 $\exp(i2\theta\sigma)$ 和二倍频下相同自旋态的 $\exp(i\theta\sigma)$ 来实现两种功能复用,此时具有相关性的超表面相位 θ 和 2θ 即是网络需要学习的参数。

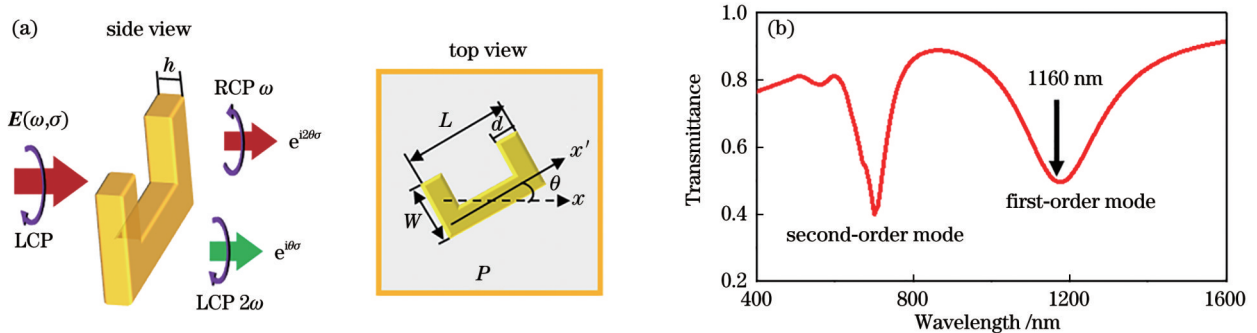


图 2 非线性超表面单元示意图和透射谱。(a) 具有定向角的 U 型谐振单元在圆偏振下产生线性和非线性几何相位; (b) 超表面单元在圆偏振入射光下的透射光谱

Fig. 2 Schematic and transmission spectrum of nonlinear metasurface units. (a) U-shaped resonant unit with directional angle generating linear and nonlinear geometric phases under circular polarization; (b) transmission spectrum of the metasurface unit under circularly polarized incident light

由于 SHG 的效率强烈依赖于 U 型谐振单元局域等离子体模的共振,因此本实验详细地设计了超表面单元的尺寸,如图 2(a)的侧视图和顶视图所示,最终的结构尺寸为 $L=180\text{ nm}$ 、 $W=100\text{ nm}$ 、 $d=40\text{ nm}$ 、厚度 $h=40\text{ nm}$ 。利用 FDTD 软件仿真了该结构在圆偏振下的透射谱[图 2(b)],以确定其共振波段。从结果可以看出,基波在 1160 nm 附近的共振最强,在 704 nm 附近存在二阶共振。由于这种高阶模态的共振波长与产生超表面高阶波信号的范围相差较大(比如二倍频波长为 580 nm),因此可以认为这种共振不会对超表面高阶波效率产生任何影响。在实际应用中,通过采用介质超表面^[28-29, 34]、BIC 增强^[35-36],或者在超表面设计中加入高度非线性材料,还可以进一步提高 SHG 转换效率。

3 仿真结果与分析

3.1 分类功能

通过基频(利用所设计的非线性超表面单元,入射

波长为 1160 nm)入射和出射相位 $\exp(i2\theta\sigma)$ 来执行第一个复用功能[图 3(a)],即使用 MNIST 训练集训练网络来识别手写数字图像,训练完成后对测试集全部 10000 个数字的识别仅需 2.48 s,从测试集[图 3(a)第一行]中选取 10 个样例数字(0, 1, ..., 9)作为输入的仿真结果。图 3(a)第二行为这 10 个输入数字通过非线性超表面后得到的基频光学输出, 28×28 维的输入被转换为 14×14 维的输出。可以看出,其振幅分布是无规律的,但具有一定的聚集性。通过这种振幅分布无法直观地判断输入数字信息的类别,但后端电子网络可以高效地学习这些差异并进行分类。图 3(a)第三行为数字识别电子网络的输出,对于每个数字它都输出一个一维概率矩阵,概率最大的标签则被认为是识别的数字结果。将识别结果与该数字的真实标签进行对比,判断识别结果是否正确,分类的准确率则是指分类正确的样本数量占总样本数的比例。从图 3(a)第三行的网络输出结果可以很明显地看出,仿真的结果非常好,相应的数字都被归到各自的类别。

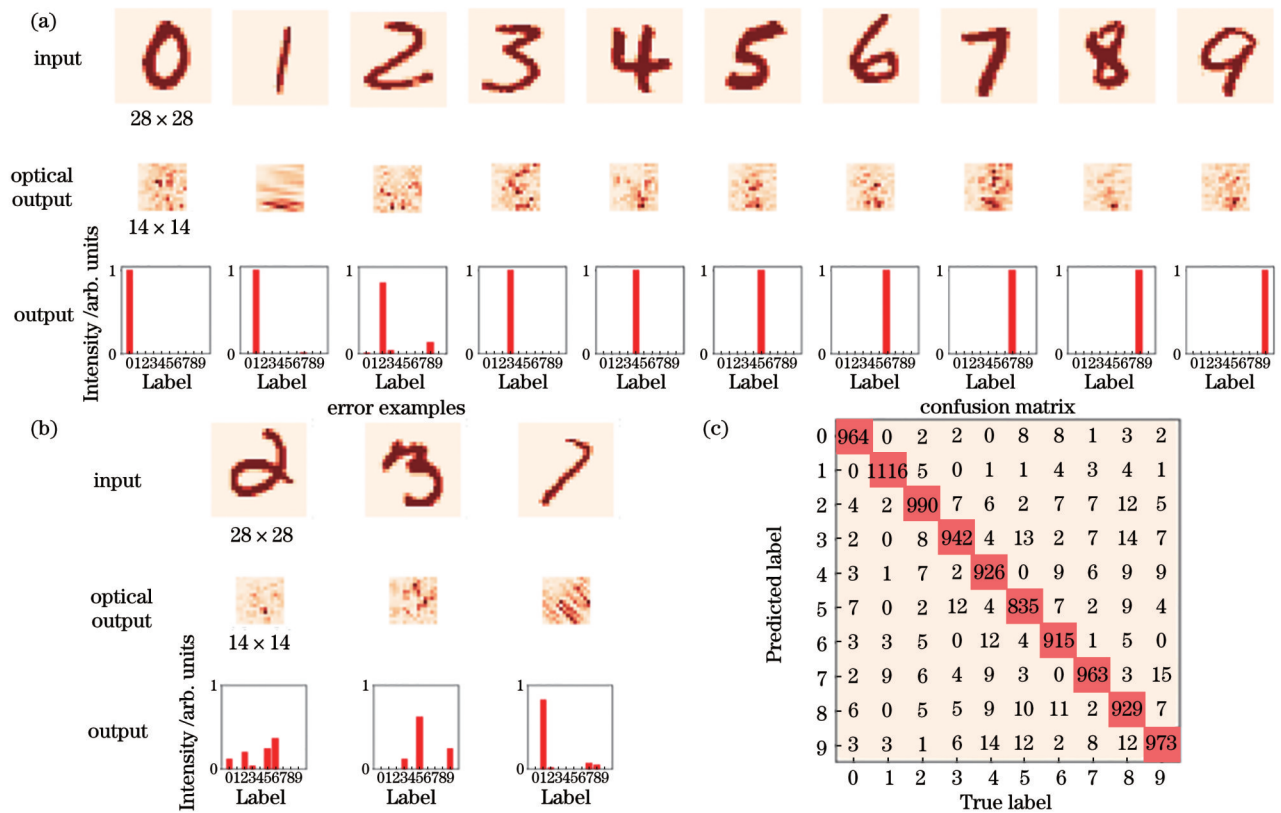


图 3 基频下所提网络分类性能的仿真结果。(a)从正确的结果中随机选取 10 个手写数字的输入、网络中间部分的光学输出和整个网络的输出;(b)一些被错误分类的案例;(c)全部测试集的混淆矩阵

Fig. 3 Simulation results of proposed network classification performance at fundamental frequency. (a) Inputs of 10 handwritten digits randomly selected from the correct results, optical output of the middle part of the network, and the output of the whole network; (b) some misclassified cases; (c) confusion matrix of the full test set

由于网络的准确率为 95.53%,仍然有一些错误的识别结果,因此选取一些例子来研究出现这种错误的原因,如图 3(b)中的数字“2”、“3”、“7”所示。数字

“2”被错误归类为“6”,数字“3”被错误归类为“5”,数字“7”被错误归类为“1”。这些错误分类的输入数字和被分类数字的类型具有一定的相似性,因而网络容

易产生误判。在 MNIST 数据集中有 1 万个未经训练的数字, 图 3(c) 所示为测试集的混淆矩阵, 从中可以发现一些数字容易被误判为与其结构相似的数字, 比如“9”被判别为“7”、“8”被判别为“3”、“5”被判别为“3”。通过调节光学输出的像素数大小可以提升分类性能^[11-12]。此外, 可以利用更加先进的电子网络作为后端来进一步提高分类准确率, 比如 CNN、LeNet 和 ResNet 等^[12, 32, 37]。

3.2 编码重建功能

网络复用的第二个功能是通过基频入射和倍频(超表面出射波长为 580 nm)相位 $\exp(i\theta\sigma)$ 来执行对输入信息的编码和重建任务。出于对比方便和可视化的目的, 仍然选取与分类任务相同的输入, 图 4(a) 所

示为重建结果, 全部测试集的图像重建时间为 9.27 s。图 4(a) 第二行是该任务的光学输出振幅分布, 可以看出, 相对于分类的结果, 这种分布更加随机, 更有利于对信息加密。换句话说, 光学前端的输出是一些无规律的乱码, 无法直观或者通过常规算法获得输入的信息, 只有通过训练好的后端电子网络才能解码, 这在光通信和加密领域具有良好的应用前景。图 4(a) 第三行是第一行输入数字的重建结果, 尽管存在一些细节上的差异, 但是图像的轮廓信息基本准确。此外, 对于数字分类任务中出现错误的数字, 比如“2”、“3”、“7”, 该模型仍然保持优异的重建性能[图 4(b)]。将错误分类的重建图像作为输入进行二次训练, 可进一步提高分类的准确率^[38]。

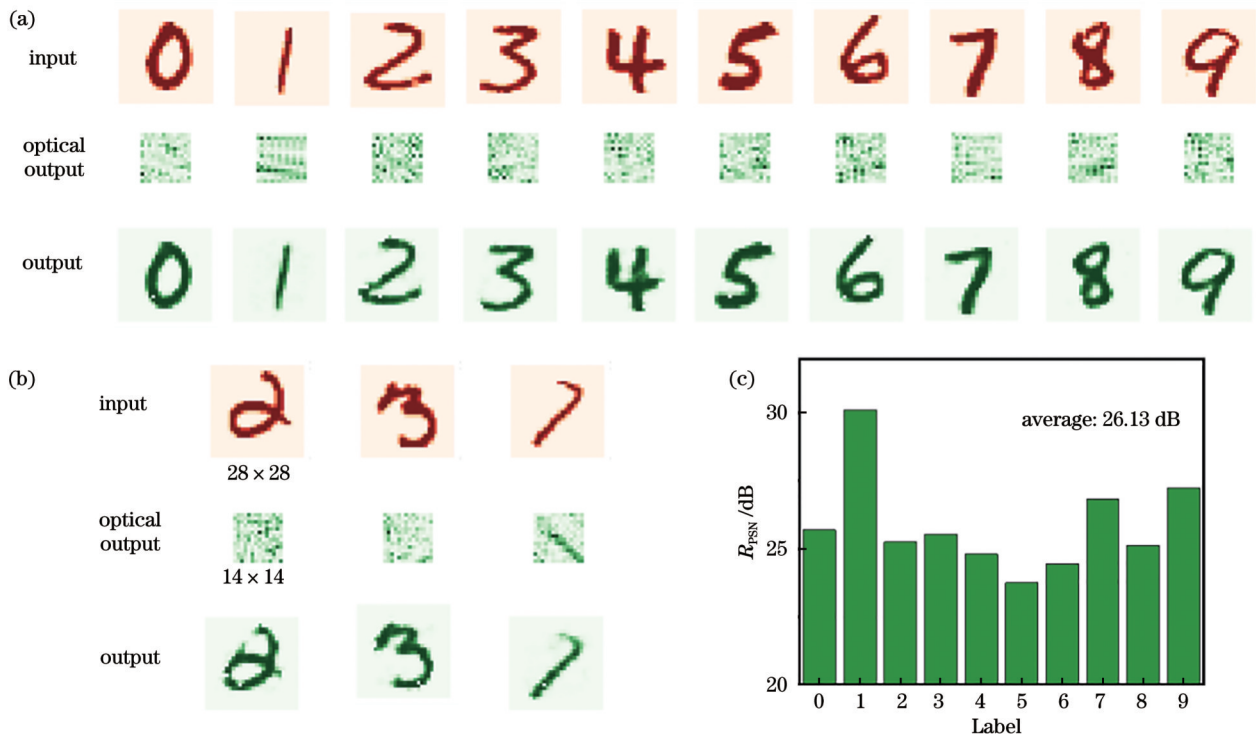


图 4 倍频下所提网络对输入图像进行编码和重建的仿真结果。(a) 10 个手写数字的输入(与分类网络的输入相同)、网络中间部分的光学输出、整个网络的输出图像即重建图像; (b) 分类网络错误识别案例的成像结果; (c) 全部测试集中 10 个手写数字的成像重建质量统计

Fig. 4 Simulation results of proposed network encoding and reconstructing the input image under multiple frequencies. (a) Input images of 10 handwritten digits (same as the input of the classification network), optical output images of the middle part of the network, and output images of the whole network, *i.e.*, the reconstructed image; (b) imaging results of the misrecognized cases of the classification network; (c) reconstruction quality statistics of the 10 handwritten digits in the full test set

为了定量地比较它们的性能, 计算了在测试数据集上重建图像的峰值信噪比, 即 $R_{\text{PSN}} = 10 \cdot \lg \left(\frac{P_{\text{max}}^2}{R_{\text{MSE}}} \right)$ ^[39], 其中 P_{max} 为图像的最大像素值, R_{MSE} 为重建图像和正确图像之间的均方误差。通常来说, $R_{\text{PSN}} > 25$ dB 时成像质量比较好, $R_{\text{PSN}} > 30$ dB 则称为“无失真”。图 4(c) 所示为数字 0, 1, ..., 9 的峰值信噪比评价结果。其中数字“1”由于结构元素简单, 峰值信噪比达到最大值

30.12 dB。全部测试集的平均峰值信噪比为 26.13 dB, 表明模型的重建性能较优, 结果是可靠的。同样地, 通过详细设计电子网络输入的像素数以及采用更先进的网络架构, 可以进一步优化重建性能^[11-12]。

4 结 论

提出一种基于非线性超表面的光电混合网络模型, 该模型在基频和二倍频两个通道下执行手写数字

分类和重建的双重任务,并通过联合优化训练方法获得了高准确率(95.53%)和高图像质量(平均峰值信噪比大于 26 dB)。所提方法结合了光学计算和电子计算的优点,同时展现了高并行、高精度和低功耗的能力,支持高通量光通信技术和多线程光电计算系统的发展。由于非线性的高频通道数量可以增加,因此可以扩展该机器视觉模型,并行执行其他学习任务,如边缘提取、图像分割和超分辨成像等。本文仅展示了一个基于非线性超表面的光电混合网络原型,并对其理论模型进行了数值测试。未来将利用电子束光刻、沉积等工艺通过实验制备该器件;通过非线性的介质超表面进一步提高网络的能量效率及对振幅的有效控制以进一步提升性能。所提方法对于大规模并行光电计算具有重要意义,在机器视觉、智能驾驶、生物成像和智慧医疗等领域具有良好的应用前景。

参 考 文 献

- [1] Wetzstein G, Ozcan A, Gigan S, et al. Inference in artificial intelligence with deep optics and photonics[J]. *Nature*, 2020, 588(7836): 39-47.
- [2] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. *Nature*, 2015, 521(7553): 436-444.
- [3] Jain A K, Mao J C, Mohiuddin K M. Artificial neural networks: a tutorial[J]. *Computer*, 1996, 29(3): 31-44.
- [4] 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.
- [5] Yan T, Wu J M, Zhou T K, et al. Fourier-space diffractive deep neural network[J]. *Physical Review Letters*, 2019, 123(2): 023901.
- [6] Liu J, Wu Q H, Sui X B, et al. Research progress in optical neural networks: theory, applications and developments[J]. *Photonix*, 2021, 2(1): 5.
- [7] Shastri B J, Tait A N, de Lima T F, et al. Photonics for artificial intelligence and neuromorphic computing[J]. *Nature Photonics*, 2021, 15: 102-114.
- [8] 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.
- [9] Luo X H, Hu Y Q, Ou X N, et al. Metasurface-enabled on-chip multiplexed diffractive neural networks in the visible[J]. *Light: Science & Applications*, 2022, 11(1): 158.
- [10] Luo X H, Dong S Y, Wei Z Y, et al. Full-fourier-component tailorable optical neural meta-transformer[J]. *Laser & Photonics Reviews*, 2023, 17(12): 2300272.
- [11] Qu G Y, Cai G Y, Sha X B, et al. All-dielectric metasurface empowered optical-electronic hybrid neural networks[J]. *Laser & Photonics Reviews*, 2022, 16(10): 2100732.
- [12] Mengü D, Luo Y, Rivenson Y, et al. Analysis of diffractive optical neural networks and their integration with electronic neural networks[J]. *IEEE Journal of Selected Topics in Quantum Electronics*, 2020, 26(1): 3700114.
- [13] Caulfield H J, Dolev S. Why future supercomputing requires optics[J]. *Nature Photonics*, 2010, 4: 261-263.
- [14] Yu N F, Genevet P, Kats M A, et al. Light propagation with phase discontinuities: generalized laws of reflection and refraction[J]. *Science*, 2011, 334(6054): 333-337.
- [15] Khorasaninejad M, Chen W T, Devlin R C, et al. Metalenses at visible wavelengths: diffraction-limited focusing and subwavelength resolution imaging[J]. *Science*, 2016, 352(6290): 1190-1194.
- [16] Luo X G. Principles of electromagnetic waves in metasurfaces[J]. *Science China Physics, Mechanics & Astronomy*, 2015, 58(9): 594201.
- [17] 罗翔豪,董思禹,王占山,等.超表面VR/AR显示技术研究进展[J].*激光与光电子学进展*, 2022, 59(20): 2011002.
Luo X H, Dong S Y, Wang Z S, et al. Research progress of metasurface-based VR/AR display technology[J]. *Laser & Optoelectronics Progress*, 2022, 59(20): 2011002.
- [18] Yu N F, Capasso F. Flat optics with designer metasurfaces[J]. *Nature Materials*, 2014, 13(2): 139-150.
- [19] Rubin N A, D'Aversa G, Chevalier P, et al. Matrix Fourier optics enables a compact full-Stokes polarization camera[J]. *Science*, 2019, 365(6448): eaax1839.
- [20] He T, Liu T, Xiao S Y, et al. Perfect anomalous reflectors at optical frequencies[J]. *Science Advances*, 2022, 8(9): eabk3381.
- [21] Dong S Y, Zhang Z Y, Xie L Y, et al. Broadband depolarized perfect Littrow diffraction with multilayer freeform metagratings[J]. *Optica*, 2023, 10(5): 585-593.
- [22] Feng C, He T, Shi Y Z, et al. Diatomic metasurface for efficient six-channel modulation of Jones matrix[J]. *Laser & Photonics Reviews*, 2023, 17(8): 2370040.
- [23] Li G X, Zhang S, Zentgraf T. Nonlinear photonic metasurfaces[J]. *Nature Reviews Materials*, 2017, 2(5): 17010.
- [24] Walter F, Li G X, Meier C, et al. Ultrathin nonlinear metasurface for optical image encoding[J]. *Nano Letters*, 2017, 17(5): 3171-3175.
- [25] Ye W M, Zeuner F, Li X, et al. Spin and wavelength multiplexed nonlinear metasurface holography[J]. *Nature Communications*, 2016, 7: 11930.
- [26] Wang L, Kruk S, Koshelev K, et al. Nonlinear wavefront control with all-dielectric metasurfaces[J]. *Nano Letters*, 2018, 18(6): 3978-3984.
- [27] Minovich A E, Miroshnichenko A E, Bykov A Y, et al. Functional and nonlinear optical metasurfaces[J]. *Laser & Photonics Reviews*, 2015, 9(2): 195-213.
- [28] Gao Y S, Fan Y B, Wang Y J, et al. Nonlinear holographic all-dielectric metasurfaces[J]. *Nano Letters*, 2018, 18(12): 8054-8061.
- [29] Tseng M L, Semmlinger M, Zhang M, et al. Vacuum ultraviolet nonlinear metalens[J]. *Science Advances*, 2022, 8(16): eabn5644.
- [30] 邓俊鸿,李贵新.非线性光学超构表面[J].*物理学报*, 2017, 66(14): 147803.
Deng J H, Li G X. Nonlinear photonic metasurfaces[J]. *Acta Physica Sinica*, 2017, 66(14): 147803.
- [31] 胡子贤,唐宇涛,李贵新.非线性光学超构表面[J].*物理*, 2021, 50(5): 285-292.
Hu Z X, Tang Y T, Li G X. Nonlinear optical metasurfaces[J]. *Physics*, 2021, 50(5): 285-292.
- [32] 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.
- [33] Zeng Y, Hoyer W, Liu J J, et al. Classical theory for second-harmonic generation from metallic nanoparticles[J]. *Physical Review B*, 2009, 79(23): 235109.
- [34] Devlin R C, Khorasaninejad M, Chen W T, et al. Broadband high-efficiency dielectric metasurfaces for the visible spectrum[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2016, 113(38): 10473-10478.
- [35] Koshelev K, Kruk S, Melik-Gaykazyan E, et al. Subwavelength dielectric resonators for nonlinear nanophotonics[J]. *Science*, 2020, 367(6475): 288-292.
- [36] Xu L, Zangeneh Kamali K, Huang L J, et al. Dynamic nonlinear image tuning through magnetic dipole quasi-BIC ultrathin resonators[J]. *Advanced Science*, 2019, 6(15):

- 1802119.
- [37] He K M, Zhang X Y, Ren S Q, et al. Identity mappings in deep residual networks[EB/OL]. (2016-03-16)[2024-01-05]. <http://arxiv.org/abs/1603.05027v3>.
- [38] Li J X, Mengu D, Yardimci N T, et al. Spectrally encoded single-pixel machine vision using diffractive networks[J]. *Science Advances*, 2021, 7(13): eabd7690.
- [39] Dou H K, 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.

Multi-Task Optoelectronic Hybrid Neural Network Based on Nonlinear Metasurface (Invited)

Luo Xuhao^{1,2,3,4}, Dong Siyu^{1,2,3,4*}, Wei Zeyong^{1,2,3,4}, Wang Zhanshan^{1,2,3,4}, Cheng Xinbin^{1,2,3,4**}

¹*Institute of Precision Optical Engineering, School of Physics Science and Engineering, Tongji University, Shanghai 200092, China;*

²*Key Laboratory of Advanced Micro-Structure Materials, Ministry of Education, Shanghai 200092, China;*

³*Shanghai Frontiers Science Center of Digital Optics, Shanghai 200092, China;*

⁴*Shanghai Professional Technical Service Platform for Full-Spectrum and High-Performance Optical Thin Film Devices and Applications, Shanghai 200092, China*

Abstract

Objective Optical and electronic computing systems attempt to understand large amounts of visual data originating from scenarios such as autonomous driving, machine vision, medical diagnostics, remote sensing, defense, and the Internet of Everything. These data are interpreted by artificial intelligence (AI) algorithms, where electronic deep neural networks swiftly emerge as the standard algorithm for visual data processing. Optical neural networks that utilize photons as computational carriers feature high speed, low power, high throughput, and massive parallelism. With the deep integration of optics, optoelectronics, and electronics, the optoelectronic hybrid network model combines the bandwidth of optical computation with the flexibility of electronic computation, providing an order of magnitude increase in the density, speed, and energy of computational systems. However, existing architectures are usually designed for a single task and lack the ability to process multiple tasks in parallel. Therefore, we propose a novel optoelectronic hybrid neural network computational model for multidimensional modulation of linear and nonlinear light waves via nonlinear metasurfaces to simultaneously process information from multiple channels.

Methods The nonlinear metasurface structure employs a one-fold symmetric U-shaped resonant unit to realize the nonlinear multiplexing function of the network and high second harmonic efficiency. Based on the geometrical phase, continuous phase modulation is realized by rotating its spatial angle. The transmission spectrum of the structure at circular polarization is simulated using FDTD. Its resonance wavelength is identified and selected as 1160 nm. Meanwhile, an optoelectronic hybrid network architecture for nonlinear metasurfaces is constructed to realize item classification and reconstruction of coding by phase multiplexing of fundamental and two-fold frequencies. The front-end optical network and the two back-end electronic networks are trained as a package, and the co-objective loss function for recognition and reconstruction is optimized using the gradient descent algorithm and the error back-propagation algorithm. After the networks are trained, unique metasurface structures are identified by extracting the multiplexed phases.

Results and Discussions The results after network training are shown in Fig. 1, and the classification part achieves a blind test accuracy of 95.53% on the MNIST test dataset, with the normalized loss function of the reconstruction of the coding converging to 0.031. The design results and simulated transmission spectra of the nonlinear metasurface are shown in Fig. 2, where its fundamental mid-wave resonates most intensely near 1160 nm and a second-order resonance exists near 704 nm. The simulation results of the network classification performance at fundamental frequency are shown in Fig. 3. The distribution of the fundamental frequency output behind the nonlinear metasurface is irregular, but the back-end electronic network can efficiently learn these differences and classify them. Figure 4 reveals the reconstruction results of the network, where the more random optical output is more favorable for encrypting the information. The PSNR of the reconstructed digit "1" reaches a maximum of 30.12 dB, and the average PSNR of the whole test set is 26.13 dB, which indicates that the model yields high reconstruction performance.

Conclusions We demonstrate an optoelectronic hybrid network model based on a nonlinear metasurface that achieves the dual tasks of performing handwritten digit classification and reconstruction in both fundamental and twofold channels. High accuracy (95.53%) and high image quality (average PSNR >26 dB) are obtained by joint training optimization. The proposed method combines the advantages of optical and electronic computations and simultaneously achieves high parallelism, high accuracy, and low power consumption. Since the nonlinear high-frequency channels can be increased, the machine vision concept can also be extended to perform other learning tasks in parallel, such as edge extraction, image segmentation, and super-resolution imaging. In future work, we will experimentally prepare the device using processes such as electron beam lithography and deposition. The proposed methodology is important for massively parallel optoelectronic computing and will find new applications in machine vision, smart driving, bio-imaging, and smart medicine.

Key words physical optics; optoelectronic computing; artificial intelligence; nonlinear metasurfaces; parallel processing; digital classification; image reconstruction