· 专题一:双清论坛“智能算网的基础理论与核心技术”
DOI:10.16262/j.cnki.1000-8217.20250425.001

面向天地一体化算力网的类脑智能计算研究现状与展望*

张 伟[1] 马征宇[1**] 田永鸿[1,2,3**]

1. 鹏城实验室 智能计算研究部,深圳 518000
2. 北京大学 计算机学院,北京 100871
3. 北京大学 信息工程学院,深圳 518055

[摘 要] 面向国家天地一体智能计算网络的重大战略需求,本文回顾了国内外在星载智能芯片与系统、天地协同智能计算方面的发展历程,并分析了其存在的问题及其未来发展的趋势;在此基础上,本文进一步总结了类脑模型与算法、类脑芯片系统的发展历程,并探讨了类脑智能计算在星载应用中的现实状况;最后,结合类脑智能的最新进展与未来天地一体化算力网络的发展需求,提出了星载低功耗类脑计算芯片与系统、大规模高效能类脑脉冲神经网络、天地协同增量学习架构等未来值得深入探索的研究方向。

[关键词] 天地一体化;类脑计算;高能效;协同计算

在国家“东数西算”工程与全国一体化算力网络的布局下,中国算力网[1]实现了全国异构计算资源的高效互联,为全国数字经济的发展提供了首个地面算力网络。然而,由于地面算力设施的固定性,地面算力网络在算力覆盖范围上存在局限性,往往难以覆盖到偏远地区,通过将算力网络向天基扩展,可以突破地面算力网络局限,使得山区、海洋、极地都能够接入网络,享受算力服务。

建设泛在智能的网络设施是国家的重大战略需求,天地一体化算力网络融合天基算力的广域覆盖与地面算力的深度计算能力,可以为广域空间范围内的各种应用提供泛在智能计算服务。例如,天基算力可以支持遥感数据的“天数天算”,减少无效遥感数据传输,亦可为月球/火星探测器的数据处理提供算力服务,降低深空通信延迟;“天地同算”则可以为自动驾驶、低空经济、应急通信等应用场景提供计算供给,赋能地面生产生活。因此,构建天地一体化算力网是建设泛在智能基础设施的必要手段,是算力网络未来发展的必然趋势。

天地一体化智能算力网络的构建尚有诸多技术有待突破,高效的智能计算系统与智能计算方法是其中两项必备核心技术,但这些技术面临着卫星载荷限制与太空辐照环境、数据通信带宽动态受限等特殊条件带来的挑战。面对这些挑战,类脑智能计算通过模仿人脑的计算架构与方法,以其低功耗和高能效的计算优势,为实现上述技术突破提供了潜在途径。为此,本综述将围绕现有天地一体化智能计算与类脑智能计算,重点讨论分析其计算芯片系统与计算方式方法的发展现状,并面向天地一体化算力网络发展,展望未来结合类脑智能计算值得探

收稿日期:2024-10-31;修回日期:2025-03-27
* 本文根据国家自然科学基金委员会第 371 期“双清论坛”讨论的内容整理。
** 通信作者,Email: mazhy@pcl.ac.cn;yhtian@pku.edu.cn
本文受到鹏城实验室重大攻关项目(PCL2024AS204,PCL2021A13)和国家自然科学基金项目(62027804,62425101)的资助。

引用格式:张伟,马征宇,田永鸿. 面向天地一体化算力网的类脑智能计算研究现状与展望. 中国科学基金,2025,39(2):218—228.
Zhang W, Ma ZY, Tian YH. Current state and future perspectives of brain-inspired intelligent computing for integrated space-ground computing networks. Bulletin of National Natural Science Foundation of China, 2025, 39(2): 218-228. (in Chinese)

索的方向。

1　天地一体智能计算发展现状

当前，低轨卫星互联网络发展迅猛，截至 2024 年 6 月，美国 SpaceX“星链”计划已发射 6 000 余颗卫星，并搭载激光通信终端，具备了形成规模化智能计算网络的潜力；同时其面向国防推出“星盾”计划，以支撑新一代国防太空体[2]。我国亦提出了“星网”计划，预计发射超过 12 000 颗卫星，目前该计划仍处于初期启动阶段；2024 年，之江实验室提出“三体计算星座”，计划建立 1 000 POPS 级天基算力网络；2024 年成都国星宇航科技股份有限公司提出“星算计划”，拟建设 2 800 颗算力卫星组网，并与地面算力网络互联。可以看出，依托低轨卫星星座，通过灵活动态组网实现规模化的智能算力上天，是构建天地一体化算力网络的重要途径。设计星载智能计算系统和实现天地协同智能计算，是构建天地一体化智能计算网络面临的两大核心问题。目前，国内外这两方面均处于实验探索阶段。

1.1　星载智能芯片与系统

宇航级处理器芯片在太空电磁环境下具备稳定安全工作能力，但其算力相对较低，很难满足复杂智能计算对算力的需求[3]；相比之下，由于商用货架智能处理芯片通常具有更高计算性能，通过抗辐照加固后具有满足低轨卫星设计寿命（通常小于 5 年）的潜力，因此，利用智能处理芯片打造高性能在轨智能计算系统成为国内外研究的热点[4]。

图 1 所示为在轨星载智能计算芯片与单元的发展情况。2019 年，美国国家航空航天局（National Aeronautics and Space Administration，U. S.，NASA）[5] 基于工业级 Xilinx FPGA 芯片研制出 Space Cube 智能计算单元，专为星载智能计算任务设计，其功耗小于 32 W。2020 年，欧洲航天局（European Space Agency，ESA）[6] 在 PhiSat-1 卫星上部署了基于 Intel 公司商业 Movidius Myriad 2 VPU 芯片制成的智能计算单元 EoT Board，该单元的算力超过 1TOPS@INT8，而功耗维持在 1 W 以下。2023 年，波兰 KP Labs 公司[7] 在 Intuition-1 卫星上搭载了采用工业级 Xilinx FPGA 芯片研制的智能计算单元 LEOPARD DPU，其算力达到 3TOPS@INT8，功耗控制在≤40 W。2024 年，爱尔兰 UBOTICA 公司在 CogniSAT-6 卫星上搭载基于 Intel Movidius Myriad X VPU 芯片制成的 CogniSAT-XE2 智能计算单元，其算力超过 4TOPS@INT8，功耗＜2W。

国内，2019 年长光卫星技术公司在吉林一号卫星上部署了基于 DSP 芯片制成的在轨智能单元，其功耗＜75 W。2022 年，天仪研究院[8] 在巢湖一号卫星上搭载包含 Nvidia GPU 芯片的 Jeston Xavier NX 智能计算单元，该单元算力达到 21TOPS@INT8，功耗≤20 W。2023 年，北京邮电大学[9]、武汉大学[10] 和中国科学院[11] 分别在天智 2 号、珞珈 3 号和北邮 1 号卫星上搭载了商业智能计算芯片（华为 NPU、Nvidia GPU 和 Nvidia GPU），相应智能计算单元（Atlas 200AI、Jeston Xavier 和 Jeston Orign Nano）的计算能力分别达到了 22 TOPS@INT8、32 TOPS@INT8 和 40 TOPS@INT8，功耗分别控制在 11 W、30 W 和 15 W 以内。

图 1　在轨星载智能计算芯片与单元的发展

Fig. 1　Development of Intelligent Computing Chips and Units on Orbit Spacecraft

从目前在轨试验的智能计算系统来看，国外机构更多地采用了商业化的 FPGA 和 VPU，在功耗和算力之间进行了较好的平衡，尤其是 VPU，功耗极低，同时提供较高的算力；国内研究则更倾向于选择商业化的 GPU 和 NPU，其具有更高算力，能支撑更复杂的计算任务；本质上，国内外芯片选择差异性主要是因实验任务对算力需求差异造成的。虽然国内外星载智能计算芯片选择存在差异，但是通过商用货架产品降低研发成本与周期、提升卫星智能计算性能的技术路线是一致的，高性能星载智能计算系统正逐渐向具有低功耗高算力的工业级处理芯片与计算单元演进。

星载智能计算系统需要满足卫星载荷限制、太空辐照环境等特殊条件，虽然商用产品经过抗辐照加固之后可以克服太空环境影响，但是其长期稳定可靠性仍需要通过大量的在轨实践证实。值得注意的是，虽然国内已有较高算力的在轨试验系统，但多以国外生产的 GPU 芯片为主，长此以往，对国外商业化芯片的依赖性会逐渐增大，在未来构建大规模算力网络时极易遭受遏制。因此，一方面有必要围绕现有国产商业智能芯片与模组开展星载环境适应性的筛选与在轨验证研究，例如华为 Atlas 200（22 TOPS@INT8，11 W）、寒武纪 MLU220（16TOPS @INT8，15 W）、比特大陆 SophonSM5（17.6TOPS @INT8，20 W）以及地平线征程 5（128TOPS@ INT8，30 W）等，以提升我国在星载智能计算系统领域的自主化能力；另一方面有必要开展支持新型智能计算范式的计算芯片与系统研发，例如类脑计算芯片与系统，以突破现有计算能效瓶颈，为星载低功耗大规模计算提供更优方案。

1.2 天地协同智能计算

得益于星载智能计算芯片与系统的发展，近年来国内外亦开展了各类星载智能计算的技术验证（图 2）。

国际上，PhiSat-1 卫星[6, 12]通过在轨部署 CloudScout 图像分割深度神经网络（Deep Neural Network，DNN），针对高光谱图像实现了云层覆盖区域的高效检测与分割，准确率达到 96%，这一成果不仅验证了卫星在轨智能推理计算的实用性，也为后续相关技术的发展提供了重要的参考。进一步，Intuition-1 卫星[7]也针对高光谱图像，通过在轨部署深度神经网络进行裸露土壤检测等智能推理任务，从而实现过滤和处理轨道上的数据，以绕过高传输带宽消耗以及向地球发送原始数据的滞后问题。2024 年，CogniSAT-6 卫星[13, 14]在轨运行集成了人工智能与实时通信功能的 CogniSatApp 控制软件，该软件能够通过持续星间链路接收地面移动设备（如手机）信息，并据此控制启动不同的深度神经网络进行在轨推理，从而实现卫星对观测图像的动态智能决策推理；通过实际验证，在 90 秒内，CogniSAT-6 卫星可以检测到 20×40 公里区域内的每艘船只；这一创新实践验证了将卫星由静态智能计算转变为动态智能计算的可行性，也标志着天地动态交互式智能计算新范式的形成。

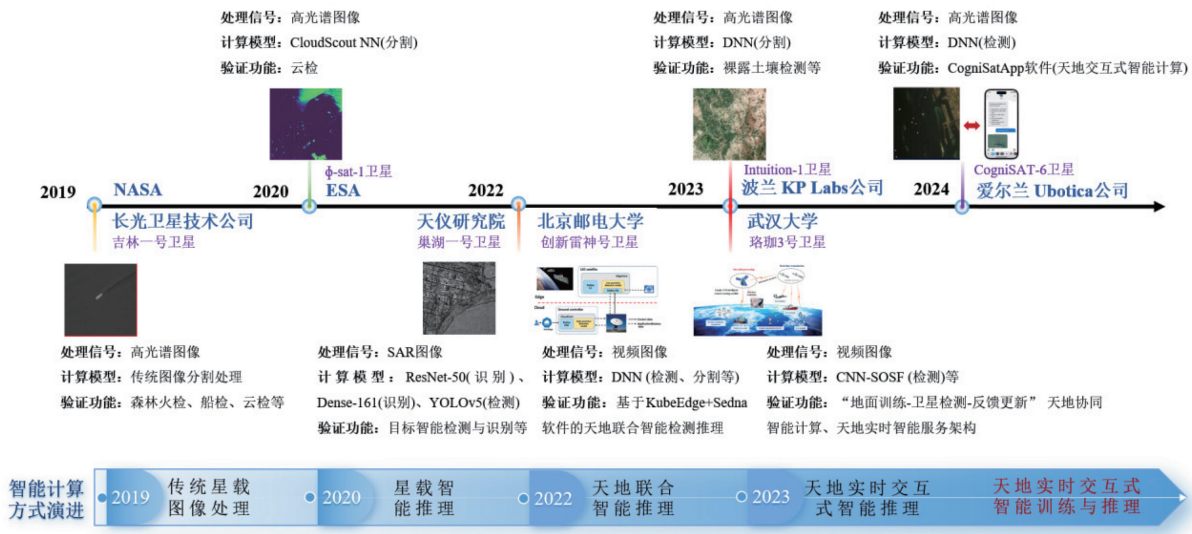

图 2 天地协同智能计算方式的发展

Fig. 2 The Development of Space-Ground Collaborative Intelligent Computing Methods

国内，2019 年吉林 1 号[15]卫星针对高光谱图像，通过在轨部署传统图像检测分割模型（图像二值化＋域连通＋分割），成功实现了森林火检、船检、云检等功能。2022 年，巢湖一号[8]卫星针对合成孔径雷达（Synthetic Aperture Radar，SAR）图像，通过在轨部署深度神经网络模型，如 ResNet-50、Dense-161、YOLOv5 等，进行了目标智能检测与识别验证。同年，创新雷神号卫星[16, 17]基于 KubeEdge（地面）＋Sedna（星载）软件，先通过卫星低精度智能检测模型获得感兴趣区域，然后采用地面高精度智能检测模型下载感兴趣区域图像进行精确计算的方式，实现了农业用地面积检测，验证了天地联合智能推理的可行性。2023 年，珞珈 3 号卫星[18, 19]在遥感视频图像处理方面，通过“地面训练—卫星检测—反馈更新”的循环机制，提高了在轨智能目标检测的精度和可迁移性；同时，提出结合星间、星地、5G 移动通信网络的实时智能服务架构，验证了从指令上行、数据采集、在轨实时处理到遥感信息下发到终端的全过程，响应时间不到 8 分钟，为天地实时智能服务应用提供了新模式。

从目前的在轨实验可以看出，星载智能计算方式正沿着“传统星载图像处理”“星载智能推理”“天地联合智能推理”“天地实时交互式智能推理”的路线发展。相对之下，国际技术更追求“星上动态自治”，国内则更注重“天地资源互补”，但本质上都是为了解决通信带宽限制带来的信息获取时延问题。值得注意的是，现有的计算方式主要是针对预先设定的任务。而卫星若要完成及时发现和识别新类别目标等突发性任务，如果卫星不具备学习能力，受限于动态通信带宽以及天地交互等待周期的限制，目前的天地协同计算方式难以进一步提升信号获取的时效。因此，未来有必要研究星端自主学习、天地协同进化等智能计算方式，这对于我国全球布站难、卫星观测信号难落地的特殊国情而言尤为重要。

另一方面，目前在轨智能计算主要采用了常用的深度神经网络模型，这主要是因为现有星载计算算力＜100TOPS，尚难以支持采用大模型进行智能处理；未来，随着更大星载算力的建设，结合大模型实现天地协同智能计算也必将成为研究趋势。

2　类脑智能计算发展现状

传统的人工智能计算已经在卫星应用中取得了重大进展；然而，传统的摩尔计算技术，如基于半导体架构的 CPU 和 GPU，预计将在 2025 年面临“数字墙”，这可能迫使计算技术范式发生转变[20]。2024 年诺贝尔物理学奖得主 Geoffrey Hinton 指出，突破当前人工智能技术的关键在于在人工智能和人脑之间建立桥梁。人脑由 10^{11} 个神经元构成，但仅消耗 10～20 W 功率。类脑计算受生物脑网络启发，采用脉冲序列进行数据的表征处理，能够以极低的能耗处理复杂数据[20]。其中，类脑计算模型又称为脉冲神经网络（Spiking Neural Networks，SNN），具备事件驱动和时空编码特征，同时具备生物合理和计算高效的特性，被称为第三代人工神经网络，可以有效支持低功耗计算。这些技术的主要优点是节能高效和设备上自适应能力，尤其适用于对于能耗敏感、需要持续信号处理的卫星系统。近年，国内外在类脑计算模型算法、类脑计算芯片系统、类脑智能在星载计算中的应用均取得了一定进展（图 3），下面将对这三方面的发展现状分别进行分析讨论。

2.1　类脑模型与算法

人工神经网络（Artificial Neural Networks，ANN）通过训练优化参数和构建适当架构来适应特定任务需求，其中模型算法和网络架构的选择对其性能至关重要。当前，基于反向传播的优化方法已成为大规模 ANN 训练的主流方法，该方法通过链式法则逐层计算参数的梯度并更新权重。相比之下，SNN 由于采用离散脉冲信号进行信息传递，其表示能力受限且无法微分，这导致深度 SNN 在训练过程中面临梯度消失和爆炸等问题，使其长期局限于简单任务和浅层网络结构，发展进程相对缓慢[21]。当前 SNN 研究主要集中在神经元模型、训练算法以及模型结构三个方面。

在神经元模型方面，作为神经网络的基本信息单元，脉冲神经元模型可根据抽象程度和功能特性进行分类。随着模型复杂度增加，其生物可信度和计算开销也随之上升。其中，Leaky Integrate-and-Fire（LIF）神经元模型虽然经典，但其单一性限制了 SNN 的表达能力。为此，研究者提出了多种增强型神经元模型：PLIF（parametric leaky integrate and fire）模型通过引入可学习的膜电位时间常数 τ 来平衡记忆和遗忘过程；BDETT（bioinspired dynamic energy-temporal threshold）模型通过采用动态阈值机制提高计算效率[22]；Diet-SNN 模型则结合了可学习的时间常数和阈值参数[23]；GLIF（gated leaky integrate and fire）模型通过引入门控单元增强表达能力；Augmented Spiking Neuron 模型通过扩展脉

冲信息维度以提升计算性能[24]；PSNN（parallel SNN）模型则通过并行计算架构大幅提高了训练效率。值得注意的是，虽然 Hodgkin-Huxley 模型具有较高的生物可信度，但由于其计算复杂度较高，主要应用于脑模拟而难以适用于 SNN[25]。

在训练算法方面，相关研究主要可分为无监督学习方法和直接监督训练算法。

早期 SNN 主要采用基于生物启发的无监督学习方法，其核心是神经突触可塑性机制，突触可塑性描述了神经元连接随突触脉冲的动态变化特性。深度 SNN 无监督学习算法通常融合多种生物启发机制，如多类型神经元组合、STDP（Spike-Timing-Dependent Plasticity）规则、自组织反向传播等，但这些方法适用的网络深度受限，在复杂数据集上的有效性尚需验证。为解决深度 SNN 直接训练难题，研究者提出了 ANN-SNN 转换方法，通过调整训练好的 ANN 权值来生成高效的 SNN，从而避免 SNN 直接训练中不可微的问题。目前，ANN-SNN 转换的仿真步长已从数千步降至不足百步，且性能损失几乎可以忽略[21]。然而，该方法在速度和能耗方面仍存在瓶颈，且难以发挥 SNN 的时间编码优势和生物合理性。知识蒸馏方法则利用预训练 ANN 教师模型指导 SNN 学生模型的学习过程，如 Local Tandem Learning 利用 ANN 中间特征监督 SNN 学习，SNN 蒸馏则基于特征和逻辑信息进行模型压缩[26]。

近年来，直接监督训练算法已在 SNN 中也取得了显著进展。针对脉冲时刻不可导的问题，代理梯度算法通过连续函数近似阶跃函数实现梯度反向传播，有效缓解了脉冲不可微问题，但其近似梯度与真实梯度之间的差异仍会影响模型性能。为此，Hybrid Macro/Micro Level Backpropagation 方法[27]将误差分解为微观电位变化和宏观损失函数的反向传播；Spatio-temporal Backpropagation 方法[28]则利用近似导数实现 SNN 的时间与空间信用分配，这些方法显著提高了时序动态性能。在处理时空事件流数据时，直接训练的 SNN 展现出优于 ANN 的性能，并具有更低的计算开销。

在网络架构方面，相关研究主要集中在卷积 SNN、Transformer-based SNN 等方面。

卷积 SNN 架构常借鉴 ANN 中的残差连接思想来解决深层网络中的梯度问题，其残差连接形式逐步演化为加后激活型、加前激活型和预激活型三类。其中，加后激活型基于经典 ResNet 结构，使用脉冲神经元替代浮点激活函数，如 Spike-Norm 模型首次实现了基于 ResNet 的 ANN-SNN 转换算法，将深度拓展至 50 层[29]。加前激活型采用加前激活残差连接，能显著缓解深度 SNN 梯度退化问题，如 SEW-ResNet 模型首次实现了超过 100 层的 SNN[21]。预激活型在卷积层前进行脉冲激活，实现了完全事件驱动的低功耗计算，如 MS-ResNet 模型成功训练了 400 层深度的 SNN[30]，有效解决了梯度爆炸和消失问题。

与卷积 SNN 侧重局部特征提取不同，Transformer 架构通过注意力机制关注全局信息，减小了归纳偏置。随着 Transformer 在多个领域的成功应用，SNN 领域也开始探索自注意机制的应用。Spikformer 模型引入了一种脉冲自注意机制（Spiking Self-attention，SSA），显著提高了网络在多任务环境下的学习效率和准确性[31]；Spike-driven Transformer 模型提出了基于事件驱动的自注意机制（Spike-Driven Self-attention，SDSA），显

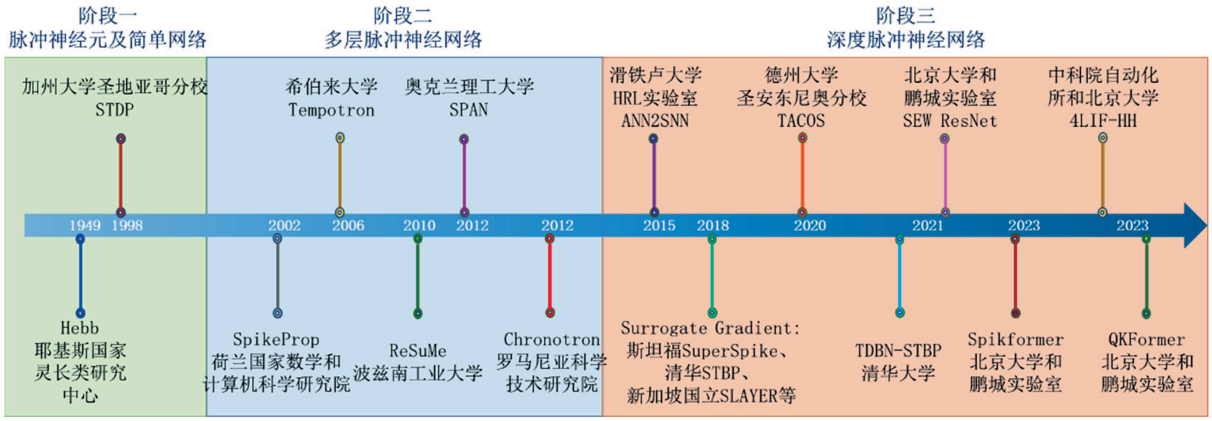

图 3 脉冲神经网络的发展

Fig. 3 The Development of Spiking Neural Networks

著降低了计算成本的同时，在多项复杂任务中保持了优异的性能，展示了 SNN 在大规模数据处理中的潜力[32]；QKFormer 模型在 token 和 channel 维度分别计算注意力，成为首个在 ImageNet-1K 数据集上准确率超过 85%的直接训练 SNN 模型，超越了绝大多数 Transformer 的 ANN 模型，其 Top-5 准确率达到 97.74%，超越人类水平的 95%[33]。

2.2　类脑计算芯片系统

高效硬件对加速大规模类脑模型的训练与推理，以及实际应用具有关键作用。与通用处理器相比，专用深度学习芯片和类脑芯片通过架构优化显著提升计算效率，能实现更好的功耗/性能/面积比。现有深度学习芯片多基于冯·诺伊曼架构，其计算与存储分离，而类脑计算芯片通过设计高效的存储与计算结构，可以极大提升并行计算效率。

类脑计算芯片受到大脑神经系统计算与存储并行特性的启发，多采用近存计算或存算一体架构。近存计算芯片在每个计算核中紧密耦合计算和存储资源，通过高效的片内外互联降低功耗。典型代表包括 Akida 1000、IBM 的 TrueNorth[34]、Intel 的 Loihi[35]以及清华大学的天机芯[36]等。这些芯片充分利用脉冲计算的稀疏性与事件驱动特性，实现了超低功耗计算，并支持 STDP 等局部学习规则。其中，天机芯融合 SNN 与 ANN 的优势，大幅提升了准确率与鲁棒性。另一类重要架构是存算一体架构，其通过将计算单元和存储单元物理融合，利用突触内存直接执行矩阵运算。存算一体芯片可分为基于传统存储器（如 SRAM、DRAM）[37]和新兴存储器（如忆阻器）两类。忆阻器芯片通过在突触内实现计算和存储的深度融合，能支持超低功耗的 SNN 计算，但目前因制造工艺限制，仍处于小规模试验阶段。

从电路设计的角度来看，类脑计算芯片可分为模拟电路、模拟—数字混合电路和全数字电路。早期芯片如 ROLLS 主要基于模拟电路，功耗低但难以编程，且对环境敏感。混合电路芯片如 DYNAPs[38]结合了模拟电路的仿真性与数字电路的稳定性，具有更强的适应性，但设计复杂度高。全数字电路芯片（如 SpiNNaker[39]、TrueNorth、Akida 1000）则提供了更高的灵活性与稳定性，但功耗相对较高，在生物仿真性方面存在一定局限性。

2.3　类脑智能在星载计算中的应用

类脑计算通过事件驱动的脉冲信号处理来实现高效计算，具备显著的功耗优势，尤其适合能源受限、实时性要求高的星载计算任务。目前，国内外已经初步在地球观测、卫星通信等应用任务上开展了类脑智能计算试验探索。

在地球观测任务方面，SNN 可以用于低功耗的卫星图像识别、检测等任务。2019 年，ESA 发射了 OPS-SAT 卫星，在 FPGA 上部署了基于 SNN 的 SPLEAT 架构，用于云层分类任务，实验表明，CNN 的 FPGA 资源占用为 SNN 的 6.62 倍，CNN 的功耗为 SNN 的 5.91 倍，同时 SNN 保持了与 CNN 相当的分类性能[40]。2023 年，印度 TCS 公司在 Akida 1000 芯片上部署了用于云层覆盖检测的 SNN，相对于在 GPU 部署能效比提升 35 倍[41]。2024 年，哈工大提出基于 SNN 的 RepSViT 模型，用于卫星遥感图像的目标识别，并基于嵌入式 GPU 上实现了 8.33 毫秒的推理延迟和 95%的识别准确率，计算成本为 0.6 GFLOPs，功耗较传统 CNN 降低了约 30%[42]。

在卫星通信方面，2022 年，NASA 发射 TechEdSat-13 卫星，首次搭载了含有 Loihi 芯片的类脑计算模块 Kapoho Bay，并实现了一种自主 SNN 的软件定义网络认知代理，在动态网络环境中实现了约 90%的最优数据路径选择[43]。同年，ESA 也发射了 Satcom 通信卫星，基于神经形态计算芯片 Loihi2，验证了 SNN 在动态波束成形任务中的能效比 CNN 提升超过 100 倍[44]。此外，SNN 在信道预测上也表现出色，可以帮助提前优化波束方向，从而提升连接质量和通信效率。针对高频段，SNNs 的高效时序特性可以实现低功耗的干扰检测和频谱分配，尤其在低地轨道卫星（Low Earth Orbit，LEO）等需要高动态频谱管理的场景中表现出色[45]。

从国内外应用探索的现状来看，以 SNN 为主的类脑计算技术在卫星通信和地球观测领域已初步展现出应用潜力。国外围绕类脑智能计算系统和模型算法已开展了多项星载实验，取得了重要进展；相比之下，国内的研究仍主要集中于地面算法模型的实验验证阶段。尽管类脑计算技术目前仍处于发展初期，其软硬件生态系统尚未完全成熟，但随着技术的不断突破与优化，这一领域有望为卫星系统的高效化与可持续性发展提供革命性的解决方案。

3　未来的探索方向

类脑智能计算作为一种创新的计算范式，对于推动当前天地一体化智能计算的变革具有重大的潜力，以下几个方向在未来研究中极有探索价值：星载

低功耗类脑计算芯片与系统、大规模高效能类脑脉冲神经网络，以及天地协同增量学习新模式。

3.1 星载高性能类脑计算芯片与系统

星载智能计算系统的性能，尤其是在算力和能耗方面，极大程度上取决于其搭载的智能计算芯片的能力。目前，已经验证的星载类脑计算芯片，如Loihi2，虽然具有一定的潜力，但其性能尚不足以支持复杂层次的 SNN 模型。因此，有必要进一步研究和开发具有高算力且低功耗特性的类脑计算芯片，以支持深度 SNN 网络模型的训练与推理。近年，H2Learn[46] 和 SATA[47] 为深度 SNN 的训练设计了特定的架构，为支持深度学习的低功耗类脑芯片设计奠定了基础。

系统层面，研究耐辐射的类脑计算系统架构显得尤为重要，这包括探索冗余设计、错误检测和动态配置等关键技术，以确保在部分部件损坏的情况下，计算系统仍能维持正常运行。这些技术研究有助于减轻空间辐射对类脑计算系统的影响，从而提高其在空间应用中的可靠性。例如通过 FPGA 监控 Akida 1000 芯片状态，使得 Akida 芯片在遭受单粒子翻转后，通过 FPGA 的故障管理机制迅速恢复功能，保障不出现永久性故障，同时芯片输出的脉冲信号信噪比仍保持在了 3.355 以上[48]。

3.2 大规模高效能类脑脉冲神经网络

当前，部署于卫星的深度神经网络模型通常是为了特定任务而训练的，这导致它们的泛化能力受限。通过利用大型预训练模型的强大通用泛化能力，可以为卫星上多任务处理提供一种高效的解决方案[49]。然而，这些大型预训练模型在训练和推理过程中需要消耗大量的能量，这对于在能量受限的卫星环境中进行部署构成了挑战。

脉冲神经网络 SNN 与人工神经网络 ANN 相比，因其更高程度的生物合理性而在理论上具有显著的能耗降低潜力[50]，例如最近报道的 QKFormer SNN 模型实践证实，在同等精度下，ANN Transformer 能耗为 SNN Transformer 的 2～8 倍[33]。因此，研究和发展面向卫星信号处理的大规模类脑脉冲神经网络模型，为解决卫星上的高能耗计算问题提供了一种有前景的方案，未来的研究应进一步结合具体星载计算任务需求探索其实际应用效能。

3.3 天地协同增量学习新模式

在应急救灾、军事目标监测等关键应用领域，卫星系统能迅速适应新目标的能力显得尤为关键。然而，当前的天地协同计算模式主要依赖于地面预训练模型来执行星载推理任务。由于天地交互等待周期长、数据通信带宽有限，这些限制因素导致缺乏学习能力的卫星难以对新目标进行快速适应。此外，由于卫星上的数据通常没有标签，这使得卫星难以采用地面环境中常用的有监督学习方法进行模型训练。

为了解决上述挑战，提出一种“地面上载轻量模型—卫星无监督学习—卫星检测下载数据—地面有监督学习—星地模型同步更新”的天地协同学习计算架构，如图 4 所示。以目标检测识别任务为例，该架构包括以下几个关键步骤：

(1) 地面计算中心收集一个包含 N 类目标的大样本池，并利用大样本池进行大模型训练（例如 RepsViT[42]、QKFormer [33]等 SNN 模型），采用模型压缩技术（例如 Gradient Rewiring[51]、软阈值剪枝[52]等 SNN 压缩技术），将大型模型转化为轻量级模型，并将其上载至卫星。

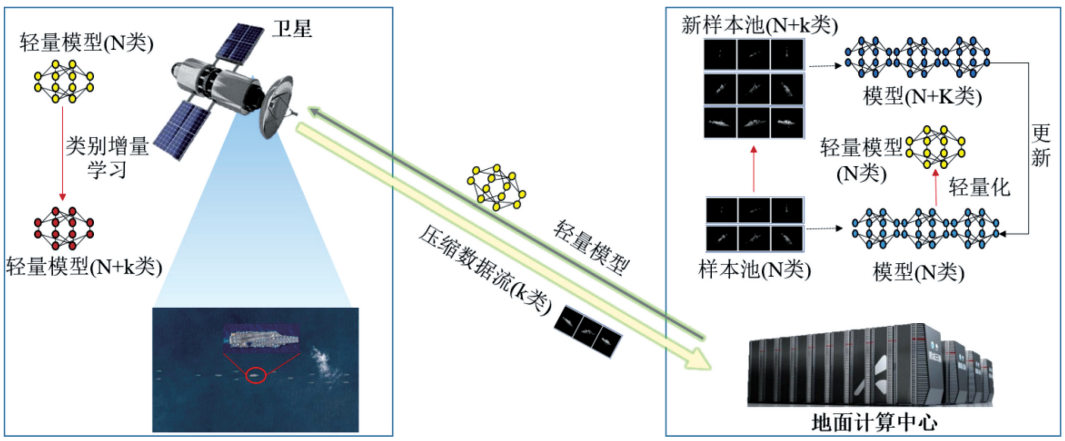

图 4 天地协同增量学习计算框架

Fig. 4 Space-Ground Collaborative Incremental Learning Computing Framework

（2）卫星接收到轻量模型后，获取小样本观测数据，采用类别增量学习方法对轻量模型进行无监督训练（例如 DSD[53]、SOR[54]、HLOP[55]等 SNN 增量学习方法），这一过程使得卫星能够快速获得一个对 N＋k 类目标具有初步检测与识别能力的轻量模型。

（3）随后，卫星使用该（N＋k）类目标检测与识别模型进行新目标的观测，并将包含新目标信息的局部观测数据压缩（例如 SNN-SC[56]等 SNN 压缩编码方法）后下载至地面计算中心。

（4）地面计算中心对下传的数据进行整理，构建新的样本池，并在此基础上进行有监督学习，以获得一个准确性更高的（N＋k）类检测与识别模型。

（5）最终，基于（N＋k）类检测与识别模型，同时对星地模型进行同步更新。

通过上述具有持续学习能力的天地协同增量学习机制，理论上卫星系统能够显著提升对新目标和新任务的适应性和响应速度。当然，要在实践中验证上述机制，还可能面临诸多挑战，例如数据传输、计算资源等限制。数据通信带宽限制可能导致数据传输延迟，因此，可以进一步考虑进行观测数据特征提取，通过天地数据特征流传输[57]，降低传输延时；此外，计算与存储开销往往会随着增量学习任务数的增多而线性增长，因此，需要面向卫星长期运行环境，进一步考虑设计适应长数据流的类别增量学习模型结构[58]。

4　结　语

本文面向天地一体化算力网的发展，从天地协同智能计算技术以及类脑智能计算技术的前沿发展与应用进行了讨论与总结，并对类脑智能＋天地协同的未来发展范式进行了展望构想，提出星载低功耗类脑计算芯片与系统、大规模高效能类脑脉冲神经网络以及天地协同增量学习架构等值得探索方向，这不仅对于推动类脑智能计算技术的发展具有重要意义，对于实现天地一体化智能计算网络的高效协同和智能化升级也有望产生深远影响。

参　考　文　献

[1] 昌道励，许宁宁，范永敬，等，“像用电用水一样便捷地用算力”，南方日报，2023-05-30(A03).

[2] 蒋忠元，王森，王启舟，等. 低轨卫星多星协同及星地协同遥通算一体化技术. 天地一体化信息网络，2024，5(1)：60—75.
Jiang ZY，Wang S，Wang QZ，et al. LEO multi-satellite coordination and satellite-ground coordination remote sensing communication and computation integration technology. Space-Integrated-Ground Information Networks，2024，5(1)：60—75.（in Chinese）

[3] 虞志刚，冯旭，陆洲，等. 宇航级处理器发展现状与趋势. 天地一体化信息网络，2023，4(1)：50—58.
Yu ZG，Feng X，Lu Z，et al. Development status and trends of space processor. Space-Integrated-Ground Information Networks，2023，4(1)：50—58.（in Chinese）

[4] 虞志刚，冯旭，戴天，等. 空间边缘计算：需求、架构及关键技术. 电子与信息学报，2022，44(12)：4416—4425.
Yu ZG，Feng X，Dai T，et al. Space edge computing：Requirement，architecture and key technique. Journal of Electronics & Information Technology，2022，44(12)：4416—4425.（in Chinese）

[5] Brewer C，Franconi NG，Ripley R，et al. NASA SpaceCube intelligent multi-purpose system for enabling remote sensing，communication，and navigation in mission architectures. [2024-10-22]. https://ntrs.nasa.gov/citations/20205005819.

[6] Giuffrida G，Fanucci L，Meoni G，et al. The Φ-sat-1 mission：The first on-board deep neural network demonstrator for satellite earth observation. IEEE Transactions on Geoscience and Remote Sensing，2021，60：5517414.

[7] Wijata AM，Lakota T，Cwiek M，et al. Intuition-1：Toward in-orbit bare soil detection using spectral vegetation indices// IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium. Athens，Greece：IEEE，2024：1708—1712.

[8] Wang MQ，Qiu H，Xu LN，et al. A case for application-aware space radiation tolerance in orbital computing. [2024-10-22]. https://arxiv.org/abs/2407.11853v1.

[9] Li Q，Wang SG，Ma X，et al. Towards sustainable satellite edge computing// 2021 IEEE International Conference on Edge Computing（EDGE）. Chicago，IL，USA. IEEE，2021：1—8.

[10] Dai RF，Wang M，Chen R，et al. Target-driven real-time geometric processing based on VLR model for LuoJia3-02 satellite. IEEE Geoscience and Remote Sensing Letters，2024，21：6016605.

[11] 胡珉琦，软件定义卫星“天智二号”D 星成功发射，中国科学报，2023-01-17(001).

[12] Giuffrida G，Diana L，de Gioia F，et al. CloudScout：A deep neural network for on-board cloud detection on hyperspectral images. Remote Sensing，2020，12(14)：2205.

[13] Rijlaarsdam D, Hendrix T, González PTT, et al. The next era for earth observation spacecraft: An overview of CogniSAT-6. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2025, 18: 2450—2463.

[14] Rijlaarsdam D, Hendrix T, González PTT, et al. Autonomous operational scheduling on cognisat-6 based on onboard artificial intelligence. 17th Symposium on Advanced Space Technologies in Robotics and Automation, 2023.

[15] Yu SH, Yu YB, He XJ, et al. On-board fast and intelligent perception of ships with the "Jilin-1" spectrum 01/02 satellites. IEEE Access, 2020, 8: 48005—48014.

[16] Wang C, Zhang YR, Li Q, et al. Satellite computing: A case study of cloud-native satellites// 2023 IEEE International Conference on Edge Computing and Communications (EDGE). Chicago, IL, USA. IEEE, 2023: 262—270.

[17] Wang SG, Li Q. Satellite computing: Vision and challenges. IEEE Internet of Things Journal, 2023, 10(24): 22514—22529.

[18] Wang M, Wu QY, Xiao J, et al. Luojia 3-01 satellite—real-time intelligent service system for remote sensing science experiment satellite. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024, 17: 8250—8257.

[19] Li YJ, Wang M, Hwang K, et al. LEO satellite constellation for global-scale remote sensing with on-orbit cloud AI computing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 16: 9369—9381.

[20] Mead C, Ismail M. Analog VLSI implementation of neural systems. Massachusetts: Kluwer, 2012.

[21] Fang W, Yu ZF, Chen YQ, et al. Deep residual learning in spiking neural networks. [2024-10-22]. https://www.semanticscholar.org/paper/Deep-Residual-Learning-in-Spiking-Neural-Networks-Fang-Yu/35fcc65db264d9c73ed19d0f3f9d53a43991bdbf.

[22] Fang W, Yu ZF, Chen YQ, et al. Incorporating learnable membrane time constant to enhance learning of spiking neural networks// 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada. IEEE, 2021: 2641—2651.

[23] Rathi N, Roy K. DIET-SNN: A low-latency spiking neural network with direct input encoding and leakage and threshold optimization. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(6): 3174—3182.

[24] Yao XT, Li FR, Mo ZT, et al. GLIF: A Unified Gated Leaky Integrate-and-Fire Neuron for Spiking Neural Networks. [2024-10-22]. https://proceedings.neurips.cc/paper_files/paper/2022/hash/cfa8440d500a6a6867157dfd4eaff66e-Abstract-Conference.html.

[25] He LX, Xu YH, He WH, et al. Network model with internal complexity bridges artificial intelligence and neuroscience. Nature Computational Science, 2024, 4(8): 584—599.

[26] Yang Q, Wu JB, Zhang ML, et al. Training spiking neural networks with local tandem learning. [2024-10-22]. https://proceedings.neurips.cc/paper_files/paper/2022/hash/523caec7832a47fb19b8471dbfeec471-Abstract-Conference.html.

[27] Jin Y, Zhang WR, Li P. Hybrid macro/micro level backpropagation for training deep spiking neural networks. [2024-10-22]. https://arxiv.org/abs/1805.07866v6.

[28] Wu YJ, Deng L, Li GQ, et al. Spatio-temporal backpropagation for training high-performance spiking neural networks. Frontiers in Neuroscience, 2018, 12: 331.

[29] Zheng HL, Wu YJ, Deng L, et al. Going deeper with directly-trained larger spiking neural networks. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(12): 11062—11070.

[30] Yao M, Gao HH, Zhao GS, et al. Temporal-wise attention spiking neural networks for event streams classification// 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada: IEEE, 2021: 10201—10210.

[31] Zhou ZK, Zhu YS, He C, et al. Spikformer: When spiking neural network meets transformer. [2024-10-22]. https://arxiv.org/abs/2209.15425.

[32] Yao M, Hu JK, Zhou ZK, et al. Spike-driven transformer. [2024-10-22]. https://proceedings.neurips.cc/paper_files/paper/2023/hash/ca0f5358dbadda74b3049711887e9ead-Abstract-Conference.html.

[33] Zhou CL, Zhang H, Zhou ZK, et al. QKFormer: Hierarchical spiking transformer using QK attention. [2024-10-22]. https://arxiv.org/abs/2403.16552.

[34] Akopyan F, Sawada J, Cassidy A, et al. TrueNorth: design and tool flow of a 65 mW 1 million neuron programmable neurosynaptic chip. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2015, 34(10): 1537—1557.

[35] Davies M. Taking neuromorphic computing to the next level with Loihi2. [2024-10-22]. https://download.intel.com/newsroom/2021/new-technologies/neuromorphic-computing-loihi-2-brief.pdf.

[36] Pei J, Deng L, Song S, et al. Towards artificial general intelligence with hybrid Tianjic chip architecture. Nature, 2019, 572(7767): 106—111.

[37] Wu MH, Huang MS, Zhu ZF, et al. Compact probabilistic poisson neuron based on back-hopping oscillation in STT-MRAM for all-spin deep spiking neural network// 2020 IEEE Symposium on VLSI Technology. Honolulu, HI, USA. IEEE, 2020: 1—2.

[38] Moradi S, Qiao N, Stefanini F, et al. A scalable multicore architecture with heterogeneous memory structures for dynamic neuromorphic asynchronous processors (DYNAPs). IEEE Transactions on Biomedical Circuits and Systems, 2018, 12(1): 106—122.

[39] Painkras E, Plana LA, Garside J, et al. SpiNNaker: A 1-W 18-core system-on-chip for massively-parallel neural network simulation. IEEE Journal of Solid-State Circuits, 2013, 48(8): 1943—1953.

[40] Abderrahmane N, Miramond B, Kervennic E, et al. SPLEAT: SPiking Low-power Event-based ArchiTecture for in-orbit processing of satellite imagery// 2022 International Joint Conference on Neural Networks (IJCNN). Padua, Italy: IEEE, 2022: 1—10.

[41] Kadway C, Dey S, Mukherjee A, et al. Low power & low latency cloud cover detection in small satellites using on-board neuromorphic processors// 2023 International Joint Conference on Neural Networks (IJCNN). Gold Coast, Australia: IEEE, 2023: 1—8.

[42] Pang YH, Yao LB, Luo YP, et al. RepSViT: An efficient vision transformer based on spiking neural networks for object recognition in satellite on-orbit remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5610916.

[43] Rahman N, Yakopcic C, Lent R, et al. Neuromorphic hardware in outer space: Software defined networking executed on an in-orbit loihi spiking processor// 2023 IEEE Cognitive Communications for Aerospace Applications Workshop (CCAAW). Cleveland, OH, USA. IEEE, 2023: 1—5.

[44] Ortiz F, Skatchkovsky N, Lagunas E, et al. Energy-efficient on-board radio resource management for satellite communications *via* neuromorphic computing. IEEE Transactions on Machine Learning in Communications and Networking, 2024, 2: 169—189.

[45] Ye ZL, Liu R, Taggart JL, et al. Evaluation of radiation effects in RRAM-based neuromorphic computing system for inference. IEEE Transactions on Nuclear Science, 2019, 66(1): 97—103.

[46] Liang L, Qu Z, Chen ZD, et al. H2Learn: High-efficiency learning accelerator for high-accuracy spiking neural networks. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2022, 41(11): 4782—4796.

[47] Yin RK, Moitra A, Bhattacharjee A, et al. SATA: Sparsity-aware training accelerator for spiking neural networks. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2023, 42(6): 1926—1938.

[48] Mena Morales R, Miralles P, Devani D, et al. Brainsat: Hardware development of a neuromorphic on-board computer applied to methane detection from low earth orbit//IAF Space Systems Symposium. Milan, Italy: International Astronautical Federation (IAF), 2024: 1—15.

[49] Wang D, Zhang J, Xu MQ, et al. MTP: Advancing remote sensing foundation model *via* multitask pretraining. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024, 17: 11632—11654.

[50] 马征宇，田永鸿. 大规模类脑神经网络理论与方法. 科学通报，2023，68(35)：4764—4781.
Ma ZY, Tian YH. Theories and methods for large-scale brain-inspired neural networks. Chinese Science Bulletin, 2023, 68(35): 4764—4781. (in Chinese)

[51] Chen YQ, Yu ZF, Fang W, et al. State transition of dendritic spines improves learning of sparse spiking neural networks// 39th International Conference on Machine Learning, 2022: 3701—3715.

[52] Chen YQ, Ma ZY, Fang W, et al. A unified framework for soft threshold pruning. [2024-10-22]. https://arxiv.org/abs/2302.13019.

[53] Han B, Zhao FF, Zeng Y, et al. Enhancing efficient continual learning with dynamic structure development of spiking neural networks. [2024-10-22]. https://arxiv.org/abs/2308.04749.

[54] Han B, Zhao FF, Pan WX, et al. Adaptive reorganization of neural pathways for continual learning with spiking neural networks. [2024-10-22]. https://arxiv.org/abs/2309.09550.

[55] Xiao MQ, Meng QY, Zhang ZP, et al. Hebbian learning based orthogonal projection for continual learning of spiking neural networks. [2024-10-22]. https://arxiv.org/abs/2402.11984.

[56] Wang MY, Li JH, Ma MY, et al. SNN-SC: A spiking semantic communication framework for collaborative intelligence. IEEE Transactions on Vehicular Technology, 2025, 74(4): 5883—5896.

[57] 高文，田永鸿，王坚. 数字视网膜：智慧城市系统演进的关键环节. 中国科学：信息科学，2018，48(8)：1076—1082.
Gao W, Tian YH, Wang J. Digital retina: Revolutionizing camera systems for the smart city. Scientia Sinica (Informationis), 2018, 48(8): 1076—1082. (in Chinese)

[58] 周大蔚，汪福运，叶翰嘉，等，基于深度学习的类别增量学习算法综述. 计算机学报，2022，46(8)：1577—1605.
Zhou DW, Wang FY, Ye HJ. Deep learning for class-incremental learning: A survey. Chinese Journal of Computers, 2022, 46(8): 1577—1605. (in Chinese)

Current State and Future Perspectives of Brain-Inspired Intelligent Computing for Integrated Space-Ground Computing Networks

Wei Zhang [1] Zhengyu Ma[1*] Yonghong Tian[1, 2, 3*]

1. *Peng Cheng Laboratory, Department of Intelligent Computing, Shenzhen 518000, China*
2. *School of Computer Science and Technology, Peking University, Beijing 100871, China*
3. *School of Information Engineering, Peking University, Shenzhen 518055, China*

Abstract Oriented to the national significant strategic demands of the integrated space-ground intelligent computing network, this paper first reviews the evolution of intelligent chips and systems deployed on satellites, and the advancement of space-ground collaborative intelligent computing. Additionally, the corresponding problems and future development trends in these domains are analyzed. Further, the paper summarizes the historical progression of brain-inspired models and algorithms, and brain-inspired chip systems. Also, the current state of brain-inspired intelligent computing within satellite applications is discussed. Finally, combining the latest advancements in brain-inspired intelligence and the future development needs of the integrated space-ground computing network, several research directions meriting exploration in the future are proposed. These include low-power spaceborne brain-inspired computing chips and systems, large-scale high-efficiency brain-inspired spiking neural networks, and a space-ground collaborative incremental learning architecture.

Keywords integrated space-ground; brain-inspired computing; high energy efficiency; collaborative Computing

马征宇 鹏城实验室智能计算部副研究员，博士生导师，深圳市高层次人才，主要从事脉冲神经网络、计算神经科学等研究。主持国家自然科学基金青年科学基金项目(C 类)、重点项目课题等，发表学术论文 30 余篇。

田永鸿 北京大学博雅特聘教授，博士生导师，电气电子工程师学会会士，北京大学深圳研究生院信息工程学院院长，鹏城实验室智能计算部副主任。主持国家自然科学基金青年科学基金项目(A 类)等项目。主要研究方向为分布式机器学习、神经形态视觉和视频大数据。

张 伟 博士，鹏城实验室智能计算部副研究员，深圳市高层次人才后备级，主要从事高速观测技术与仪器、神经形态计算视觉、激光干涉测量等研究，主持国家自然科学基金项目、博士后面上基金项目等，发表学术论文 30 余篇，获国家发明专利 10 余项。

(责任编辑 陈 鹤 张 强)

* Corresponding Authors, Email: mazhy@pcl.ac.cn; yhtian@pku.edu.cn