

基于 10.6 μm 波长的小型化非线性全光衍射深度神经网络建模方法

孙一宸, 董明利*, 于明鑫**, 夏嘉斌, 张旭, 白雨晨, 鹿利单, 祝连庆

北京信息科技大学光电测试技术及仪器教育部重点实验室, 北京 100192

摘要 提出了一种基于 10.6 μm 波长的小型化非线性全光衍射深度神经网络建模方法。采用波长为 10.6 μm 的二氧化碳(CO_2)激光光源, 其对应的神经网络物理尺寸为 1 mm \times 1 mm, 依据相关的光学物理参数特性, 构建了基于 10.6 μm 波长的非线性全光衍射深度神经网络模型框架, 使用网格搜索法确定最优的神经网络模型超参数, 并选择交叉熵损失函数和 Adam 优化器对神经网络进行了优化。分别在 MNIST 手写数字数据集和 Fashion-MNIST 数据集上对该方法进行了测试, 其分类结果分别达到了 0.9630 和 0.8743。所提方法为制备小型化的全光衍射光栅提供了理论参考。

关键词 光计算; 非线性全光衍射深度神经网络; 激活函数; 10.6 μm ; 深度学习

中图分类号 TP183

文献标志码 A

doi: 10.3788/LOP202158.0820001

Modeling Method of Miniaturized Nonlinear All-Optical Diffraction Deep Neural Network Based on 10.6 μm Wavelength

Sun Yichen, Dong Mingli*, Yu Mingxin**, Xia Jiabin, Zhang Xu, Bai Yuchen,
Lu Lidan, Zhu Lianqing

*Key Laboratory of the Ministry of Education for Optoelectronic Measurement Technology and Instrument,
Beijing Information Science and Technology University, Beijing 100192, China*

Abstract One method used for modeling a miniaturized nonlinear all-optical diffraction deep neural network based on 10.6 μm wavelength is proposed. First, a carbon dioxide (CO_2) laser light source with a wavelength of 10.6 μm is used, and the corresponding physical size of the neural network is 1 mm \times 1 mm. Second, the model framework of the nonlinear all-optical diffraction deep neural network based on 10.6 μm wavelength is constructed according to the characteristics of relevant optical physical parameters. Finally, the grid search method is used to determine the hyper-parameters of the optimal neural network model, and the cross entropy loss function and the Adam optimizer are selected to optimize the neural network. The proposed method is tested on the MNIST handwritten digital dataset and the Fashion-MNIST dataset, respectively, and the classification results reach 0.9630 and 0.8743, respectively. The proposed method provides theoretical reference for the preparation of miniaturized all-optical diffraction gratings.

Key words optics in computing; nonlinear all-optical diffraction deep neural network; activation function; 10.6 μm ; deep learning

OCIS codes 200.4260; 050.1950; 200.4700

收稿日期: 2020-09-29; 修回日期: 2020-10-21; 录用日期: 2020-11-05

基金项目: 北京市自然科学基金-市教委联合基金(KZ201911232044)、高等学校学科创新引智计划(先进光电子器件与系统学科创新引智基地)(D17021)

* E-mail: dongml@bistu.edu.cn; ** E-mail: yumingxin@bistu.edu.cn

1 引言

深度学习在图像分类^[1]、自然语言处理^[2]和语音识别^[3]等方面取得了一定的成功。深度神经网络具有增强模型表达的能力^[4]。相比于数字实现的神经网络,光计算具有高带宽、高速计算、固有的并行性和低功耗等优势。LED 阵列的 Hopfield 网络^[5]、光电实现的储备池计算^[6-10]、含有微米共振器的尖峰循环网络^[11-12]以及利用马赫-曾德尔干涉仪(MZI)的前馈神经网络^[13]等统称为光神经网络(Optical Neural Network, ONN)。光神经网络具有并行处理、高密度布线以及可直接处理图像等独特的优势^[14],它可由空间光互连和波道光互连实现。

在光神经网络研究中, Lin 等^[15-16]通过芯片将神经网络应用于现实世界,该芯片依靠光的传播实现几乎零消耗和零延迟的深度学习,即全光衍射深度神经网络(all-optical diffraction deep neural network, D²NN)。该结构的物理模型由输入层、5 个隐藏层和输出层组成。输入层由太赫兹波段光源进行照明,输入表面的相位或者振幅以编码光信息。入射光束在输入层衍射后,隐藏层对光束的相位或者振幅进行调制。输出层的光电探测器阵列检测输出光的强度,并根据 10 个不同区域中的光强差异来识别手写数字。通过更新的相位对衍射光栅进行建模,并以 3D 打印的方式制作衍射光栅。2019 年,研究者提出了一种基于上述架构的宽带衍射神经网络^[17]。该模型对光源的要求不再局限于单色相干光,且拓展了此框架的应用范围。然而,这两种方案存在一定的缺陷:1)太赫兹光源价格昂贵;2)衍射光栅尺寸过大而不利于集成。本课题组使用一种衍射光栅代替 3D 打印的衍射光栅^[18-19],由于使用了二氧化碳(CO₂)激光器发射 10.6 μm 的红外

激光和 HgCdTe 检测阵列器,每个神经元的尺寸可以缩小至 5 μm,因此 1 mm×1 mm 的衍射光栅包含 200×200 个神经元。选择 10.6 μm 作为波长可节约成本,整个衍射光栅是以锗为基底,且锗的工作波长为 10.6 μm,这种小型集成化的衍射光栅获得了更广泛的应用。然而,之前的工作中并未考虑模型的非线性效应。

非线性激活函数通过学习输入与输出之间的复杂映射,在神经网络模型中起着至关重要的作用。如果神经网络中没有激活函数,无论有多少神经网络,输出都是输入的线性组合,非线性拟合能力较弱。目前,非线性激活函数主要包括 Sigmoid、Tanh、ReLU 等。其中,ReLU 具有以下三个优势:1)解决梯度爆炸和梯度消失的问题;2)加速收敛;3)使得一些神经元的输出为 0,从而导致网络的稀疏。ReLU 激活函数的改进函数包括 Leaky ReLU、PReLU、RReLU 等。这些函数提高了分类的速度和正确率。ReLU 激活函数允许网络本身引入稀疏性。该方法相当于无监督学习的预训练,大大缩短了学习周期。

本文提出了一种基于 10.6 μm 波长的非线性全光衍射深度神经网络建模方法,相比于文献^[15-16],将神经网络的物理尺寸缩小了 1/80,进一步提高了神经网络的非线性表征能力,并在仿真状态下验证了该模型分类的正确率。研究结果为微米波长的非线性全光衍射深度神经网络模型框架的研究提供了理论依据,为进一步实现大规模集成化、小型化光子计算芯片提供了参考。

2 非线性全光衍射深度神经网络

2.1 MNIST 数据集和 Fashion-MNIST 数据集

MNIST 数据集是由数字 0~9 组成的手写数字数据集,如图 1(a)所示。该数据集由四部分组

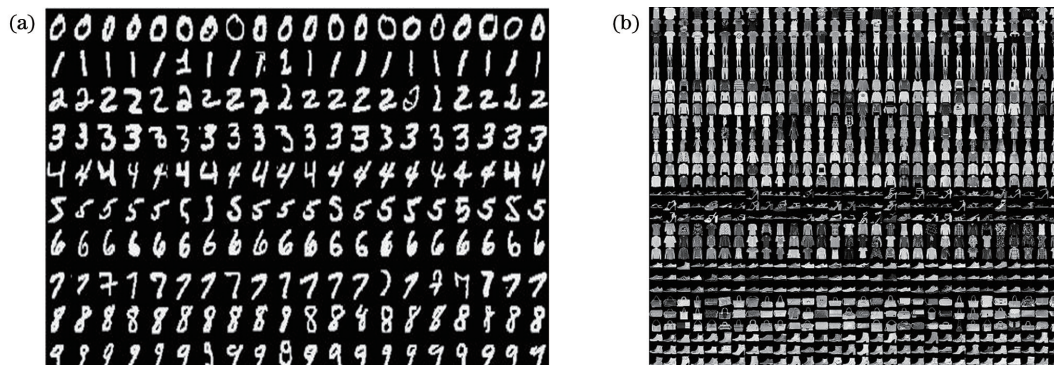


图 1 数据集示例。(a) MNIST 数据集;(b) Fashion-MNIST 数据集

Fig. 1 Dataset examples. (a) MNIST dataset; (b) Fashion-MNIST dataset

成:训练集图像、训练集标签、测试集图像和测试集标签。MNIST 数据集来自美国国家标准与技术研究院(NIST)。训练集和测试集均是由两个数据库混合组成的手写数字组成,其中一个数据库来自高中生,另一个数据库来自人口普查局。MNIST 手写数据集是由包含 60000 个示例的训练集和 10000 个示例的测试集组成。MNIST 数据集集中的每张图片都是由 28×28 个像素点组成,数字均被归一化且以固定方式居中。

Fashion-MNIST 数据集是一种代替 MNIST 手写数字集的分类服饰数据集,具有和 MNIST 数据集同样的特征,即包含 60000 张训练图像和标签及 10000 张测试图像和标签,每张图像的分辨率为 $28 \text{ pixel} \times 28 \text{ pixel}$ 等。相比于 MNIST 数据集, Fashion-MNIST 数据集不再是抽象的数字符号,而是更加具体化的服饰类型,如图 1(b) 所示。Fashion-MNIST 数据集中每个训练样本和测试样本都按照表 1 中的种类进行标注。

表 1 Fashion-MNIST 数据集中的标签号码和种类

Table 1 Label numbers and categories in Fashion-MNIST dataset

Label number	Category
0	T-shirt
1	Trousers
2	Pullover
3	Dress
4	Coat
5	Sandal
6	Shirt
7	Sneaker
8	Bag
9	Ankle boot

2.2 网络模型

图 2(a)所示为物理系统模型,图 2(b)所示为光路模型。系统采用波长为 $10.6 \mu\text{m}$ 的 CO_2 激光器作为光源,激光通过基于输入数据制作的模板后照射到衍射光栅上,经过多层衍射光栅的衍射后在探测器平面形成结果图像,探测器接收到的结果图像即为输入数据的分类结果。通过每个光栅的光经不同厚度的光栅像元调制后,再由次级光栅上的所有光栅像元接收。这种网络连接方式类似于全连接神经网络,第一层光栅接收图像,经过中间多层光栅后到达检测平面。在神经网络中,第一层光栅相当于输入层,中间层光栅相当于隐藏层,检测平面相当于

输出层。输入光的相位调制效应随着光栅高度的变化而变化,在神经网络中对应着不同的权重,如图 2(c)所示。

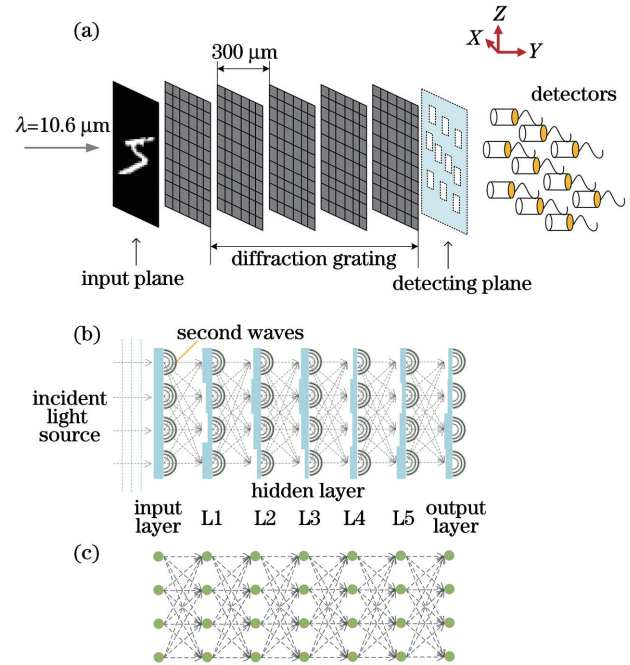


图 2 非线性全光衍射深度神经网络的结构示意图。

(a)系统物理模型;(b)光路模型;(c)神经网络模型

Fig. 2 Structural diagrams of nonlinear all-optical diffraction deep neural network. (a) Physical model of system; (b) optical path model; (c) neural network model

根据 Rayleigh-Sommerfeld 衍射方程,非线性全光衍射深度神经网络中每层衍射光栅的神经元可由次级波方程^[20-21]进行计算:

$$w_i^l(x, y, z) = \frac{z - z_i}{r_2} \left(\frac{1}{2\pi r} + \frac{1}{j\lambda} \right) \exp\left(\frac{j2\pi r}{\lambda}\right), \quad (1)$$

式中: w 为神经网络层中的神经元; (x, y, z) 为该神经元在神经网络层的位置坐标; l 表示网络的第 l 层; i 表示第 l 层的第 i 个神经元,且坐标位置为 (x_i, y_i, z_i) ; $r = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}$ 表示第 l 层节点 i 与第 $l+1$ 层节点的欧氏距离; λ 为入射波长。

光在衍射光栅之间的前向传播过程可表示为

$$\begin{cases} n_{i,p}^l = w_{i,p}^l \cdot g \\ m_i^l = \sum_k n_{k,i}^{l-1} \\ t_i^l = a_i^l \exp(j\phi_i^l) \\ g = \phi_i^l(m_i^l \cdot t_i^l) \end{cases}, \quad (2)$$

式中: $n_{i,p}^l$ 表示位于 (x, y, z) 的第 l 层第 i 个神经元

的输出; p 为下一层的一个神经元, 通过光学衍射与神经元 i 相连; m_i^l 表示第 $l-1$ 层第 k 个神经元的输出总和; t_i^l 为复调制; $|A| = a_i^l(x_i, y_i, z_i)$ 为二次波的相对幅值, 对于只有相位的全光衍射深度神经网络结构, 将振幅 $|A|$ 视为常数, 忽略光损耗, 理想状态为 1; $j\varphi_i^l$ 为每个神经元上输入波 $\sum_k n_k^{l-1}(x_i, y_i, z_i)$ 和复值神经元调制函数 t_i^l 所增加的相位延迟; $\varphi_i^l(m_i^l \cdot t_i^l)$ 表示神经网络中的非线性激活函数, 其作用是经过调制后的二次波神经元通过非线性单元传输到下一层, g 表示经过激活函数的输出值。

相比于文献[15-16], 本文在基于微米波长的非线性全光衍射深度神经网络模型中加入了非线性激活函数 $\varphi_i^l(m_i^l \cdot t_i^l)$, 以提高神经网络模型的非线性拟合能力和泛化能力。ReLU 激活函数具有计算速度、收敛速度快等优势, 并出现了 ReLU 改进函数,

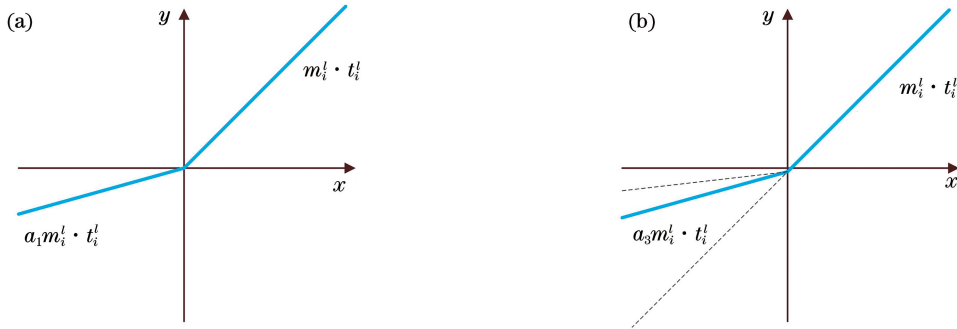


图 3 不同激活函数的数学模型。(a) Leaky-ReLU 和 PReLU; (b) RReLU

Fig. 3 Mathematical models of different activation functions. (a) Leaky-ReLU and PReLU; (b) RReLU

2.3 模型训练

输入光在经过多层光栅的衍射之后会在探测平面输出一幅结果图像, 探测器对结果图像中的探测区域进行探测以得到网络分类结果。因此需要在参数训练阶段对数据标签进行处理, 设计不同标签以对应结果图像中的标志。如图 4 所示, 通过判断结果图像中探测区域内光强最大的区域即可得到结果图像所表征的标签。

对于包含 N 个隐藏层的非线性全光衍射深度神经网络, 其输出层的光强可以表示为

$$I_i^{N+1} = |m_i^{N+1}|^2. \quad (6)$$

对探测器在输出平面测量的强度进行归一化处理, 使它们位于每个样本的 $(0, 9)$ 区间, 用 I_1 表示入射到输出层第 1 个探测器上的光信号总量, 则归一化强度 I_i' 为

$$I_i' = \frac{I_i}{\max\{I_i\}} \times 10. \quad (7)$$

如 Leaky-ReLU、PReLU 和 RReLU^[22-24] 等。模型中三种 ReLU 改进函数与 $\varphi_i^l(m_i^l \cdot t_i^l)$ 的关系为

$$\varphi_L(m_i^l \cdot t_i^l) = \begin{cases} m_i^l \cdot t_i^l, & m_i^l \cdot t_i^l \geq 0 \\ a_1 m_i^l \cdot t_i^l, & m_i^l \cdot t_i^l < 0 \end{cases}, \quad (3)$$

$$\varphi_P(m_i^l \cdot t_i^l) = \begin{cases} m_i^l \cdot t_i^l, & m_i^l \cdot t_i^l \geq 0 \\ a_2 m_i^l \cdot t_i^l, & m_i^l \cdot t_i^l < 0 \end{cases}, \quad (4)$$

$$\varphi_R(m_i^l \cdot t_i^l) = \begin{cases} m_i^l \cdot t_i^l, & m_i^l \cdot t_i^l \geq 0 \\ a_3 m_i^l \cdot t_i^l, & m_i^l \cdot t_i^l < 0 \end{cases}, \quad (5)$$

式中: $\varphi_L(m_i^l \cdot t_i^l)$ 和 $\varphi_P(m_i^l \cdot t_i^l)$ 分别表示模型中的激活函数 Leaky-ReLU 和 PReLU, 在 Leaky-ReLU 中 a_1 表示一个很小的固定常数, 而在 PReLU 中 a_2 表示随模型数据变化的常数; $\varphi_R(m_i^l \cdot t_i^l)$ 表示模型中的激活函数 RReLU, 在 RReLU 中 a_3 表示在 $(0, 1)$ 范围中随机选取的常数。不同激活函数的数学模型如图 3 所示。

将非线性全光衍射深度神经网络的输出结果与输入值进行对比, 使用误差反向传播对光栅参数进行迭代, 并根据目标特性对非线性全光衍射深度神经网络的输出定义损失函数, 使用交叉熵函数作为神经网络中的损失函数。交叉熵函数为

$$H(p, q) = - \sum_i^K p_i^l(x) \ln q_i^l(x), \quad (8)$$

式中: $p_i^l(x) = \frac{e^{I'}}{\sum_i^K e^{I'}}$ 表示神经网络中 Softmax 层的输出值, Softmax 回归可以看作是学习算法优化的分类结果; K 为分类图像标签总数; $q_i^l(x)$ 表示实际图像的输出值; $e^{I'}$ 表示输出平面的归一化强度。

为了将非线性全光衍射深度神经网络模型训练成数字分类器, 采用 MNIST 手写数字数据集作为输入层。

图 5 表示各层衍射光栅的输出图像, 最后一层

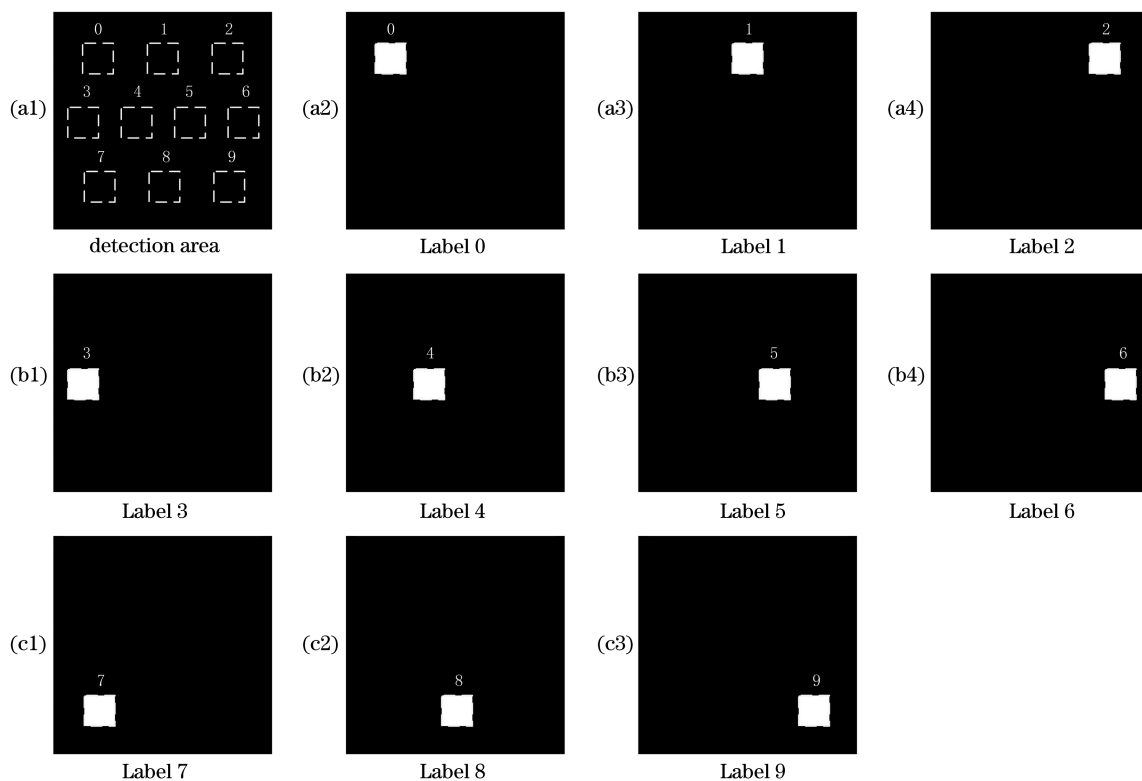


图 4 图像标签设计

Fig. 4 Image label design

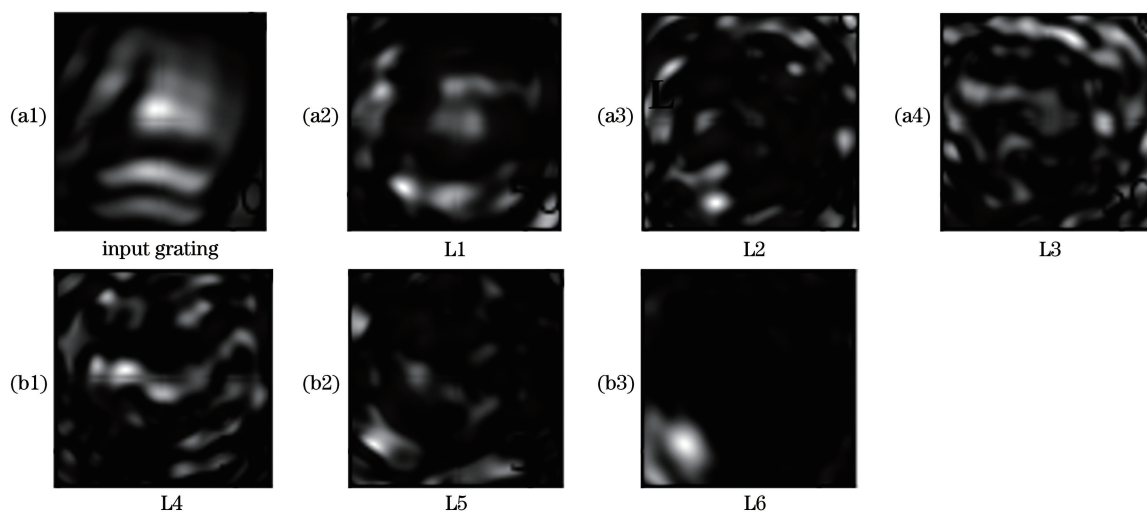


图 5 训练完成后各层光栅的输出图像

Fig. 5 Output image of each layer of grating after training

光栅的输出图像即为最终的结果图像。为了判断结果图像的正确性,首先除去探测区域对背景信息的影响,然后使用探测区域模板对结果图像进行提取,以得到预测标签。本文的输入图像以 MNIST 数据集中数字 7 为例,入射光经过输入光栅和各层光栅 L1~L6 衍射后,得到的最后一层光栅结果图像中光强最大的区域与图 4 中探测区域标签 7 的位置一致。

3 实验结果

3.1 评估方法

在仿真过程中,在 MNIST 数据集和 Fashion-MNIST 数据集中分别选取 55000 张图像和 5000 张图像作为训练集和验证集,剩下的 10000 张保留给测试集。采用混淆矩阵作为评估方法。首先,需要对混淆矩阵 $C_{i'}$ ($i'=0\sim 9$) 进行计算。对于每一个

单类, 评估由 $T_{i'}$ 、 $F_{i'}$ 、 $N_{i'}$ 和 $P_{i'}$ 定义, 分类器的正确率可以表示为

$$A = \frac{T_{i'} + N_{i'}}{T_{i'} + N_{i'} + P_{i'} + F_{i'}} \quad (9)$$

式中: $T_{i'}$ 表示混淆矩阵 $C_{i'}$ 中真实值为真且模型预测为真的数量; $N_{i'}$ 表示混淆矩阵 $C_{i'}$ 中真实值为假且模型预测为假的数量; $P_{i'}$ 表示混淆矩阵 $C_{i'}$ 中真实值为假而模型预测为真的数量; $F_{i'}$ 表示混淆矩阵 $C_{i'}$ 中真实值为真而模型预测为假的数量。

3.2 模型超参数优化结果

全光衍射深度神经网络模型中的超参数包括光栅物理参数和神经网络训练参数。入射波长、光栅像元尺寸和光栅间距属于光栅参数; 光栅层数、光栅像元规模、每批数据量、循环次数和学习率属于训练参数。其中, 确定入射波长 (λ) 为 $10.6 \mu\text{m}$, 神经元尺寸为 5 mm 。基于微米波长的全光衍射深度神经网络模型识别 MNIST 数据集的超参数选择如表 2 和表 3 所示。

表 2 MNIST 数据集中神经网络光栅的物理参数

Table 2 Physical parameters of neural network grating in MNIST dataset

Grating parameter	Value
Wavelength	$10.6 \mu\text{m}$
Cell size	$5 \mu\text{m}$
Grating spacing	70λ

使用网格搜索法对神经网络的超参数进行选择, 首先对光栅层数进行选择。在仿真状态下, 选择神经网络模型中的每批数据量为 100, 为了减少仿真时间, 循环次数为 10, 损失函数为交叉熵函数, 优化器为 Adam 优化器, 学习率选择为 0.01。由图 6 可

表 4 不同像元规模和衍射光栅间距对应的 MNIST 数据集分类正确率

Table 4 Classification accuracy of MNIST dataset corresponding to each pixel size and diffraction grating spacing

Spacing	Pixel size of 30×30	Pixel size of 40×40	Pixel size of 50×50	Pixel size of 60×60	Pixel size of 70×70
30λ	0.8427	0.8642	0.8736	0.8694	0.8664
40λ	0.8218	0.8623	0.8707	0.8744	0.8667
50λ	0.7545	0.8614	0.8594	0.8759	0.8712
60λ	0.6499	0.8327	0.8710	0.8714	0.8741
70λ	0.6190	0.8304	0.8683	0.8765	0.8696

由表 4 可知, 当神经网络模型中的衍射光栅间距一定时, 分类正确率一般随着像元规模的增大而升高; 当神经网络模型中的像元规模 $\leq 50 \times 50$ 时,

表 3 MNIST 数据集中神经网络的训练参数

Table 3 Neural network training parameters in MNIST dataset

Training parameter	Value
Number of grating layers	6
Number of neurons per layer	60×60
Batch size	100
Epoch	50
Learning rate	10^{-2}

知, 当光栅层数 ≤ 5 时, 随着光栅层数的增加, 神经网络模型的正确率增加; 当光栅层数 > 5 时, 分类正确率达到饱和。考虑到物理实现和工艺精度, 光栅层数选择 6。

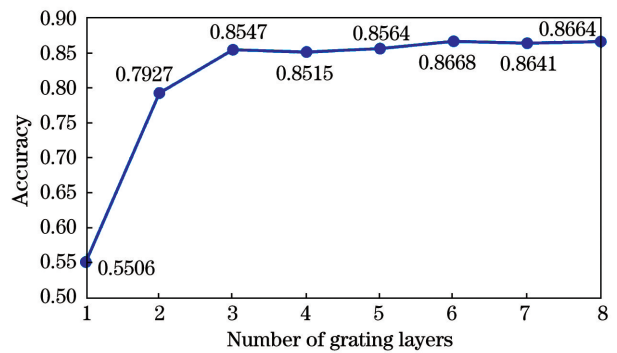


图 6 MNIST 数据集中不同光栅层数对应的分类正确率

Fig. 6 Classification accuracy corresponding to each number of grating layers in MNIST dataset

确定神经网络模型中的光栅层数后, 对模型训练参数中的像元规模和衍射光栅间距等进行优化。不同像元规模和衍射光栅间距对应的分类正确率如表 4 所示。

分类正确率随着光栅间距的增加而减小, 当像元规模 $> 50 \times 50$ 时, 分类正确率维持恒定。其中, 当像元规模为 60×60 , 衍射光栅间距为 70λ 时, 神经网络

络的分类正确率最高。

基于微米波长的全光衍射深度神经网络模型识别 Fashion-MNIST 数据集的超参数选择方式与上述方法一致,参数如表 5 和表 6 所示。

表 5 Fashion-MNIST 数据集中神经网络光栅的物理参数
Table 5 Physical parameters of neural network grating in Fashion-MNIST dataset

Grating parameter	Value
Wavelength	10.6 μm
Cell size	5 μm
Grating spacing	30 λ

表 6 Fashion-MNIST 数据集中神经网络的训练参数
Table 6 Neural network training parameters in Fashion-MNIST dataset

Training parameter	Value
Number of grating layers	6
Number of neurons per layer	70 \times 70
Batch size	100
Epoch	50
Learning rate	10 ⁻²

表 7 不同像元规模和衍射光栅间距对应的 Fashion-MNIST 数据集分类正确率

Table 7 Classification accuracy of Fashion-MNIST dataset corresponding to each pixel size and diffraction grating spacing

Spacing	Pixel size	Pixel size	Pixel size	Pixel size	Pixel size
	of 30 \times 30	of 40 \times 40	of 50 \times 50	of 60 \times 60	of 70 \times 70
30 λ	0.7012	0.7797	0.7943	0.7969	0.7994
40 λ	0.6569	0.7539	0.7882	0.7903	0.7947
50 λ	0.6137	0.7419	0.7664	0.7849	0.7937
60 λ	0.6098	0.7411	0.7574	0.7831	0.7935
70 λ	0.6069	0.7246	0.7539	0.7735	0.7809

3.3 全光衍射深度神经网络的性能评估

在基于微米波长的标准全光衍射深度神经网络模型中未加入激活函数的情况下, MNIST 数据集的分类结果如图 8 所示,仿真得到的 MNIST 数据集的分类正确率为 0.8678。

由图 8 可知,标准全光衍射深度神经网络模型对 MNIST 数据集中各个数字的分类正确率并不相同,模型对数字 1 的分类正确率高达 0.98;然而,模型对数字 8 和数字 9 的分类正确率分别为 0.73 和 0.81。其中,将数字 8 误分为数字 3 和数字 5 的概率为 0.05,误分为数字 0、4 和 7 的概率为 0.03;将数字 9 误分为数字 4 的概率为 0.07,误分为数字 7 的概率为 0.05。

在基于微米波长的标准全光衍射深度神经网络

利用网格搜索法对超参数进行优化,光栅层数的优化选择如图 7 所示,因此光栅层数选择 6。然后对模型训练参数中的像元规模和衍射光栅间距等进行优化。不同像元规模和衍射光栅间距对应的分类正确率如表 7 所示,因此选择像元规模为 70 \times 70,衍射光栅间距为 30 λ 。

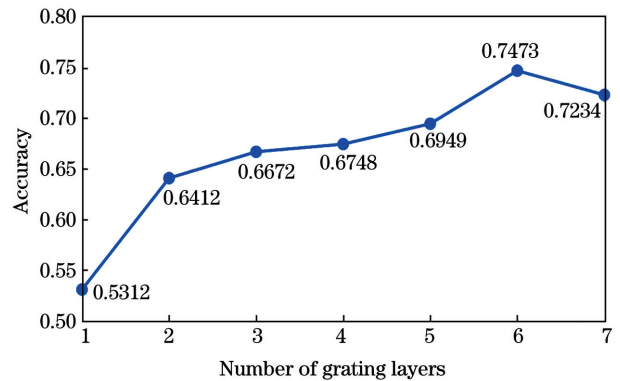


图 7 Fashion-MNIST 数据集中不同光栅层数对应的分类正确率

Fig. 7 Classification accuracy of Fashion-MNIST dataset corresponding to each number of grating layers in Fashion-MNIST dataset

模型中未加入激活函数的情况下, Fashion-MNIST 数据集的分类结果如图 9 所示,仿真得到的 Fashion-MNIST 数据集的分类正确率为 0.8110。

由图 9 可知,标准全光衍射深度神经网络模型对 Fashion-MNIST 数据集中各个数字的分类正确率并不相同,模型对标签 1 和标签 8 的分类正确率最高,分别为 0.94 和 0.95;然而,模型对标签 6 的分类正确率仅为 0.35。其中,将标签 6 误分为标签 0 的概率高达 0.23,误分为标签 2 的概率为 0.14。由此可知,未加入激活函数的标准全光衍射深度神经网络模型的非线性拟合能力和泛化能力较弱。

3.4 非线性全光衍射深度神经网络的性能评估

在传统全光衍射深度神经网络模型的基础上,选择 Leaky-ReLU、PReLU、RReLU 函数作为模型

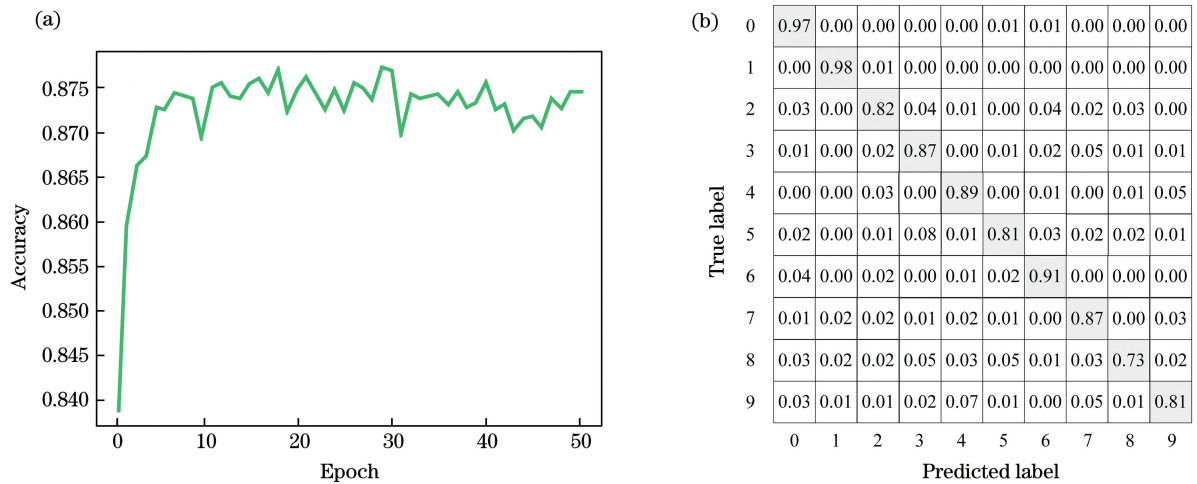


图 8 标准全光衍射深度神经网络对 MNIST 数据集的分类结果。(a)分类正确率;(b)混淆矩阵
 Fig. 8 Classification results of MNIST dataset by standard all-optical diffraction deep neural network.
 (a) Classification accuracy; (b) confusion matrix

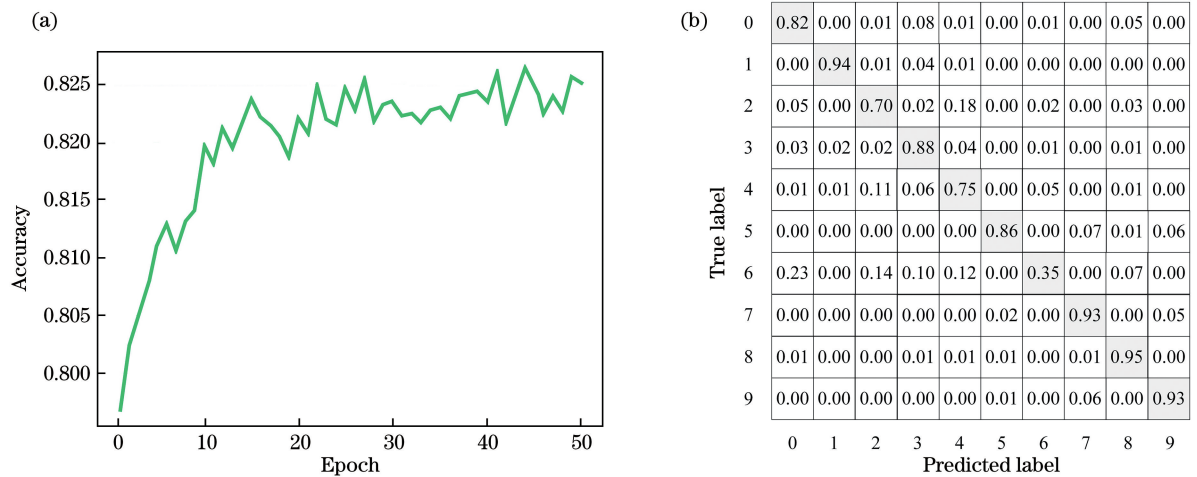


图 9 标准全光衍射深度神经网络对 Fashion-MNIST 数据集的分类结果。(a)分类正确率;(b)混淆矩阵
 Fig. 9 Classification results of Fashion-MNIST dataset by standard all-optical diffraction deep neural network.
 (a) Classification accuracy; (b) confusion matrix

的激活函数,将模型对 MNIST 数据集的分类正确率与传统全光衍射深度神经网络模型的结果进行对比。仿真得到的 MNIST 数据集的分类正确率如表 8 所示。

表 8 具有不同激活函数的非线性全光衍射深度神经网络对 MNIST 数据集的分类正确率
 Table 8 Classification accuracies of MNIST dataset by nonlinear all-optical diffraction deep neural networks with different activation functions

Activation function	Accuracy
Leaky-ReLU	0.9609
PReLU	0.9628
RReLU	0.9630

可以看出,具有 RReLU 函数的神经网络具有 0.9630 的分类正确率,与文献[15-16]的结果相比,基于微米波长的非线性全光衍射深度神经网络模型在特征尺寸缩小 1/80 的前提下,分类正确率仅相差 0.01,理论证明了模型的正确性。图 10 为具有不同激活函数的全光衍射深度神经网络对 MNIST 数据集的分类正确率和混淆矩阵图像。

由图 10 可知,具有三种激活函数的神经网络对 MNIST 数据集中每个数字的分类正确率均在 0.94 以上。其中,具有三种激活函数的模型对数字 0、1 的识别正确率均高达 0.99;然而,模型对数字 5 和数字 9 的分类能力稍差,正确率分别为 0.95 (Leaky-ReLU)、0.95 (PReLU)、0.94 (RReLU) 和 0.94 (Leaky-ReLU)、0.95 (PReLU)、0.95 (RReLU)。

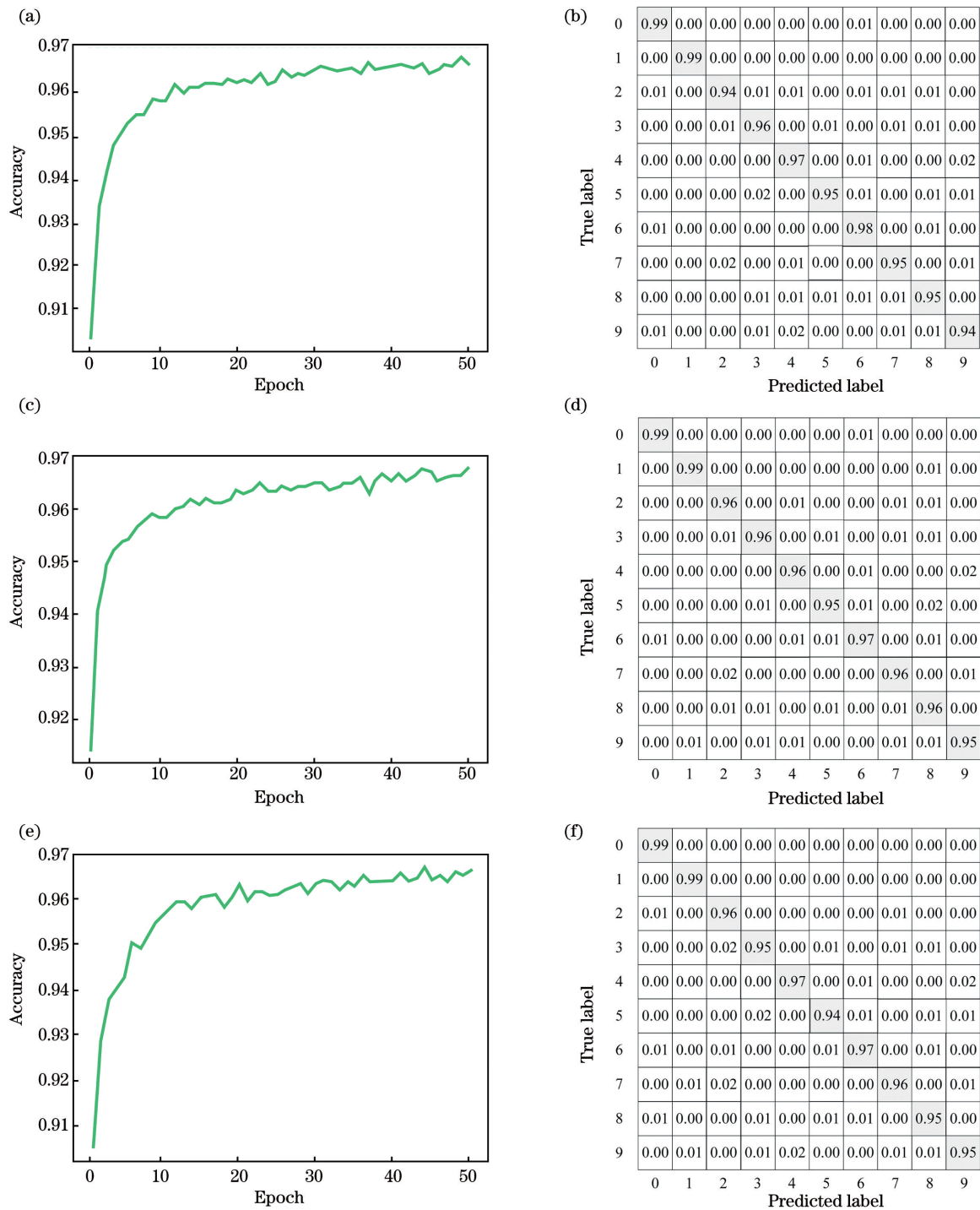


图 10 具有不同激活函数的全光衍射深度神经网络对 MNIST 数据集的分类正确率和混淆矩阵。(a)(b) Leaky-ReLU; (c)(d) PReLU; (e)(f) RReLU

Fig. 10 Classification accuracies and confusion matrixes of MNIST dataset by all-optical diffraction deep neural networks with different activation functions. (a)(b) Leaky-ReLU; (c)(d) PReLU; (e)(f) RReLU

相比于标准的全光衍射深度神经网络,分类正确率有很大的提高。例如数字 8,具有三种激活函数的神经网络的分类正确率最高提升了 0.23。图 11 为各种神经网络模型对 MNIST 数据集中各个数字的识别正确率。

在对 Fashion-MNIST 数据集进行分类时,选择与上述相同的非线性激活函数,并将模型对 Fashion-MNIST 数据集的分类正确率与传统全光衍射深度神经网络模型的结果进行对比。仿真得到的 Fashion-MNIST 数据集分类正确率如表 9 所示。

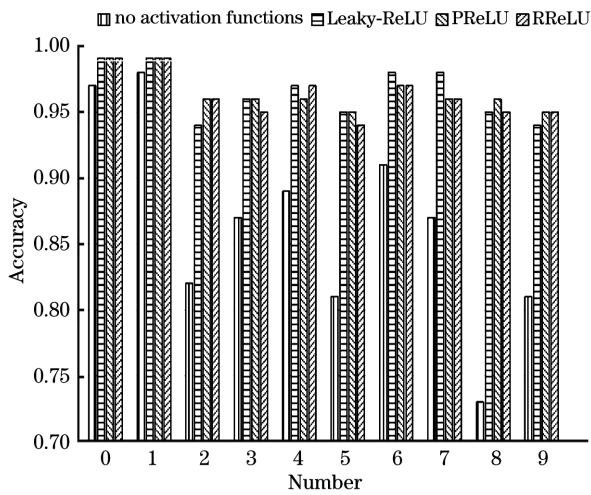


图 11 各种全光衍射深度神经网络模型对 MNIST 数据集中各个数字的识别正确率

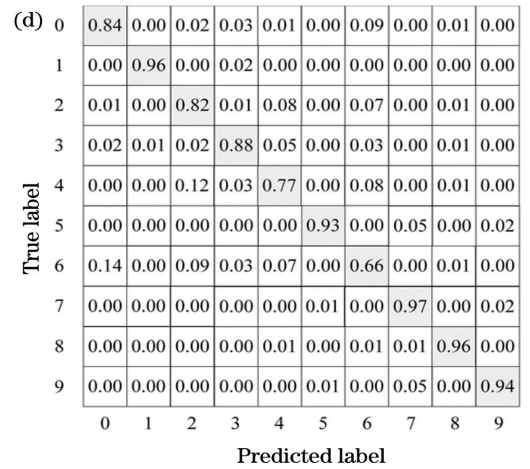
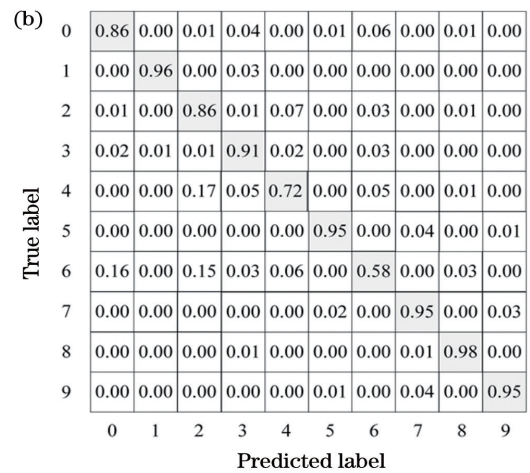
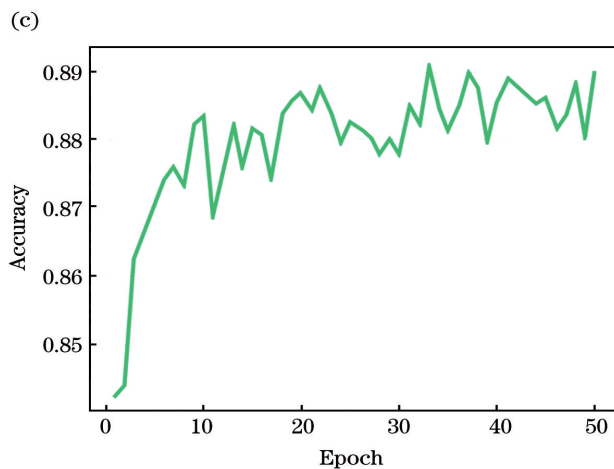
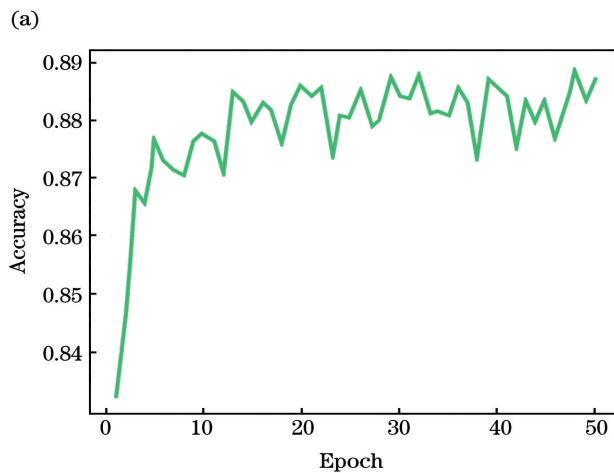
Fig. 11 Recognition accuracy of each number in MNIST dataset by each all-optical diffraction deep neural network model

表 9 具有不同激活函数的非线性全光衍射深度神经网络对 Fashion-MNIST 数据集的分类正确率

Table 9 Classification accuracies of Fashion-MNIST dataset by nonlinear all-optical diffraction deep neural networks with different activation functions

Activation function	Accuracy
Leaky-ReLU	0.8717
PReLU	0.8736
RReLU	0.8743

加入 RReLU 激活函数后,分类正确率为 0.8743。与文献[15-16]的结果相比,分类正确率仅相差 0.02,理论证明了模型的正确性。图 12 为具有不同激活函数的全光衍射深度神经网络对 Fashion-MNIST 数据集的分类正确率和混淆矩阵图像。



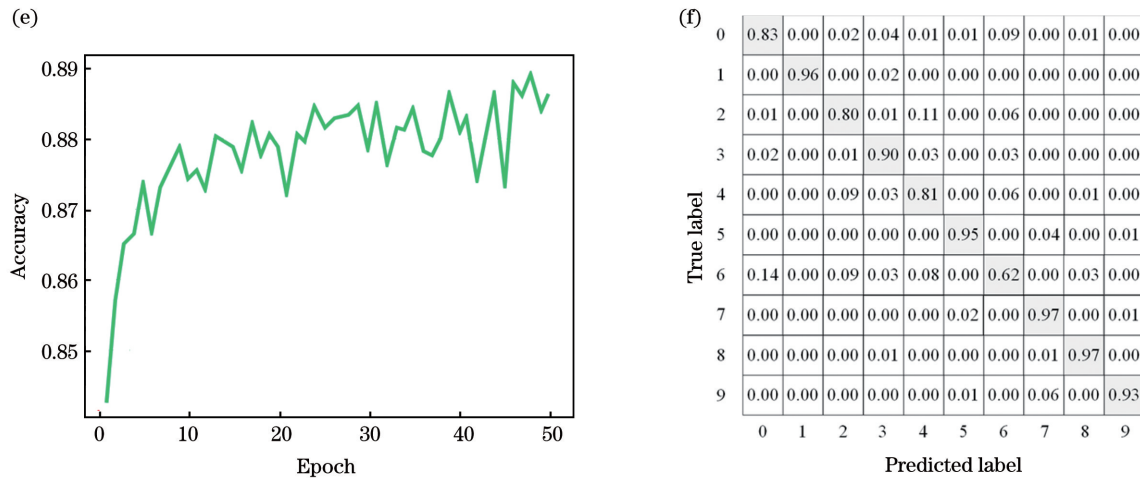


图 12 具有不同激活函数的全光衍射深度神经网络对 Fashion-MNIST 数据集的分类正确率和混淆矩阵。

(a)(b) Leaky-ReLU; (c)(d) PReLU; (e)(f) RReLU

Fig. 12 Classification accuracies and confusion matrixes of Fashion-MNIST dataset by all-optical diffraction deep neural networks with different activation functions. (a)(b) Leaky-ReLU; (c)(d) PReLU; (e)(f) RReLU

由图 12 可知,具有三种激活函数的模型对标签 7 和标签 8 的识别正确率分别为 0.95 (Leaky-ReLU)、0.97 (PReLU)、0.97 (RReLU) 和 0.98 (Leaky-ReLU)、0.96 (PReLU)、0.97 (RReLU);然而,模型对标签 6 的分类能力稍差,正确率分别为 0.58、0.66 和 0.62。相比于标准的全光衍射深度神经网络,分类正确率有很大提高。例如标签 6,具有三种激活函数的神经网络的分类正确率最高提升了 0.31。图 13 为各种神经网络模型对 MNIST 数据集中各个数字的识别正确率。

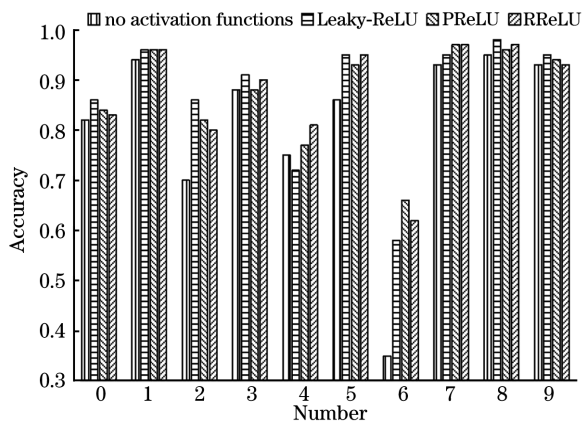


图 13 各种全光衍射深度神经网络模型对 Fashion-MNIST 数据集中各个数字的识别正确率

Fig. 13 Recognition accuracy of each number in Fashion-MNIST dataset by each all-optical diffraction deep neural network model

4 结 论

提出了一种基于微米波长光源的非线性全光衍

射深度神经网络的实现方案,通过前向传播计算神经网络输出及误差反向传播优化参数,建立了非线性全光衍射深度神经网络模型结构,针对 MNIST 数据集和 Fashion-MNIST 数据集的分类任务,对神经网络的超参数进行了优化。仿真结果表明,加入 RReLU 函数的非线性全光衍射深度神经网络对 MNIST 数据集和 Fashion-MNIST 数据集的分类正确率分别达到 0.9630 和 0.8743。在后续工作中将对神经网络模型进行进一步优化,以期为实现小型化微米波长的非线性全光衍射深度神经网络物理系统提供理论依据。

参 考 文 献

- [1] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [2] Cho K, Merriënboer B V, Gulcehre C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation [EB/OL]. (2014-06-03) [2020-09-02]. <https://arxiv.org/abs/1406.1078>.
- [3] Graves A, Mohamed A R, Hinton G. Speech recognition with deep recurrent neural networks[C]// 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, May 26-31, 2013, Vancouver, BC, Canada. New York: IEEE, 2013: 6645-6649.
- [4] Le Cun Y, Bengio Y, Hinton G. Deep learning [J]. Nature, 2015, 521(7553): 436-444.
- [5] Farhat N H, Psaltis D, Prata A, et al. Optical

- implementation of the Hopfield model [J]. *Applied Optics*, 1985, 24(10): 1469.
- [6] Paquot Y, Dupont F, Smerieri A, et al. Optoelectronic reservoir computing [J]. *Scientific Reports*, 2012, 2: 287.
- [7] Appeltant L, Soriano M C, Van der Sande G, et al. Information processing using a single dynamical node as complex system [J]. *Nature Communications*, 2011, 2: 468.
- [8] Zhou Z P, Xu P F, Dong X W. Computing on silicon photonic platform [J]. *Chinese Journal of Lasers*, 2020, 47(6): 0600001.
周治平, 许鹏飞, 董晓文. 硅基光电计算 [J]. *中国激光*, 2020, 47(6): 0600001.
- [9] Zhang B, He T, Zhong L, et al. Recent progress of terahertz wave modulator based on organic photoelectric materials [J]. *Chinese Journal of Lasers*, 2019, 46(6): 0614012.
张波, 和挺, 钟良, 等. 基于有机光电材料的太赫兹波调制器件研究进展 [J]. *中国激光*, 2019, 46(6): 0614012.
- [10] Zhang Y N, Xu A S, Feng Y L, et al. Chaos characteristics of the output from a semiconductor laser subject to optoelectronic feedback [J]. *Acta Optica Sinica*, 2020, 40(12): 1214001.
张依宁, 徐艾诗, 冯玉玲, 等. 光电反馈半导体激光器输出光的混沌特性 [J]. *光学学报*, 2020, 40(12): 1214001.
- [11] Tait A N, de Lima T F, Zhou E, et al. Neuromorphic photonic networks using silicon photonic weight banks [J]. *Scientific Reports*, 2017, 7(1): 7430.
- [12] Tait A N, Nahmias M A, Shastri B J, et al. Broadcast and weight: an integrated network for scalable photonic spike processing [J]. *Journal of Lightwave Technology*, 2014, 32(21): 4029-4041.
- [13] Shen Y C, Harris N C, Skirlo S, et al. Deep learning with coherent nanophotonic circuits [J]. *Nature Photonics*, 2017, 11(7): 441-446.
- [14] Chen H W, Yu Z M, Zhang T, et al. Advances and challenges of optical neural networks [J]. *Chinese Journal of Lasers*, 2020, 47(5): 0500004.
陈宏伟, 于振明, 张天, 等. 光子神经网络发展与挑战 [J]. *中国激光*, 2020, 47(5): 0500004.
- [15] 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.
- [16] Mengu 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): 1-14.
- [17] 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: 112.
- [18] Lu L D, Zhu L Q, Zhang Q K, et al. Miniaturized diffraction grating design and processing for deep neural network [J]. *IEEE Photonics Technology Letters*, 2019, 31(24): 1952-1955.
- [19] Niu H S, Yu M X, Zhu B F, et al. Design and implementation of diffraction grating based on 10.6 μm all-optical depth neural network [J]. *Journal of Infrared and Millimeter Waves*, 2020, 39(1): 13-18.
牛海莎, 于明鑫, 祝博飞, 等. 基于 10.6 微米全光深度神经网络衍射光栅的设计与实现 [J]. *红外与毫米波学报*, 2020, 39(1): 13-18.
- [20] Bianchi V, Carey T, Viti L, et al. Terahertz saturable absorbers from liquid phase exfoliation of graphite [J]. *Nature Communications*, 2017, 8: 15763.
- [21] Goodman J W. *Introduction to Fourier optics* [M]. New York: McGraw-Hill Companies, 2005.
- [22] Nair V, Hinton G E. Rectified linear units improve restricted Boltzmann machines [EB/OL]. (2010-06-05) [2020-09-02]. <https://dl.acm.org/doi/10.5555/3104322.3104425>.
- [23] Maas A L, Hannun A Y, Ng A Y. Rectifier nonlinearities improve neural network acoustic models [EB/OL]. [2020-09-02]. http://ai.stanford.edu/~amaas/papers/relu_hybrid_icml2013_final.pdf.
- [24] He K M, Zhang X Y, Ren S Q, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification [C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 1026-1034.