四元光计算芯片，用于并行高维数据处理

Songyue Liu[1], Qi Lu[1], Yuan Zhong[1], Yuru Li[2], Meng Xiang[3], Zhaohui Li[4], Chao Lu[5], Yikai Su[1*]，以及 Lu Sun[1*]

[1]上海交通大学电子工程系光子学与通信国家重点实验室，中国上海 200240[2]中山大学微电子科学与技术学院，中国珠海 519000[3]宁夏大学教育部集成传感与通信光子技术重点

摘要

光计算芯片凭借其高计算密度、低能耗和紧凑的占用空间，已成为一种变革性的计算技术。尽管实数和复数计算芯片已经得到充分发展，但它们在表示高维数据方面存在根本性限制，严重制约了其在现代信号处理中的适用性。四元数能够直接处理三维和四维数据，为数据分析和人工智能中的高维处理提供支持。在此，我们首次展示了一种四元数光计算芯片（QOCC），并在多个典型应用场景中对其性能进行了基准测试：三维点云处理、RGB色度

变换，以及用于彩色图像识别的四元数卷积神经网络。 QOCC通过波分复用利用光的高并行性，经由多个光波长通道同时处理高维数据。与电子计算对应方案相比，我们的QOCC实现了更高的计算保真度（均方根误差< 0.035），并显著降低了计算负载（降低2/3）。它为下一代光计算铺平了道路，克服了传统计算系统在高维数据处理方面的局限。

引言

人工智能（AI）在重塑现代生活的各个方面取得了显著成功。人工智能的训练和推理过程对计算芯片提出了巨大的计算需求。借助实数和复数计算芯片，人工智能在从Pattern Recognition[1-5]、语音处理[6-10]和决策[11-13]到雷达[14-16]、无线通信[17-21]和机器人学[22,23]等任务中取得了重大进展。然而，这些计算架构仍受限于数据的低维表示，限制了其在张量处理中的应用，例如三维（3D）点云重建和彩色图像识别。超复代数，尤其是四元数，在处理基于高维数据的任务方面具有非凡优势。四元数计算能够自然表示四维数据，显著减少处理过程中的连续乘法次数，并稳健地保留数据不同维度之间的相关性，因此已广泛应用于现代信号处理。近期，基于四元数的神经网络取得了 avances，正在革新信息技术的各个方面，包括旋转动力学建模、多通道数据融合和高维特征提取

在计算机视觉等领域[24,25], 自主导航[26,27], 先进信号处理[28,29], 等等。

尽管四元数计算能够更高效地处理高维数据，但仍需要大量乘加运算（MAC）和频繁的数据访问，这给现有电子计算硬件带来了巨大的计算负担。近年来，硅光子集成电路已成为片上光计算和神经网络加速的有前景平台[30-44]，在带宽、并行性[40]和能效[43,44]方面展现出无可比拟的优势。尽管人们已利用微环谐振器[33,45]（MRRs）或马赫-曾德尔干涉仪阵列[46,47], 实现了实数和复数计算芯片，但片上四元数计算的发展仍基本处于探索阶段。必须开发新的光计算范式和架构，以实现高精度计算，并为多样化的四元数值计算和神经网络提供高效加速。弥合硅光子学与四元数计算之间的鸿沟，对于实现能够并行处理高维数据的高性能光子处理器至关重要。

在本研究中，我们首次展示了四元数光计算芯片（QOCC）。该芯片在硅光子平台上集成了包括输入准备、权重乘法和输出生成在内的所有关键计算功能。不同于传统电子实现中繁琐的算术运算需要按顺序执行，QOCC通过利用波分复用（WDM）提高并行性，并借助光固有的低延迟特性实现高速计算。此外，我们还展示了其在多种高维数据处理应用中的可行性，例如三维点云重建、RGB色彩变换和彩色图像识别。在所有这些任务中，我们的QOCC均表现出较高的计算保真度（均方根误差（RMSE）< 0.035）和较低的

计算负载（相比电子处理器为1/3）。它还具备9.2 Gbaud的数据传输吞吐量和331.2 每秒十亿次运算的计算速度，显著超越传统电子四元数计算芯片。我们的工作为各种高影响力的实时应用开辟了新途径，应用范围从自动驾驶车辆感知[48-50] 到高精度目标检测[51-55] 以及大规模无人机群协同[56-60]。

结果

工作原理

四元数是哈密顿首次提出的复数扩展，广泛用于表示和处理高维数据。四元数独特的代数性质包括但不限于：将旋转紧凑地表示为单次运算、避免万向节锁[61], 、通过乘法高效组合四元数分量、支持平滑插值[62], ，以及天然保留维度间的相关性。其中，保留数据不同维度之间相关性的能力，对于执行三维旋转和彩色图像处理等任务至关重要。例如，传统的实数和复数方法会分别处理彩色图像的红、绿、蓝通道，从而忽略通道间相关性中蕴含的信息。然而，基于四元数的计算能够利用这类信息，提高彩色图像识别的精度。因此，四元数广泛应用于许多现代实时控制和高级信号处理系统，包括无人机姿态控制、多传感器融合辅助自动驾驶，以及动态环境中的彩色图像识别，如图 1a 示意图所示。显然，对于处理这些

任务。

在数学上，四元数 q 可表示为 $q = q_0 + q_1 i + q_2 j + q_3 k$，其中虚数单位 i、j、k 满足：$i^2 = j^2 = k^2 = ijk = -1$，$ij = -ji = k$。对于三维空间中的任意一点，都可以将其写成四元数 $p = (0, p)$，其中 p 为位置向量。下文中，除非另有说明，斜体字母表示四元数量，粗体字母表示向量或矩阵。点的旋转和缩放等变换（图 1b）可用四元数算子 q 表示。经过变换 q 后的点 p' 与原始点 p 的关系为 $p' = qpq^{-1}$。这种运算也可以用变换矩阵 R(q) 表示（更多详情请参见补充说明 1）

$$\mathbf{R}(q) = \begin{bmatrix} 1-2q_2^2-2q_3^2 & 2q_1q_2-2q_0q_3 & 2q_0q_3+2q_0q_2 \\ 2q_1q_2+2q_0q_3 & 1-2q_1^2-2q_3^2 & 2q_2q_3-2q_0q_1 \\ 2q_1q_3-2q_0q_2 & 2q_2q_3+2q_0q_1 & 1-2q_1^2-2q_2^2 \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix}, \quad (1)$$

其中，式（1）最右侧的矩阵是用于配置 QOCC 的权重矩阵 w。基于 QOCC 执行四元数计算并加速四元数卷积神经网络（QCNN）的实验装置如图 1c 所示。四元数 $p = 0 + x_i i + y_i j + z_i k$ 表示的三维信号，其分量 x_i、y_i 和 z_i 被分配到功率相等但波长不同的光载波上（λ_1、λ_2、λ_3）。随后，使用密集波分复用器（DWDM）对多波长光进行复用。信号编码由任意波形发生器（AWG）实现，它将 n 个点的坐标映射为序列 $S_i = [x_{i1}、y_{i1}、z_{i1}、x_{i2}、y_{i2}、z_{i2} \ldots x_{in}、y_{in}、z_{in}]$。编码后的电信号经电放大器（EA）放大，然后输入马赫 - 曾德尔调制器（MZM），对三波长复用光信号进行高速电光调制，从而生成数据集 Si 的三个编码副本。随后，光波通过光

时延线（OTDL），其延迟步长（相邻波长之间）等于 $\mathbf{S_i}$ 的符号持续时间，从而有效实现时分与波分交织。时延 Δt 设为调制信号 f_b 波特率的倒数（即 $\Delta t = 1/f_b$），这样，在同步各信号周期后，可确保三个波长上编码的信号实现精确的时间对齐。随后，将时分与波分交织光信号耦合至 QOCC 进行四元数处理。所提出的 QOCC 由三个微环谐振器分支构成，每个分支集成三个采用下载式结构的热调谐微环谐振器（QOCC 的设计参数详见补充说明2）。通过精确控制施加到各个微加热器的功率，可动态配置 QOCC 中每个微环的工作权重 w_{ij}。QOCC 输出的光信号经平衡光电探测器（BPD）转换回电信号，从而以高度并行的方式完成乘累加运算。QOCC 的输出序列中包含一些冗余数据，在进行特征提取等后续处理前应将其消除。将序列划分为每组包含三个时隙的子集后，只有第三个时隙中的数据会对四元数计算的最终结果产生贡献，如图 1c 中黑色虚线框所示（详情见补充说明3）。因此，我们保留第三个时隙中的数据，并将其作为 $\mathbf{S_o} = [(x_{o1}, y_{o1}, z_{o1}), (x_{o1}, y_{o1}, z_{o1}) \ldots (x_{on}, y_{on}, z_{on})]$ 输出。

图 1d 和 1e 展示了制造完成的器件以及封装后的 QOCC 的显微镜图像。三个波长承载的光信号同时通过微环谐振器阵列传播，并受到不同权重因子的作用。随后，传输信号经由平衡光电探测器非相干叠加，从而实现并行光子计算。通过调节微环谐振器组，可配置 [-1,1] 之间的权重。例如，通过控制工作波长与

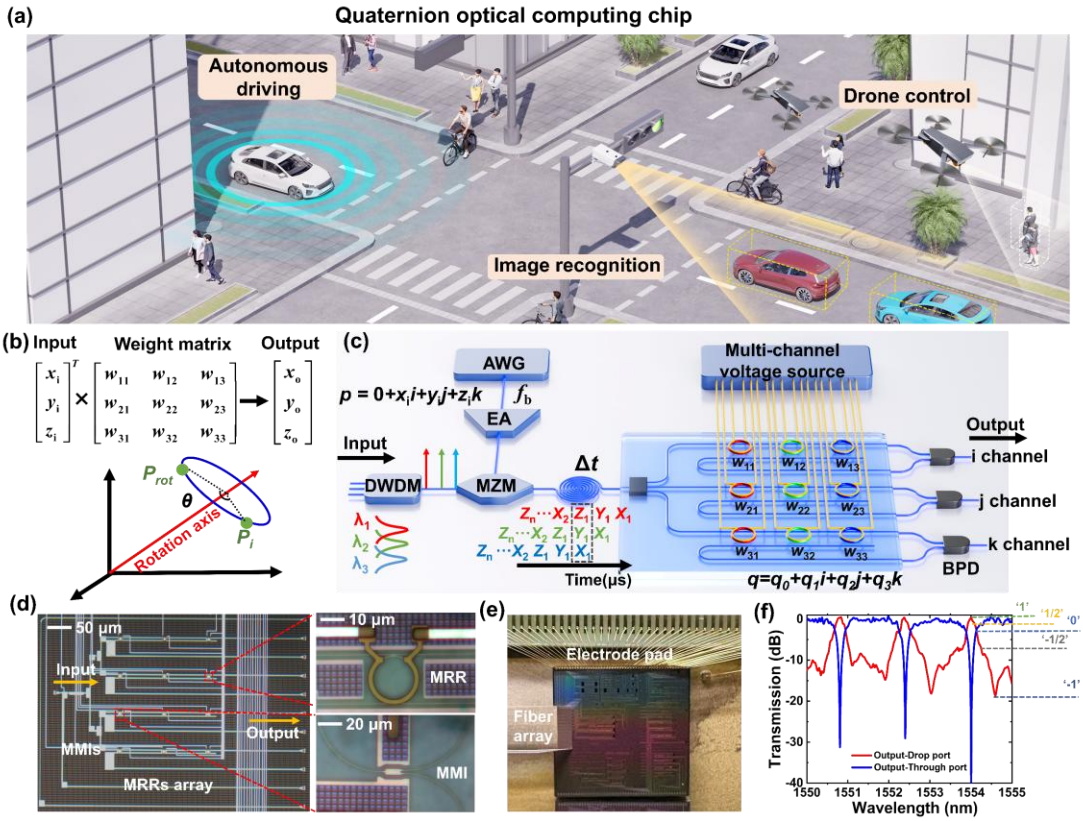


图 1| QOCC赋能高维信号处理应用。 a, QOCC的典型应用，包括自动驾驶中的地图重建、复杂环境中的彩色图像识别以及无人机实时控制。 **b,** 三维空间中由四元数表示的一个点及其等效矩阵形式上的运算。 **四元数计算实验装置示意图。密集波分复用器：密集波分复用器。任意波形发生器：任意波形发生器。马赫-曾德尔调制器：马赫-曾德尔调制器。电放大器：电放大器。平衡光电探测器：平衡光电探测器。 d、e,** 片上器件的显微镜图像（**d**）以及经过光电封装后的芯片图像（**e**）。**f,** MRR在 [-1,1] 范围内通过谐振失谐实现的权重配置。

之间的失谐，可获得 1、1/2、0、-1/2 和 1 的权重。

resMRR的共振波长，如图1f中的虚线所示（详见补充

注释2）。

三维点云处理

三维点云通过对三维空间中的空间坐标进行采样，以离散形式表示物体的几何结构。它们广泛应用于机器人学、自动驾驶、遥感、虚拟现实与增强现实等多个领域。在保持物体结构完整性的同时对其三维点云进行变换，需要先进的

计算模型和硬件加速器，以确保高保真度、高效率和可扩展性。在此，我们使用 QOCC 进行三维点云处理。实验装置如图2a所示。激光阵列产生了三个非相干光载波，其波长分别为 1550.8 nm、1552.4 nm 和 1554 nm；这些光载波经密集波分复用后，输入马赫-曾德尔调制器进行电光转换。在信号编码过程中，三维点云中从点 1 到点 n的空间坐标被串行化为序列 $\mathbf{S} = [x_1, y_1, z_1, x_2, y_2, z_2 \ldots x_n、y_n、z_n]$，随后输入任意波形发生器。任意波形发生器以 9.2 GHz 的频率运行，每个数据点的脉冲宽度为 108 ps。在此，我们选用色散为 17 ps/(km·nm) 的 8 km OTDL，在工作波长相邻的信号之间产生 1 位时间延迟（$\Delta t = 108$ ps）。随后，脉冲光信号被耦合至 QOCC，以进行四元数计算。最后，通过提取输出序列中每三位中的一位来获得计算结果，具体说明见补充说明3。

在这项工作中，我们使用两个模型展示了三维点云处理：一个包含2500个点的“SJTU”字母模型和一个包含10000个点的“狗”模型。“SJTU”字母模型和“狗”模型的变换分别由四元数 $q = 1 - 0.4\mathbf{i} - 0.1\mathbf{j} + 0.3\mathbf{k}$ 和 $q = 1 + 0\mathbf{i} + 0\mathbf{j} - 1\mathbf{k}$ 描述。由于这些四元数不是单位四元数（即 $|q| \neq 1$），这意味着变换同时包含旋转和缩放操作。图2b展示了“SJTU”字母模型的实验结果。左图显示原始三维点云。中图对比了点云 x、y、z 坐标的理想波形（红色曲线）和实验波形（蓝色曲线），二者表现出高度一致性。右图展示了点云在三维空间中经过变换后的理想结果和实验结果。可以看出，使用QOCC处理的点云形态与计算得到的

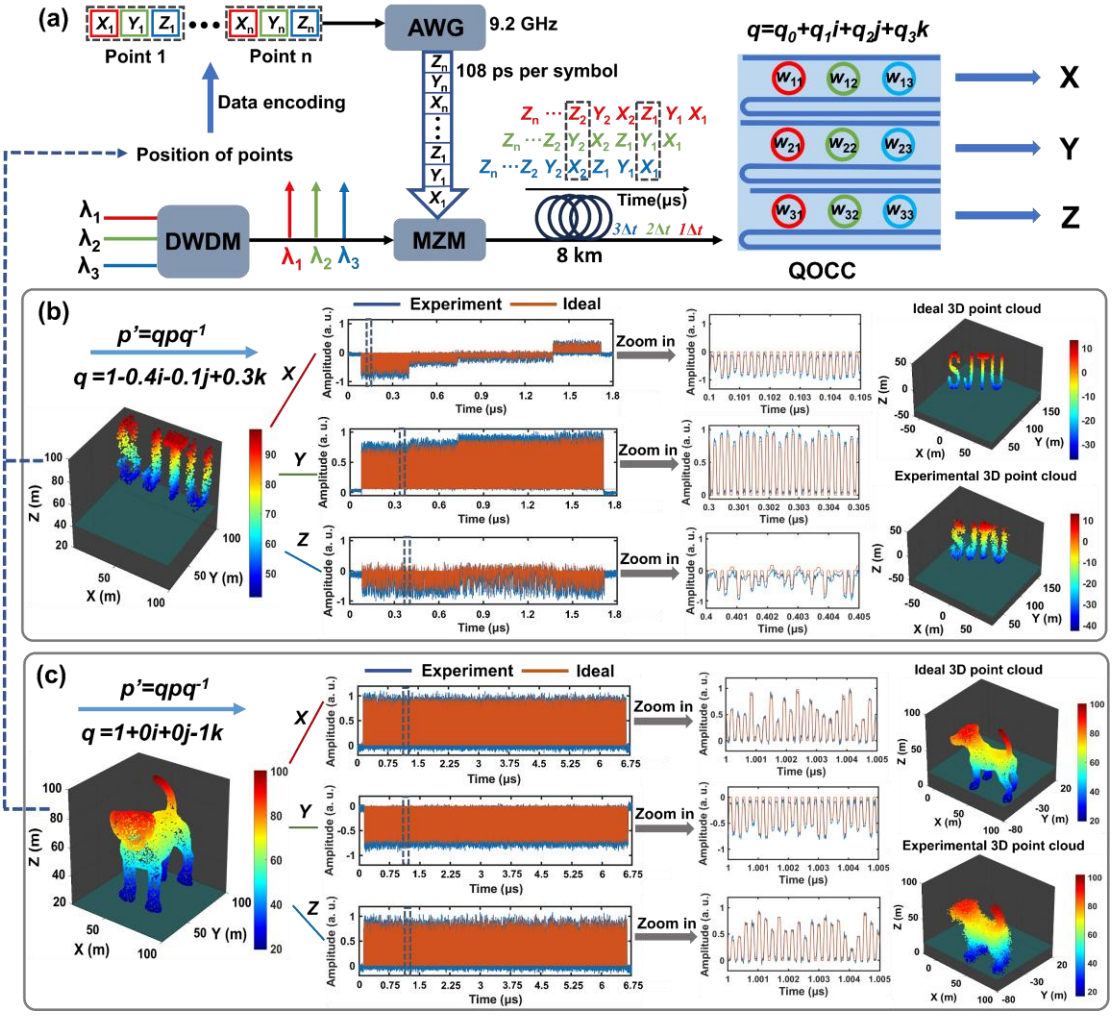


图2| 3：QOCC 进行的三维点云处理。 a, 基于四元数的三维点云处理实验中的数据流示意图。点云坐标信息编码到三个光学波长上，调制速率为 9.2 GHz。 光信号经密集波分复用器复用，并通过 8 km OTDL 进行时域波长拉伸，随后输入、QOCC 进行四元数运算。计算结果由平衡光电探测器读出。**b、c,** “b”字母模型和“狗”模型（c）的三维点云变换实验结果。左图：原始三维点云图像。中图：变换后三维点云坐标的理想波形（红色曲线）与实验波形（蓝色曲线）的对比（含放大视图）。右图：变换后输出三维点云图像的对比可视化。

预测结果。图2c展示了“狗”模型的实验结果。 与“SJTU”字母模型相比，“狗”模型具有更高的点密度，呈现出更精细的形态细节（左图）。变换后理想点云（红色曲线）和实验点云（蓝色曲线）的波形显示在中图中，而变换后的点云本身显示在右图中，证明了QOCC处理的高保真度。我们还进行了

实现了“SJTU”字母模型从 q = 1 + 0i + 0j + 0k（恒等操作）到 q = 1 - 0.4i - 0.1j+ 0.3k 的连续变换，以及“Dog”模型从 q = 1 + 0i + 0j + 0k（恒等操作）到 q = 1 + 0i + 0j - 1k 的连续变换。变换过程中三维点云的快照见补充说明4。对于这两个模型，理论结果与测量结果之间的均方根误差均小于 0.035，QOCC 的定位精度达到 5 位，证明了 QOCC 在三维点云处理中的可行性（详情见补充说明5）。

与传统电子芯片中乘累加运算按顺序执行的三维点云处理方案相比，所提出的 QOCC 通过波分复用实现了高度并行。结合四元数计算的固有优势，它将计算复杂度降低了 2/3（一次传播中执行三重运算）。QOCC 的三维点云处理能力可在无人机集群控制等先进应用中发挥关键作用；在这些应用中，实时、大规模的三维空间制图和无人机姿态估计必不可少。

RGB色彩变换

RGB 色度变换是彩色图像处理中的核心操作，它在三维色彩空间中对红、绿、蓝分量进行联合处理。将 RGB 像素表示为四元数（即$c = 0 + R\mathbf{i} + G\mathbf{j} + B\mathbf{k}$，其中$R$、$G$和$B$分别表示红、绿、蓝通道的强度）后，色彩变换便可表示为用于 RGB 色彩空间旋转的单位四元数q（$|q| = 1$）。这种方法天然保留了色彩空间中固有的通道间相关性和几何连续性，避免了传统逐通道处理造成的信息损失。最重要的是，它遵循 RGB 能量守恒定律，

即，$R^2 + G^2 + B^2$ 在变换过程中保持不变。因此，高效实现 RGB 色度变换对于颜色校正、增强和风格变换等高保真彩色图像处理任务至关重要，尤其是在高维和数据密集型场景中。

基于四元数的 RGB 色度变换实验流程如图 3a 所示。由于示波器存储深度有限，一幅包含 RGB 分量、分辨率为 512 × 512 像素的饱和彩色图像被划分为四个子集（$512 \times 128 \times 3 \times 4$ 个数据块）。输入信号 R_i、G_i、B_i 被分配给功率相同但波长不同的光载波 λ_1、λ_2、λ_3。在信号编码过程中，按照像素顺序将彩色图像的像素值循环排列为 $S = [R_1, G_1, B_1, R_2 \ldots R_n, G_n, B_n]$，并依次输入任意波形发生器以生成电信号。随后，信号被送入马赫-曾德尔调制器进行电光转换。调制后的光信号被送入 QOCC 进行四元数计算。通过调节微环谐振器的谐振，可自适应地重新配置不同颜色通道的权重因子，从而完成 RGB 色彩空间旋转。随后，对 QOCC 输出的信号进行累加，并由平衡光电探测器采集。与三维点云处理类似，在最终用于生成变换后的彩色图像之前，按照补充说明3中所述的准则移除输出序列中的冗余数据。

这里我们展示了使用 QOCC 对彩色图像进行 RGB 色度变换的两个示例。第一个变换是 RGB 通道的循环置换（即 R→G→B→R），由单位四元数 $q = 0.5 + 0.5\mathbf{i} + 0.5\mathbf{j} + 0.5\mathbf{k}$ 表示，可在色彩空间中实现 120° 旋转。图 3b 对比了彩色图像经过 120° 旋转后的理想结果与实验结果。变换后的彩色图像中红、绿、蓝通道的波形

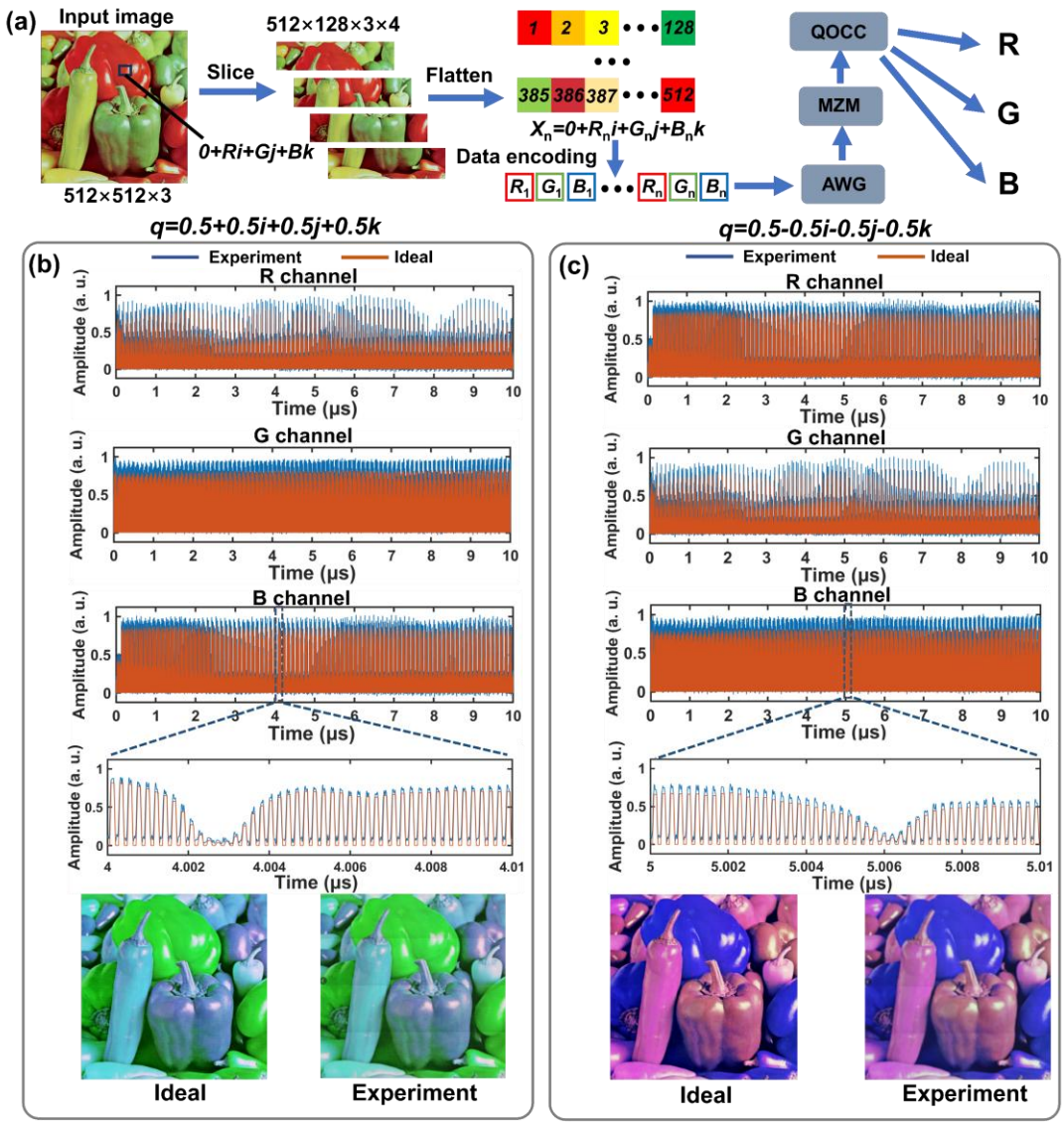


图 3| QOCC 执行的 RGB 色度变换。**a,** 基于四元数的 RGB 色度变换实验中的数据流示意图。**b、c,** 以$q = 0.5 + 0.5\mathbf{i} + 0.5\mathbf{j} + 0.5\mathbf{k}$ （**b**）和$q = 0.5 - 0.5\mathbf{i} - 0.5\mathbf{j} - 0.5\mathbf{k}$（**c**）表示的 RGB 色度变换实验结果。上图：变换后彩色图像中红、绿、蓝通道的理想波形（红色曲线）与实验波形（蓝色曲线）对比（含放大视图）。 下图：RGB 色度变换后理想彩色图像与实验彩色图像的对比可视化。

显示在上方图中，而这两幅图像的对比可视化结果则显示在下方图中。理想结果与实验结果高度吻合，证实了 QOCC 在执行 RGB 色度变换时具有很高的保真度。第二个变换是第一个变换的逆变换，可在色彩空间中实现 -120° 旋转（即 R→B→G→R），由单位四元数 $q = 0.5 - 0.5\mathbf{i} - 0.5\mathbf{j} - 0.5\mathbf{k}$ 表示。该变换的理想结果与实验结果

这些变换如图3c所示，同时也展示了QOCC的高计算保真度（详见补充说明6）。对于这两种变换，实验波形与理想波形之间的均方根误差均低于0.035，证明了基于四元数的方法适用于任何三维空间（无论是实空间还是色彩空间）。与分别处理RGB通道的传统电子处理器相比，QOCC利用四元数计算的优势以及WDM赋予的光学并行性，在单个步骤中完成色彩变换，在成功保留通道间相关性的同时，显著提高了计算效率（提升3倍）。

用于彩色图像识别的四元数值卷积

四元数卷积（QC）是四元数卷积神经网络的核心计算操作，而四元数卷积神经网络在彩色图像识别等人工智能任务中发挥着重要作用。传统的实值和复值卷积神经网络（卷积神经网络）在彩色图像处理中存在固有局限，因为卷积核通常独立聚合各通道的输出，因而无法捕捉内在的通道间相关性。相比之下，四元数卷积神经网络在输入层不将输入图像分解为独立的RGB通道，从而保留了通道之间的关联信息。更重要的是，传统卷积需要处理三个独立的颜色通道，而四元数卷积将彩色像素视为统一实体，并在一步内完成运算，大幅降低了计算负担。通过充分利用四元数计算的独特性质，四元数卷积神经网络增强了神经网络提取彩色图像特征的能力，从而能够高效执行彩色图像识别和分割等高吞吐量视觉任务。

为实现四元数卷积运算，我们将二维彩色图像的像素定义为四元数矩阵 $\boldsymbol{X}= [x_{nn'}]_{N\times N}$，其中 N 表示图像尺寸，且 $x_{nn'} =0 + R_{nn'}\mathbf{i} + G_{nn'}\mathbf{j} + B_{nn'}\mathbf{k}$ 是一个四元数。

类似地，四元数卷积核可以表示为四元数矩阵 $\boldsymbol{W}=[w_m]_{L\times L}$，其中 L 表示卷积核尺寸，且 w_m 是四元数元素。因此，四元数卷积过程可表示为

$$\boldsymbol{X}\otimes\boldsymbol{W}=\boldsymbol{Y}=[y_{kk'}]_{(N-L+1)\times(N-L+1)}, \tag{2}$$

其中，$\otimes$ 表示四元数卷积运算，$y_{kk'}$是 卷积后特征图像中的四元数值像素。经过一些繁琐但直接的推导，我们发现$y_{kk'}$ 等价于

（详见 补充说明7 ）

$$y_{kk'}=\sum_{l=1}^{L}\sum_{l'=1}^{L}\begin{bmatrix} w_{m1} & w_{m2} & w_{m3} \\ w_{m3} & w_{m1} & w_{m2} \\ w_{m2} & w_{m3} & w_{m1} \end{bmatrix} x_{(k+l)(k'+l')}, \tag{3}$$

其中，w_{m1}、w_{m2}和w_{m3} 是将$q=w_m$ 代入公式（1）后得到的旋转矩阵元素).

接下来，我们展示如何使用 QOCC 加速四元数计算。其工作原理如图4所示。在我们的实现中，2×2 卷积核用于图像卷积，不同的卷积核实现不同的卷积函数，以提升特征提取的效果。卷积核从左向右、再从上向下在输入图像上滑动以执行计算，其过程与实数和复数卷积相同，但此处采用的是四元数运算。如图4a所示，首先将输入的 32×32 像素彩色图像切分为四个 31×31 像素子图像（图4b）。随后，将四个子图像展平，生成四个四元数序列。例如，第一个和第四个子图像对应的序列分别为 $S_1=[x_1, x_2, x_3, ..., x_{31}, x_{33}, x_{34}, ..., x_{989}, x_{990}, x_{991}]$ 和 $S_4=[x_{34}, x_{35}, x_{36}, ..., x_{64}, x_{66}, x_{67}, ..., x_{1012}, x_{1023}, x_{1024}]$, 。随后，将这些四元数序列转换为 RGB格式的 Data 1 - 4，以进行光信号编码（图4c）。图4d展示了四元数卷积的实验装置。光经过密集波分复用器后，均等分为四路，分别传输至马赫-曾德尔调制器 1 - 4。子图像数据馈入任意波形发生器，以驱动马赫-曾德尔调制器

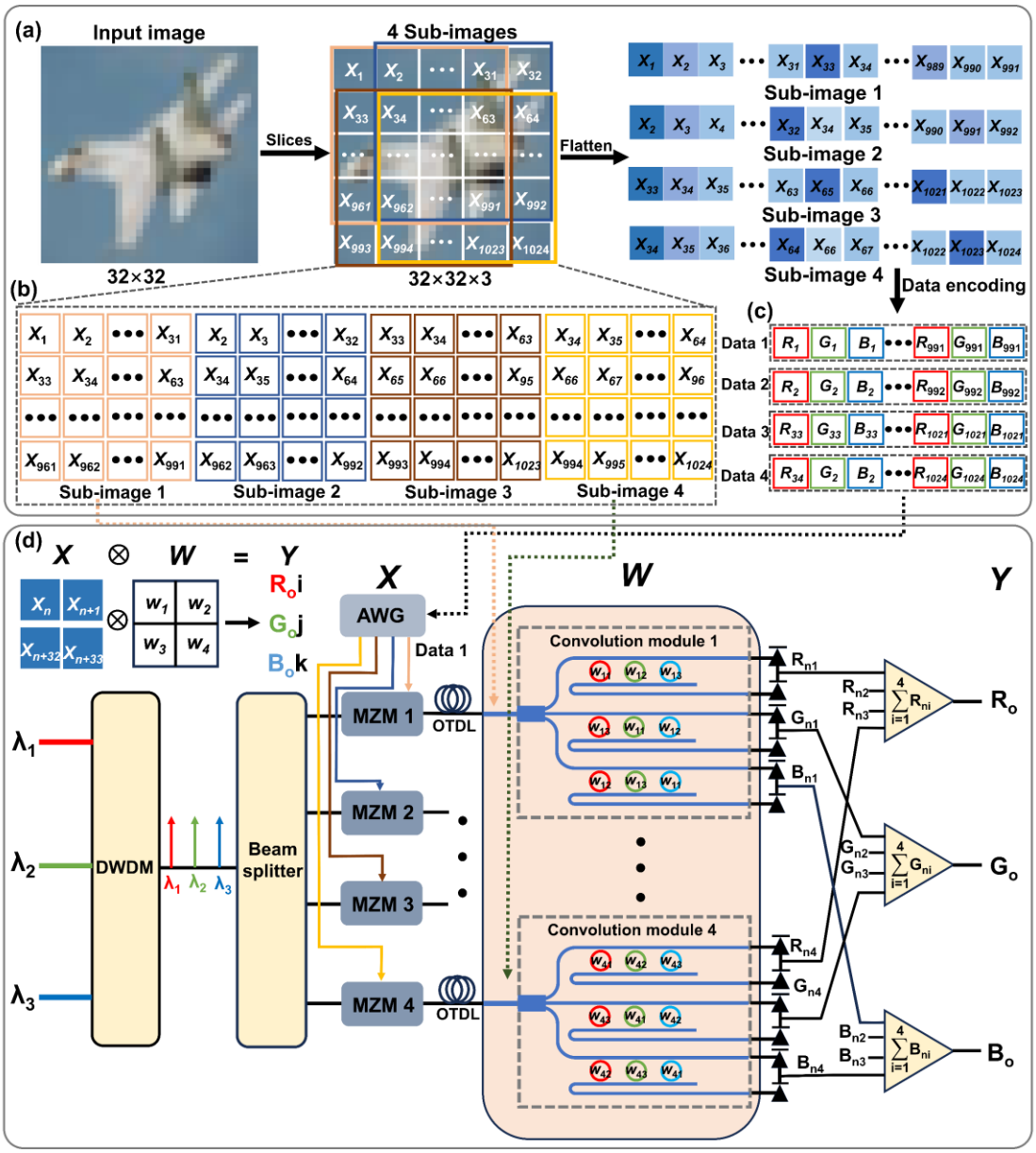


图 4| 基于量子卷积的彩色图像特征提取的工作原理。**a,** 彩色图像的预处理。将原始图像切分为四个子图像，然后展平为四元数序列以进行数据存储。**b、c,** 四个子图像采用四元数（b）和 RGB（c）格式的数据结构。 展平后的像素序列作为任意波形发生器的输入。**d,** 四元数卷积实验装置示意图。

（Data 1 - 4 分别分配给马赫-曾德尔调制器 1 - 4），其输出信号通过 OTDL 进行时间对齐。随后，信号被送入 QOCC 进行四元数卷积。量子卷积层由四个并行计算模块（卷积模块 1 - 4 ）组成，每个模块独立处理来自相应马赫 - 曾德尔调制器（马赫 - 曾德尔调制器 1 - 4 ）的数据流。根据公式（3）重新配置 QOCC 中微环谐振器的权重因子，从而定义卷积核。

计算结果由平衡光电探测器检测，并通过加权求和得到提取特征的 RGB值 $\boldsymbol{Y}$ (详见 补充说明7 ）。为评估 QOCC 的性能，我们将其用于四元数卷积神经网络，基于 CIFAR-5数据集开展彩色图像识别实验。该数据集包含五个类别：飞机、汽车、马、船和卡车。如图 5a 所示，所用的四元数卷积神经网络由一个输入层、四个卷积层和两个全连接层组成。该网络使用包含 50000 张图像的 CIFAR-5训练集进行训练，随后在从 CIFAR-5测试集随机选取的 500 张图像子集上进行评估。四元数卷积神经网络的输入层接收32×32×3 RGB彩色图像，图像随后被展平并输入第一个四元数卷积层。该量子卷积层使用一个2×2 卷积核处理32×32像素的 RGB彩色图像。第一个四元数卷积层的卷积核 W 具有相同的四元数元素（即 $w1 = w2 = w3 = w4 = w$），其变换矩阵 R(w) 如图 5b 所示。初始卷积后，获得了原始图像的31×31像素特征图。随后，这些特征经过池化激活，再输入后续的四元数卷积层进行进一步的特征提取[63]。尽管理论上增加四元数卷积层数可以提高四元数卷积神经网络的准确率，但在实际应用中，增加层数会导致光子处理器中的物理误差累积[64], ，包括相位误差、振幅误差、散射损耗和耦合损耗。为在四元数卷积神经网络的识别准确率与连续层间累积的实验误差之间取得平衡，我们采用了四层架构。其他四元数卷积层（QC 2–QC 4）的卷积核不固定，并在训练过程中更新。该四元数卷积神经网络采用基于前向传播和反向传播的端到端训练框架[65], ，并使用 Adam算法进行权重优化（详情见补充说明8）。完成所有四元数卷积运算后，应用非线性 ReLU激活

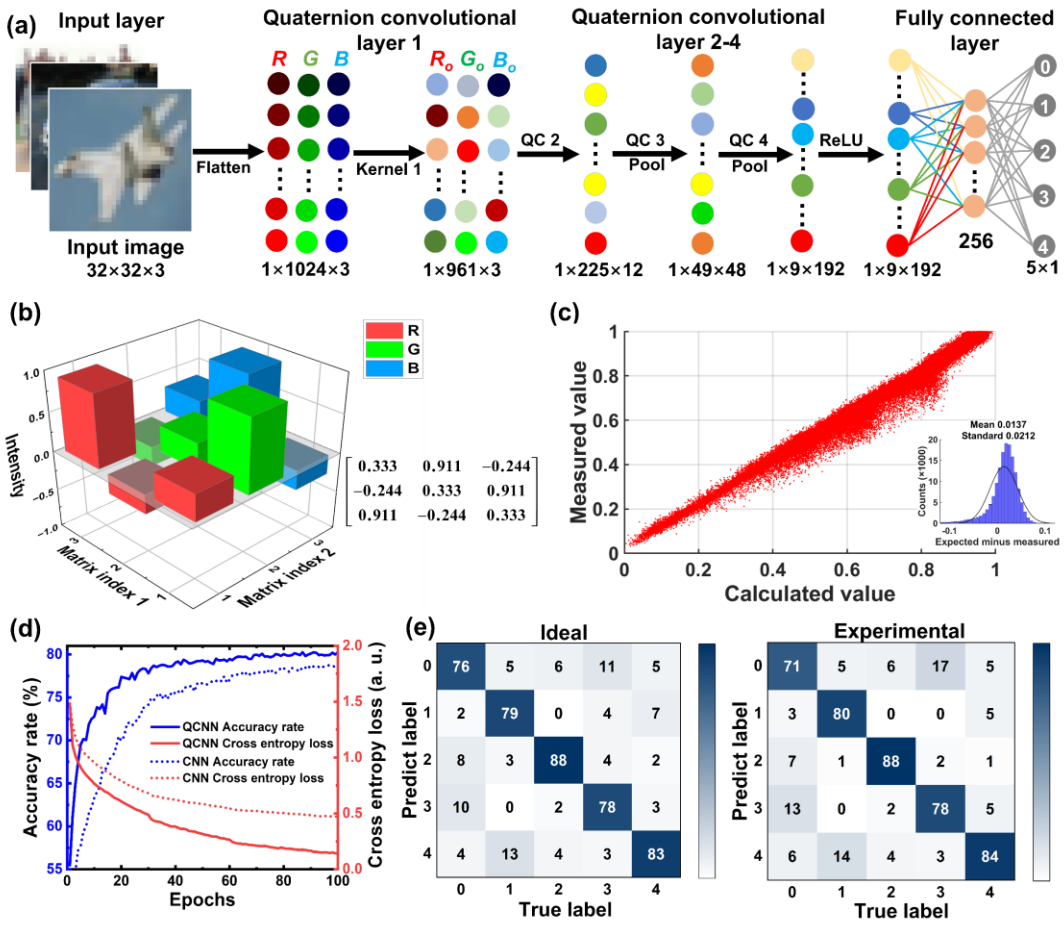


图 5| 使用 QOCC 对 CIFAR-5 数据集进行彩色图像识别。 **a,** 由四个带激活函数的卷积层和两个全连接层组成的 QCNN 架构。 QC、池化层和修正线性单元分别表示四元数卷积层、池化层 和修正线性单元。 **b,** 第一个四元数卷积核内四元数元素的权 重分布。 **c,** QOCC 的计算精度分析。插图中的直方图显示了计算误差分布，其标准差为 0.0212，对应的精度为 5 位。 **d,** CNN 和 QCNN 的识别准确率与交叉熵损失随训练轮次的变化。 **e,** 理想 QCNN 与实验 QCNN 所得混淆矩阵的比较。

函数，将1×9×192 特征图转换为 256 个特征向量，再将其输入全连接层。1×5 输出向量对应五个类别（飞机、汽车、马、船、卡车），预测类别为五个向量元素中数值最大的类别（更多详情请参阅 补充说明 9 ）。

图5c展示了电子计算机（横轴）和 QOCC（纵轴）得到的计算结果。为表征光计算方法的准确性，我们计算了计算机与

QOCC得到的计算结果（详情见 补充说明 9）。QOCC 的平均均方根误差为 0.035，平均值为 ~0.0137，标准差为 0.0212，位精度为 5 位。图5d比较了随着训练轮次增加，四元数卷积神经网络和传统卷积神经网络模型的理论识别准确率与交叉熵损失。经过 80 轮训练后，四元数卷积神经网络的准确率达到饱和值 82%，而卷积神经网络达到 78%。四元数卷积神经网络和卷积神经网络的交叉熵损失值分别收敛至 0.14 和 0.47。显然，四元数卷积神经网络的图像识别准确率比传统卷积神经网络高 ~5%，这可归因于四元数计算的通道间相关性保持特性。图5e展示了在计算机上得到的理想混淆矩阵，以及通过 QOCC 获得的 CIFAR-5数据集实验矩阵。光计算系统的分类准确率为 ~80.2%，比理想预测低约 2%。这种差异可能源于电噪声和光噪声、微环谐振器工作点漂移、掺铒光纤放大器产生的放大自发辐射噪声等（更多详情见 补充说明 9）。此外，在硅光子平台上实现的 QOCC，其计算速度优于传统电子芯片。在输入数据速率为 9.2 Gbaud 时，QOCC 对四元数计算的计算速度达到 9.2（数据 速率）× 3 （每个卷积模块的分支数）× 4 （卷积模块数）× 3 （波长数）= 331.2 每秒十亿次运算。单次四元数卷积运算的功耗约为 108 ps × 3 mW× 9 （每个卷积模块的微环谐振器数）× 4 （卷积模块数）= 11.6 pJ。QOCC 实现了高保真计算并提高了处理效率，为先进光子处理器提供了新的范式，可用于处理高维空间变换和彩色图像识别等计算密集型任务。

讨论

基于四元数的计算已成为现代人工智能中的强大工具，在无人机集群协同、彩色图像处理和目标识别等应用中展现出独特的计算优势。然而，在传统电子芯片中，四元数计算通常通过连续的乘累加运算实现，导致大量计算开销，阻碍了大规模部署。在此，我们提出并通过实验展示了一种基于MRR阵列的QOCC。借助光计算的高并行性和低延迟，该设计同时实现了高计算保真度（均方根误差 < 0.035）和低计算负载（与电子处理器相比仅为三分之一）。此外，该设计还实现了高达331.2 每秒十亿次运算的计算速度，以及四元数计算领域创纪录的低能耗——每次运算仅为11.6 pJ，在高维数据处理中展现出卓越性能。

作为概念验证演示，我们将 QOCC 应用于三种典型的四元数值计算任务，包括三维点云处理、RGB色彩变换以及基于四元数卷积神经网络的 CIFAR-5数据集分类。在所有这些应用中，QOCC 得出的结果与理想的电子计算基准仅有微小偏差，同时实现了更快的速度和更低的能耗。得益于对 RGB彩色图像通道间相关性的保留，基于 QOCC 的四元数卷积神经网络在彩色图像识别中的分类准确率成功达到 80% 以上，优于分别处理 RGB通道的传统卷积神经网络。此外，通过增加更多微环谐振器并采用波长复用，光子处理器架构可以轻松扩展。计算系统的性能仅受实验装置的限制，例如马赫-曾德尔调制器和平衡光电探测器的工作速度、任意波形发生器和示波器的带宽，以及电放大器和光放大器的噪声系数。这

光计算的并行性与四元数计算范式天然契合，使复杂的四元数运算（例如四元数计算）能够通过一次光传播和测量来执行。凭借这些优势，我们的工作展现出巨大潜力，有望在高维数据处理中超越传统电子芯片的计算能力，从而支持人工智能训练和推理等大规模计算任务。

展望未来，通过将分布式反馈（DFB）激光器、阵列波导光栅、低损耗 OTDL 和高速平衡光电探测器单片集成到 QOCC 中，可以实现更高密度的集成。采用相变材料或磁光材料对微环谐振器进行非易失性权重配置，将显著降低芯片在计算过程中的功耗。总体而言，所提出的 QOCC 为下一代光计算平台提供了理想的计算范式。随着集成光子技术日趋成熟，该系统有望实现更小的占用面积、更低的功耗和更高的带宽，达到或超过成熟电子计算平台的水平。

方法

实验装置

QOCC 由多项目晶圆代工厂（CUMEC，中国）采用标准 CSiP180Al 生产线工艺设计套件制造。氮化钛（TiN）沉积在波导顶部，作为微加热器。为抑制相邻器件之间的热串扰，制作了隔离沟槽。 QOCC 的详细信息 见 补充说明2。为表征制造出的器件，使用了可调谐连续波（CW）激光器（Santec TSL-770）和光电探测器（PD，Santec MPM-210）。首先，使用光纤偏振控制器（PC，Thorlabs FPC032）调节可调谐激光器发出的光的偏振状态。随后，横电偏振光

通过光栅耦合器耦合进出芯片。使用光功率计进行光学校准。将激光器的波长从 1500 nm 扫描至 1600 nm，以测量器件的传输光谱。

四元数计算的实验装置由商用光电器件和 QOCC 构成。DFB（IDPHOTONICS）激光阵列产生波长分别为 1550.8 nm、1552.4 nm 和 1554 nm 的光，输出功率可在 10 至 20 dBm 范围内调节。三个波长通道由 40 GHz 的马赫-曾德尔调制器（iXblue）调制，该调制器将任意波形发生器（KEYSIGHT M9502A）产生的 65 GSa/s 电波形转换为光信号。多通道可编程电源（TIME-TRANSFER）提供 64 个可调电压通道（范围为 0–10 V），用于配置 MRR阵列的权重。光信号通过 QOCC 后，由工作带宽超过 50 GHz 的平衡光电探测器（FINISAR BPDV2150RQ）转换为时域电信号。随后，使用采样率为 80 Ga/s 的高速示波器（TELEDYNE LECROY）对信号的时域波形进行采样。

QOCC 的权重配置

QOCC 的权重配置由自动化系统实现。在自动扫描算法中，控制程序使用 Python 3.12 编写，通过 socket 和 PyVISA 库分别与多通道电压源和示波器建立串行通信。该系统通过 UDP协议与多通道电压源进行通信，并同时通过 SCPI协议从示波器采集测量信号。在运行过程中，每个 MRR 均执行线性电压扫描，同时示波器记录每个

电压点对应的权重。随后，算法确定 MRR 的最优权重（见补充材料

注释 10 详情见后文）。

致谢

本研究部分经费由中国国家重点研发计划（2023YFB2905503（C. L.））、国家自然科学基金（62475146（L. S.）和62341508（Y. S.））以及教育部集成传感与通信光子技术重点实验室开放基金（L. S.）资助。

利益冲突

作者声明不存在利益冲突。

数据可用性

支持本研究结论的数据见补充

信息/源数据文件。本文附有源数据。

参考资料

1. Evans, C.G., O'Brien, J., Winfree, E. & Murugan, A. 非平衡自组装成核动力学中的模式识别。*Nature* **625**(2024). 2. Cherry, K.M. & Qian, L.L. 利用基于 DNA 的赢家通吃神经网络扩展分子模式识别。*Nature* **559**, 370-+ (2018). 3. Chen, C.S., 等。 基于双-2,6-[2-(2-氧代吲哚啉-3-亚基)丙二腈]萘衍生物/氧化锌纳米棒异质结构的光电探测器：结合机器视觉观察与人工智能模式识别。*Rsc Adv* **15**, 43763- 43773 (2025). 4. Fontanella, F., Colace, F., Molinara, M., Di Freca, A.S. & Stanco, F. 文化遗产领域的模式识别与人工智能技术。*Pattern Recogn Lett* **138**, 23-29 (2020). 5. “智能”手指通过模式识别感知路径。*Nature* **608**, 453-453 (2022). 6. Xavier, S.P. & John, S.P. 利用复值时空图卷积神经网络，从面部图像和语音信号中进行高级情感识别。*Signal Image VideoP* **19**, 4086-4099 (2025)。

7. Zaitsu, W., 等。文体计量学可以揭示人工智能作者身份，但人类难以做到：对比人类与七种日语大语言模型。PLOS ONE 20(2025). 8. Grace, B., Zhu, J.R., Dudakia, H., Ajao-Rotimi, F. & Colton, N. Ctrl plus Alt plus Conceive：人工智能时代的生育意识——大语言模型表现如何? Human Fertility 28(2025). 9. Déniz, C., 等。胸外科会诊中的人工智能：评估大语言模型与专家临床决策的一致性。Frontiers in Digital Health 7(2025). 10. Sharma, P., 等。麻醉学中的人工智能：比较 ChatGPT 与 Google Gemini 大语言模型在麻醉前教育中的效用：内容、可读性与情感分析。BMC Anesthesiology 25(2025). 11. Reichstein, M., 等。面向数据驱动地球系统科学的深度学习与过程理解。Nature 566, 195-204 (2019). 12. Mnih, V., 等。通过深度强化学习实现人类水平的控制。Nature 518, 529-533 (2015). 13. Jain, A., Tompson, J., LeCun, Y. & Bregler, C. MoDeep：一种利用运动特征进行人体姿态估计的深度学习框架。Lect Notes Comput Sc 9004, 302-315 (2015). 14. Silver, D., 等。在无需人类知识的情况下精通围棋。Nature 550, 354-+ (2017). 15. Cao, S., Zhao, C.X., Dong, J. & Fu, X.J. 基于深度学习与形态学网络，在复杂地理环境下检测合成孔径雷达图像中的船舶。Sensors 24, 4290 (2024). 16. Ghosh, T., Hazra, S., Malick, K. & Talukdar, T. 利用多时相 ALOS-2 L波段成像雷达数据进行永久散射体干涉测量，监测近年来印度西恒河三角洲的地表形变。Nat Hazards, 1-31 (2025). 17. Vieira, S.T., 等。Q-meter：一种基于深度学习情感分析的电信服务质量监测系统。Sensors 21, 1880 (2021). 18. Liu, Z. & Li, R. 无线通信与人工智能背景下基于多源传感器的青年群体注意力分析与信息传播效果。Scientific Programming 2022(2022). 19. Zhao, J.H., Gao, X., Wu, Z.Y., Zhang, Q.M. & Han, H.T. 轨道交通无线通信系统中的人工智能：现状、挑战与解决方案。Physical Communication 67(2024). 20. Gharghan, S.K. & Hashim, H.A. 利用无线通信与人工智能技术进行老年人跌倒检测的全面综述。测量 226(2024). 21. Gbenga-Ilori, A., Imoize, A.L., Noor, K. & Adebolu-Ololade, P.O. 人工智能赋能先进无线通信中的动态频谱接入：全面综述。Ai-Basel 6(2025). 22. Liu, C.X., 等。基于摩擦纳米发电机与微型热电发电机的多模态感知方法，用于水下机械臂的可靠抓取。测量 256, 118147 (2025). 23. Ma, D.X., Zhang, C., Xu, Q.F. & Zhou, G.H. 面向工业5.0人机协作装配的大规模与小规模模型融合驱动的主动式机器人操作控制。Robot Cim-Int Manuf 97, 103078 (2026). 24. Yang, L.Q., Miao, J.F., Jiang, T.X., Zhang, Y.L. & Kou, K.I. 随机四元数张量

用于彩色图像和彩色视频处理的UTV分解。Pattern Recogn 165, 111580 (2025). 25. Jiang, C., Wang, G., Zhang, D. & Guo, Z.W. 分裂四元数矩阵的奇异值分解及其在彩色图像处理中的应用。J Electron Imaging 34, 013054-013054 (2025). 26. Patil, P.V., Khan, M.K., Korulla, M., Nagarajan, V. & Sha, O.P. 使用四元数对自主水下航行器进行机动仿真。Defence Sci J 72, 290-307 (2022). 27. Zhang, H.X., Cao, Y.C., Shan, Q.H. & Sun, Y.K. 基于模型预测控制、意图数据和四元数船舶领域的海上自主水面船舶避碰。J Mar Sci Eng 13, 124 (2025). 28. Samad, M.A., et al. 四元数分数阶傅里叶变换：连接信号处理与概率论。Mathematics-Basel 13, 195 (2025). 29. Jabeen, D., Iqrar, T. & Khan, M.F. 在序列域中使用四元数复Hadamard变换进行多维信号处理。Electron Lett 54, 1435-+ (2018). 30. Meng, X.Y., et al. 基于多模干涉的紧凑型光学卷积处理单元。Nature Communications 14, 3000 (2023). 31. Wang, T.Y., et al. 一种每次乘法所需光子数少于1个的光学神经网络。Nature Communications 13, 123 (2022). 32. Huang, C.Y., Cui, L.N., Xia, H., Qiu, Y.P. & Ni, Q.Q. Twaron(R)织物在斜向冲击下低速冲击行为的数值研究。Rev Adv Mater Sci 60, 980-994 (2021). 33. Xu, X.Y., et al. 用于光学神经网络的11 TOPS光子卷积加速器。Nature 589, 44-+ (2021). 34. Chen, Y.T., et al. 用于高速视觉任务的全模拟光电芯片。Nature 623, 48-+(2023). 35. Nilforoshan, H., et al. 人类移动网络揭示大城市中日益加剧的隔离现象。Nature 624, 586-+ (2023). 36. Zhou, T.K., et al. 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). 37. Lin, X., et al. 使用衍射深度神经网络进行全光机器学习。Science 361, 1004-+ (2018). 38. Ashtiani, F., Geers, A.J. & Aflatouni, F. 用于图像分类的片上光子深度神经网络。Nature 606, 501-+ (2022). 39. Dong, B.W., et al. 使用连续时间数据和光子张量核心进行高维处理。Nat Photonics 17, 1080-+ (2023). 40. Feldmann, J., Youngblood, N., Wright, C.D., Bhaskaran, H. & Pernice, W.H.P. 具备自学习能力的全光脉冲神经突触网络。Nature 569, 208-+ (2019). 41. Feldmann, J., et al. 使用集成光子张量核心进行并行卷积处理（第589卷，第52页，2021年）。Nature 591, E13-E13 (2021). 42. Shen, Y.C., et al. 利用相干纳米光子电路进行深度学习。Nat Photonics 11, 441-+(2017). 43. Shastri, B.J., et al. 面向人工智能和神经形态计算的光子学。Nat Photonics 15, 102-114 (2021). 44. Wetzstein, G., et al. 利用深度光学和光子学进行人工智能推理。Nature

588, 39-47 (2020). 45. Bai, Y.P., et al. 用于特征提取和推理的TOPS速度复值卷积加速器。Nature Communications 16, 292 (2025). 46. Zhang, H., et al. 一种用于实现复值神经网络的光学神经芯片。Nature Communications 12, 457 (2021). 47. Huang, Y., et al. 基于可调单元和单内部移相器、易于扩展的光子张量核心。Laser & Photonics Reviews 17(2023). 48. Xu, X. & Fan, C.K. 自动驾驶汽车、风险认知与保险需求：中国的一项个人调查。Transport Res a-Pol 124, 549-556 (2019). 49. Wu, J.W. & Kim, S.T. 一种结合SOR与SCT模型的方法，用于探究中国公众对自动驾驶汽车的认知。Scientific Reports 15(2025). 50. Kang, I., Yang, J. & Kim, Z. 通过对新闻文章进行主题建模，揭示公众对自动驾驶汽车电池技术的认知。Scientific Reports 15(2025). 51. Ji, B.F., et al. BCP-YOLOv5：一种基于YOLOv5的牡丹花识别高精度目标检测模型。Technologies 13(2025). 52. Chen, H.M., Wang, H.F., Liu, Z.L., Gu, D.B. & Ye, W. HP3D-V2V：一种高精度三维目标检测车车协同感知算法。Sensors 24(2024). 53. Chen, H.M., Wang, H.F., Liu, Z.L., Gu, D.B. & Ye, W. HP3D-V2V：一种高精度三维目标检测车车协同感知算法（第24卷，2170，2024）。Sensors 24(2024). 54. Li, S.C. & Peng, X. DyAqua-YOLO：一种基于动态自适应架构的高精度实时水下目标检测模型。Front Mar Sci 12(2025). 55. Fan, R., et al. Tri-FRNet：一种用于高精度小目标检测的三路径特征融合网络。Eng Res Express 7(2025). 56. Qin, C.H. & Pournaras, E. 大规模无人机协同：用于时空感知的去中心化能量感知群体智能。Transport Res C-Emer 157(2023). 57. Tang, R.P., Tang, J.R., Talip, M.S.A., Aridas, N.K. & Xu, X.F. 榴莲园无人机群巡逻的增强型多智能体协同算法。Scientific Reports 15(2025). 58. Chen, W., Zhu, J.Y., Liu, J.J. & Guo, H.Z. 一种面向大规模无人机群的快速协同方法。J Netw Comput Appl 221(2024). 59. Yang, B., Ma, C.F. & Xia, X.F. 一种用于异构无人机群协同的关联模仿学习方法。IEEE Transactions on Emerging Topics in Computing 10, 1704-1716 (2022). 60. Yan, L.H., Zhou, J.Y. & Yang, K.D. 杂乱环境中用于无通信无人机群协同的控制感知轨迹预测。P Amer Contr Conf, 3520-3526 (2025). 61. Özcan, B., Kinli, F. & Kiraç, F. 四元数胶囊网络。Int C Patt Recog, 6858-6865 (2021). 62. Barr, A.H., Currin, B., Gabriel, S. & Hughes, J.F. 使用四元数在角速度约束下对方向进行平滑插值。Comp Graph 26, 313-320 (1992). 63. Glorot, X., Bordes, A. & Bengio, Y. 深度稀疏整流神经网络。载于《第十四届人工智能与统计国际会议论文集》，第15卷（编者：Geoffrey, G., David, D. & Miroslav, D.），315--323页（PMLR，机器学习研究论文集，2011）。64. Zhang, X., et al. 解锁高速与能效：集成卷积

基于薄膜铌酸锂的处理。*Laser & Photonics Reviews* **19**，2401583（2025）。65。
Rumelhart, D.E.、Hinton, G.E. 和 Williams, R.J. 通过反向传播误差学习表示。*Nature* **323**，533-536（1986）。