面向卫星图像分类的非相干多光谱衍射光学网络设计

作者

Xilin Yang[1,2,3], Md Sadman Sakib Rahman[1,2,3], Ryan Lee[1], Aydogan Ozcan[1,2,3,*]

[1]加利福尼亚大学洛杉矶分校电气与计算机工程系，美国加利福尼亚州洛杉矶，邮编 90095。

[2]加利福尼亚大学洛杉矶分校生物工程系，美国加利福尼亚州洛杉矶，邮编 90095。

[3]加利福尼亚大学洛杉矶分校加利福尼亚纳米系统研究所（CNSI），美国加利福尼亚州洛杉矶，邮编 90095。

* ozcan@ucla.edu

摘要

光学神经网络和衍射光计算为执行具有高并行性且电子后处理负担较轻的特定任务推理提供了一条很有前景的途径。这些优势对于多光谱卫星成像尤其具有吸引力，因为传统工作流程依赖将高维图像立方体传输至地面，以进行后续电子分析。在此，我们提出了一种计算高效的代理模型引导框架，用于设计非相干多光谱衍射光学网络，使其能够在星上直接输出类别概率，从而无需采集并传输完整的多光谱数据立方体以进行地面端推理。我们首先训练一个物理兼容代理模型，该模型由能够适应广义探测器读出的非负全连接模型表示，并生成可传输至光学处理器的、面向特定任务的非负多光谱强度变换。随后，我们在空间非相干照明条件下，通过无数据的基于点扩散函数的优化合成其衍射对应模型。利用这一方法，我们通过数值方式演示了一种用于三分类卫星场景识别的六光谱通道衍射分类器；对于人工地表、水体和植被覆盖地表图像，其盲测准确率达到79.8%，接近其代理模型的81.8%准确率。我们进一步展示了一种更加紧凑的三光谱通道设计，其数值盲测准确率达到72.7%，而对应代理模型的准确率为75.1%。这一代理模型引导策略为面向资源受限的遥感推理，设计紧凑、特定任务的非相干多光谱光学处理器提供了一条切实可行的路径。

引言

光计算利用光传播的并行性，为减轻信息处理的电子负担提供了一条颇具吸引力的途径[1–7]。衍射光学网络尤其能够通过无源结构层变换光学输入，并在探测器平面上生成针对特定任务的输出[8–20]。这一能力对于遥感颇具吸引力，因为多光谱成像能够为地球监测提供丰富的空间和光谱信息[21–23]。然而，传统的卫星工作流程通常需要先采集高维多光谱图像立方体，将其传输到地面，再利用外部计算资源进行处理。这一数据传输与推理流程会带来带宽、延迟和硬件方面的负担，因此亟需紧凑的星载推理架构，以便在采集现场直接处理多光谱场景[24–27]。

卫星多光谱场景天然采用强度编码，因此更适合在空间非相干照明下建模；而要通过光学处理对这类场景进行分类，光学处理器需要联合处理多个波长通道。这些要求使得非相干多光谱衍射网络的设计相比相干或单波长光学分类器[28]更具挑战性。此外，直接

在物理上准确的非相干传播下进行端到端优化可能计算成本高昂[28], , 从而限制了对网络架构、超参数扫描、波长组合和探测器平面设计的系统探索。基于代理模型的框架通过先构建一个物理兼容的计算代理模型，再将其学习到的变换迁移至光学处理器，为应对这一挑战提供了一种实用方法。

本文提出一种用于卫星图像分类的非相干多光谱衍射光学网络设计高效框架。系统优化完成后，设计出的系统可直接根据星载监测视场输出类别概率，无需采集完整的多光谱数据立方体并将其传输至外部计算资源进行推理。为设计该系统，我们首先训练一个带有广义探测器读出的非负全连接代理模型，以学习特定任务所需的强度兼容多光谱变换。随后，该代理模型作为目标，用于基于点扩散函数、无需数据的多层非相干衍射处理器优化。利用这一框架，我们通过数值实验展示了一种用于三分类卫星场景识别的六光谱通道衍射分类器；对于人工地表、水体和植被覆盖地的图像，其盲测准确率达到79.8%，接近其代理模型的81.8%准确率。我们还展示了一种紧凑的三光谱通道设计，其数值盲测准确率达到72.7%，而对应代理模型的准确率为75.1%，凸显了该策略在实现光谱高效的星载光学推理方面的灵活性。

图 1 展示了我们对多光谱卫星场景进行分类的整体构想。在传统遥感工作流程中（图 1a–c），监测卫星通过望远成像系统和多光谱传感器采集多光谱图像立方体，随后将生成的高维数据下传至地面，利用外部计算资源和大规模电子神经网络进行推理。为免去采集大量数据并将其传输至地面的需求，我们提出一种星载非相干多光谱衍射光学分类器（图 1d）：多光谱光学信息由紧凑型衍射网络和探测器平面处理，从而直接输出类别概率，仅需极少的电子后处理。

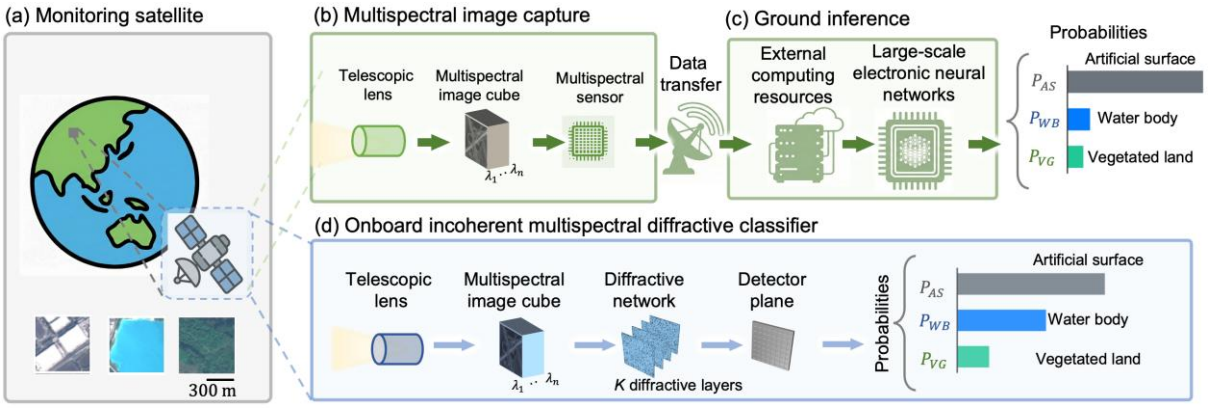


图 1. 用于卫星成像的星载多光谱衍射分类。 (a) 监测卫星。(b) 使用望远镜镜头和多光谱

平面

结果

S非相干多光谱衍射光学网络的替代模型

为高效设计用于卫星图像分类的非相干多光谱衍射光学网络，我们首先建立一个代理模

该代理模型旨

的三分类性能，

移至

如图2a所示，我们首先采用交叉熵目标函数，通过数据驱动优化对非负全连接（NNFC）模型进行计算机模拟训练。卫星图像被分为三个目标类别：（i）人工地表，（ii）水体，以及（iii）植被覆盖地。非负全连接模型将输入的多光谱图像立方体映射为一组紧凑的有效输出传感器像素读数，其中每一对像素组专用于一个语义类别。探测器平面设计如图2c所示。每个类别分配一个空分复用探测器对，相应的强度读数用于构建类别分数和预测概率。为提升这一紧凑探测器读出的表达能力，我们引入了一种广义探测器形式，其中每一对强度测量值都经过一种轻量级数字变换，该变换被参数化为学习得到的有理函数。这些广义探测器与非负全连接模型联合训练，从而使代理模型能够同时学习一种物理上可迁移的非负线性映射以及一个

针对探测器平面架构定制的紧凑决策规则。广义探测器读出的详细数学表述见方法部分。

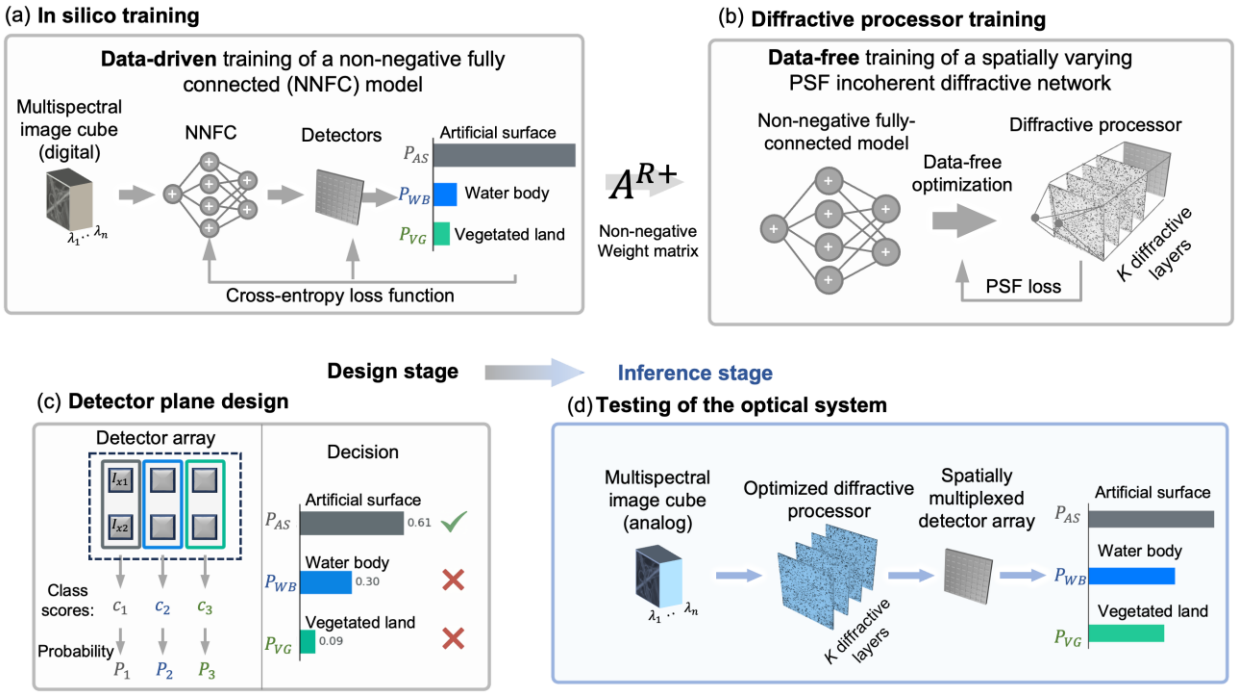


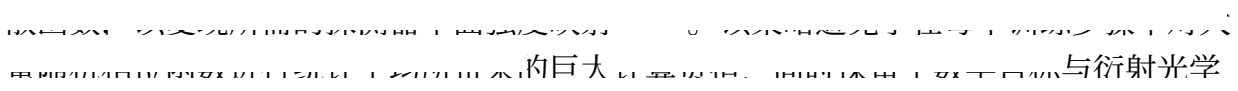
图2. 代理模型引导的多光谱衍射光学分类器设计。 (a) 使用带标签的多光谱卫星图像对 [illegible] 用于 [illegible]

训练完成后，NNFC代理模型提供由非相干衍射光学处理器实现的目标非负强度变换。如图2b所示，这种学习到的线性映射随后通过基于点扩散函数（PSF）的数据无关优化，转化为非相干衍射网络设计。所得光学系统与广义探测器读出共同构成最终的多光谱衍射分类器，如图2d所示，并将在以下各节中进行介绍。

非相干多光谱衍射光学网络设计

代理模型训练完成后，我们将其学习到的非负线性变换作为目标空间和光谱变化的点扩 [illegible] 设计阶 [illegible]

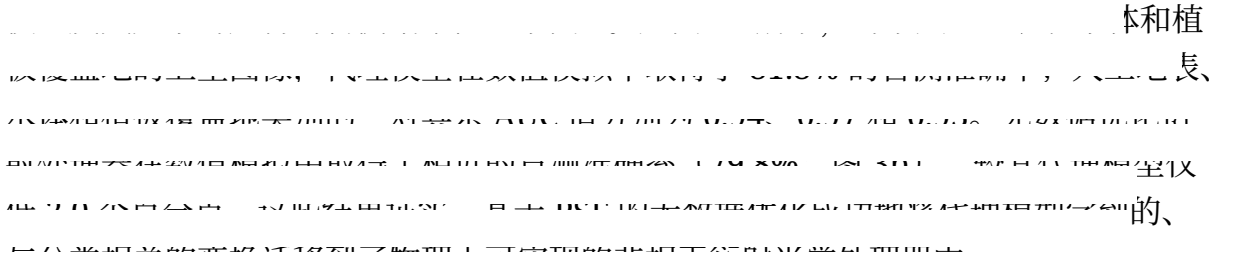
传播模型逼近代理模型的目标变换。我们不直接使用带标签的多光谱图像训练光学网络，



训练完成后，将优化后的衍射处理器与空间复用探测器阵列及广义探测器读出相结合，



我们首先对一个具有六个光谱通道的非相干多光谱衍射分类器进行数值评估。相应代理



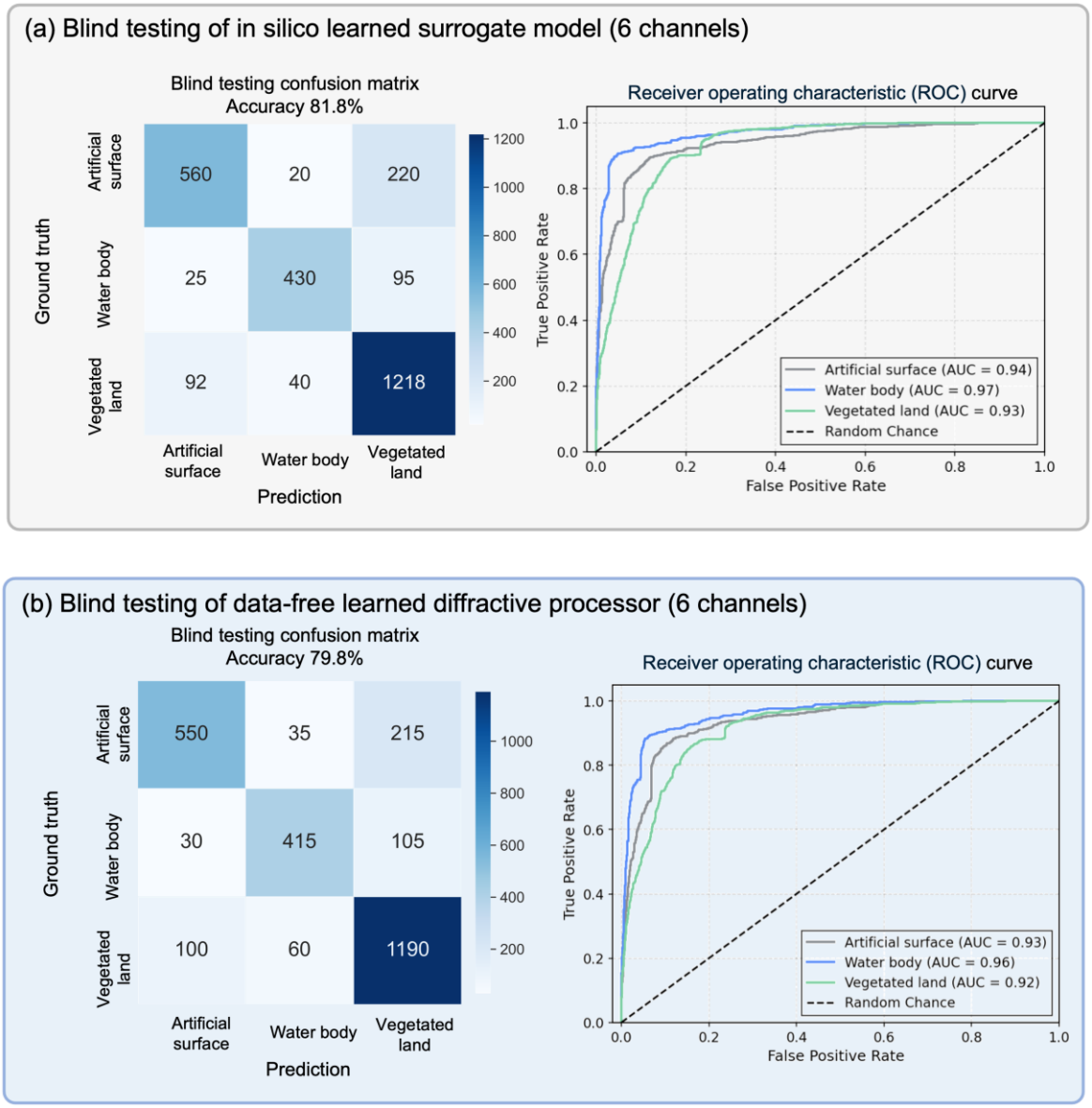


图 3. 六光谱通道数字与光学分类器的盲测性能。展示了数值分类结果。(a) 六光谱通

图 3 中的混淆矩阵进一步表明，衍射光学网络保留了代理模型整体的类别判别行为。在

及将人工

准确保留代理模型的主要决策结构，而不仅仅是匹配总体准确率。

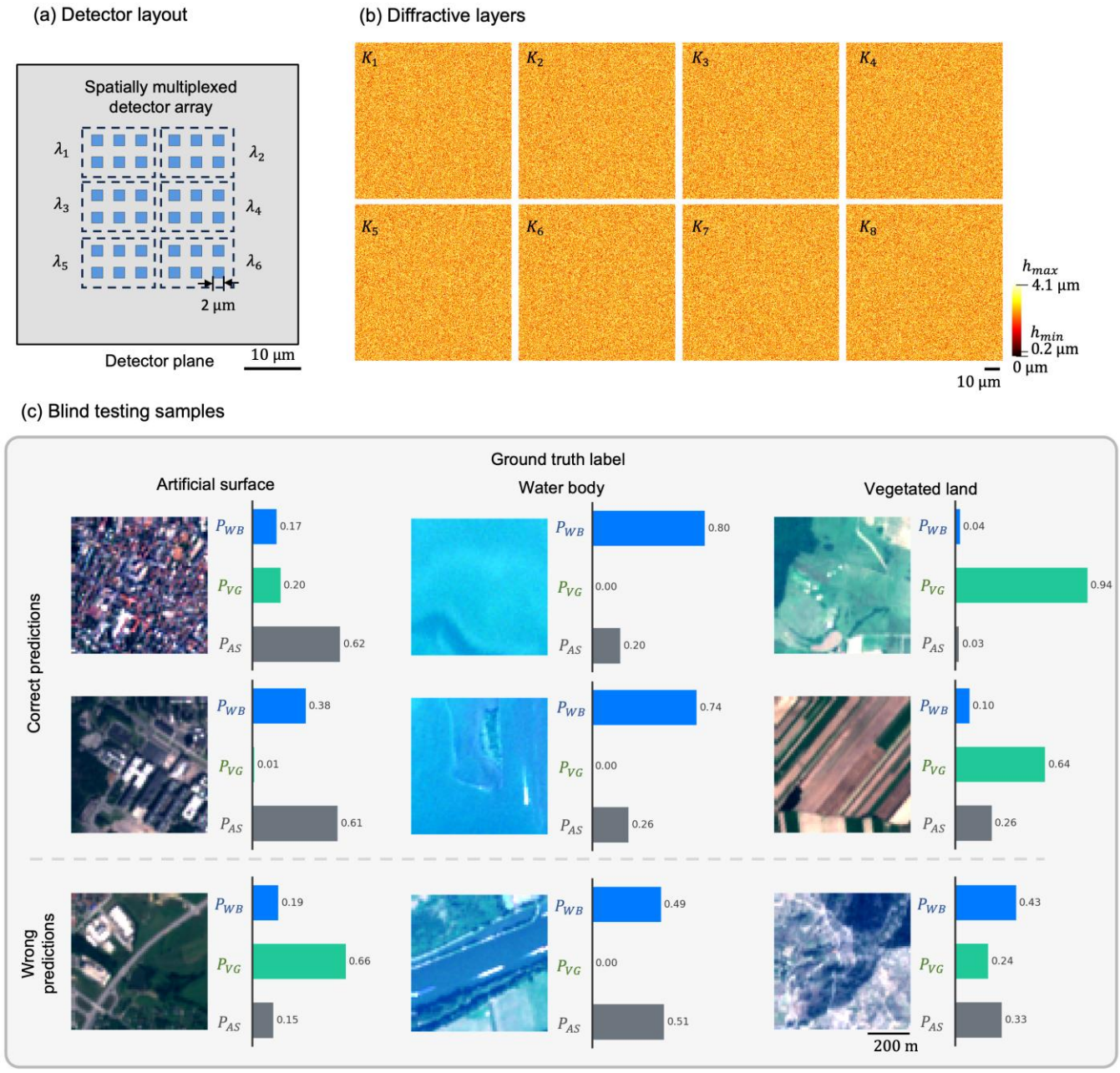


图4。六光谱通道衍射光学分类器及代表性预测结果。可视化展示了光学设计。(a) 六通

类

图4进一步展示了六波长通道光学分类器的详细信息以及具有代表性的盲测示例。图4a

波长通道的探测器区域。图4b展示了光学处理器中所用衍射层的优化高度分布。图4c给

接下来，我们研究了基于代理模型的框架如何通过减少光谱通道数量来指导更紧凑的系统设计。我们利用代理模型为六通道和三通道光谱设计选择波长（详情请参见方法部分）。具体而言，我们设计了一种紧凑型三通道非相干多光谱衍射分类器，如图5所示。相应的代理模型达到了75.1%的盲测准确率（图5c），而无数据优化衍射处理器达到了72.7%的准确率（图5d），使数字和光学实现之间存在2.4个百分点的差距。这些数值结果表明，即使在光谱维度大幅降低的情况下，所展示的设计策略仍然有效，从而能够实现更紧凑的光学分类器，同时保持具有实际意义的三类分类性能。

讨论

在本研究中，我们提出了一种代理模型引导的框架，用于设计面向卫星图像分类的非相干多光谱衍射光学网络。我们没有直接基于标注图像数据优化衍射处理器，而是首先训练了一个物理兼容代理模型，使其学习非负全连接强度映射，以及针对目标分类任务的紧凑型广义探测器读出。随后，通过对多层非相干衍射处理器进行基于点扩散函数的无数据优化，将所学习的变换迁移至其光学对应物。利用这一策略，我们设计了多光谱衍射分类器，其性能接近相应代理模型的性能，其中包括一种仅使用三个光谱通道的紧凑型设计，同时保持了具有实际意义的分类性能。

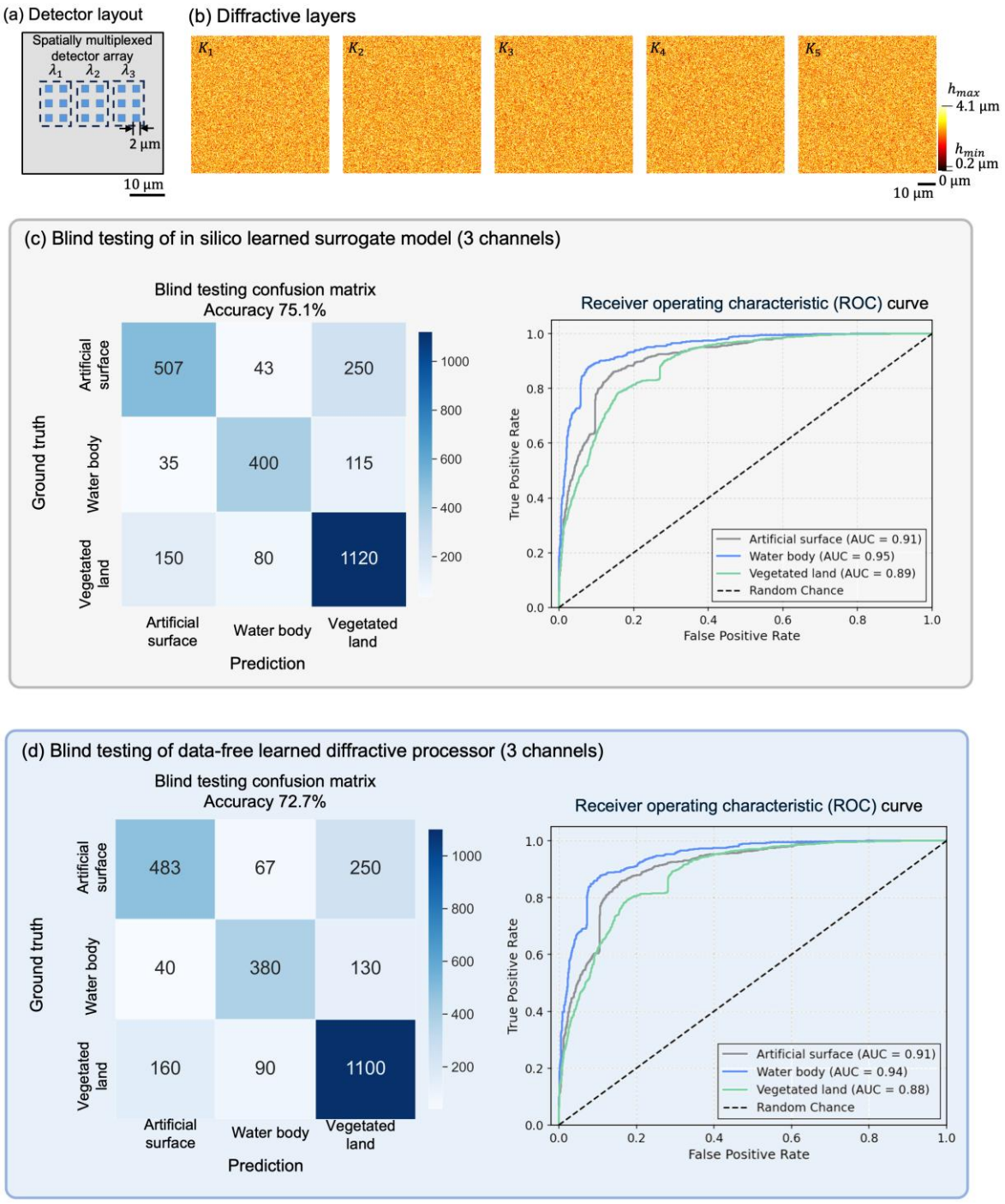


图5. 紧凑型三光谱通道衍射光学处理器。 (a) 三通道系统的探测器平面布局。(b) 五个衍

本工作的核心贡献之一，是将代理模型用作任务驱动学习与受物理约束的衍射光学合成之间的实用设计接口。在这一框架中，代理模型不仅是数值参考模型，还是一种计算效率高的工具，可在投入计算成本高昂的衍射优化之前，快速探索光学系统设计方案。例如，可以先在计算机模拟中利用代理模型考察波长选择策略、探测器读出方案和紧凑的架构配置。这种“思想实验式设计”方法[31]为推理可实际实现的光学处理器提供了一条灵活途径：首先，探究在所需读出约束下，哪一类非负映射能够完成任务；随后，再设法合成能够实现该映射的衍射光学系统。当光学训练计算成本高昂，或可用设计空间包含光谱、探测器和衍射自由度之间的诸多耦合选择时，这种策略尤其有价值。

本文采用的基于点扩散函数的无数据优化策略，对于使这一设计流程在计算上切实可行也至关重要。当通过对随机相位实现取平均来模拟空间非相干传播时，直接进行任务驱动优化需要反复执行场传播步骤；对于较大的输入孔径和多光谱运行，这些步骤的计算成本会很高。相比之下，基于点扩散函数的策略利用非相干强度变换的线性特性，对系统响应点状输入激励的过程进行优化，从而在光学训练期间无需带标签图像样本，即可匹配所学习的目标变换。这为将数字映射迁移至光学硬件提供了一种计算效率高且模块化的途径。同时，衍射处理器近似任意多波长强度变换的能力，仍从根本上受其可用自由度[4,32–34]的限制。在此背景下，对于波长复用线性强度变换，所需的可训练衍射特征数量随$N \sim 2N_w N_i N_o$变化，其中N_w、N_i和N_o分别表示光谱通道数、输入像素数和输出像素数。这一关系凸显了紧凑任务表述、缩减光谱预算以及探测器平面协同设计的重要性，有助于将非相干衍射分类器扩展至更复杂的遥感问题。

本研究仍存在若干局限，也有一些未来研究方向。首先，本研究侧重于在模拟的空间非相干照明下进行数值设计和盲测。实验实现还需进一步考虑制造精度、对准容差、材料吸收、探测器噪声，以及光学系统在实际卫星运行条件下的稳定性。通过将这些实验非理想因素纳入训练模拟，可以设计出能够抵御此类因素的衍射处理器[35]。尽管当前模型假设采用空间非相干照明，但实际遥感场景可能会因场景、成像几何和光学架构而涉及部分相干或混合相干条件。已有研究表明，衍射处理器能够在部分相干照明下运行[30]。这为将我们的分类器设计扩展到部分相干输入提供了一条途径，我们将其留待未来研究。最后，当前系统

以光学方式实现非负线性变换，并通过广义探测器在数字端保留轻量级非线性读出。未来工作可以探索更丰富且与物理规律相容的非负函数、混合级联光学—数字推理方案，或其他探测器平面表述，在保持衍射光学处理器低功耗和紧凑特性的同时，进一步提升决策能力。总体而言，这些方向指向一类更广泛的数字化设计非相干光学系统，可用于机载遥感、资源受限分类以及其他多光谱推理任务。

材料与方法

数据集与预处理

我们以 EuroSAT[29] 数据集作为基础遥感基准，通过将 EuroSAT 的原始类别归并为水体、人工地表和植被覆盖地表图像，构建了一个三分类多光谱分类任务。具体而言，River 和 SeaLake 归为水体；Highway、Industrial 和 Residential 归为人工地表；其余所有类别归为植被覆盖地。完整数据集包含 13 个光谱波段，其中心波长分别为 443、490、560、665、705、740、783、842、865、945、1375、1610 和 2190 nm。我们使用固定的随机种子生成了训练集、验证集和测试集比例为 80/10/10 的数据划分，并在所有数字和光学数值实验中复用相同的数据划分。

我们使用在所有训练样本和光谱通道上计算得到的同一个最小值和最大值，将训练输入缩放至 $[0, 1]$ 。验证集和测试集输入也使用相同的数值进行归一化。每个样本表示为一个包含 N_w 光谱通道的多光谱图像立方体，其中 N_w 取决于所研究的具体设计。主衍射光学分类器使用六个光谱通道，其中心波长分别为 490、665、783、842、1610 和 2190 nm；而紧凑型设计使用三个光谱通道，其中心波长分别为 490、842 和 2190 nm。每幅多光谱图像都作为一幅 64×64像素图像放置在衍射光学网络的输入平面上，物理像素尺寸为 400 nm。

使用非负全连接模型设计代理模型

为获得紧凑且与物理特性兼容的代理模型，我们采用受约束线性模型训练了一个非负全连接分类器。设多光谱输入为 $I_{in} \in \mathrm{R}_{+}^{H\times W\times C}$。我们将每个光谱通道展平为非负向量 $x_i \in \mathrm{R}_{+}^{N_i}$，其中 H、W 和 C 分别表示图像高度、宽度和光谱通道数，且 C=N_w，$N_i = H \times W$ 表示输入孔径处的有效像素总数。我们将经偏置调整的探测器读数记为 $y \in \mathrm{R}^{N_o}$，其中 N_o=$2N_{class}$，N_{class} 表示类别总数（三类，除非另有说明）。NNFC 模型在各光谱通道上应用分组线性变换，并对权重施加非负约束：

$$y = \sum_{i=1}^{N_w} A(W_i)x_i + b$$

其中，W_i 和b 分别表示与第 i 个光谱通道关联的权重矩阵和偏置项，$A(\cdot)$是用于强制权重非负的约束函数。衍射光学处理器仅实现了对应于第一项的非负线性分量。学习到的偏置不受约束，并以固定偏移量的形式通过数字方式加到非负光学读数上，然后计算类别分数。我们考虑了 $A(\cdot)\in\{\mathrm{ReLU}(\cdot), \mathrm{SoftPlus}(\cdot)\}$，其中：

$$\mathrm{ReLU}(x) = \max(x, 0)$$

$$\mathrm{SoftPlus(x)} = \log(1 + \exp(x))$$

下标 i 表示相应的光谱通道，权重从均匀分布 $U[10^{-4}, 1]$中初始化。

随后，使用下一小节所述的广义探测器读出，将探测器输出转换为类别得分 C_n 。再通过 softmax 函数将这些类别得分转换为类别概率：

$$P_n = \frac{\exp(C_n)}{\sum_{m=1}^{N_{\mathrm{class}}} \exp(C_m)}$$

其中，P_n 表示第 n 类的预测概率。代理模型使用多分类交叉熵损失进行训练：

$$\mathcal{L}_{\mathrm{CE}} = -\sum_{n=1}^{N_{\mathrm{class}}} t_n \log(P_n)$$

其中，t_n 是第 n 类的独热编码真实标签。

我们使用 Adam[36] 优化器训练最多 200 个训练轮次，并围绕批次大小、学习率和非负激活约束进行网格搜索。选择验证准确率最高的训练轮次所对应的模型，用于后续光学设计。性能最佳的配置采用批次大小 64、ReLU[37] 激活函数和 1×10^{-4} 的学习率。六通道波长集合是通过依次从原始 13 波段集合中移除两个、两个、两个和一个光谱通道进行穷举搜索获得的。对于三通道设计，我们穷举评估了双通道组合，固定验证损失最低的通道对，然后评估加入每个剩余通道的效果。根据验证损失选择波长组合。训练和测试在一台配备 NVIDIA GeForce RTX 3090 Ti 图形处理器、256 GB 随机

存取存储器和 Intel Core i9 中央处理器的工作站上进行。代理模型和衍射网络仿真使用 Python 3.9 和 PyTorch[38] 2.1.0，并借助 CUDA 12.1 实现。

广义探测器读出

基于差分探测器设计[39]，，为每个类别 n 分配两个探测器输出：$y_{n,1}$和 $y_{n,2}$。差分探测器对按如下方式计算相应的类别得分：$C_n = \frac{y_{n,1} - y_{n,2}}{y_{n,1} + y_{n,2}}$。

为增强探测器平面的读出表达能力，同时保留轻量级的后处理计算，我们将这一形式推广为一对探测器输出的有理函数：

$$C_n = \frac{\sum_{p=0}^{d} \sum_{q=0}^{d} a_{n,p,q} y_{n,1}^{p} y_{n,2}^{q}}{\sum_{p=0}^{d} \sum_{q=0}^{d} b_{n,p,q} y_{n,1}^{p} y_{n,2}^{q}}$$

其中，C_n 是第 n类的类别得分，$a_{n,p,q}$ 以及 $b_{n,p,q}$ 为可训练系数。除非另有说明，参数 d 设为 2。该广义形式放宽了传统差分探测器对固定函数形式的限制，所有系数均与 NNFC 参数使用同一优化器联合优化。将 $a_{n,1,0} = b_{n,1,0} = b_{n,0,1} = 1$ 和$a_{n,0,1} = -1$设为相应数值且其余系数均设为零时，广义形式可在 $d = 1$ 的条件下还原为传统差分探测器这一特例。广义读出以传统差分形式初始化，并在两种读出形式的分母中均加上 FP32 机器精度（$\varepsilon \approx 1.19 \times 10^{-7}$）。训练过程中未观察到数值不稳定性。对于差分探测器基线，可保持其他模型和训练设置不变，同时将读出系数固定为传统差分形式，而非进行联合优化。

非相干多光谱衍射光学网络设计

该衍射网络由多个纯相位衍射层组成，各层之间由自由空间传播距离隔开，并使用标量衍射模型通过计算机模拟进行优化。除非另有说明，主要的多光谱设计包含八个衍射层，每层有 512×512 个衍射特征；紧凑型三光谱通道设计包含五个衍射层，每层有 384×384 个衍射特征。两种设计的衍射特征间距均为 200 nm，输入平面与第一层、相邻衍射层之间，以及最后一层与探测器平面之间的间隔均为 10 μ m。

我们首先描述给定波长 λ 下的相干光传播模型，该模型构成了空间非相干传播建模的基础。对于多光谱场景，不同波长使用相同的传播模型，并对各波长通道产生的强度进行积分。每个选定波段均以其中心波长建模，未对有限的光谱通带进行积分。所有选定的波长贡献均以相同权重进行积分，未进行波长相关增益校准。在数值

读出中，每个波长指定的探测器区域仅保留指定波长的贡献，这对应于理想的带通选择。每个探测器区域通过对其面积内保留的光强进行积分，产生一个读数。按类别和探测器索引对保留的贡献求和，然后使用各设计最后一次训练迭代中的通用归一化因子进行缩放。随后，我们添加学习到的偏置并计算类别分数。无滤波器设计也可以在点扩散函数优化期间，对波长指定探测器区域内的带外功率进行惩罚，以抑制光谱串扰。所选波长按从短到长的顺序编号为 λ_1、…、 λ_{N_w}。这种特定的空间分配并非目标变换的根本限制，因为衍射处理器可以通过训练实现相应的输出排列[40]。

每个衍射层通过空间变化的相位延迟调制入射光场：

$$t^k(x,y) = \exp\{j\phi^k(x,y)\}$$

其中，$\phi^k(x,y)$表示第k个衍射层引入的相位调制。连续两个平面之间的自由空间传播采用角谱法[1]进行建模：

$$u(x,y,z+d_k) = \mathcal{F}^{-1}\{\mathcal{F}\{u(x,y,z)\}\cdot H(f_x,f_y;d_k)\}$$

其中，$u(x,y,z)$表示传播前的光场，$u(x,y,z+d_k)$表示传播距离d_k后的光场。F和F^{-1} 分别表示二维傅里叶变换及其逆变换，而f_x 和f_y 表示空间频率。相应的自由空间传递函数为：

$$H(f_x,f_y;d_k) = \begin{cases} \exp\left\{j\frac{2\pi}{\lambda}d_k\sqrt{1-(\lambda f_x)^2-(\lambda f_y)^2}\right\}, & f_x^2+f_y^2 < \frac{1}{\lambda^2} \\ 0, & f_x^2+f_y^2 \geq \frac{1}{\lambda^2} \end{cases}$$

其中，λ 是照明波长。在数值传播过程中，将计算窗口沿横向填充至原始尺寸的两倍，以提高传播精度。

为模拟空间非相干照明，我们采用直接公式[28]，通过对20,000个相互独立的随机相位实现中相干传播光场的幅度平方取平均，来估计输出强度。对于每次实现，200 nm输入平面仿真网格中每个点的相位均独立地从$[-\pi,\pi]$上的均匀分布中采样。该直接非相干传播模型用于衍射光学网络的所有盲测评估。

然而，为了训练光学处理器，我们采用一种无需数据、基于点扩散函数的优化策略[28]，并假设来自不同输入像素的光彼此非相干。

目标非负线性变换首先由训练好的代理模型获得。在点扩散函数前向传播中，逐个照亮每个输入像素，并将其表示为多个点光源；随后优化得到的点扩散函数，使其逼近相应的目标输出强度分布。通过遍历所有输入像素，优化衍射层，使其共同逼近目标非相干线性变换，而光学训练过程中无需带标签图像样本。每个训练轮次中，随机抽取输入像素且不重复，并将其分成每批 16 个。在每个批次中，我们联合评估所有选定波长，并对各波长的点扩散函数损失取平均。对于各波长的点扩散函数损失，我们采用带归一化因子的均方误差损失：

$$\mathcal{L}_{PSF,\lambda} = \frac{1}{N_i N_o}\sum_{p=1}^{N_i}\sum_{q=1}^{N_o}\left(L_\lambda[p,q] - \sigma\widehat{L_\lambda}[p,q]\right)^2$$

其中，每个训练批次内的所有波长通道共享一个归一化因子 σ，该因子用于补偿与衍射效率相关的全局强度失配。对于汇总最小二乘归一化，其定义如下：

$$\sigma = \frac{\sum_{\lambda,p,q} L_\lambda[p,q]\widehat{L_\lambda}[p,q]}{\sum_{\lambda,p,q} \widehat{L_\lambda}[p,q]^2}$$

这些求和是在当前训练批次中包含的各个波长和点扩散函数条目上进行的。此处，L_λ 和 $\widehat{L_\lambda}$ 分别表示波长 λ处的目标光学非负线性变换和输出光学非负线性变换，相应的索引则表示输入像素和输出像素。每次参数更新后，都会使用当前批次重新计算公共归一化因子 σ。在反向传播过程中，σ 保持固定，梯度不会通过其计算过程反向传播。

为考虑材料色散，衍射层采用其物理高度分布进行参数化，而不是直接优化相位调制：

$$h_k = h_{min} + \frac{h_{max} - h_{min}}{2}\left(\sin\left(h_k^{latent}\right) + 1\right)$$

此处，k 表示衍射层的索引，h_k^{latent} 是可训练的潜变量，用于计算物理高度 h_k。潜变量通过从标准正态分布（均值为 0，标准差为 1）中采样进行初始化。此参数化将衍射层的高度限制在 $\mathrm{h}_{min} = 0.2\ \mu\mathrm{m}$ 至 $\mathrm{h}_{max} = 4.1\ \mu\mathrm{m}$ 的制造范围内。随后，根据 N-BK7玻璃[41] 的色散特性及其折射率，计算相应的波长相关相位延迟：

$$\phi_{k,\lambda}(x,y) = \frac{2\pi}{\lambda}(n_\lambda - 1)h_k(x,y)$$

衍射光学模型使用 AdamW[42] 优化器进行优化，学习率为 1×10^{-3} ，最多训练 20 轮或直至收敛。我们选择点扩散函数损失最低的光学检查点进行盲测。训练和测试均在用于训练代理模型的同一台工作站上进行。

数据可用性

多光谱 EuroSAT 数据集可公开获取，网址为 https://zenodo.org/records/7711810

参考文献:

1. Lin, X. 等 使用衍射深度神经网络实现全光机器学习。《科学》**361**, 1004–1008 (2018).

2. Wetzstein, G. 等 利用深度光学和光子学实现人工智能推理。*Nature***588**, 39–47 (2020).

3. Chen, A., Wang, Y., Rahman, M. S. S., Li, Y. & Ozcan, A. 利用非相干光通过线性衍射系统实现大规模非线性光学计算。预印本, https://doi.org/10.48550/arXiv.2603.29131（2026）。

4. Rahman, M. S. S., Li, Y., Yang, X., Chen, S. & Ozcan, A. 使用衍射处理器大规模并行实现非线性函数的通用逼近。*eLight* **5**，32（2025）。

5. Zhang, D. 等。利用电调谐光电探测器实现光谱核机器。《科学》**390**, eady6571（2025）。

6. Momeni, A. 等。物理神经网络的训练。*Nature* **645**，53–61（2025）。

7. Chen, S., Li, Y., Wang, Y., Chen, H. & Ozcan, A. 光学生成模型。*Nature* **644**, 903–911（2025）。

8. Li, J., Hung, Y.-C., Kulce, O., Mengu, D. & Ozcan, A. 偏振复用衍射计算：通过偏振编码衍射网络实现一组线性变换的全光学实现。《光：科学与应用》 **11**，153（2022）。

9. Kulce, O., Mengu, D., Rivenson, Y. & Ozcan, A. 衍射表面的全光信息处理能力。《光：科学与应用》 **10**, 25 (2021)。

10. Kulce, O., Mengu, D., Rivenson, Y. & Ozcan, A. 利用衍射表面全光合成任意线性变换。《光：科学与应用》 **10**, 196 (2021)。

11. Yang, X., Rahman, M. S. S., Bai, B., Li, J. & Ozcan, A. 利用空间非相干衍射网络实现复值通用线性变换与图像加密。*Adv.Photonics Nexus* **3**, 016010 (2024)。

12. He, C. 等 用于高容量光学推理的轨道角动量复用衍射神经网络。*Nat. Photonics* **20**, 896–904 (2026)。

13. Zhou, T. 等 利用可重构衍射处理单元实现大规模神经形态光电计算。*Nat. Photonics* **15**，367–373（2021）。

14. Xu, Z. 等 大规模光子芯粒 Taichi 赋能 160-TOPS/W 通用人工智能。《科学》 **384**，202–209（2024）。

15. Chen, Y. 等 用于高速视觉任务的全模拟光电芯片。*Nature* **623**，48–57（2023）。

16. Yu, H. 等 利用远端端面上的微型化衍射神经网络，通过多模光纤实现全光图像传输。*Nat. Photonics* **19**，486–493（2025）。

17. Yildirim, M., Dinc, N. U., Oguz, I., Psaltis, D. & Moser, C. 线性光学中的非线性处理。*Nat. Photonics* **18**, 1076–1082 (2024).

18. Liu, C. 等 基于数字编码超表面阵列的可编程衍射深度神经网络。《自然·电子学》**5**, 113–122 (2022).

19. Luo, X. 等 可见光波段的超表面赋能的片上复用衍射神经网络。《光：科学与应用》*11*，158（2022）。

20. Bezuidenhout, H.、Peters, C.、Kumar, R. N.、Forbes, A. 和 Nape, I. 用于检测光的斯格明子拓扑及其实时训练的深度衍射光学神经网络。《自然·通讯》**17**, 10490（2026）。

21. Shaw, G. A. 和 Burke, H. K. 用于遥感的光谱成像。《林肯实验室杂志》**14**，3–28（2003）。

22. Thomas, C.、Ranchin, T.、Wald, L. 和 Chanussot, J. 将多光谱图像合成为高空间分辨率图像：基于遥感物理的融合方法综述。《IEEE地球科学与遥感汇刊》**46**，1301–1312（2008）。

23. Horstrand, P. 等 一种基于高光谱传感器的无人机平台，用于图像采集和机载处理。《IEEE开放期刊》**7**, 66919–66938 (2019).

24. Li, J. 等 使用衍射网络实现光谱编码单像素机器视觉。《科学进展》 **7**, eabd7690 (2021).

25. Yang, X. 等 通过离焦编码和傅里叶成像网络实现快照式多光谱成像。载于《光学前沿》 + *2025*年激光科学会议（*FiO, LS*） FM3C.5（Optica出版集团，2025）。doi:10.1364/FIO.2025.FM3C.5.

26. Lin, C.-H., Huang, S.-H., Lin, T.-H. & Wu, P. C. 基于凸/深度（CODE）小样本学习理论的超表面赋能快照式高光谱成像。《自然·通讯》**14**, 6979 (2023).

27. Yako, M. 等 基于兼容CMOS的法布里–珀罗滤波器随机阵列的视频速率高光谱相机。*Nat. Photonics* **17**, 218–223 (2023).

28. Rahman, M. S. S., Yang, X., Li, J., Bai, B. & Ozcan, A. 利用空间非相干衍射处理器实现通用线性强度变换。《光：科学与应用》 **12**, 195 (2023).

29. Helber, P., Bischke, B., Dengel, A. & Borth, D. EuroSAT：一个用于土地利用和土地覆盖分类的新型数据集与深度学习基准。《*IEEE*地球观测与遥感应用精选主题期刊》 **12**, 2217–2226 (2019).

30. Ma, G. 等 利用部分相干光实现单向成像。*Adv. Photonics Nexus***3**, 066008 (2024).

31. Huang, L., Chen, H., Liu, T. & Ozcan, A. 利用物理一致性进行全息图重建的自监督学习。《自然·机器智能》 **5**, 895–907 (2023).

32. Shen, C.-Y. 等 使用波长复用衍射光学处理器进行多平面定量相位成像。《先进光子学》 **6**, 056003 (2024)。

33. Shen, C.-Y. 等 波长复用衍射光学信息存储与图像投影。《先进光子学》 **8**, 046011 (2026)。

34. Li, J. 等 使用波长复用衍射光学网络实现大规模并行通用线性变换。《先进光子学》 **5**, 016003 (2023)。

35. Mengu, D. 等 抗失调衍射光学网络。《纳米光子学》 **9**, 4207–4219 (2020)。

36. Kingma, D. P. & Ba, J. Adam：一种随机优化方法。预印本, https://doi.org/10.48550/arXiv.1412.6980（2017）。

37. Nair, V. & Hinton, G. E. 修正线性单元可改进受限玻尔兹曼机。载于 第 *27* 届国际机器学习会议论文集，807–814（2010）。

38. Paszke, A. 等 PyTorch：一种命令式、高性能深度学习库。载于神经信息处理系统进展第 32 卷，8024–8035 页（2019）。

39. Li, J., Mengu, D., Luo, Y., Rivenson, Y. & Ozcan, A. 衍射光学神经网络中的类别特定差分检测可提高推理准确率。《先进光子学》 **1**, 046001（2019）。

40. Mengu, D., Zhao, Y., Tabassum, A., Jarrahi, M. & Ozcan, A. 衍射互连：利用衍射网络实现全光置换运算。《纳米光子学》 **12**, 905–923（2023）。

41. SCHOTT AG. N-BK7 色散数据，SCHOTT Zemax 目录 2017-01-20b。*RefractiveIndex.INFO* https://refractiveindex.info/download/data/2017/schott_2017-01- 20b.agf（2017）。

42. Loshchilov, I. 与 Hutter, F.《解耦权重衰减正则化》。预印本，https://doi.org/10.48550/arXiv.1711.05101（2019）。