

---

# 时空稀疏小样本学习大规模神经形态数据集发布

作者：writer 来源：科学网

本文原地址：<https://www.iikx.com/news/progress/21156.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！**

时空稀疏小样本学习大规模神经形态数据集发布。近日，中科院自动化研究所研究员曾毅团队提出了一个用于时空稀疏小样本学习的大规模神经形态数据集——N-Omniglot，为脉冲神经网络的学习与训练提供了一个更具挑战性的基准。相关研究成果发表于自然出版社旗下期刊《科学数据》。

深度学习的成功在很大程度上归功于像ImageNet和COCO这样的数据集的引入。然而针对类脑智能，特别是脉冲神经网络(SNN)的发展，目前广泛使用的数据集并不十分适用。

为此，曾毅团队提出了第一个用于小样本学习的神经形态数据集：N-Omniglot。原始的Omniglot数据集是在小样本学习领域被广泛使用的数据集。N-Omniglot由来自50种不同语言的44个手写字符的1623个字母组成，每个字符只有20个不同的样本。它通常被认为是静态文字图像，而忽略了书写过程中丰富的时间信息。

论文的第一作者、中科院自动化所博士生李杨介绍说，为了证明N-Omniglot和编码的Omniglot的区别，所有实验都在N-Omniglot和编码Omniglot上进行。由于获取神经形态数据的DVS相机具有较高的时间分辨率，而过长的时间轴对于目前的时钟驱动的SNN算法来说是一个巨大的负担。因此，我们以或运算及脉冲发放率对事件数据进行处理。

为了证明N-Omniglot的有效性和为SNN算法的训练提供新挑战的潜力，研究团队对四种SNN算法进行了实验，包括对两种通用的经典模式分类方法和两种小样本学习算法的适配或SNN版本。

实验结果表明，四种方法在N-Omniglot上的性能都低于在Omniglot上的结果，一方面是因为所提出的数据集在空间维度上更加稀疏，另一方面是因为缺乏神经形态数据集的预处理方法。

值得注意的是，研究团队同时测试了两种经典的小样本学习方法在不同仿真时间下的识别精度。结果表明，模拟时间越长，精度越低。这是因为模拟时间越长，事件被划分的帧数越多，帧与帧之间的信息连接就越困难。这说明数据描述符对于提高脉冲神经网络SNN提取更重要时空特征的能力至关重要。因此，N-Omniglot可以被认为对研发未来的脉冲神经网络是一个有效的、鲁棒的、具有挑战性的神经形态数据集。

曾毅表示，此次研究成果为类脑脉冲神经网络在小样本学习领域提供了一个基准，并为探索高效的脉冲神经网络训练方法提供了新的挑战。该研究也是团队长期推进的科学探索项目类脑认知智能引擎Brain-inspired Cognitive Intelligence Engine (智脉BrainCog)的一部分和阶段性成果。在面向未来的通用人工智能的基础计算平台方面，团队研制的揭示智能的计算本质、面向人类水平的、基

---

于脉冲神经网络的人工智能引擎智脉(BrainCog)已开源开放发布。(来源：中国科学报 田瑞颖)

相关论文信息：<https://doi.org/10.1038/s41597-022-01851-z>

作者：曾毅等 来源：《科学数据》

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