
深圳先进院等提出忆阻器神经网络的高能效权重更新方案

作者：writer 来源：中国科学院

本文原地址：<https://www.iikx.com/news/progress/15300.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

近日，中国科学院深圳先进技术研究院生物医学与健康工程研究所医学人工智能研究中心副研究员黄明强研究团队，在基于忆阻器阵列的卷积神经网络的研究上，提出了硬件友好型的随机自适应学习方法。区别于具有复杂外围电路设计的梯度下降方法（SGD）和分段线性（PL）的传统方法，该方法具有良好的硬件友好特性和高能效特性。相关研究成果以Hardware-Friendly Stochastic and Adaptive Learning in Memristor Convolutional Neural Networks为题，发表在Advanced Intelligent Systems上。

神经网络应用于各种人工智能技术，如计算机视觉和语音识别。然而，在传统互补金属氧化物半导体CMOS数字平台上运行的人工智能算法，在一定程度上受到摩尔定律即将结束或冯·诺依曼架构限制的限制。因此，学术界提出了基于忆阻器阵列的存储计算模型（computing in memory）。传统的基于忆阻器阵列的权重更新，需要计算出具体的脉冲数，这需要构建复杂的外围电路来计算脉冲数。在新的学习方案中，随机提取训练样本，基于训练样本更新整个卷积神经网络的权重，计算梯度值获取权重更新的方向，引入忆阻器非理想特性，给予忆阻器单脉冲进行权重更新，该方法不需要复杂的外围电路设计，节省了芯片面积和功耗。研究基于新的学习方案提出了四种权重更新方式，将传统的SGD和PL作为基准，在基于LeNet5网络对MNIST图像识别中取得了良好的识别率。该方法可应用到更加复杂的神经网络结构中，如resnet8网络和VGG网络识别cifar10数据集，具有良好的应用前景和优化空间。

该成果是科研团队近年来在新型微电子器件高能效计算领域发表的第3篇研究型论文。2020年，团队先后提出柔性铁电突触晶体管神经网络方案（Applied Physics Letters 117 (9), 092903, 2020）、功耗可调型1T1DM低比特神经网络方案（IEEE Electron Device Letters 42 (1), 106-109）。未来，该团队将结合新型微电子器件和集成电路设计技术，在更大规模的忆阻神经网络应用上拓展，为高能效人工智能和终端智能的硬件架构提供新思路。

研究工作得到国家重点研发计划、国家自然科学基金和深圳市科技计划等的资助。

[论文链接](#)

(a) 卷积神经网络模型；(b) 忆阻神经网络的忆阻-权重映射方案；(c) 本论文提出的高能效权重更新方案（红色线）与传统忆阻权重更新方案（蓝色线）的性能对比

研究团队单位：深圳先进技术研究院

更多 科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发