

基于光电忆阻器的双向一维压缩与量化 CNN 识别^{*}

盛喜乐 林引瑞 张 章

(合肥工业大学微电子学院 合肥 230601)

摘 要: 冯·诺伊曼式体系中感知与计算的物理分离导致大量数据在存储与处理单元间搬运,带来高延时与能耗,制约了边缘视觉应用的发展。为降低数据传输开销并提升前端计算效率,提出了一种基于光电忆阻器阵列与卷积神经网络(CNN)的混合识别方案,用于 64×64 手写字母的低带宽特征采集与分类。利用两组 4k 光电忆阻器阵列分别沿行、列方向对图像进行并行一维卷积压缩,得到两路 64 维特征并拼接为 128 维联合特征;在后端采用轻量级 CNN 并结合混合精度量化。实验结果显示,行列联合特征在未量化时可稳定达到约 $91\% \sim 93\%$ 的识别率,量化并微调后模型峰值识别率约为 88.8% 。

关键词: 光电忆阻器;感存算一体;图像识别;卷积神经网络

中图分类号: TN36 **文献标识码:** A **国家标准学科分类代码:** 510.4030

Bidirectional one-dimensional compression and quantization for
CNN recognition based on optoelectronic memristors

Sheng Xile Lin Yinrui Zhang Zhang

(School of Microelectronics, Hefei University of Technology, Hefei 230601, China)

Abstract: The physical separation of perception and computation in von Neumann architectures necessitates extensive data movement between storage and processing units, resulting in high latency and energy consumption that constrain the development of edge vision applications. To reduce data transmission overhead and enhance front-end computational efficiency, this study proposes a hybrid recognition scheme based on optoelectronic memristor arrays and convolutional neural networks (CNN) for low-bandwidth feature acquisition and classification of 64×64 handwritten characters. Two sets of 4k optoelectronic memristor arrays perform parallel one-dimensional convolutional compression along row and column directions respectively, yielding two 64-dimensional feature streams that are concatenated into a 128-dimensional joint feature. A lightweight CNN is employed in the backend, combined with mixed-precision quantisation. Experimental results demonstrate that the combined row-column features achieve stable recognition rates of approximately $91\% \sim 93\%$ without quantisation. Following quantisation and fine-tuning, the model attains a peak recognition rate of approximately 88.8% .

Keywords: optoelectronic memristor; in-sensor computing; image recognition; convolutional neural network

0 引 言

随着人工智能与大数据技术的快速发展,传统基于冯诺依曼架构的计算系统在处理高并行度、低功耗的智能感知任务时逐渐暴露出瓶颈^[1-3]。该架构将存储单元与计算单元物理分离,数据在两者间频繁传输形成“冯诺依曼瓶颈”,导致能耗高、延迟大、并行度受限等问题。尤其在图像识别、自然语言处理等深度学习任务中,模型参数规模呈指数级增长,传统处理器的能效已难以满足边缘智能设备的

发展需求^[4-9]。

为了突破这一限制,感存算一体成为近年的研究热点^[10-12]。该思路通过在同一器件中实现光电信号感知、存储与计算,大幅减少数据搬运,提高系统的计算效率。在众多实现方式中,忆阻器因其非易失性、模拟权值调控、低功耗与高密度集成等优势,被认为是构建神经形态硬件的重要候选器件^[13]。特别是光电忆阻器,能够同时响应光信号与电信号,可直接将外部光学刺激转化为可调制通态,实现光感知和权值映射,为构建新型的类脑感知芯片提供了可能^[14-17]。

收稿日期:2025-12-24

* 基金项目:安徽省自然科学基金(2308085MF207)项目资助

近年来,围绕将忆阻器及其阵列用于卷积运算、矩阵乘加速与类脑记忆等神经形态计算的研究取得了显著进展^[18-24]。Chari 等^[20]提出的全并行卷积神经网络(convolutional neural network, CNN)利用 1 个忆阻器和 1 个晶体管固定电路阵列在单周期内实现整张特征图的并行卷积,从系统层面实现了高吞吐与低功耗;Kim 等^[21]通过可配置核映射在阵列内调节权重复制与读取周期,在时间效率与阵列面积之间提供灵活折中;Zhang 等^[22]提出的原位卷积忆阻器则将卷积计算搬到单个忆阻器的时域动态上,通过导电丝生长与切换时间实现器件级原位卷积与噪声抑制;Tan 等^[23]通过 1 个晶体管和 3 个忆阻器交叉阵列与全尺寸并行卷积算法实现了从算法到硬件的协同优化,显著降低了计算复杂度;Byun 等^[24]则利用动态忆阻器的短期记忆与脉冲易化特性,在单器件层面实现了时序信号的高维非线性映射,构建了物理储备池混合计算框架。上述工作从不同角度展示了忆阻器体系在提升并行度、能效与计算密度方面的潜力,但也暴露出若干亟待攻克的问题:1)器件动态特性与神经网络模型的协同设计不足,尤其是如何将光电忆阻器的多态电导可靠且可重现地映射为网络权重;2)量化、有限位宽与器件噪声等硬件约束对模型精度的影响尚缺乏系统性的量化与补偿策略;3)现有工作多集中于器件级或阵列级功能验证,尚未对从光学感知、前端特征提取到后端分类识别的完整流程进行系统评估,从而限制了其在实际视觉识别任务中的应用潜力。

基于上述背景,本文利用已制备的光电忆阻器阵列作为感知前端,对手写字母图像进行特征采集,并将器件输出数据输入 CNN 进行分类识别。本研究的主要工作包括:1)基于两组光电忆阻器阵列构建手写字母特征采集方案;2)设计适用于器件输出特征的 CNN;3)开展权重训练与量化实验;4)分析网络深度与分类任务对模型性能的影响。

1 光电忆阻器阵列

1.1 光电忆阻器特性

如图 1 所示,本文设计所使用的光电忆阻器采用 ITO/PEDOT: PSS/PVK/C60 + BCP/Ag 的层状结构,其中 $\text{FA}_{0.8}\text{Cs}_{0.2}\text{Sn}_{0.5}\text{Pb}_{0.5}\text{I}_3$ (PVK) 作为光敏层,用于将入射光信号转换为电学响应。器件在光照与暗态条件下表现出明显不同的光电响应特性,在相同偏压下,光照状态下的输出电流显著高于暗态,表明器件具有良好的光电响应能力和较低的暗电流水平。该特性使器件能够有效感知光强变化,并将光学信息稳定映射为电流信号,适合作为神经形态视觉感知与特征提取的前端光电转换单元。

1.2 特征提取

光电忆阻器单元能够在接收光照的同时提取图像的光学特征,并在器件内部完成卷积运算。当不同像素位置的光信号照射在器件上时,其光强会调制器件的光生载流子浓度,进而影响输出的光电流幅值。与此同时,器件预设的

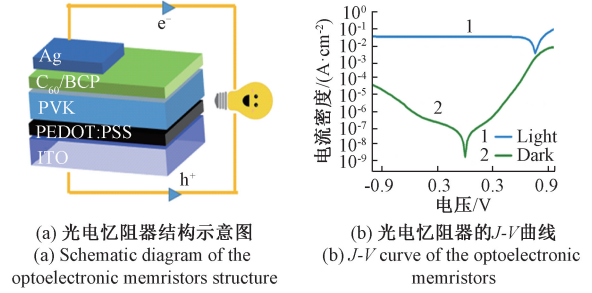


图 1 光电忆阻器结构示意图和 $J-V$ 曲线

Fig. 1 Schematic diagram of the optoelectronic memristors structure and its $J-V$ curve

电导状态可等效为卷积核权重,使得光照强度与电导值在器件内部发生乘积运算。最终,光生电流即表示卷积计算后的特征响应,实现了将图像感知与卷积特征提取集成在同一器件中的原位计算过程。

如图 2 所示,图像不同像素点的光照强度 P 作为输入,器件光电响应产生的光电流 I 作为输出。阵列中每列光电忆阻器的光电流可自动累加作为输出电流,实现列卷积运算。计算公式为:

$$I_n = \sum_{m=1}^M P_{mn} R_{mn} \quad (1)$$

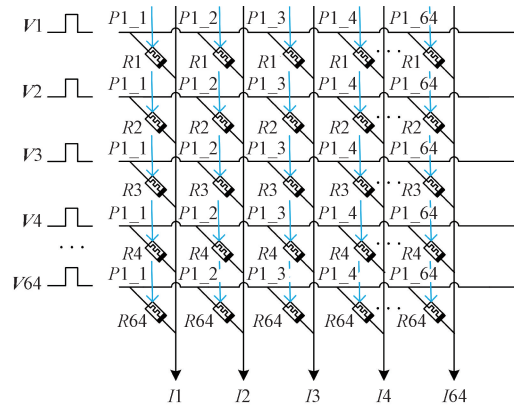


图 2 光电忆阻器的一维卷积压缩

Fig. 2 One-dimensional convolutional compression of optoelectronic memristors

本文要识别的图像的尺寸为 64×64 ,卷积核的权重 W 预先映射为器件阵列中每一行的光响应度 R ,阵列的每行包含 64 个对应权重的光响应单元。在计算过程中,图像像素的光强作为输入照射到器件表面,各像素光强与相应的光响应度在单元内部完成乘法运算,产生与输入加权相关的光电流。阵列按列对这些光电流进行累加,得到 64 路输出电流,构成图像的列方向特征向量。随后,该特征向量被输入到 CNN 中,实现对图像类别的进一步判别。

本文在前端实验中对光电忆阻器阵列的噪声、电导漂移、写入误差及阵列不一致性等非理想因素做了理想化假

设。为提升系统在实际部署中的鲁棒性,后续在算法层面采取了多项针对性策略,在后端 CNN 训练中引入权重衰减正则化并结合通道级混合精度量化进行微调,该组合在数学上等效于一种隐式的噪声注入训练,能够抑制模型对理想仿真数据的过拟合并提高对前端参数扰动的容错能力。另一方面,所用数据集本身包含手写样本的自然变异与噪声,在训练阶段还施加了随机平移、旋转及亮度与对比度扰动等数据增强,进一步扩展了输入分布,从而增强模型对真实前端输入的适应性。

1.3 权重映射

在 CNN 的前两层中,卷积核权重被映射为光电忆阻器阵列的光响应度,具体通过调节阵列中各行交叉开关的外加电压实现权重的存储与写入。器件的 32 态光响应度如图 3 所示,其电导-光响应关系呈现出良好的线性特性。基于这一特性,可采用线性映射方式将网络权重量化后对应到器件状态。

由于器件可提供 32 个可区分的电导态,权重被量化为 5 bit,并与 32 态光响应度一一对应,从而实现量化权重到光响应度的线性映射,映射关系如图 4 所示。图 5 所示为卷积核权重向量到光响应度的线性映射。该映射方式保证了权重在器件中的准确表示,为后续的卷积计算奠定基础。

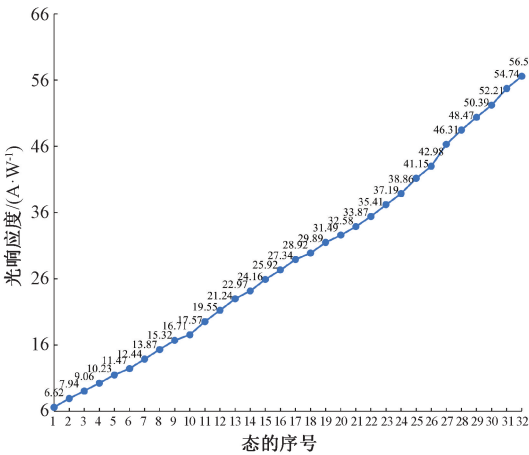


图 3 光电忆阻器的 32 态光响应度

Fig. 3 32-state photoresponsivity of the optoelectronic memristor

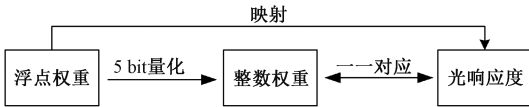


图 4 权重与光响应度映射关系

Fig. 4 Mapping relationship between weight and light responsivity

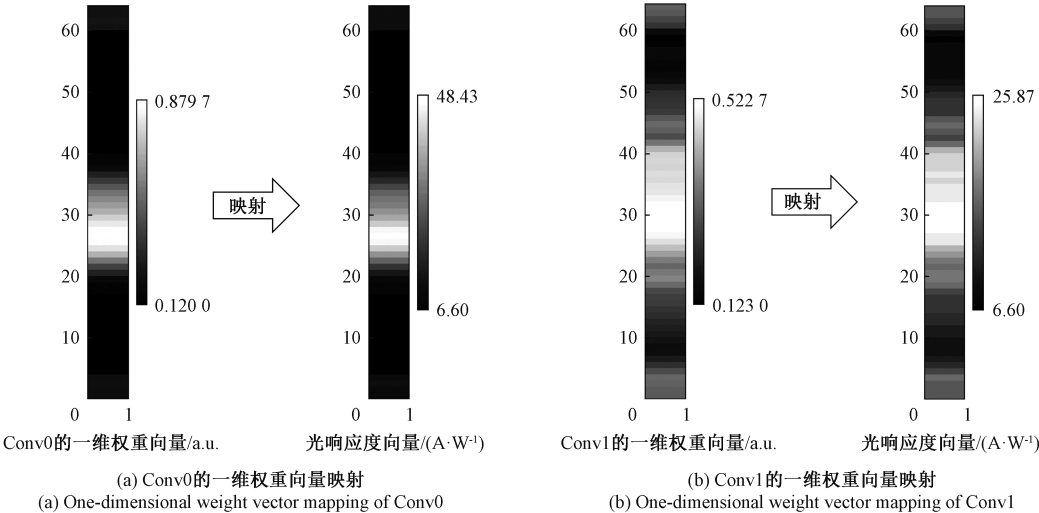


图 5 卷积核权重向量到光响应度的线性映射示意图

Fig. 5 Schematic representation of the linear mapping from convolutional kernel weight vectors to optical responsivity

2 网络算法设计

本文提出一种双阵列一维压缩与联合特征识别的混合计算方案。使用两组 4k 光电忆阻器阵列分别沿行、列方向对 64×64 图像进行并行一维卷积压缩,产生两路 64 维特征向量并拼接为 128 维联合特征,随后将其输入后端 CNN 完成分类。该方法在显著降低前端数据量与卷积计算负担的同时通过行列互补保留关键判别信息,后文实验结果表明联合方案在识别率和收敛性方面均优于单向压缩方案。

该双向一维压缩方法的有效性,可以从矩阵奇异值分解(singular value decomposition, SVD)的理论框架中获得深刻解释。任何输入图 $I \in \mathbf{R}^{64 \times 64}$ 均可通过 SVD 分解为一系列秩-1 矩阵的加权和,计算公式为:

$$I = \sum_{i=1}^{64} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \tag{2}$$

其中,奇异值 $\sigma_1 \geq \sigma_2 \geq \dots \geq 0$ 表征了不同成分的重要性,左奇异向量 $\mathbf{u}_i$ 张成了图像的列空间,右奇异向量 $\mathbf{v}_i$ 张成了图像的行空间。本文的双向压缩可视为一种高效

的、物理实现的 SVD 截断近似。行方向卷积旨在提取图像在最重要的若干左奇异向量 $\{u_i\}$ 所张成的空间上的投影,捕获图像的纵向模式;列方向卷积则对应应在最重要的若干右奇异向量 $\{v_i\}$ 空间上的投影,捕获图像的横向模式。拼接后的 128 维联合特征,相当于构建了一个由行、列两个最主要方向张成的联合特征子空间,近似保留了 SVD 分解中前几个奇异值最大的分量所携带的主要判别信息。从计算效率角度看,该方案与可分离卷积的数学思想同源。传统二维卷积核若其 SVD 近似为秩-1 矩阵,则可分离为两个一维卷积核的连续运算,将计算复杂度从 $O(k^2)$ 降至 $O(2k)$ 。本文方法通过硬件阵列直接实现双向一维投影,从根本上继承了这种计算效率优势。

2.1 网络结构

本文设计的网络算法结构如图 6 所示。整个神经网络

推理流程可以分为两个部分:1)部署在光电忆阻器阵列上的图像特征提取;2)由 CNN 完成的高维特征分类。在阵列端,输入的手写字母图像被分别投射到两个光电忆阻器阵列中,两个阵列分别对应网络前两层卷积核的权重。图像以光照形式作用于阵列单元后,不同像素的光强与对应器件的光响应度在单元内部产生加权光电流,各列单元的光电流经累加形成长度为 64 的电流特征向量。来自两个阵列的电流特征向量进一步拼接,构成网络的输入图像联合特征。随后,这一联合特征向量被送入后端的 CNN 结构中。网络首先经过两级卷积-下采样模块,用于进一步提取局部空间结构信息并压缩特征维度。经卷积层提取的深层特征随后输入两层全连接网络,完成类别空间的映射,并最终输出手写字母的分类结果。

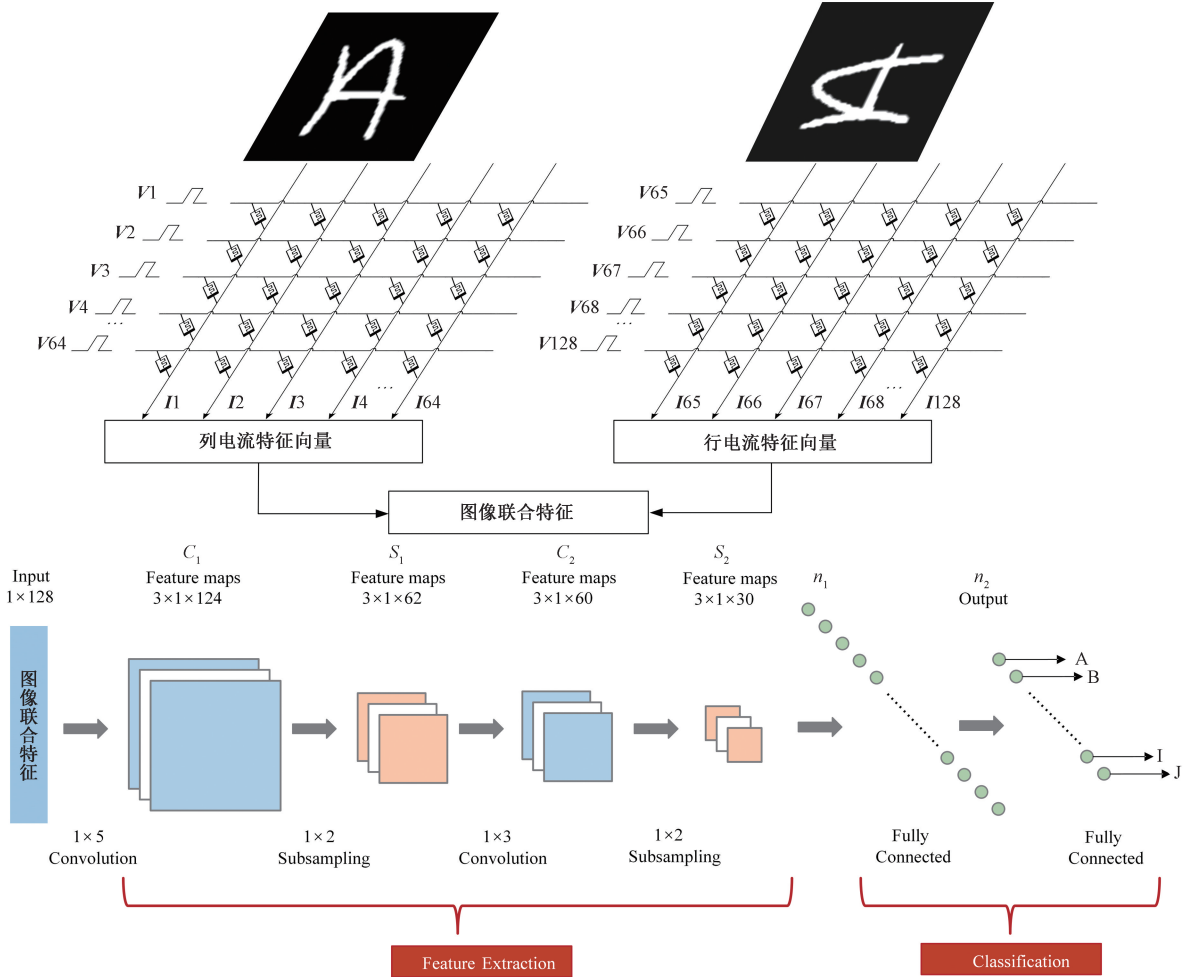


图 6 网络算法结构
Fig. 6 Network algorithm architecture

2.2 网络训练

为对来自光电忆阻器阵列的联合特征向量进行分类,本文在 PyTorch 框架下构建并训练了 CNN。所有实验均采用统一超参数设置以保证可比性,主要训练配置如表 1

所示。训练使用交叉熵损失函数衡量分类误差;优化器选用 Adam 以兼顾收敛速度与稳定性;为抑制过拟合引入权重衰减。训练过程中采用逐轮学习率衰减策略以在后期精细调整模型参数。

表 1 神经网络重要参数

Table 1 Key parameters of neural networks

参数	值
Epochs	20
Batch Size	64
Learning Rate	0.001
Weight Decay	0.000 1
Criterion	Cross Entropy Loss
Optimizer	Adam
Gamma	0.9

在实验实现方面,所用数据来自 EMNIST 数据集的字母子集,仅选取字母 A~J 共 10 类。在进入训练流程前,对所有样本进行了统一的预处理,灰度归一化并按需要重采样至与前端感知尺寸一致。随后按 8:1:1 的比例划分为训练集、验证集和测试集。模型在每个 Epoch 结束后均在验证集上进行性能评估,并根据验证准确率自动保存当前表现最佳的模型权重,以确保最终测试结果的可靠性与代表性。训练阶段持续记录训练损失、验证损失以及对应的

分类准确率,并绘制收敛曲线用于分析模型的训练动态与稳定性。

2.3 权重量化

为了在保持模型精度的同时进一步压缩网络规模,本文对训练后的权重进行混合精度量化。该方法根据不同层的权重分布特性和参数规模,为卷积层与全连接层分配不同的量化位宽,以在模型压缩率与性能之间取得平衡。

从图 7 可以观察到,卷积层的权重数量相对较少,但权重值呈现更强的离散性与峰值分布;而全连接层权重数量庞大,分布接近零均值、较为平滑且呈近似高斯形态。基于这两类层的分布差异,卷积层采用 5 bit 有符号对称量化。原因是卷积层参数量少,但对前端特征提取影响较大,需要更高精度来避免量化误差放大;同时卷积核权重的值域较宽,需要更多离散点来保证表示能力。全连接层采用 4 bit 有符号对称量化。全连接层参数量大,具备更强冗余性,适合使用更低精度进行压缩,同时有限的量化损失也可以通过后续训练补偿。卷积层高精度和全连接层低精度的配置是混合精度量化中较为常用的结构性策略,能够在节省存储与计算开销的同时最大程度减少性能损失。

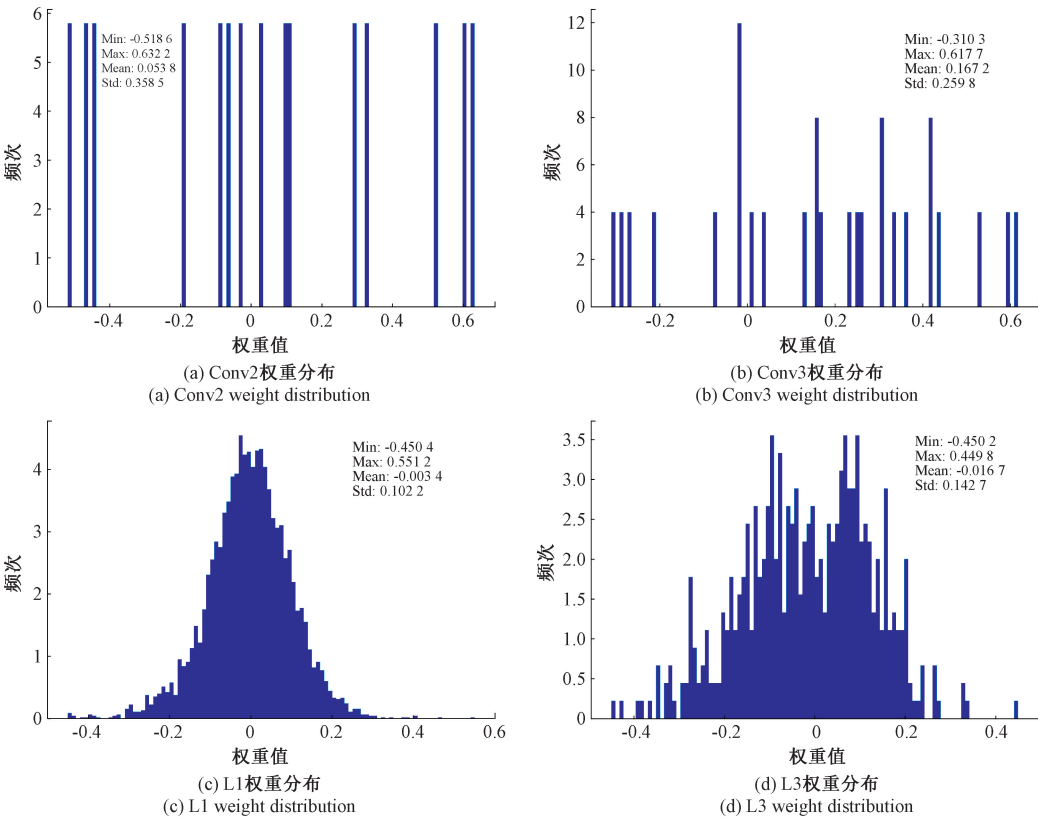


图 7 神经网络权重分布

Fig. 7 Neural network weight distribution

3 结 果

图 8 所示为在 4 种特征提取策略下随训练轮次的识别

精度对比,仅列向压缩、仅行向压缩、行列联合压缩以及纯数字 CNN,都使用相同的网络深度与卷积核通道数。行列联合方案在训练后期稳定达到约 91%~93%的准确率,明

显优于仅列向的 79%~80% 和仅行向的 67%~68%。与纯数字 CNN 相比,行列联合在最终精度上仅有微小差距,差值在 1% 之内,但它将前端特征维数从 4 096 降至 128,数据传输量和后端计算负担均显著降低。除此之外,行列联合方案收敛更快且曲线更平滑,表明多维投影的互补性有助于提升训练稳定性与鲁棒性。综合来看,行列联合一维压缩在带宽受限的工程场景中实现了在精度与资源消耗之间的良好折中。

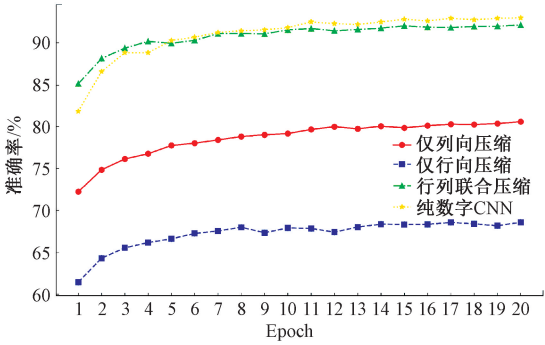


图 8 4 种策略在训练过程中的准确率对比

Fig. 8 Accuracy comparison of four strategies during training

如图 9 所示,随着训练 Epoch 的增加,训练损失与验证损失均出现稳定下降趋势;同时训练与验证准确率持续上升并在后期趋于平稳。最终训练准确率超过 92%,验证准确率接近 92%,训练与验证曲线之间无明显发散,说明模型具有良好的收敛性且未表现出严重过拟合。由损失曲线可见,模型在前 5 Epoch 内完成大幅下降,随后进入精

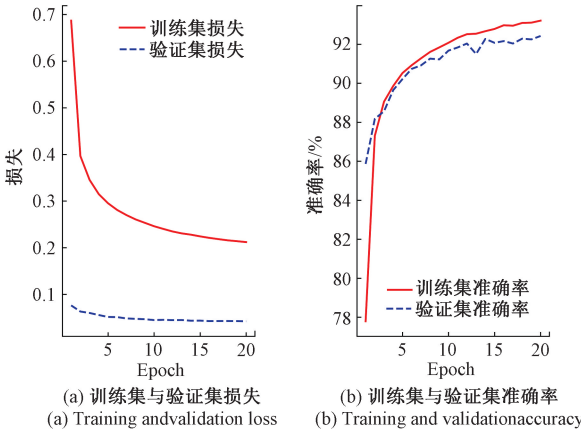


图 9 神经网络在 20 个训练周期内的识别精度与损失
Fig. 9 Recognition accuracy versus loss of the neural network over 20 training cycles

细调整阶段,表明所选超参数能够支持稳定训练。

采用混合精度量化策略并在量化后进行微调的实验结果如图 10 所示。量化模型在前 6 Epoch 训练阶段波动较大,随后性能逐步回升并稳定。量化并微调后模型在验证集上第 19 Epoch 取得最高识别率约 88.8%,且在大多数后期保持在约 86%~89%。相较于未量化的约 92%~93% 浮点基线,混合精度量化带来了可接受的精度下降,同时显著降低了模型的表示位宽和存储开销。这一结果表明所采用的混合精度与量化感知微调策略能在压缩比与性能之间实现良好折中。

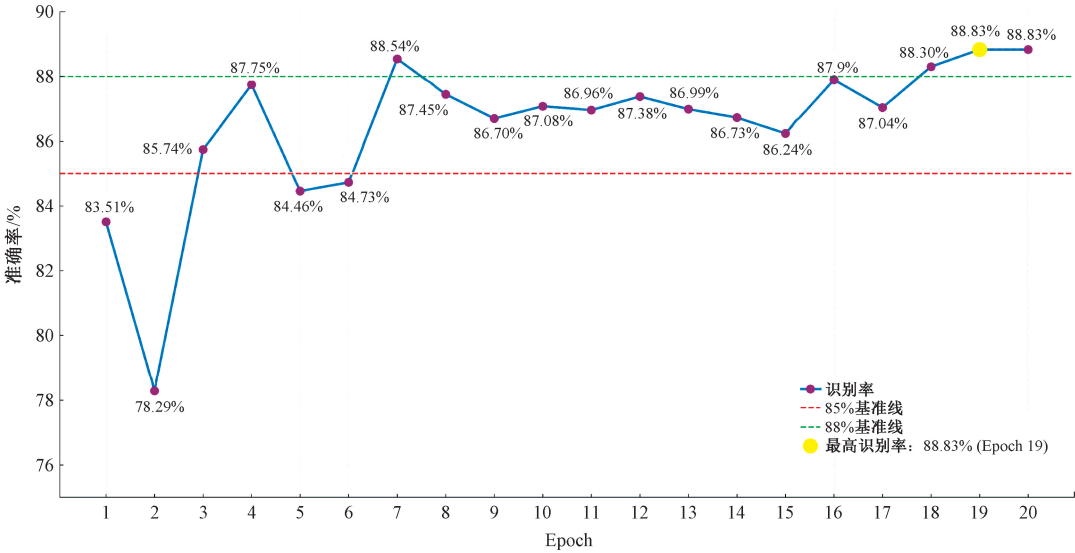


图 10 混合精度量化模型在 20 个训练周期内的识别率变化

Fig. 10 Recognition rate variation of the mixed-precision quantisation model over 20 training cycles

在采用混合精度量化的前提下,图 11 直观地反映了网络深度与任务规模对识别性能和训练行为的影响。图 11(a)比较了在相同 10 类任务下 3 种深度随训练轮次

的准确率曲线,4 层网络收敛较慢且最终精度显著低于更深网络,而 5 层与 6 层均能在较少轮次内达到高精度,且曲线更平滑稳健,说明适度加深网络能够显著提升对压缩后

投影特征的判别能力。图 11(b)固定为 5 层结构,展示其在类别数递增时的表现,随着类别数增加,整体准确率呈下降趋势,且波动略增,表明任务复杂度上升会放大模型

对细粒度判别能力的需求。在本文所用的前端强降维设置下,采用 5 层中等深度的轻量后端在性能、收敛速度与稳定性之间取得了最佳折中。

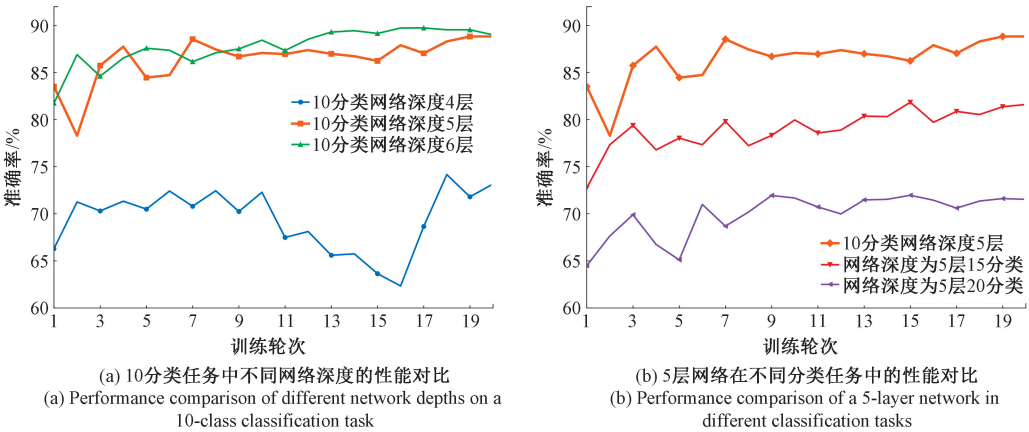


图 11 网络深度与分类任务对模型性能的影响分析
Fig. 11 Analysis of the impact of network depth and classification tasks on model performance

基于最终测试量化微调后性能较佳的模型绘制的混淆矩阵如图 12 所示,大多数类别的识别率较高,主对角线上的计数占多数,表明模型对绝大多数字母具有良好判别能力。然而,仍能观察到若干特定类别间的系统性混淆:某些字母如 G 容易被误判为 A 或 B;这是由于这些字母在手写样式中与 A 和 B 在笔画分布或整体轮廓上存在相似性,尤其在笔画连接或局部形状被压缩后更易混淆。类似 I 与 J 以及部分邻近字符之间存在相互错分的情况,说明这些字母的局部细节对分类器十分敏感。总体来看混淆矩阵中对角线条目显著高于非对角线,大多数类的准确率仍处于较高水平,但在某些易混淆对上存在统计上可观的误分类数量。

本研究还比较了部分相关的工作,如表 2 所示。与文献[19-20]侧重于高态数器件与阵列并行化实现不同,本文的贡献在于针对前端资源受限与带宽受限场景,在前端通过两组光电阻器阵列将 4k 像素压缩为 128 维联合特征,

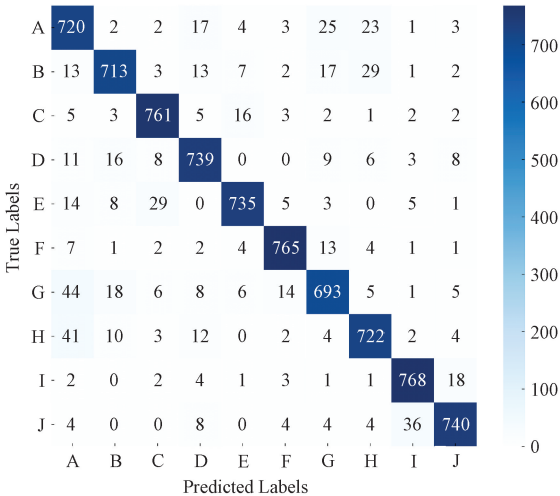


图 12 量化后模型在测试集上的混淆矩阵
Fig. 12 Confusion matrix of the quantified model on the test set

表 2 与其他相关研究的比较

Table 2 Comparison with other relevant studies											
方法	核心器件	数据集	识别目标	忆阻器电导态数	前端是否压缩	前端特征压缩率/%	后端是否量化	量化方式	量化位数	训练轮次	最大识别率/%
文献[19]	动态忆阻器	自制	校园地标(10类)	128	否	—	否	—	—	>25	95.13
文献[20]	普通忆阻器	MNIST	0~9	128	否	—	否	—	—	100	98.63
本文	光电忆阻器	EMNIST	A~J	32	是	97	是	混合精度	卷积层 5 bit 全连接层 4 bit	20	93

显著降低了从传感端到数字域的传输与处理数据量;本文仅使用器件可实现的 32 态电导,说明方案对器件位宽具

有更好的适配性和鲁棒性。相较于需要较多训练迭代或依赖高分辨率器件的工作,本文后端采用轻量级网络并在

较少训练轮次即能收敛,训练开销和部署复杂度更低。尽管在绝对峰值识别率上相比较低,但在极端压缩、低位宽与端到端验证这组严格工程约束下,本文展示了更好的“压缩—精度—可实现性”折中,证明了在带宽和能耗受限的边缘视觉场景中,本方法具有更高的工程可用性与推广潜力。

4 结 论

本文提出并实现了基于光电忆阻器阵列的前端特征采集与后端 CNN 分类的混合识别框架。通过在器件端并行执行一维投影式卷积压缩并将特征拼接输入 CNN,本方法在大幅减少需传输与数字化的特征量的同时,保留了足够的判别信息。为适配器件的有限位宽与实际部署需求,本文采用混合精度量化并进行量化感知微调,量化后峰值识别率约为 88.8%,显示了精度与存储节省之间的良好折中。需要指出的是,本文并不涉及忆阻器的制备工作,所用器件作为现成硬件被用作前端传感单元。总体上,本文验证了光电忆阻器作为感知计算前端与量化 CNN 协同工作的实用性,为资源受限场景下实现高效神经形态视觉感知提供了可行路径。

参考文献

- [1] 蓝龙英, 宋程霖, 左石凯, 等. 存算一体技术研究进展与挑战[J]. 半导体技术, 2025, 50(9): 873-884.
LAN L Y, SONG CH L, ZUO SH K, et al. Research progress and challenge of compute-in-memory technology [J]. Semiconductor Technology, 2025, 50(9): 873-884.
- [2] MOLAS G, NOWAK E. Advances in emerging memory technologies: From data storage to artificial intelligence[J]. Applied Sciences, 2021, 11(24): 11254.
- [3] MARCHESIN A, NACLERIO A, RIENTE F, et al. Beyond von Neumann architectures: Exploring algorithmic opportunities via Octantis [J]. IEEE Access, 2024, 12: 120005-120022.
- [4] 张国静, 李承家, 韩敬伟, 等. 基于深度学习的多帧瞳孔检测算法[J]. 电子测量技术, 2025, 48(9): 168-176.
ZHANG G J, LI CH J, HAN J W, et al. Multi-frame pupil detection algorithm based on deep learning[J]. Electronic Measurement Technology, 2025, 48(9): 168-176.
- [5] JHANG C J, XUE CH X, HUNG J M, et al. Challenges and trends of SRAM-based computing-in-memory for AI edge devices[J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2021, 68(5): 1773-1786.
- [6] 赵满庆. 基于深度学习的水下声目标智能识别方法[J]. 国外电子测量技术, 2025, 44(3): 44-48.
ZHAO M Q. Intelligent recognition method for underwater acoustic targets based on deep learning[J]. Foreign Electronic Measurement Technology, 2025, 44(3): 44-48.
- [7] 未争超, 乔强, 郎俊杰, 等. 基于深度学习与聚类的开集个体识别技术[J]. 电子测量技术, 2025, 48(5): 111-117.
WEI ZH CH, QIAO Q, LANG J J, et al. An open set SEI method based on deep learning and clustering[J]. Electronic Measurement Technology, 2025, 48(5): 111-117.
- [8] 陈斯睿, 孙兴伟, 杨赫然, 等. 基于改进支持向量机的砂带磨损程度识别[J]. 电子测量与仪器学报, 2024, 38(2): 10-18.
CHEN S R, SUN X W, YANG H R, et al. Abrasion degree recognition of abrasive belt based on improved support vector machine [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(2): 10-18.
- [9] 李响, 赖本涛, 张媛霖, 等. 基于图像识别的铁路机车转向架螺栓紧固状态检测[J]. 电子测量与仪器学报, 2023, 37(12): 143-155.
LI X, LAI B T, ZHANG AI L, et al. Detection of bolt fastening state of locomotive bogie based on image recognition[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(12): 143-155.
- [10] JIANG N Y, XIE J P, JIANG M X, et al. Memristor-based multilayer neural network with edge learning and weight update [J]. Neurocomputing, 2025, 647: 130549.
- [11] AN J J, WANG L F, YE W, et al. Design memristor-based computing-in-memory for AI accelerators considering the interplay between devices, circuits, and system[J]. Science China Information Sciences, 2023, 66(8): 182404.
- [12] 李锟, 曹荣荣, 孙毅, 等. 基于忆阻器的感存算一体技术研究进展[J]. 微纳电子与智能制造, 2019, 1(4): 87-102.
LI K, CAO R R, SUN Y, et al. Research progress on the fused technology of sensing, storage and computing based on memristor [J]. Micro/nano Electronics and Intelligent Manufacturing, 2019, 1(4): 87-102.
- [13] 胡伟, 王又弘, 康开进, 等. 忆阻式传感器的研究进展与展望[J]. 仪器仪表学报, 2024, 45(12): 1-11.
HU W, WANG Y H, KANG K J, et al. Progress and prospect of memristive sensor[J]. Chinese Journal

- of Scientific Instrument, 2024, 45(12): 1-11.
- [14] ALAM M S, YAKOPCIC C, HASAN R, et al. Memristor-based neuromorphic system for unsupervised online learning and network anomaly detection on edge devices[J]. Information, 2025, 16(3): 222.
- [15] RAMIREZ-MORALES R R, PONCE-PONCE V H, MOLINA-LOZANO H, et al. Analog implementation of a spiking neuron with memristive synapses for deep learning processing[J]. Mathematics, 2024, 12(13): 2025.
- [16] BIAN J Y, LI Z Z, LIANG S, et al. 2D tellurene-based optoelectronic memristor with temporal dynamics for multimodal reservoir computing system[J]. Advanced Science, 2025, 12(43): e13647.
- [17] TIAN Q L, SUI X Y, ZHAO X N, et al. An optoelectronic memristor based on proton-involved photoreduction for bimodal sensing, memory, and processing[J]. Advanced Intelligent Systems, 2025, 7(7): 2400968.
- [18] XIA Z P, SUN X, WANG Z L, et al. Low-power memristor for neuromorphic computing: From materials to applications [J]. Nano-Micro Letters, 2025, 17(1): 217.
- [19] ZHANG Q S, CHE Q, WU D C, et al. Dual redox-active covalent organic framework-based memristors for highly-efficient neuromorphic computing[J]. Angewandte Chemie International Edition, 2024, 63(46): e202413311.
- [20] CHARI A Y, GHOLIPOUR M, HASSANZADEH M. Efficient fully parallel convolutional neural network architecture using 1-memristor and 1-transistor(1MIT)[J]. The Journal of Supercomputing, 2025, 81(13): 1275.
- [21] KIM G, YOUN S, PARK J, et al. Configurable kernel map implementation in memristor crossbar for convolution neural network[J]. Advanced Intelligent Systems, 2025, 7(11): 2500032.
- [22] ZHANG X H, QIN C Y, PENG W H, et al. Memristor-based in-situ convolutional strategy for accurate braille recognition[J]. Science China Materials, 2024, 67(12): 3986-3993.
- [23] TAN J P, SHEN S Y, DUAN S K, et al. An efficient full-size convolutional computing method based on memristor crossbar[J]. Artificial Intelligence Review, 2024, 57(6): 158.
- [24] BYUN Y, SO H, KIM S. Convolutional neural network for high-performance reservoir computing using dynamic memristors[J]. Chaos, Solitons & Fractals, 2024, 188: 115536.

作者简介

盛喜乐, 硕士研究生, 主要研究方向为光电探测与图像传感技术。

E-mail: 2023110937@mail. hfut. edu. cn

林引瑞, 硕士研究生, 主要研究方向为图像超分辨率神经网络加速器。

E-mail: 2024110930@mail. hfut. edu. cn

张章(通信作者), 博士, 教授, 主要研究方向为混合信号芯片、基于混合器件的 AI 神经形态芯片等。

E-mail: zhangzhang@hfut. edu. cn