

---

# 模拟芯片大幅削减AI模型运行能耗

作者：writer 来源：科学网

本文原地址：<https://www.iikx.com/news/progress/23998.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！**

模拟芯片大幅削减AI模型运行能耗。

模拟计算机芯片运行人工智能(AI)语音识别模型的效率是传统芯片的14倍，有可能为AI研究中巨大且不断增长的能源消耗以及全球范围内通用数字芯片短缺问题提供解决方案。

这一新产品由IBM Research开发，概述这项工作的论文发表于《自然》。研究人员声称，模拟芯片可以减少人工智能发展的障碍。

GPU是一种图形处理器，最初设计用于运行视频游戏，传统上也用于训练和运行AI模型，目前全球都在抢购这种芯片，供不应求。研究还表明，AI的能源使用量正在迅速增长，从2012年到2021年增长了100倍，其中大部分能源来自化石燃料。这些问题导致有人认为，AI模型发展将很快陷入僵局。当前AI硬件的另一个问题是，它必须在内存和处理器之间来回传输数据。一种解决方案是模拟存算一体(CIM)芯片，它直接在内存中执行计算，IBM现在已经大规模演示了这种芯片。

IBM的新设备包含3500万个所谓的相变存储单元，这是CIM的一种形式，可以设置为两种状态中的一种，就像计算机芯片中的晶体管一样，也可以在两种状态之间进行不同程度的设置。

这一特征至关重要，因为这些不同的状态可以用于表示神经网络中人工神经元之间的突触权重——神经元之间的连接强度。神经网络是一种AI，它模拟人类大脑中神经元之间的连接在学习新信息或技能时的变化，这些信息或技能通常以数位值的形式存储在计算机内存中。这使得新芯片能够存储和处理这些突触权重，而无需在远程存储芯片中进行数百万次调用或存储数据的操作。

美国英特尔公司相关负责人评价说，这种芯片还远不是一个成熟的产品，但实验表明，它可以在当今常用的AI神经网络上有效工作，并有潜力支持ChatGPT等流行应用。就像GPU无法覆盖标准计算机处理器CPU可以执行的所有任务一样，模拟AI芯片或模拟CIM芯片也有其局限性。但如果AI的趋势能够继续并保持当前的发展态势，高度定制化的芯片肯定会变得更加普遍。

尽管该芯片是专门定制的，但它在IBM语音识别任务之外可能有其他用途。正如展示的那样，模拟AI或模拟CIM具有更高的功率和硅使用效率，与CPU或GPU相比，这可能会降低成本。上述负责人说。(来源：中国科学报 文乐乐)

相关论文信息：<https://doi.org/10.1038/s41586-023-06337-5>

作者：S. Ambrogio 来源：《自然》

---

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](http://www.iikx.com)转发