面向物理的衍射光学神经网络粗糙度优化

Shanglin Zhou*[1], Yingjie Li*[2], Minhan Lou[2],Weilu Gao[2],ZhijieShi[1],Cunxi Yu[1],Cai wen Ding[11]康涅狄格大学， [2]犹他大学 *两位作者对本研究的贡献相同

arXiv:2304.01500v1 [cs.LG] 4 Apr 2023

摘要——**作为一种代表性的、超越互补金属氧化物半导体（CMOS）的下一代器件/电路技术，衍射光学神经网络（DONNs）凭借极快的计算速度（光速）和低能耗，相较于传统深度神经网络展现出显著优势。然而，由于衍射层内存在像素间相互作用，DONN 数值建模与物理光学器件的实际部署之间存在**不匹配，**即预测准确率显著下降。本文提出一种物理感知的衍射光学神经网络训练框架，以缩小数值建模与实际部署之间的性能差距。具体而言，我们在训练过程中提出粗糙度建模正则化，并结合物理感知稀疏化方法，在相位掩模中引入稀疏性，以减小衍射层中相邻像素之间的相位变化幅度。我们还进一步开发了 2π 周期性优化方法，以降低相位掩模的粗糙度，同时保持 DONN 的性能。实验结果表明，与当前先进方法相比，我们的物理感知优化方法在 MNIST、FMNIST、KMNIST 和 EMNIST 上分别降低了**35.7%、34.2%、28.1%**和**27.3%**的粗糙度，且仅造成准确率损失。**

索引词—**衍射光学神经网络、权重稀疏化、粗糙度建模**

I. 引言

深度神经网络（DNN）的高计算量和高存储需求给传统冯·诺依曼架构带来了巨大挑战 [1]–[3]，导致存储层次结构中产生大量数据传输。对超高效深度神经网络加速器的需求推动了对多种冯·诺依曼架构的研究。我们迫切需要：（i）超越互补金属氧化物半导体（CMOS）的新一代器件/电路技术；（ii）相应的定制化算法与技术协同设计，以在各种应用中实现超低功耗、实时的深度神经网络处理。

作为一种代表性方案，衍射光学神经网络（DONNs）备受关注。DONNs 基于物理现象实现全光处理，即光衍射和光信号相位调制，这些现象天然以光速发生。权重以复数透射系数的形式编码在衍射层中，并采用自由空间传播函数 [4]与光波函数相乘以实现计算。多层衍射表面包含数百万个神经元，并在物理上构成神经网络，模拟人工神经网络的形态；与在数字平台上运行传统深度神经网络相比，其能耗显著更低。因此，我们能够实现较高的系统吞吐量和计算速度；衍射光学神经网络系统中的相位掩模一经制造或组装，便无需额外能耗来维持计算单元（相位掩模中的像素）的功能 [5]–[9]。

当前用于全光推理的衍射光学神经网络系统是在数字平台上通过对衍射光学神经网络系统进行数值建模来训练的。然而，数值建模与物理部署之间存在不匹配，即预测精度显著下降。这种不匹配的潜在原因包括：光学器件中的离散控制级别、

表 I：不同方法的比较 .

方法	粗糙度感知	稀疏性	2π 周期性优化
[5], [16]	×	×	×
[6], [8]	×	×	✓
Ours	✓	✓	✓

器件的制造误差 [6], 、不准确的仿真核 [9]–[11], 等。在这些因素中，我们认为相位掩模内的像素间串扰 [12] 最为关键 [13]–[15]，因为它会破坏数值建模所仿真的计算光学响应：相邻像素之间的突变会引入快速变化的入射场。因此，可以用相邻像素厚度差来量化像素间串扰的影响，也就是粗糙度。

为缩小数字仿真与硬件部署之间的准确率差距，本文提出了一种适用于衍射光学神经网络系统的物理感知粗糙度优化流程。具体而言，我们提出的物理感知优化流程包括三个步骤（表 I）：**(1)** 在训练过程中将粗糙度正则化纳入衍射光学神经网络损失函数。详情见第 III-B 节，其中在训练损失中引入正则化因子，用于进行粗糙度感知的衍射光学神经网络训练；**(2)** 通过块稀疏化压缩模型，在活跃像素之间留出更多空间，以减弱像素间相互作用（第 III-C 节）；**(3)** 进一步利用 2π 周期性相位调制平滑掩模。我们的工作优于现有方法 [5], [6], [8] ，原因有二。首先， [5] 未对衍射光学神经网络中的像素间串扰建模，因此没有考虑粗糙度。相比之下，我们将像素间串扰的影响纳入衍射光学神经网络损失函数，从而使优化具备准确率感知能力 [13]。其次，以往研究 [6], [8]利用相位调制的周期性特征，在现有硬件上部署负相位参数；也就是说，通过添加 2π 或其整数倍，相位调制在物理上等效。然而，这种周期性特征尚未被用于减小像素间串扰的影响。

我们的贡献总结如下：

- 我们提出了一种物理感知粗糙度优化框架，将粗糙度建模和权重稀疏化整合到衍射光学神经网络的训练过程中，以有效缩小数字仿真的衍射光学神经网络与物理硬件部署之间的性能差异。
- 我们提出了一种粗糙度建模方法，可在衍射光学神经网络的训练过程中平滑相位掩模，并提供一种量化性能差距的方法。
- 我们将物理感知块稀疏化方法融入优化过程，以提高衍射光学神经网络系统的稀疏性并减少像素间相互作用。
- 我们应用 2π 周期性相位调制，进一步平滑

相位掩模，并减小性能差异。

评估结果表明，我们提出的物理感知优化框架最多可将粗糙度降低 35.7%。据我们所知，这是首项解决 DONNs 数字仿真与硬件部署之间性能不匹配问题的工作。

II. 背景与相关工作

A. 衍射光学神经网络（*DONNs*）

衍射光学神经网络系统具有两方面的低碳足迹优势：（1）信息编码光信号的传播是物理过程的自然属性，即数据传输无需额外成本；（2）衍射光学神经网络系统通常由无源器件构成，例如 3D打印相位掩模（如图 1(d) 所示）、未发生损坏的液晶阵列和相变材料。这意味着器件一经制造并组装完成，便能以零额外能耗对光信号进行相位调制，也就是说，计算同样无需额外成本。

衍射光学神经网络包含两组参数。第一组（可训练）用于控制衍射层中各衍射像素提供的相位调制。入射光波的相位不同，会导致系统末端的光强分布不同。第二组（不可训练）用于衍射近似，用于描述光波在各衍射层发生衍射时的传播过程，并连接各层之间的神经元。这组参数通过对现实世界物理现象进行数学近似而计算得出 [8]。衍射光学神经网络系统的最终输出（例如分类任务中的预测标签）可表示为探测器观测到的最后一层输出光强中的最大能量值。因此，与传统深度神经网络的训练过程类似，可通过最小化常用的机器学习损失函数 [5]，获得衍射光学神经网络中衍射层相位调制的最优权重。

B. 衍射光学神经网络的硬件部署挑战

将数字训练得到的衍射光学神经网络模型部署到实际光学器件上，可能会导致性能显著下降 [6], ，尤其是在采用亚波长结构时 [17], [18]。相位掩模（衍射层）中的像素间相互作用会使入射场快速变化，从而破坏计算得到的光学响应，导致衍射光学神经网络系统的数值建模与硬件部署之间出现显著的失相关差距。例如，Zhou等人 [6] 声称 $\geq$ 30% 的准确率下降，这是在未考虑掩模粗糙度的情况下，直接将模型部署到物理光学系统所导致的。

C. 权重稀疏化

已有研究采用权重稀疏化来减小模型大小并加速计算 [19]–[21]。然而，本文旨在利用稀疏化技术优化权重分布 [22] ，并强制衍射光学神经网络中相位掩模的平滑度。

非结构化幅值权重稀疏化 [23] 已被证明不仅能够使模型具有较高的稀疏性，还能保持准确率。然而，经过非结构化稀疏化 [24], [25]后，稀疏权重矩阵中的非零元素呈随机分布。这会增加电压控制模式的设计工作量，因为在将衍射光学神经网络系统部署到实体光学器件时，衍射层上的权重（相位调制）分布不规则。结构化稀疏化，例如分组平衡稀疏化 [26], [27] 和块循环矩阵稀疏化 [28] 的发展更加注重稀疏化模式，使权重矩阵中的非零元素分布更加规则。然而，这些方法仍然关注逐元素模式，且未优化权重的分布和粗糙度，因此无法提升神经网络权重矩阵的平滑度。

III. 物理感知粗糙度优化

我们从两个方面缩小 DONN 数值建模与实际部署在像素间相互作用方面的差异：**(1)** 平滑相位掩模，即减小邻近像素之间剧烈的相位变化。降低掩模的粗糙度后，入射场的变化会更平滑，从而与训练期间计算得到的光学响应保持一致。**(2)** 减少起作用的衍射像素。我们将衍射层中重要性较低的像素置为精确的零，使相位掩模更加稀疏，并为剩余像素之间留出更多空间，从而减少像素间相互作用。整体示意图见图 1(b–c)。

A. *DONNs* 的可微建模

衍射光学神经网络系统由三个部分组成（图 1(a)）：(1) 用于对输入图像进行编码的激光光源，(2) 用于提供可训练相位调制的衍射层，以及 (3) 用于捕获正向传播衍射图样的探测器。具体而言，首先使用相干激光对输入图像进行编码。携带编码信息的光信号在衍射层之间的自由空间中发生衍射，并在每一层通过相位调制进行调制。最后，在探测器平面捕获光传播后形成的、对应于光强分布的衍射图样，以用于预测。

输入（例如图像）编码在激光光源发出的相干光信号上，其波函数可表示为 $f_0(x_0,\ y_0)$。光信号经过距离为 z 的光衍射到达第一衍射层后，其波函数是输入平面上各输出的求和，即

$$f_1(x,y)=\iint f_0(x_0,y_0)h(x-x_0,y-y_0,z)dx_0dy_0 \qquad (1)$$

其中，$(x,\ y)$ 是接收平面上的坐标。 h 是自由空间的脉冲光学响应函数，它是光衍射的数学近似，例如瑞利–索末菲近似、菲涅耳近似、夫琅禾费近似 [29]。

方程 1 可通过谱算法计算，其中我们采用快速傅里叶变换（FFT）进行快速且可微的计算，即 $U_1(\alpha,\beta)=U_0(\alpha,\beta)H(\alpha,\beta,z)$，其中 U 和 H 分别是 f 和 h 的傅里叶变换。波函数$U_1(\alpha,\beta)$ 通过快速傅里叶逆变换（iFFT）转换到时域，以进行相位调制：$f_2(x,y)=\text{iFFT}(U_1(\alpha,\beta))\times \mathrm{W}_1(x,y)$，其中 $\mathrm{W}_1(x,y)$ 是第一衍射层中的相位调制。$f_2(x,y)$ 是第二衍射层光衍射对应的波函数。

我们将一个衍射层上的一轮光衍射和相位调制封装为一个计算模块，称为 **DiffMod**，即 $\text{DiffMod}(f(x,y),\mathbf{W})=L(f(x,y)\ z)\times\mathbf{W}(x,y)$，其中 $f(x,y)$ 是输入波函数。$\mathbf{W}(x,y)$ 是相位调制。$L(f(x,y)\ z)$ 是在时域内经过距离 z 的光衍射后的波函数。对于一个三层衍射光学神经网络系统，其前向函数可表示为：

$$\begin{aligned} I(f_0(x,y),\mathbf{W})=\text{DiffMod}(\text{DiffMod}(\text{DiffMod}(f_0(x,y),\\ \mathbf{W}_1(x,y)),\mathbf{W}_2(x,y)),\mathbf{W}_3(x,y)) \end{aligned} \qquad (2)$$

公式 2 中关于光强 I 的最终衍射图样在探测器平面上采集。预先定义了探测器

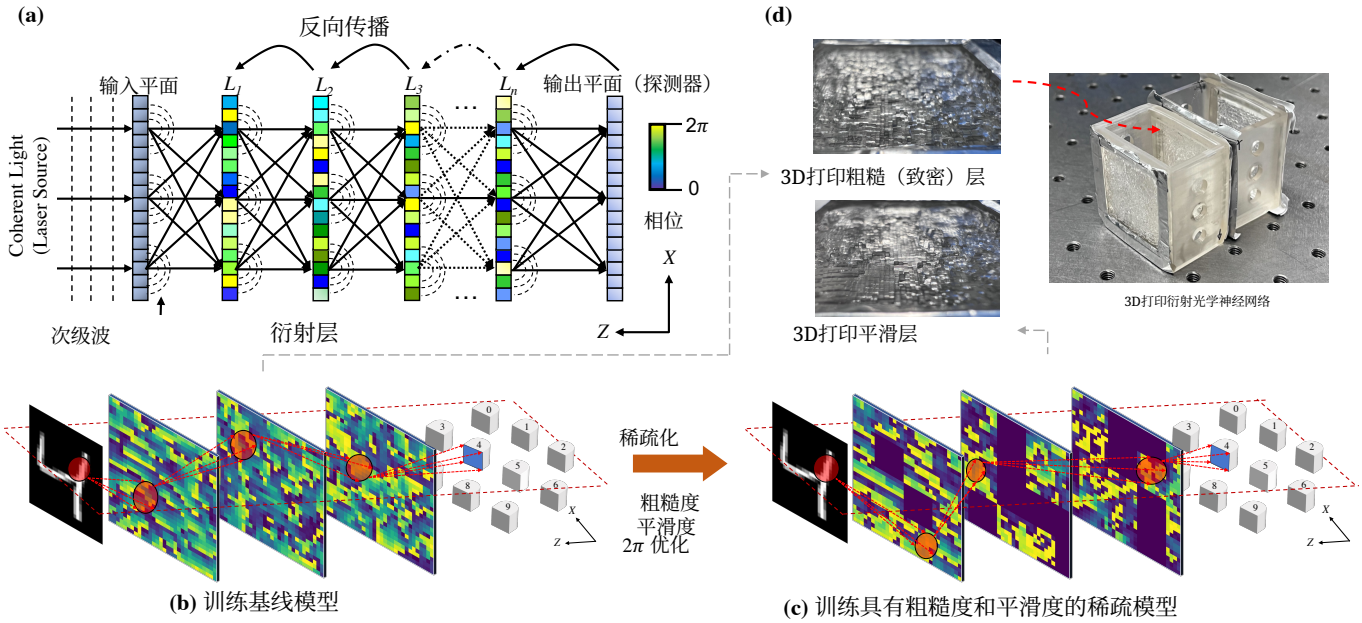


图 1：所提出的物理感知粗糙度优化流程概览。(a) 衍射光学神经网络系统示意图，包括输入平面、多个衍射层和探测器（输出）平面。(b) 基线模型训练过程示意图。(c) 粗糙度感知稀疏化过程示意图。三个衍射层中的黑色块均已稀疏化，其中所有数值均为零。(d) 3D 打印的密集层、平滑层和衍射光学神经网络的图像。

区域，以模拟传统神经网络的类别预测输出。通过最大值索引，选取对应探测器区域内光强总和最高的类别作为预测结果 [5], [30]。设真实类别为 t，使用 **均方误差损失** 的损失函数 ℓ 为 $\ell = \| \text{Softmax}(I) - t \|_2$。因此，整个系统旨在实现可微分，并与传统自动微分引擎兼容。

B. 粗糙度建模

我们提出了一种粗糙度建模方法，可独立应用于每个衍射层，并可在训练过程中进行微分，旨在用于衍射光学神经网络模型的粗糙度感知训练。如图 2 所示，像素的粗糙度根据其邻近像素的值计算。为计算边界像素的粗糙度，我们对原始相位掩模进行一维零填充。我们取该像素的值与其邻居像素值之间的 L_2-范数差的平均值，作为该像素的粗糙度。例如，图 2 中像素 p_{11} 的 k-邻域粗糙度计算如下：

$$R(p_{11}) = \frac{1}{k} \times \Sigma_{ij} \| p_{ij} - p_{11} \|_2, p_{ij} \in \text{neighbor}(p_{11}) \quad (3)$$

其中 k 可以是 4 或 8。对于四邻域（图 2a），$\text{neighbor}(p_{11}) = \{p_{01}, p_{10}, p_{12}, p_{21}\}$；而对于 8-邻域粗糙度（图 b），$\text{neighbor}(p_{11}) = \{p_{00}, p_{01}, p_{02}, p_{10}, p_{12}, p_{20}, p_{21}, p_{22}\}$。粗糙度-



(a) 四邻域。　　(b) 8 邻域。

图 2：尺寸为 3×3 – 的相位掩模粗糙度建模示意图：（a）采用四邻域，（b）采用 8 邻域。

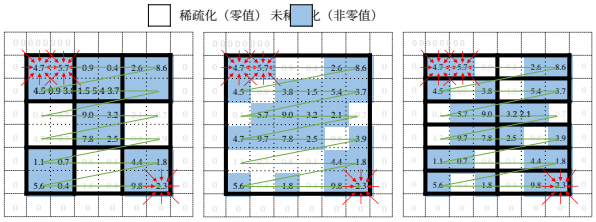


图 3：块稀疏化与另外两种稀疏化方法的比较。三者的稀疏化率均为 0.33。

整个相位掩模的粗糙度可通过对每个像素的粗糙度求和计算，即：

$$R(\mathbf{W}) = \Sigma_{ij} R(p_{ij}), i \in [0, N-1], j \in [0, N-1] \quad (4)$$

其中，相位掩模 $\mathbf{W}$ 的尺寸为 $N \times N$。我们将所有像素的粗糙度求和，作为衍射层的粗糙度分数。我们进一步将粗糙度矩阵纳入训练，即：

$$L = \| \text{Softmax}(I) - t \|_2 + p \times R(\mathbf{W}) \quad (5)$$

其中，p 是正则化因子，$R(\mathbf{W})$ 是各衍射层的粗糙度结果。

C. 物理感知稀疏化：块稀疏化

1）粗糙度感知稀疏化：为考虑不同稀疏化方法（即存储体平衡稀疏性、非规则稀疏性和块稀疏性）的情况，图 3 展示了在相同稀疏率（即 0.33）下粗糙度的影响。粗糙度根据含 8 邻域的公式 4 计算。对于块稀疏化（图 3(a)），我们将权重矩阵划分为大小相同的块。根据预先设定的阈值或百分位数，将 L_2-范数小于该阈值或百分位数的块内所有权重设为零。对于非结构化稀疏化（图 3 (b)），绝对值小于预先设定的阈值或百分位数的权重将被稀疏化 [23]。对于分组平衡稀疏化

（图 3(c)），权重矩阵的行被划分为大小相同的分组，并在稀疏化 [27]时保持各分组中的稀疏率相同。在所有方法中，块稀疏化更侧重于权重的整体块，而非逐元素模式，因此其粗糙度最低。

2) SLR 正则化优化：我们将块权重稀疏化方法 [31] 与基于代理拉格朗日松弛（SLR）的模型压缩优化技术 [32] 相结合，以减少有效衍射像素的数量，并进一步提升衍射光学神经网络中的相位掩模平滑度（图 1(c)）。对于具有 N 层的衍射光学神经网络，其中 $i \in 1, 2, \ldots, N$，将每个衍射层的权重记为 $\mathbf{W}_i$。块稀疏化的目标是在最小化衍射光学神经网络损失函数（公式 5）的同时，减少每个 $\mathbf{W}_i$ 中非零权重块的数量，即 $\min_{\mathbf{W}_i}\{\ell(\mathbf{W}_i) + \ell_r(\mathbf{W}_i)\}$，满足以下约束：$\mathbf{W}_i$ 中非零块行的数量小于 r_i，且 $\mathbf{W}_i$ 中非零块列的数量小于 c_i；其中，r_i 和 c_i 分别是期望保留的非零块行数和块列数，且 $i = 1, \ldots, N$。无约束形式可写为公式 6。

$$\min_{\mathbf{W}_i} \ell(\mathbf{W}_i) + \ell_r(\mathbf{W}_i) + \sum_{i=1}^{N} g_i(\mathbf{W}_i) \quad \text{where } g_i(\mathbf{W}_i) = \begin{cases} 0 & \text{if ① and ② are satisfied} \\ +\infty & \text{otherwise} \end{cases} \tag{6}$$

$\ell(\cdot) + \ell_r(\cdot)$ 表示带有粗糙度正则化的衍射光学神经网络系统损失。$g_i(\cdot)$ 是示性函数，用于表示每个衍射层的不可微惩罚项。为求解该问题，引入副本变量 $\mathbf{Z}_i$ [32], [33]。该损失函数等价于：$\min_{\mathbf{W}_i}\{\ell(\mathbf{W}_i)+\ell_r(\mathbf{W}_i)+\sum_{i=1}^{N} g_i(\mathbf{Z}_i)\}$，$s.t. \mathbf{W}_i = \mathbf{Z}_i$，$i = 1, \ldots, N$。由此得到的增广拉格朗日函数为

$$\mathcal{L}(\mathbf{W}_i, \mathbf{Z}_i, \mathbf{\Lambda}_i) = \ell(\mathbf{W}_i) + \ell_r(\mathbf{W}_i) + \sum_{i=1}^{N} g_i(\mathbf{Z}_i) + \sum_{i=1}^{N} \mathrm{tr}\left[\mathbf{\Lambda}_i^T(\mathbf{W}_i - \mathbf{Z}_i)\right] + \sum_{i=1}^{N} \frac{\rho}{2} \|\mathbf{W}_i - \mathbf{Z}_i\|_F^2 \tag{7}$$

H 此处，L_i 是用于放松约束的拉格朗日乘子；对约束违反施加系数为 ρ 的二次惩罚项。$\|\cdot\|_F$ 是弗罗贝尼乌斯范数。$tr(.)$ 表示迹。我们可以将上述问题分解为两个子问题，并迭代求解直至收敛。在第 k 次迭代中，第一个子问题为：给定上一轮迭代得到的$\mathbf{\Lambda}_i^k$和 Z_i，求解损失函数 $\min_{\mathbf{W}_i} \mathcal{L}\left(\mathrm{W}_i, \mathrm{Z}_i^{k-1}, \mathbf{\Lambda}_i^k\right)$。此时，我们检查代理最优性条件 $\mathcal{L}\left(\mathrm{W}_i^k, \mathrm{Z}_i^{k-1}, \mathbf{\Lambda}_i^k\right) < \mathcal{L}\left(\mathrm{W}_i^{k-1}, \mathrm{Z}_i^{k-1}, \mathbf{\Lambda}_i^k\right)$。若满足条件，则将 $\mathbf{\Lambda}_i^k$更新为 $\mathbf{\Lambda}'^k_i = \mathbf{\Lambda}_i^k + s'^k\left(\mathrm{W}_i^k - \mathrm{Z}_i^{k-1}\right)$。第二个子问题为：固定第一个子问题得到的 W_i^k，求解损失函数 $\min_{\mathbf{Z}_i} \mathcal{L}\left(\mathrm{W}_i^k, \mathrm{Z}_i, \mathbf{\Lambda}'^k_i\right)$。类似地，我们需要

用于检查代理最优性条件

$$\mathcal{L}\left(\quad \right) < \mathbf{W}, \ \mathbf{Z}, \ \mathbf{W}, \ \mathcal{L}\left(\quad \right) \qquad \mathbf{\Lambda}_i^{k+1} = \mathbf{\Lambda}'^k_i + s^k\left(\quad \right)$$

$\mathbf{Z}$，以更新 $\mathbf{WZ}$。s'^k和 s^k 均为步长 [32]。

D. 平滑度

1) 块内平滑性：粗糙度建模会在训练过程中提升衍射层中整个权重矩阵的平滑度，而块稀疏化则将权重矩阵划分为多个块，其中稀疏化块中的所有权重均设为零，未稀疏化块中的权重则呈不规则分布。为减少未稀疏化块中的剧烈相位变化，我们采用块内平滑性，并以其可在训练过程中进行微分为前提。如图 4 所示，衍射层的权重矩阵被划分为九个 2×2 块。计算每个块的方差并求和

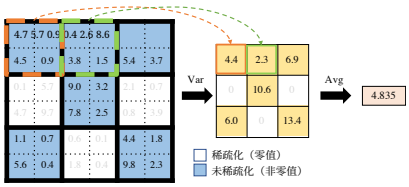


图 4：稀疏化率为 0.33、块大小为 2 时稀疏化相位掩模的内部平滑度示意图。

。所得总和作为粗糙度项加入衍射光学神经网络系统的损失函数中，如公式 8 所示。

$$L = \| \mathrm{Softmax}(I) - t \|_2 + p \times R(\mathbf{W}) + q \times R_{intra}(\mathbf{W}) \tag{8}$$

其中，q 为正则化因子，$R_{intra}(\mathbf{W})$ 是每个衍射层中各块方差之和。在这种情况下，$\ell_r(\mathbf{W}) = p \times R(\mathbf{W}) + q \times R_{intra}(\mathbf{W})$。

2) 使用 2π 进行平滑：由于稀疏化块中的像素均为零，而稀疏化块周围的其他像素可能具有较高的正值，因此块稀疏化像素及其周围像素的粗糙度会更大，如表 II–V 和图 5 所示。因此，我们需要通过后处理来恢复平滑度。我们观察到，光信号的相位调制特征具有 2π 周期，即对于训练后的相位调制 $c \in [0, 2\pi]$，有 $f(c + 2\pi) = f(c)$，其中 f 是衍射光学神经网络系统的前向函数。这使我们能够通过调整衍射光学神经网络模型中具有 2π 的衍射层像素来平滑相位掩模。得益于这一特性，我们无需重新训练即可优化相位掩模的粗糙度，只需有选择地为每个像素添加 2π，因为模型推理的性能不会发生变化。

具体而言，我们将为衍射层中的每个像素添加 2π 的选择建模为一个组合优化（CO）问题，并使用 Gumbel-Softmax（GS） [34]，通过梯度下降算法求解。在前向过程中，对于尺寸为 200×200的相位掩模，我们使用尺寸为 $200 \times 200 \times 2$ 的选择掩模，为相位掩模中的每个像素选择添加相位 0 或 2π；每个像素处的选择均以 one-hot 形式表示。因此，给定一个 $[[0], [2\pi]]$矩阵，通过将 $200 \times 200 \times 2$ 选择掩模与尺寸为 2×1的向量 $[[0], [2\pi]]$ 进行矩阵乘法，得到的附加相位掩模为 $200 \times 200 \times 1$，其中附加相位为 0 或 2π。损失函数是使用附加相位和原始相位掩模计算得到的粗糙度，具体如第 III-B 节所述。

在反向传播过程中，我们最小化粗糙度损失函数。GS算法通过对离散样本进行连续近似，使离散的独热选择变得可微 [34]。具体而言，该算法从损失函数出发，用连续概率选择近似独热选择，并利用反向传播梯度更新各选择选项的概率。因此，在下一次迭代中，损失函数会根据新优化的选择概率进行更新。

IV. 实验

A. 系统参数和训练设置

1) 衍射光学神经网络系统参数：我们设计的系统包含三个衍射层，层的尺寸以及总共十个探测器区域的尺寸为 200×200。像素尺寸为 $36\ \mu m$，因此每个制成的衍射层的尺寸为 $720\ \mu m \times 720\ \mu m$。为适配光学系统，我们将 MNIST、FMNIST、EMNIST 和 KMNIST 数据集中的原始输入图像插值至 28×28 200×200，并使用波长为 $532\ nm$ 的激光光源进行编码。层与层之间、第一层与光源之间，以及最后一层与探测器之间的物理距离均设为 $27.94\ cm$。

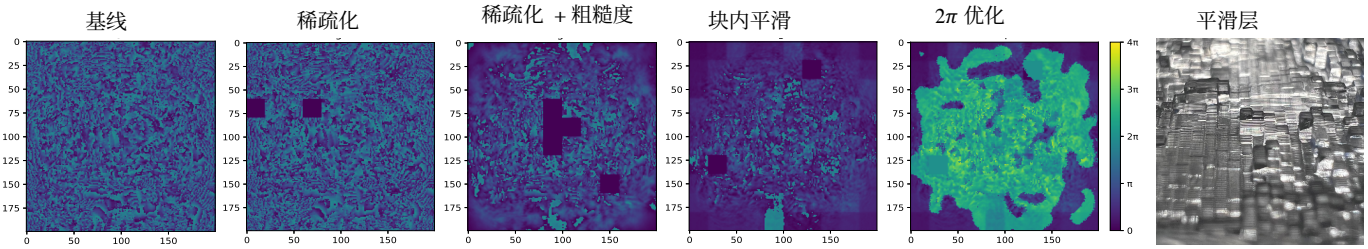


图 5：EMNIST 数据集下第二衍射层相位掩模的对比。黑色块表示其中的权重已稀疏化并设为零。第五个是经过 2π 优化的相位掩模，训练时采用了稀疏化、粗糙度和块内平滑性。最后一个是 3D 打印的平滑衍射层。

表 II：MNIST 结果。基线经过 50 个训练轮次。稀疏化模型采用大小为 25 的块进行训练。

模型	准确率 (%)	$R_{overall}$ 之前 2π 优化	$R_{overall}$ 之后 2π 优化
[5], [6], [8]	96.67	466.39	460.85
Ours-A	96.18	416.07	–
Ours-B	96.38	538.78	400.38
Ours-C	96.47	409.41	299.87
Ours-D	95.90	375.35	**280.32**

表 III：FMNIST 结果。基线模型训练了 150 个训练轮次。稀疏化模型采用块大小为 20 进行训练。

模型	准确率 (%)	$R_{overall}$ 之前 2π 优化	$R_{overall}$ 之后 2π 优化
[5], [6], [8]	87.98	464.78	461.98
Ours-A	86.99	421.49	–
Ours-B	87.88	488.11	438.53
Ours-C	86.79	350.67	305.86
Ours-D	85.76	450.73	**229.70**

表 IV：KMNIST 结果。基线模型训练 100 轮。稀疏化模型采用块大小为 20 进行训练。

模型	准确率 (%)	$R_{overall}$ 之前 2π 优化	$R_{overall}$ 之后 2π 优化
[5], [6], [8]	86.92	460.61	445.57
Ours-A	85.26	462.7	–
Ours-B	86.83	473.08	432.26
Ours-C	85.01	396.84	331.22
Ours-D	83.19	327.48	**288.42**

表 V：EMNIST 结果。基线训练了 100 个训练轮次。稀疏化模型的训练块大小为 20。

模型	准确率 (%)	$R_{overall}$ 之前 2π 优化	$R_{overall}$ 之后 2π 优化
[5], [6], [8]	92.30	463.42	458.48
Ours-A	91.61	435.58	–
Ours-B	92.36	465.85	443.91
Ours-C	91.16	349.61	336.75
Ours-D	90.74	312.17	**298.09**

不同类别对应的独立探测器区域均匀分布在探测器平面上，尺寸为 20×20；这些区域的强度总和构成一个 float32 类型的向量。最终预测结果将通过最大值索引生成。

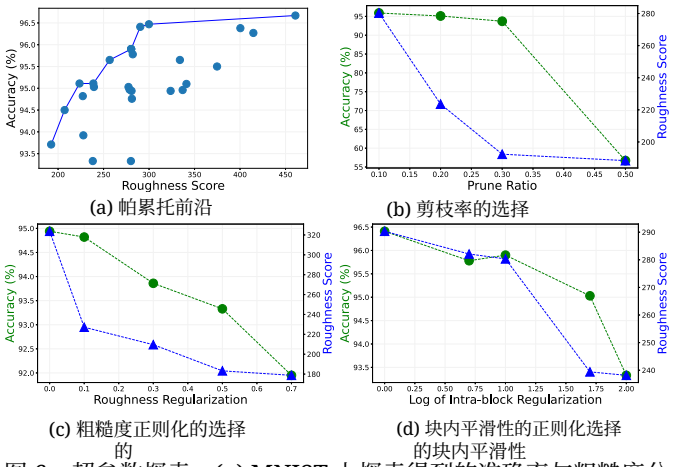
2) 训练设置：我们在四个数据集上使用学习率为 0.2、批量大小为 200 的 Adam 优化器训练基线模型。对于稀疏化，SLR 参数设置为 ρ 0.001 ,,,。学习率设置为使用 Adam，批量大小为 200。稀疏化率设置为 0.1。所有实现均使用 PyTorch v1.8.1 构建，实验结果在 Nvidia 2080 Ti 图形处理器上获得。 $0.1\ M = 300\ r = 0.1\ s_0 = 0.01$

B. 准确率与粗糙度评估

我们展示了将所提出的物理感知粗糙度优化应用于四个数据集后得到的结果。在所有实验中，**Ours-A** 指粗糙度感知训练模型；**Ours-B** 指采用稀疏化训练的模型；**Ours-C**指采用稀疏化和粗糙度训练的模型；**Ours-D** 指采用稀疏化、粗糙度和块内平滑性训练的模型。系统粗糙度评分按所有相位掩模的粗糙度平均值计算：$R_{overall} = R(\mathbf{W})$，其中 $R(\mathbf{W})$ 来自公式 4。$R_{overall}$ 用于量化衍射光学神经网络系统中所有相位掩模的整体像素间相互作用。分数越低，表示像素间相互作用越弱，数值建模与实际部署之间的不匹配程度越小。

表 II、III、IV 和 V 比较了我们提出的优化方法中各个组件在四个数据集上的准确率和粗糙度分数。在所有实验中，[5] 的粗糙度分数最高。采用 2π 优化后，粗糙度分数的降幅不到 2%，这意味着在 2π 范围内，相位掩模中像素之间的相位变化非常剧烈。粗糙度感知训练（Ours-A）使粗糙度分数下降约 10%，但分数仍然相对较高。与粗糙度感知训练相比，稀疏化（Ours-B）取得了更高的准确率，但粗糙度分数也更高，这是因为稀疏化像素与其周围像素之间存在剧烈的相位变化，正如第 III-D2 节所述。借助 2π 优化，相位间隙通过 2π 周期性特征得到平滑，粗糙度分数降至与粗糙度感知训练相同的水平。因此，结合稀疏化与粗糙度建模（Ours-C）取得了显著改进。在四个数据集上，优化前粗糙度分数分别下降了 12.2%、24.6%、13.8% 和 24.6%；优化后分别下降了 35.7%、34.2%、28.1% 和 27.3%，而对准确率影响很小。引入块内平滑性（Ours-D）后，粗糙度分数进一步降低。在允许准确率有略多下降空间（所有数据集平均约 2%）的情况下，2π 优化后四个数据集的粗糙度分别降低了 40%、50.6%、37.4% 和 35.7%。

图 5 比较了 EMNIST 数据集下不同模型中衍射光学神经网络（DONN）的第二个相位掩模（衍射层）。前四个相位掩模未进行 2π 优化。黑色块为稀疏化区域，与周围颜色形成鲜明对比。最后一个是经过稀疏性、粗糙度和块内平滑性优化训练得到的模型所使用的相位掩模，并应用了 2π 后置-



(a) 帕累托前沿 (b) 剪枝率的选择

(c) 粗糙度正则化的选择的 (d) 块内平滑性的正则化选择的块内平滑性

图 6：超参数探索。(a) MNIST 上探索得到的准确率与粗糙度分数帕累托前沿。(b)–(d) 探索稀疏化率、粗糙度正则化和块内平滑度与准确率及粗糙度分数之间的关系。

平滑优化。通过向整个相位掩模的每个像素有选择地添加 2π，先前的黑色区域便与周围颜色融合，整个掩模也变得更加平滑。此外，训练过程中考虑掩模粗糙度后，衍射层中有效的相位调制会更多地聚集在掩模中央，而边缘部分则更可能提供边际调制。这与光学特性相符：光信号强度从中心向外衰减。因此，预计最有效的信息和相位调制会出现在掩模中心。平滑后的相位掩模将更易于制造，并能更好地使衍射光学神经网络的数值建模与其物理部署相吻合。

C. 设计探索

我们进一步在图 6 中进行了超参数探索。图 6(a) 展示了粗糙度分数与准确率之间的帕累托前沿。显然，准确率提高时，粗糙度分数也随之增加，这意味着衍射光学神经网络系统在数字仿真与硬件部署之间的性能差异会更大。在这种情况下，需要调整超参数以应对这一权衡。对于图 6(b)–(d)，我们进一步探索了准确率和粗糙度分数与稀疏化率、粗糙度正则化及块内平滑度正则化之间的关系。我们观察到，随着稀疏化率和正则化强度增加，准确率和粗糙度分数均会下降，但下降幅度不同。对于图 6(c) 中的粗糙度正则化，两项指标在 0.1 处均出现拐点；此后准确率迅速下降，而粗糙度变得更加平滑。类似地，对于图 6(d) 中的块内正则化，当正则化强度的对数为 1 时出现拐点。我们还观察到，这些趋势在其他三个数据集上同样成立。

V. 结论

本文提出了一种面向衍射光学神经网络的物理感知粗糙度优化框架，旨在缩小数字仿真衍射光学神经网络部署到实际光学器件时的性能差异。我们通过块稀疏化在相位掩模中引入稀疏性，并将粗糙度正则化纳入衍射光学神经网络损失函数，以减少衍射层内的像素间相互作用。 2π 采用周期性相位调制和块内平滑度，以进一步提升平滑度。结果表明我们的物理感知粗糙度优化方法在 MNIST、FMNIST、KMNIST 和 EMNIST 上分别将粗糙度降低了 35.7%、34.2%、28.1% 和 27.3%，而准确率仅略有下降。若对准确率损失的容忍度稍高一些，采用所提出的块内平滑度方法还可进一步将粗糙度分别降低 40%、50.6%、37.4% 和 35.7%。

致谢

本研究部分得到了美国国家科学基金会（NSF）的资助，资助项目编号为 NSF-2047176、NSF-2008144 和 NSF-2019336。

参考文献

[1] Sahidul Islam 等人。《Eve：面向低功耗能量采集系统的环境自适应神经网络模型》。发表于国际计算机辅助设计会议，2022年。 [2] Hongwu Peng 等人。面向现场可编程门阵列的二值复值神经网络加速。发表于 IEEE ASAP会议，第 85–92 页，2021年。 [3] Sahidul Islam 等人。为微型能量采集物联网设备实现快速深度学习。发表于 IEEE DATE会议，第 921–926 页，2022年。 [4] Okan K Ersoy。《衍射、傅里叶光学与成像》。John Wiley & Sons，2006年。 [5] Xing Lin 等人。利用衍射深度神经网络实现全光机器学习。《科学》，361(6406):1004–1008，2018年。 [6] Tiankuang Zhou 等人。利用可重构衍射处理单元实现大规模神经形态光电计算。《自然·光子学》，15(5):367–373，2021年。 [7] Yingjie Li 和 Cunxi Yu。最新研究成果：衍射深度神经网络的物理对抗攻击。发表于设计自动化会议，2021年。 [8] Yingjie Li 等人。面向衍射光学神经网络的物理感知可微离散协同设计。国际计算机辅助设计会议，2022年。 [9] Ruiyang Chen 等人。可重构衍射全光神经网络中的物理感知复值对抗机器学习。arXiv 预印本 arXiv:2203.06055，2022年。 [10] Yingjie Li、Weilu Gao 和 Cunxi Yu。Rubikonns：利用 Rubik 衍射光学神经网络进行多任务学习。2022年。 [11] Ruiyang Chen 等人。使用经济型空间光调制器的复值可重构衍射光学神经网络。发表于 CLEO：应用与技术。Optica出版集团，2022年。 [12] Amnon Yariv。正交（波长复用）全息图中的页间与像素间串扰。《光学快报》，1993年。 [13] Minhan Lou 等人。层间反射和像素间相互作用对衍射光学神经网络的影响。《光学快报》，2023年。 [14] Marco Vallone 等人。用于碲镉汞焦平面阵列中像素间串扰的扩散概率模型。IEEE电子器件学会期刊，2018年。 [15] Donald C O'Shea 等人。《衍射光学：设计、制造与测试》，第 62 卷。SPIE出版社，2004年。 [16] Deniz Mengu 等人。尺度、平移和旋转不变衍射光学网络。《ACS光子学》，8(1):324–334，2020年。 [17] Mahdad Mansouree 等人。利用伴随优化进行大规模参数化超表面设计。《ACS光子学》，8(2):455–463，2021年。 [18] Jinhie Skarda 等人。用于大面积超表面仿真与优化的低开销分布式策略。《npj计算材料》，8(1):1–6，2022年。 [19] Tianyun Zhang 等人。使用交替方向乘子法的系统化深度神经网络权重剪枝框架。发表于欧洲计算机视觉国际会议，2018年。 [20] Shanglin Zhou 等人。自动驾驶中利用嵌入式GPU进行端到端多任务目标检测。发表于国际质量电子设计研讨会，2021年。 [21] Shanglin Zhou 等人。结合多视图集成与代理拉格朗日松弛，在边缘设备上实现实时三维生物医学图像分割。《神经计算》，512:466–481，2022年。 [22] Tianyun Zhang 等人。使用重加权优化方法的统一深度神经网络权重剪枝框架。发表于设计自动化会议，第 493–498 页。IEEE，2021年。 [23] Song Han 等人。学习权重与连接，实现高效神经网络。神经信息处理系统大会，28，2015年。 [24] Song Han 等人。EIE：压缩深度神经网络上的高效推理引擎。ACM计算机体系结构特别兴趣小组，44(3):243–254，2016年。 [25] Wei Wen 等人。在深度神经网络中学习结构化稀疏性。《神经信息处理系统进展》，29，2016年。 [26] Zhuliang Yao 等人。面向图形处理器上高效深度神经网络推理的平衡稀疏性。发表于人工智能促进协会年会，第 33 卷，第 5676–5683 页，2019年。 [27] Shijie Cao 等人。采用存储体平衡稀疏性的现场可编程门阵列高效稀疏长短期记忆网络。发表于现场可编程门阵列会议，第 63–72 页，2019年。

[28] Caiwen Ding 和 等人。Circnn：利用块循环权重矩阵加速并压缩深度神经网络。发表于*MICRO*，2017年。 [29] Mary S Tobin。傅里叶光学导论。《美国科学家》，85(6):581–584，1997年。 [30] Yingjie Li 和 等人。通过软硬件协同设计实现衍射深度神经网络中的多任务学习。《科学报告》，2021年。 [31] Franc¸ois Lagunas 和 等人。用于加速 Transformer 的块剪枝。发表于 *EMNLP*，第 10619–10629 页，2021年。 [32] Deniz Gurevin 和 等人。利用代理拉格朗日松弛实现无需重新训练的深度神经网络剪枝。*IJCAI*，2020年。 [33] Stephen Boyd 和 等人。通过交替方向乘子法实现分布式优化与统计学习。《机器学习基础与趋势》，3(1):1–122，2011年。 [34] Eric Jang、Shixiang Gu 和 Ben Poole。使用 Gumbel-Softmax 的类别重参数化。*arXiv* 预印本 *arXiv:1611.01144*，2016年。 [35] Diederik P Kingma 和 Jimmy Ba。Adam：一种随机优化方法。*arXiv* 预印本 *arXiv:1412.6980*，2014年。