

科技前沿信息

第 15 期 (总第 135 期)

上海科技成果转化促进会
上海科学技术情报研究所 主办

2019 年 8 月 上旬刊

【专题报道】

类脑芯片的最新突破及国内外研发进展

[导读] 类脑芯片是指参考人脑神经元结构和人脑感知认知方式来设计的芯片，代表了人工智能芯片研究的最新方向。

类脑芯片的最新突破及国内外研发进展

近日，我国清华大学教授施路平团队研发的类脑计算芯片“天机芯”，登上世界顶级学术杂志《自然》封面。据论文介绍，“天机芯”是一款新型人工智能芯片，结合了类脑计算和基于计算机科学的人工智能，是世界首款异构融合类脑芯片。与世界先进的 IBM 公司 TrueNorth 芯片相比，“天机芯”功能更全、灵活性和扩展性更好，密度提升 20%，速度提高至少 10 倍，带宽提高至少 100 倍。

一、什么是类脑芯片？

类脑芯片是指参考人脑神经元结构和人脑感知认知方式来设计的芯片，代表了人工智能芯片研究的最新方向。具体而言，类脑芯片架构就是模拟人脑的神经突触传递结构，众多的处理器类似于神经元，通讯系统类似于神经纤维，每个神经元的计算都是在本地进行的，从整体上看神经元们分布式进行工作的，也就是说整体任务进行了分工，每个神经元只负责一部分计算。

对于人工智能产业而言，随着数据量的不断增加，想要提升运算效率最常见的方向就是优化算法和提升芯片算力。然而在摩尔定律下，处理器集成器件数量越发接近极限，提升芯片算力显然不能依靠传统方法。想要突破数据量激增所带来的计算瓶颈，就需要从芯片底层架构需求重构和变化。

目前，围绕着人工智能芯片的创新路径主要有以下三个方向：

1、基于 GPU、FPGA 等通用芯片打造半定制的方案。例如英伟达公司针对各类应用场景开发对应的 GPU 芯片，并通过打造 CUDA 平台提升其编辑效率、开放性和丰富性，并建立了完整的算法平台，使得人工智能算法可以和多种类型智能终端相融合。

2、针对深度学习算法开发 ASIC 芯片。目前谷歌推出的 TPU 和寒武纪推出的 NPU 都是基于 ASIC 架构研发的深度学习芯片。

3、开发类脑计算芯片。其本质上借鉴人脑的工作原理来实现深度学习，进而解决极其复杂的计算问题，有望带来计算体系革命与架构变革。

然而从本质上讲，目前主流的 GPU、FPGA 以及 ASIC 芯片都属于冯·诺依曼结构，在这种结构中，计算模块和存储模块是分开的。CPU 在执行命令时需要先从存储单位读取数据，也就是人们常说的“内存墙”，这就产生延时及大量的功耗浪费，从而限制了计算性能的提升。

相比之下，人脑却没有此类问题出现。据研究表明，人类大脑平均每秒可执行 1 亿亿次操作，所需能量只有 10~25 瓦特。因而研究员们正转向模拟人类大脑研究，试图通过模拟人脑运转机制，通过研发类脑芯片，使计算机能低能耗高功效地进行计算，甚至使计算机优于类人的智能。除此之外，目前主流的人工智能系统都需要提前的进行人工建模，转化为计算问题进行处理再进行深度学习。而采用类脑芯片的系统，理论上可以自动感知，进行问题分析和求解，决策控制等。

二、类脑芯片国内外进展情况

基于类脑芯片的数据计算领域的革新性，吸引了大量公司投入研发。而为了保持技术优势，美国先发起类脑计算芯片的相关研究工作，通过模仿人脑工作原理，使芯片能够进行异步、并行、低速和分布式处理信息数据，并具备自主感知、识别和学习的能力。

1、IBM 公司 TrueNorth 芯片

2011 年，IBM 率先在类脑芯片上取得进展，研发出两个具有感知、认知功能的硅芯片原型，IBM 戏称第一代 TrueNorth 为“虫脑”。2014 年 IBM 又开发了第二代 TrueNorth 芯片，采用了三星 28nm 工艺，共用了 54 亿个晶体管，其性能相比第一代有了不少提升。功耗每平方厘米消耗仅为 20 毫瓦，是第一代的百分之一，直径仅有几厘米，是第一代的十五分之一。每个核都简化模仿了人类大脑神经结构，包含 256 个“神经元”（处理器）、256 个“轴突”（存储器）和 64000 个突触（神经元和轴突之间的通信）。总体来看，TrueNorth 芯片由 4096 个内核，100 万个“神经元”、2.56 亿个“突触”集成。

2、英特尔公司 Loihi 芯片

2017 年 9 月，英特尔公司发布“Loihi”的神经模拟原型芯片，这是其推出的首款自学习神经元芯片。芯片包含的数字电路能模拟人脑 13 万个神经元和 1.3 亿个突触连接。Loihi 芯片采用了一种“异步激活”的全新计算方式，来模仿大脑运作模式，组成它的 128 个计算核心每个都包含 1024 个神经元，总共模拟出人脑 13 万个神经元和 1.3 亿个突触连接，组成异步神经形态的多核网状结构，每个神经元都能与成千上万个其他神经元同时通讯。据英特尔测试结果，Loihi 的学习效率比其他智能芯片高 100 万倍，而且完成同一个任务所消耗的能源可节省近 1000 倍。

3、高通公司 Zeroth 芯片

早在 2013 年高通就公布了 Zeroth 芯片，名称来源于 IsaacAsimov 的机器人学第零定律：“机器人不会伤害人类，而它们也不会允许人类来伤害它们。” Zeroth 芯片基于一款新的硬件结构，其与主宰计算机领域多年的硬件结构有很大区别。它模仿人类大脑结构，由数十亿协力合作的神经细胞组成。

4、西井科技公司 DeepSouth 芯片

上海西井科技已推出了自主研发的拥有 100 亿规模的神经元人脑仿真模拟器（Westwell Brain）和可商用化的 5000 万类脑神经元芯片（DeepSouth）两款产品。Westwell Brain 可以通过接受医学上大脑神经元脉冲放电数据，以更直观方式呈现人脑的脉冲形态，用于检验药物及治疗的作用区域及效果。可商用化的 DeepSouth 则能模拟出高达 5000 万级别的“神经元”，总计有 50 多亿“神经突触”，在同一任务下的功耗仅为传统芯片的几十分之一到几百分之一。

5、“达尔文”类脑芯片

2015 年来自浙江大学与杭州电子科技大学的年轻的 researchers 研发出一款成为达尔文的类脑芯片。这款芯片是国内首款基于硅材料的脉冲神经网络类脑芯片。“达尔文”芯片面积为 25 平方毫米，比 1 分硬币还要小，内含 500 万个晶体管。芯片上集成了 2048 个硅材质的仿生神经元，可支持超过 400 万个神经突触和 15 个不同的突触延迟。

6、aiCTX 公司 DynapCNN 处理器

类脑芯片公司 aiCTX 于 2017 年 2 月底在瑞士苏黎世创立，随后便发布了世界首款完全基于事件触发运算的动态视觉 AI 处理器 DynapCNN。该处理器是一款纯异步、高可配置性、可拓展性的神经形态处理器。单芯片集成超过 100 万脉冲神经元和 400 万可编程参数，支持多种 CNN 架构，其芯片架构所具有的可拓展性适合于实现大规模脉冲卷积神经网络。