

基于光学可调非线性函数的目标分类

汪嘉伟^{1,2}, 李一鸣^{1,2}, 张启明^{1,2*}

¹上海理工大学光子芯片研究院, 上海

²上海理工大学光电信息与计算机工程学院, 人工智能纳米光子学中心, 上海

收稿日期: 2025年4月23日; 录用日期: 2025年5月16日; 发布日期: 2025年5月26日

摘要

非线性函数在神经网络中扮演着重要的角色, 通过对神经元引入非线性因素, 从而使得神经网络得以逼近任何非线性函数, 但随着网络本身规模的增加也导致训练时间和算力成本过高。尽管目前已经提出基于电子计算的优化方法, 却仍存在可拓展性差、计算自由度单一等问题。因此, 文章提出了一种利用大量可编程光学神经元的光子计算作为解决方法, 以高鲁棒性、高精度度方法进行Tanh和Sigmoid非线性计算。实验结果表明, 所提出的基于衍射所构建的非线性计算模块在神经网络的应用方面具有相对较高的泛化能力。对于 5×5 数据输入维度, Sigmoid函数实现了 9.66×10^{-4} 的误差损失, Tanh函数实现了 1.02×10^{-3} 的损失。在MNIST手写数字分类任务上, 通过实验验证了架构在面对传输中高斯噪声等现实因素干扰时具有相当的稳定性和鲁棒性。

关键词

光学计算, 非线性计算, 衍射神经网络

Target Recognition Based on Optically Tunable Nonlinear Functions

Jiawei Wang^{1,2}, Yiming Li^{1,2}, Qiming Zhang^{1,2*}

¹Institute of Photonic Chips, University of Shanghai for Science and Technology, Shanghai

²Center of Artificial Intelligence Nanophotonics, School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 23rd, 2025; accepted: May 16th, 2025; published: May 26th, 2025

Abstract

Nonlinear functions play a crucial role in neural networks by introducing nonlinearity into neurons,

*通讯作者。

enabling the network to approximate arbitrary nonlinear mappings. However, as the scale of neural networks continues to grow, the associated training time and computational costs have become increasingly prohibitive. Although several optimization strategies based on electronic computing have been proposed, they still suffer from limited scalability and restricted computational flexibility. To address these challenges, this paper proposes a photonic computing approach based on a large number of programmable optical neurons to perform nonlinear computations with high robustness and precision. Specifically, nonlinear functions such as Tanh and Sigmoid are implemented using a diffractive optical structure. Experimental results demonstrate that the proposed diffraction-based nonlinear computing module exhibits strong generalization capabilities when applied to neural network tasks. For a 5×5 input dimension, the Sigmoid function achieved a loss of 9.66×10^{-4} , while the Tanh function yielded a loss of 1.02×10^{-3} . Furthermore, evaluation on the MNIST handwritten digit classification task shows that the proposed architecture maintains considerable stability and robustness in the presence of practical disturbances such as Gaussian noise during transmission.

Keywords

Optical Computing, Nonlinear Computing, Diffractive Neural Network

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

非线性激活函数对于神经网络学习理解非线性函数有着重要作用,神经网络通过在神经元的输出应用激活函数引入非线性因素,使得神经网络可以逼近任何非线性关系[1]。对于神经网络模型,一个合适的非线性函数不仅可以显著提高其泛化能力以及性能,也可以稳定其训练过程,通过避免梯度消失或爆炸来使模型有效收敛[2]。Sigmoid 和 Tanh 函数是近年来经常被应用于各种神经网络模型中的非线性激活函数[3][4]。Tanh 函数在 RNN 等网络的归一化与对称数据处理中起到重要的作用[5], Sigmoid 函数在卷积神经网络(CNN) You Only Look Once (Yolo)模型的多标签预测 logistics 分类器[6]、递归神经网络(LSTM)的门控机制(如输入门、遗忘门、输出门等)[6][7]中起到重要作用。但由于传统电子计算中需要晶体管参与执行,所以根据摩尔定律,其计算的速度和能耗不可避免地受到晶体管带来的限制,无法满足神经网络规模扩张的需求[8]-[10]。

光学计算区别于传统电子电路计算结构,凭借其高带宽、低损耗、高速、可拓展性强与兼具集成化、模块化等特点,近年来在信号处理、数学计算和神经网络方面都有着卓越的进展[11][12]。目前研究者们提出了许多光学计算方法,诸如线性卷积希尔伯特变换[13],以及非线性光学逻辑门等[14][15]。衍射神经网络(Diffractive Neural Network, D²NN),是一种采取机器学习训练方法的全光神经网络框架[16],根据计算需求的不同,其通过衍射调控光场的相位和幅值,拟合出对应关系。基于衍射的光学计算方法有着很高的模块性、灵活性以及可拓展性[11]。近年来针对 D²NN 的研究也产生了许多成果,不仅可以减少传统计算中的误差、优化计算流程,同时也在计算机视觉、信息处理等领域单独训练完成某些分类任务[17]-[20],为光学计算的发展提供了一种新的思路。

本文设计了一种基于机器学习方法的衍射处理单元(Diffractive Processing Unit, DPU) [21]的光电非线性计算模块。这个模块使用三个线性光学层和最后的非线性强度激活函数,通过可编程光学神经元来拟合 Sigmoid 和 Tanh 两种非线性激活函数。实验结果表明,对于 5×5 数据输入维度, Sigmoid 函数实现了 9.66×10^{-4} 的误差损失, Tanh 函数实现了 1.02×10^{-3} 的损失。此外,随着输入数据维度从 3×3 增加到 $9 \times$

9, Sigmoid 函数的误差损失从 8.46×10^{-4} 增加到了 1.18×10^{-3} 。同样, 对于 Tanh 函数, 随着 MU 尺寸从 3×3 增加到 9×9 , 误差损失从 9.12×10^{-4} 增加到了 1.33×10^{-3} 。并且, 在 MNIST 手写数字分类任务上, 通过实验验证了架构在面对传输中高斯噪声等现实因素干扰时具有相当的稳定性和鲁棒性。

2. 使用须知

2.1. 实验参数设置

本实验的模型基于 PyTorch 1.12.1 框架、Python 3.9 以及 CUDA 11.3.58。使用配置 GPU 为 nvidia GeForce 3050, Ram 容量为 12 GB, 处理器为 12 th Gen Inter (R) Core (TM) i7-12700H。在实验过程中, 经过测试, 使用均方误差(Mean Square Error, MSE)作为损失函数, 其计算公式是:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (X_i - Y_i)^2 \quad (1)$$

其中 X 、 Y 分别为衍射神经网络的输出与实验设置标签值。并且为了进一步降低训练中的损失率, 实验中使用 Adam 优化器和阶梯式降低学习率。在 Sigmoid 训练中, 学习率设置为 20 轮下降, $L_{\text{learn}} = L_{\text{current}} \times 10^{-3}$, 其中 L_{current} 是当前学习率, 每 20 个 epoch 后进行一次下降; 而在 Tanh 训练中, 学习率同样设置为每 20 轮下降, 学习率设置为 $L_{\text{learn}} = L_{\text{current}} \times 10^{-2}$ 。

模拟时, 每层所使用的可编程光学神经元使用随机初始化。关于其他相关的参数, 主要和实际实验设备中参数相匹配, 其中模拟时, 每层所使用的可编程光学神经元使用随机初始化。在实验中, 使用波长为 532 nm 的激光器作为光源, 设备中的像素大小设置为 $16 \mu\text{m} \times 16 \mu\text{m}$, 其中数据在 DPU 中各种像素范围设置为 5×5 , 单层衍射层中的光电子神经元数量配置为 200×200 个, 光场在被探测器检测前自由传播 15 cm。在衍射层的最后对接收输出的能量分布结果进行平均掩码, 步长需要根据输入数据零填充的位置进行调整。

2.2. 数据集与预处理方法

在训练中, 可编程的光学神经元由误差反向传播算法优化, 损失函数通过对比输出图像的值与设置的标签图像得到。作为训练光电神经网络模块的训练图像集, 使用任意的高斯随机向量做成图像矩阵。为了测试不同输入向量大小下的模块性能, 图像矩阵大小分别设置为 3^2 、 5^2 、 7^2 、 9^2 , 训练、验证和测试分别有 16000、3000、1000 个样本。图像数据输入时为了提高衍射时模拟的数值精度, 在预处理中将图像进行限带光谱的零填充映射处理[22], 其过程如图 1 所示。

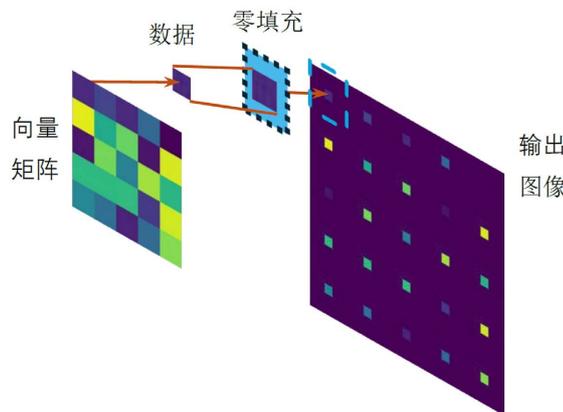


Figure 1. Schematic diagram of zero-padding
图 1. 零填充过程示意图

零填充(Zerp Padding)在模拟自由空间传播中是常用的减少计算混叠误差的方式,但是由于引入了更多的数据点,这种方法也会导致计算负担的增加。研究者为了解决这一问题提出自适应采样角谱方法,通过自适应的零填充与采样参数避免循环卷积误差,从而减少计算复杂度。

在测试非线性计算模式性能的仿真中主要使用 MNIST 手写数字数据集,该数据集包含 60,000 张训练图像和 10,000 张测试图像,每张图像的分辨率为 28×28 像素。为了使这些图像适配 DPU 中衍射层的分辨率,需要对图像进行插值缩放处理,在保持图像细节的同时,减少图像由于放大而导致的模糊现象。

3. 非线性计算与衍射神经网络传播原理

3.1. 非线性计算模式

Tanh 函数和 Sigmoid 函数主要用于神经网络的分类任务中。Sigmoid 公式为:

$$\text{Sigmoid}(x) = \frac{1}{1 + \exp(-x)} \quad (2)$$

其中 x 为输入向量。Sigmoid 函数的值域限定在 $(0, 1)$ 之间,具有明显的概率解释意义,因此非常适合用于处理具有概率性质的数据,尤其是在二分类问题中常作为输出节点的激活函数。Tanh 函数计算公式:

$$\text{Tanh}(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} \quad (3)$$

其中 x 为输入向量。Tanh 函数的值域限定在 $(-1, 1)$ 之间,具有对称性和零均值的特点,非常适合用于表示具有正负关系的信号或特征数据。与 Sigmoid 函数相比, Tanh 函数通过将输入映射到负值与正值的区间,能够更有效地表示数据的方向性和幅度,从而在特征表达上更具优势,并且能够有效缓解因梯度过小而导致的梯度消失问题。在实际实验的设置中通常随着神经网络模型的需求而设置。

3.2. 衍射层正向传播原理

衍射层的搭建基于 Rayleigh-Sommerfeld 衍射积分公式,将对应位置作为神经元,其光学波为:

$$w(x, y, z) = \frac{z - z_i}{r^2} \left(\frac{1}{2\pi r} + \frac{1}{j\lambda} \right) \exp\left(\frac{j2\pi r}{\lambda}\right) \quad (4)$$

其中 λ 为波长, $r = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}$, $j = \sqrt{-1}$ 。输入光被调制到入射光场,其振幅和相位为复值函数,最后从输出光场的平均振幅能量分布中提取结果,将结果转化为实数,通过光和电的综合作用保证模块的非线性运算。

DPU 的整体正向传播过程同样遵循远场区域的 Rayleigh-Sommerfeld 衍射积分方程。在此过程中,光电神经元对输入节点的加权光场进行求和,并经过复合激活函数产生相应的输出。第 i 个光电神经元的输出可以表示为:

$$u_{output}^i = \varphi(\tilde{u}_{output}^i) = \varphi\left(\sum_m \tilde{w}^{im} \tilde{u}_x^i \exp(j\theta^m)\right) \quad (5)$$

其中 m 代表第 m 个神经元, $\exp(j\theta^m)$ 表示第 m 个 SLM 相位调制的可训练相位值。

3.3. 神经网络中的误差反向传播

衍射层中具体的参数通过计算机中的正向传播与反向误差传播方法进行训练和修正,最后将训练好的参数部署到 DPU 中。具体而言,正向传播通过将数据输入神经网络后,每一层神经元根据前一层的输出与权重对数据进行加权计算后传递给下一层,最后由输出层输出网络的预测结果,过程可以表示为:

$$y^l = f^l(W^l y^{l-1} + b^l) \quad (7)$$

其中, y^l 代表第 l 层的输出, y^{l-1} 代表上一层的输入, W^l 是第 l 层的权重矩阵, b^l 为第 l 层的偏置向量, f^l 是第 l 层的激活函数, 如 ReLU、Sigmoid 等。对应到衍射神经网络中则是光场的振幅信息通过角谱衍射理论进行正向传播, 表达式为:

$$A(f(x), f(y)) = \mathcal{F}^{-1}(\mathcal{F}(A_0(f(x), f(y))) \cdot H(f(x), f(y))) \quad (7)$$

其中, $A_0(f(x), f(y))$ 表示输出光场的复振幅, $H(f(x), f(y)) = \exp\left(jkz\sqrt{1 - (\lambda f(x))^2 - (\lambda f(y))^2}\right)$, 而反向传播一般通过利用损失函数相对于网络中参数(例如权重和偏置)的梯度来不断迭代调整参数, 从而将预测结果误差最小化。具体到 DPU 的训练过程中则是通过输出图像与标签图像的均方误差来计算神经网络中参数的梯度。反向传播具体过程可以表示为:

$$\delta^l = \nabla_{y^l} L \odot f'^l(z^l) \quad (8)$$

其中 $\nabla_{y^l} L$ 为损失函数对输出的梯度, $f'^l(z^l)$ 为输出层激活函数的导数, δ^l 为输出层的误差项。利用公式(1)和公式(8), 可以得到损失函数对第 l 层的误差表达式:

$$\frac{\partial E(\phi_i^l)}{\partial \phi_i^l} = \frac{4}{k} \sum_k (x_k - \hat{x}_k)^2 \cdot \text{Real}\left\{ (m_k)^* \cdot \frac{\partial m_k}{\partial \phi_i^l} \right\} \quad (9)$$

其中, $\frac{\partial E(\phi_i^l)}{\partial \phi_i^l}$ 代表输出复数光场的梯度, m_i 表示前一层的输出。

4. DPU 级联网络

4.1. DPU 级联网络性能分析

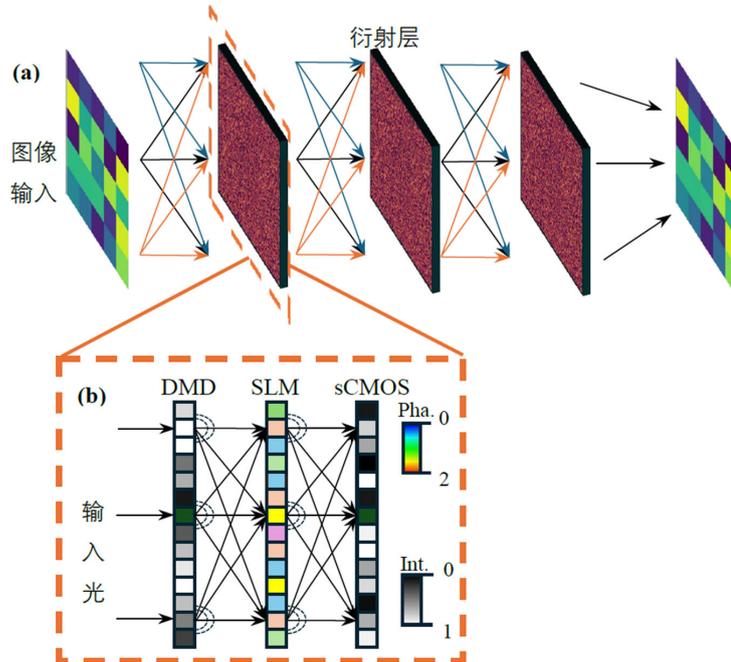


Figure 2. (a) Schematic diagram of DPU cascade diffraction neural network structure; (b) Schematic diagram of DPU structure
图 2. (a) DPU 级联衍射神经网络结构示意图; (b) DPU 结构示意图

DPU 模块可以被视为一个可以对光场相位和振幅信息进行调制的衍射层，前一级的光学神经元作为下一级的光源逐级进行前向传播，层间神经元通过光的自由空间传播相连。DPU 依赖光场的并行传播进行信息处理，其运算不依赖于串行指令，因而可以通过光速进行推理。在三层的衍射层中，每层有 200×200 个光学神经元，保证了 DPU 相比于传统电子神经网络在时间效率方面显著的优势。DPU 级联衍射神经网络结构和 DPU 的结构如图 2 所示。

在单个的 DPU 中，首先数字微镜(Digital Micro-Mirror Device, DMD)将输入信息编码为相干光场的振幅，然后空间光调制器(Spatial Light Modulator, SLM)调制其相位分布。输入信息和神经元之间的加权连接是通过自由空间光学衍射建立的，从 SLM 平面到探测器平面的衍射定义了每个神经元的感受野。在探测器平面上，科学互补金属氧化物半导体传感器(Scientific Complementary Metal Oxide Semiconductor, sCMOS)测量复光场振幅分量的平方，通过光电效应从入射场中提取强度信息。

关于不同输入图像大小下 Sigmoid 和 Tanh 训练对比结果和学习曲线如图 3、图 4 所示，输入尺寸较小时，在训练的前 20 Batches 中有着较好的效果。

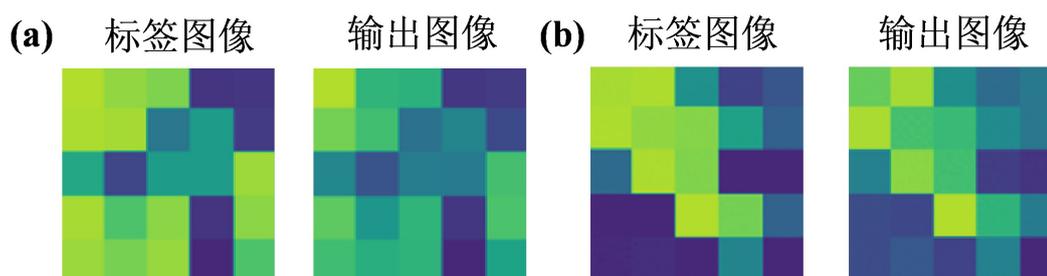


Figure 3. Comparison of training results for 5×5 image sizes: (a) Sigmoid; (b) Tanh

图 3. 5×5 图像尺寸训练结果对比: (a) Sigmoid; (b) Tanh

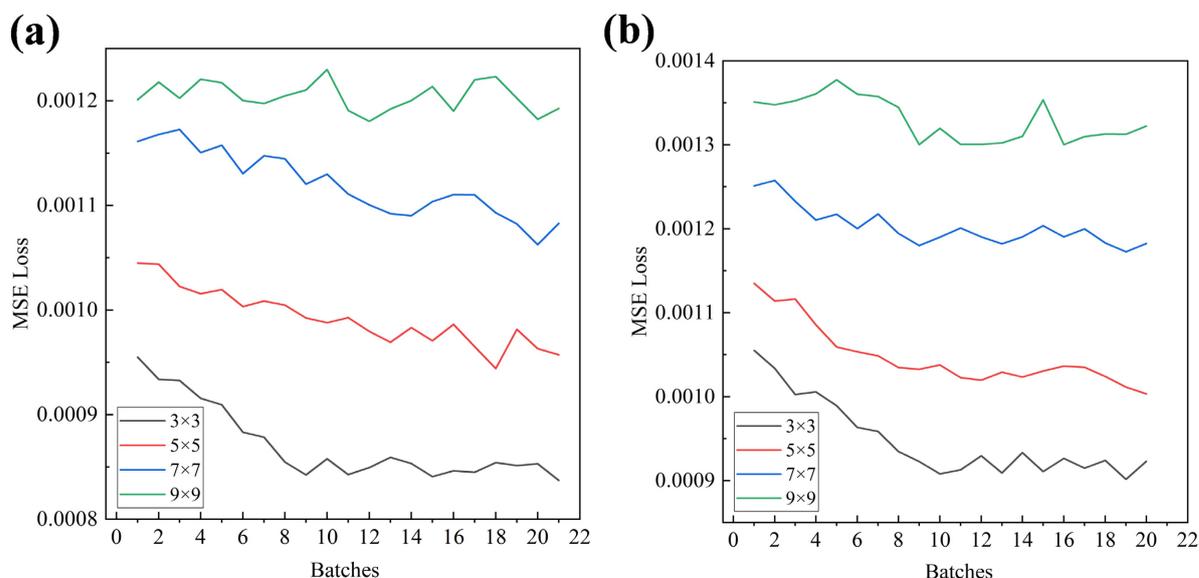


Figure 4. (a) Sigmoid mode training and validation MSE curve; (b) Tanh mode training and validation MSE curve

图 4. (a) Sigmoid 模式训练的 MSE 曲线; (b) Tanh 训练的 MSE 曲线

在 3×3 和 5×5 输入数据维度下，模型的输出图像与标签图像之间具有较高的相似度，表明系统在较小输入维度时能够较准确地完成非线性映射，保持良好的重构效果，Sigmoid 模式最终训练的损失为

8.46×10^{-4} 和 9.66×10^{-4} , Tanh 模式分别为 9.12×10^{-4} 和 1.02×10^{-3} 。随着输入数据维度增大至 7×7 和 9×9 , 模型仍能保持主要的数量关系特征, 但整体图像质量相较于小尺寸输入有所下降, 损失也变得难以收敛, 训练过程中波动较大, Sigmoid 模式损失为 1.08×10^{-3} 和 1.08×10^{-3} , Tanh 模式损失提升至 1.19×10^{-3} 和 1.33×10^{-3} 。

为了进一步量化不同非线性映射模式的性能表现, 实验对 Sigmoid 和 Tanh 函数在不同输入维度下的关键指标进行了对比分析。具体而言, 通过计算均方误差、结构相似性指数和感知损失的评估指标, 衡量不同图像尺寸在图像重构精度、结构保持能力及整体稳定性方面的差异。实验数据表明, 在不同指标下, DPU 级联非线性计算模块均展现出较高的准确性, 验证了系统在不同输入维度下的优越性能。验证结果如表 1 所示。

Table 1. System resulting data of standard experiment

表 1. 标准试验系统结果

函数	评价指标	均方误差	结构相似性	感知损失
Sigmoid	3×3	8.46×10^{-4}	0.9083	0.1435
	5×5	9.66×10^{-4}	0.8779	0.1633
	7×7	1.08×10^{-3}	0.8402	0.1796
	9×9	1.18×10^{-4}	0.8179	0.2041
Tanh	3×3	9.12×10^{-4}	0.8913	0.1562
	5×5	1.02×10^{-3}	0.8530	0.1703
	7×7	1.19×10^{-3}	0.8133	0.1920
	9×9	1.33×10^{-3}	0.8029	0.2107

出现这种现象可能是由于衍射层的分辨率在较大输入数据尺寸时会受到显著影响。尤其是当需要精细调制更多像素时, 衍射层内的响应速度降低, 导致信号处理效率下降。这种情况在训练过程中尤为明显, 因为模型性能会随着输入数据尺寸的增加而降低, 从而影响模型的学习效率。具体表现为模型的收敛速度变慢以及训练误差的增加。相比之下, 较小尺寸的输入数据能够在光学器件的最佳工作范围内操作, 确保输入数据的质量更高, 从而加速模型的收敛过程, 并提高训练结果的稳定性。此外, 较大尺寸的输入数据不仅在自由空间传播中可能会出现相邻采样区域的串扰, 并且在光学器件上还可能存在严重的边缘失真效应。

4.2. 不同层数对网络非线性能力的影响

不同层数 DPU 的堆叠对于非线性函数的映射同样有影响。实验结果显示, 对于 Sigmoid 和 Tanh 函数, 在 DPU 层数较少时, 网络在训练过程中难以有效地学习和准确映射复杂的非线性关系, 导致输出结果中误差较大。适当增加光学处理层数能够显著提升特征提取能力与表达能力, 更深的光学衍射网络通过引入更多的计算层次与相位调制自由度, 增加了网络的非线性表达能力与计算自由度, 从而获得更精细的拟合效果。实验结果表明, 随着 DPU 层数的增加, 系统对非线性函数的拟合精度显著提升。当 DPU 层数较少时, 模型在拟合复杂非线性特征时的输出结果与标签表现出更大的误差。这一现象也反映出, 通过增加光学处理层数, 模型的计算自由度得到扩展, 光场的相位和振幅调制能力也相应增强, 系统可

以具有更复杂的函数逼近能力，从而在拟合非线性函数时表现出更好的结果。通过测试 1 至 3 层 DPU 在拟合上述非线性函数关系时的性能表现，以分析光学处理层数对非线性函数拟合效果的影响。不同层数衍射层对非线性函数在不同指标下的仿真结果表 2 所示。

Table 2. Effects of nonlinear ability of different diffraction layers

表 2. 不同层数衍射层对非线性能力影响

函数	层数	均方误差	结构相似性
Sigmoid	1 层	1.50×10^{-3}	0.8757
	2 层	8.97×10^{-4}	0.8834
	3 层	8.46×10^{-4}	0.9083
Tanh	1 层	1.66×10^{-3}	0.8521
	2 层	1.02×10^{-3}	0.8745
	3 层	9.12×10^{-4}	0.8913

5. MNIST 分类任务测试

为了测试 DPU 级联衍射神经网络的泛化能力，实验中搭建了一个同时使用 Sigmoid 和 Tanh 算子完成的简单手写数字分类任务，并研究其在实际分类任务中的表现效果。分类网络使用 3 层线性层，在网络前 2 层的线性层后分别连接 Sigmoid 算子和 Tanh 算子，在网络第三层线性层后连接电子 SoftMax 进行处理。网络结构如图 5 所示。

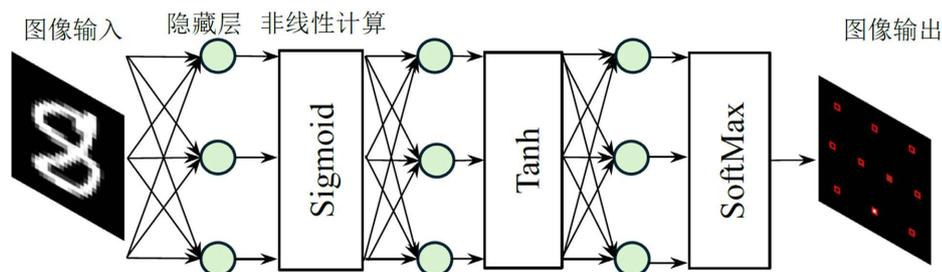


Figure 5. Schematic diagram of MNIST classification task structure based on diffractive neural network

图 5. 基于衍射神经网络的 MNIST 分类任务结构示意图

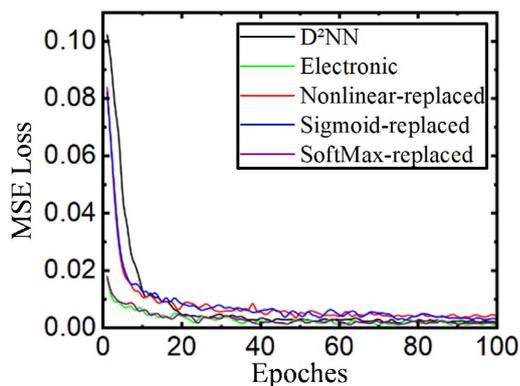


Figure 6. Performance comparison of MNIST classification tasks with different replacement schemes

图 6. 不同替换方案 MNIST 分类任务性能对比

实验结果表明, 在使用复数层光电条件下, 经过训练的网络能够在较小误差($MSE \approx 4 \times 10^{-2}$)的情况下完成对 MNIST 数据集的分类任务。同时我们为了对比使用不同数量光电函数模块对网络能力带来的影响, 设计了只替换 Sigmoid 函数和只替换 Tanh 函数的网络作为对照。实验结果表明, 在使用复数层光电条件下, 经过训练的网络能够在较小误差($MSE \approx 4 \times 10^{-3}$)的情况下完成对 MNIST 数据集的分类任务。同时为了对比使用不同数量光电函数模块对网络能力带来的影响, 设计了分别替换 Sigmoid 函数和 Tanh 函数算子的网络作为对照, 只替换 Tanh 模块的效果相比之下会存在更多的劣化($MSE \approx 3 \times 10^{-3}$)。各类方案对比结果如图 6 所示。

6. 网络稳定性与鲁棒性测试

光电模块在实际应用中不可避免地存在现实干扰, 例如设备物理位置错位、层间误差堆叠、电子设备量化误差、传输中的高斯噪声等, 这些因素对于计算模块会导致产生计算精度、计算速度劣化。所以在鲁棒性测试中, 分别引入高斯噪声、白噪声、椒盐噪声和不同比特深度等影响因素测试对计算模块性能的影响。白噪声和高斯噪声主要通过扰乱光场相位与强度分布, 影响干涉与衍射效果, 从而降低系统对特征边缘和灰度梯度的响应能力。而由于这种误差具有连续性和累积性, 所以会引起数据误差扩大, 从而导致模型的稳定性下降。在白噪声的测试中, 通过调整白噪声的信噪比(SNR)来控制噪声的强度, 分别在高、中、低信噪比区间内测试噪声对输出图像的影响程度。高斯噪声的分布被设定为 $(\mu, \sigma) \in \mathbb{R}^{200 \times 200}$, 其中期望值(μ)和标准差(σ)根据输入数据的尺寸和 DPU 的特性动态调整, 以确保在不同输入数据尺寸下噪声影响的一致性。比特深度作为影响图像表示精度和细节保留的关键因素, 增加其深度可以扩大图像的灰度级范围, 从而更细腻地捕捉亮度变化, 但同时也会增加存储和计算的复杂度。训练阶段使用的图像均为 8 位比特深度, 测试阶段则分别对 4 位、6 位、10 位、12 位和 16 位比特深度的图像进行了量化效果测试。

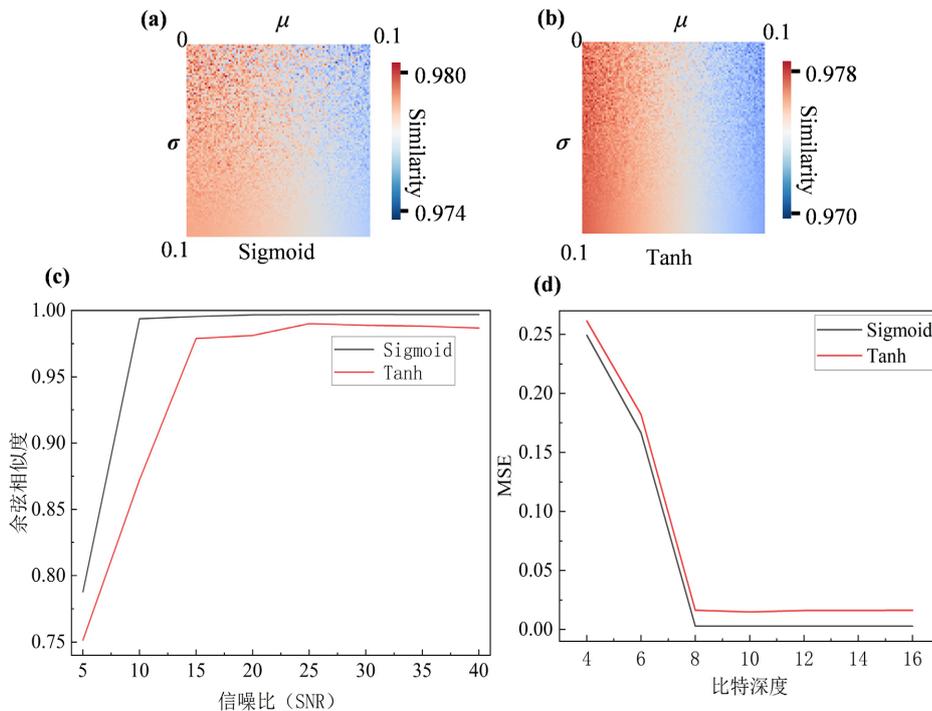


Figure 7. (a) Gaussian noise heat map in Sigmoid mode; (b) Gaussian noise heat map in Tanh mode; (c) Comparison of two calculation modes with different SNR for white noise; (d) Comparison of two calculation modes with different bit depths
图 7. (a) Sigmoid 模式下高斯噪声热度图; (b) Tanh 模式下高斯噪声热度图; (c) 2 种计算模式下白噪声不同信噪比对比; (d) 2 种计算模式下不同比特深度对比

高斯噪声的测试结果表明, 过大的期望值(μ)和标准差(σ)会导致系统性能的显著下降, 但在一定范围内的高斯噪声对整体架构的影响相对较小。系统整体对高斯噪声具有一定的耐受能力。在白噪声测试中, 系统的分类性能随着信噪比的变化呈现出明显的差异性。如图 7(c)所示, 在添加信噪比范围为 40 dB 到 5 dB 的限带白噪声时, 尽管噪声强度逐渐增加, 但由于架构中引入了池化计算等降噪机制, 系统在低信噪比环境下仍能保持较高的可靠性。比特深度实验结果表明, 当比特深度超过 8 位时, 系统架构表现出较高的可靠性。这一现象也说明, 该系统对不同的比特深度具有一定的适应性, 并且在较高比特深度下仍能保持稳定的性能。高斯噪声、白噪声和比特深度的测试结果如图 7 所示。

为了测试 MNIST 分类光电神经网络的鲁棒性与稳定性, 为全非线性替换模型实验中添加高斯噪声与椒盐噪声等现实干扰因素进行测试实验。椒盐噪声为图像的突发性异常像素, 在图像中形成孤立扰动点, 对图像的边结构造成破坏, 会导致衍射层中局部光场突变, 破坏衍射层中相位的连续性, 从而对整个级联网络中光场的正向传播产生严重干扰。测试中将椒盐噪声分别设置 1%、5%、10% 的噪声比例(Salt-Pepper Ratio, SPR)进行测试, 高斯噪声调整标准差 σ 条件噪声强度, 测试范围分别为 0.03、0.06、0.10。实验结果如表 3 所示。

Table 3. System resulting data of noise test

表 3. 噪声测试结果

噪声	噪声强度	均方误差	结构相似性	准确率/%
无噪声	无	0.0041	0.9632	93.2
	$\sigma = 0.03$	0.0040	0.9593	93.0
高斯噪声	$\sigma = 0.06$	0.0053	0.9516	91.1
	$\sigma = 0.1$	0.0072	0.9478	89.6
椒盐噪声	1%	0.0046	0.9412	91.4
	5%	0.0072	0.9352	88.7
	10%	0.0129	0.8933	86.3

根据实验数据分析, 在不同噪声强度下, 模型的性能表现出现了明显的变化。引入高斯噪声后, 模型的性能随噪声强度的增加而逐渐下降。对于椒盐噪声, 模型性能的下落趋势与噪声强度呈现出类似的变化规律, 但是椒盐噪声对模型性能的影响随着噪声强度的增加更加明显, 特别是在噪声强度较高时, 模型的准确率下降更加显著。测试数据说明, 该系统在对抗不同的噪声中均表现出相对的鲁棒性, 并且在较高噪声影响下仍能保持基本的性能。

7. 总结与展望

本文基于可重构光电元器件设计了 DPU 级联系统来映射不同光电计算的非线性函数。通过测试证明, 本文所提出的光电非线性运算模块在不同非线性函数需求的计算场景下, 可以以高精度、低能耗的方式实现 Sigmoid 和 Tanh 运算。模块所表现的映射关系与灵活性也证明了衍射神经网络在不同计算需求下均可以作为替换部分代替传统电子神经网络非线性计算的模块。仿真利用自生成的图像数据集, 通过用不同规模的输入数据对比测试了光学非线性计算模块的性能, 在 Sigmoid 函数的训练过程中, 输入尺寸 3×3 和 5×5 时, 最终损失分别为 8.46×10^{-4} 和 9.66×10^{-4} , 且较早达到稳定状态, 而当输入尺寸增至 7×7 和 9×9 时, 损失值升至 1.08×10^{-3} , 并伴随较大波动; 对于 Tanh 函数, 当输入尺寸为 3×3 和 5×5 时, 损失分别为 9.12×10^{-4} 和 1.02×10^{-3} , 但在更大输入尺寸 7×7 和 9×9 时, 损失增至 1.19×10^{-3} 和 1.33×10^{-3} 。

10^{-3} 。并且根据实验,验证了高斯噪声、系统比特深度等现实因素干扰对运算模块整体性能的影响。这种方法探究了衍射神经网络在映射不同非线性计算的灵活性,并进一步发掘了其在光学计算方面所表现出的潜力。最后实验设计使用 MNIST 数据集分类神经网络替换非线性计算实验进行泛化性能测试,最终替换模型在手写数字分类任务中呈现低损失率($MSE < 10^{-2}$),且出于其模块化设计的结构特征,也可以作为各类光电计算方案中的插入模块作为参考。

在未来的工作中,可以优化 DPU 的结构和后处理算法来优化整体模块的速度并减少电子运算的处理步骤。高精度的光学元件、接口通信的优化、相机的帧率提高等同样也可以在实际应用中提高模块的性能。同时,对于其他非线性函数的光学实现也可以拓展光学计算的应用范围。

参考文献

- [1] Lecun, Y., Chopra, S., Hadsell, R., Ranzato, M. and Huang, F.J. (2006) A Tutorial on Energy-Based Learning. In: Bakir, G., Hofman, T., Schölkopf, B., Smola, A., Taskar, B., Eds., *Predicting Structured Data*, The MIT Press. <https://doi.org/10.7551/mitpress/7443.003.0014>
- [2] Vaswani, A., Shazeer, N., et al. (2017) Attention Is All You Need. <https://ui.adsabs.harvard.edu/abs/2017arXiv170603762V>
- [3] LeCun, Y., Bengio, Y. and Hinton, G. (2015) Deep Learning. *Nature*, **521**, 436-444. <https://doi.org/10.1038/nature14539>
- [4] Kitaev, N., Kaiser, L. and Levskaya, A. (2020) Reformer: The Efficient Transformer. <https://ui.adsabs.harvard.edu/abs/2020arXiv200104451K>
- [5] Zaremba, W., Sutskever, I. and Vinyals, O. (2014) Recurrent Neural Network Regularization. <https://ui.adsabs.harvard.edu/abs/2014arXiv1409.2329Z>
- [6] Liu, F., Zhang, B., Chen, G., Gong, G., Lu, H. and Li, W. (2021) A Novel Configurable High-Precision and Low-Cost Circuit Design of Sigmoid and Tanh Activation Function. 2021 *IEEE International Conference on Integrated Circuits, Technologies and Applications (ICTA)*, Zhuhai, 24-26 November 2021, 222-223. <https://doi.org/10.1109/icta53157.2021.9661606>
- [7] Pogiri, R., Ari, S. and Mahapatra, K.K. (2022) Design and FPGA Implementation of the LUT Based Sigmoid Function for DNN Applications. 2022 *IEEE International Symposium on Smart Electronic Systems (iSES)*, Warangal, 18-22 December 2022, 410-413. <https://doi.org/10.1109/ises54909.2022.00090>
- [8] Moposita, T., Trojman, L., Crupi, F., Lanuzza, M. and Vladimirescu, A. (2022) Voltage-to-Voltage Sigmoid Neuron Activation Function Design for Artificial Neural Networks. 2022 *IEEE 13th Latin America Symposium on Circuits and System (LASCAS)*, Puerto Varas, 1-4 March 2022, 1-4. <https://doi.org/10.1109/lascas53948.2022.9789075>
- [9] Xing, S. and Wu, C. (2020) Implementation of a Neuron Using Sigmoid Activation Function with CMOS. 2020 *IEEE 5th International Conference on Integrated Circuits and Microsystems (ICICM)*, Nanjing, 23-25 October 2020, 201-204. <https://doi.org/10.1109/icicm50929.2020.9292239>
- [10] Kagalkar, A. and Raghuram, S. (2020) CORDIC Based Implementation of the Softmax Activation Function. 2020 *24th International Symposium on VLSI Design and Test (VDATE)*, Bhubaneswar, 23-25 July 2020, 1-4. <https://doi.org/10.1109/vdat50263.2020.9190498>
- [11] Yang, Z., He, J., Yan, Z., Hu, Y., Li, X., Dong, N., et al. (2024) Inverse-Designed Integrated All-Optical Nonlinear Activators for Optical Computing. *Optics Express*, **32**, Article 34001. <https://doi.org/10.1364/oe.531679>
- [12] Yan, T., Wu, J., Zhou, T., Xie, H., Xu, F., Fan, J., et al. (2019) Fourier-Space Diffractive Deep Neural Network. *Physical Review Letters*, **123**, Article 023901. <https://doi.org/10.1103/physrevlett.123.023901>
- [13] Zhao, Y., Chen, Z., Ding, Z., Ren, H. and Nelson, J.S. (2002) Real-Time Phase-Resolved Functional Optical Coherence Tomography by Use of Optical Hilbert Transformation. *Optics Letters*, **27**, 98-100. <https://doi.org/10.1364/ol.27.000098>
- [14] Zhao, Z., Wang, Y., Guan, C., Zhang, K., Wu, Q., Li, H., et al. (2022) Deep Learning-Enabled Compact Optical Trigonometric Operator with Metasurface. *Photonix*, **3**, Article No. 15. <https://doi.org/10.1186/s43074-022-00062-4>
- [15] You, J., Han, Z., Zhang, N., Zhang, Q., Zhang, Y., Liu, Y., et al. (2024) All-Optic Logical Operations Based on the Visible-Near Infrared Bipolar Optical Response. *Advanced Science*, **11**, Article 2404336. <https://doi.org/10.1002/advs.202404336>
- [16] Lin, X., Rivenson, Y., Yardimci, N.T., Veli, M., Luo, Y., Jarrahi, M., et al. (2018) All-Optical Machine Learning Using Diffractive Deep Neural Networks. *Science*, **361**, 1004-1008. <https://doi.org/10.1126/science.aat8084>
- [17] Wang, T., Sohoni, M.M., Wright, L.G., Stein, M.M., Ma, S., Onodera, T., et al. (2023) Image Sensing with Multilayer

-
- Nonlinear Optical Neural Networks. *Nature Photonics*, **17**, 408-415. <https://doi.org/10.1038/s41566-023-01170-8>
- [18] Li, Y., Li, J. and Ozcan, A. (2024) Nonlinear Encoding in Diffractive Information Processing Using Linear Optical Materials. *Light: Science & Applications*, **13**, Article No. 173. <https://doi.org/10.1038/s41377-024-01529-8>
- [19] Dou, H., Deng, Y., Yan, T., Wu, H., Lin, X. and Dai, Q. (2020) Residual D²NN: Training Diffractive Deep Neural Networks via Learnable Light Shortcuts. *Optics Letters*, **45**, 2688-2691. <https://doi.org/10.1364/ol.389696>
- [20] Li, J., Gan, T., Bai, B., Luo, Y., Jarrahi, M. and Ozcan, A. (2023) Massively Parallel Universal Linear Transformations Using a Wavelength-Multiplexed Diffractive Optical Network. *Advanced Photonics*, **5**, Article 016003. <https://doi.org/10.1117/1.ap.5.1.016003>
- [21] Zhou, T., Lin, X., Wu, J., Chen, Y., Xie, H., Li, Y., *et al.* (2021) Large-Scale Neuromorphic Optoelectronic Computing with a Reconfigurable Diffractive Processing Unit. *Nature Photonics*, **15**, 367-373. <https://doi.org/10.1038/s41566-021-00796-w>
- [22] Matsushima, K. and Shimobaba, T. (2009) Band-Limited Angular Spectrum Method for Numerical Simulation of Free-Space Propagation in Far and Near Fields. *Optics Express*, **17**, 19662-19673.