附件一：2021 年 CCF-海康威视斑头雁基金申报主题

目录

申报主题

（以下主题均不限于给定的建议研究方向，可基于研究者背景和兴趣确定）

课题一：大规模多智能体强化学习算法研究

【研究背景】

多智能体强化学习算法在下棋、游戏等领域已经取得了令人瞩目的成绩。然而在一些大规模多智能体协作的场景，如仓储环境的多智能体路径规划，智能体的数量可达成百上千。如何让如此大规模的智能体之间高效协作，对强化学习的算法框架设计提出了新的挑战。本课题基于海康威视自研的多智能体导航虚拟环境，以多智能体路径规划为切入点，研究大规模多智能强化学习算法，提升系统的整体运行效率。

【研究内容】

1. 面向大规模多智能体的强化学习策略协作研究；
2. 面对未知或动态变化环境下的强化学习策略泛化性研究；
3. 研究在相似环境之间快速迭代的迁移学习算法。

课题二：开放世界下的视觉感知与理解

【研究背景】

大规模 ImageNet 图像分类数据集的出现推动了深度学习革命，目前深度神经网络已经可以很好地解决传统的图像分类问题。然而,现实中开放世界的视觉环境配置远不如 ImageNet 那般理想化。现实世界是一个未经结构化梳理的、长尾分布的、开放类别的复杂场景。在这样的视觉环境下如何设计 AI 算法进行视觉感知与理解是一个非常有挑战性的长期课题。本课题旨在探索下一代智能视觉系统中的一些关键性问题，如长尾、噪声、灾难遗忘、概念漂移、无监督语义发现等，使得学习算法能够在开放环境中很好地完成视觉感知和语义理解。

【研究内容】

1. 开放世界下的图像识别；
2. 开放世界下的目标检测。

课题三：深度学习中的数据隐私保护和模型保护

【研究背景】

数据和模型是深度学习中最重要的两个资源。在数据方面，如何妥善地获取和使用数据已经越来越受产业界关注，尤其当深度学习的训练需要用到敏感的用户私人信息时，不恰当的使用方式将造成极大的隐私泄露隐患。在模型方面，目前训练一个性能优异的模型需要消耗大量的数据资源和计算资源，这些模型具备极大的经济效益甚至是一些企业的核心技术。当模型被盗用时很容易通过迁移学习或模型压缩等技术转为竞争对手所用。因此，模型受侵权将极大损害模型拥有者的利益。本课题旨在改进已有的深度学习算法，提升对数据的隐私保护和对模型的知识产权保护。

【研究内容】

1. 深度学习中的数据隐私泄露或模型窃取；
2. 深度学习中的数据隐私保护；
3. 深度学习中的模型保护。

课题四：基于生物启发的高效脉冲神经网络研究

【研究背景】

深度神经网络在认知任务上获得了广泛的成功，然而其计算效率及性能与生物脑仍有不小差距。脉冲神经网络作为一种具有较强生物基础理论支撑的计算模型，本身具有高效计算，时空信息表征，异步事件信息处理等能力，被业界普遍认为是新一代人工智能的基本形态。本课题旨在借鉴生物脑处理方式，研究高效脉冲神经网络学习方法及优化策略，最终实现具有更快学习速度，更小能量消耗，更强适应性的脉冲神经网络。

【研究内容】

1. 高效、鲁棒脉冲信息编码方式研究，基于生物学理论，探索具有更低能耗开销，更高鲁棒性的脉冲信息编码方法；
2. 基于生物多尺度可塑性的弱监督、无监督脉冲神经网络学习方法研究，探索更高精度及效率的弱监督、无监督脉冲网络学习方法及优化策略；
3. 人工神经网络与脉冲神经网络高效迁移方法研究，探索更高精度及效率的脉冲网络模型转换方法。

课题五：基于信息论的神经网络表达研究

【研究背景】

寻找适当的数学工具去建模深度神经网络的表达能力，将调参式深度学习模式过渡为以客观评测指标为指导的学习范式，是新一代人工智能需要面对的课题，也是在当前深度学习大背景下一个重要的突破方向。从本质上说，神经网络推理计算可以看作为对数据的一种信息传输、编码及解码过程，信息论作为一种研究信息传递和信息处理的基础理论方法，为研究神经网络表达提供了一种新的思路，结合信息论理论基础，对神经网络中信号传递及处理方式进行研究，探索出描述神经网络表达能力的参数化工具，从而为设计开发具有低算力，高表达能力的网络计算模型提供了坚实理论基础。

【研究内容】

1. 基于信息论的模型表示方法及建模研究，探索网络模型表达能力的参数化表示方法；
2. 模型表示方法在模型压缩、高效模型生成领域的研究，结合模型表示方法相关理论，研究新的模型压缩方式及其高效模型生成方法。

课题六：深度学习模型可解释性研究

【研究背景】

本项目期望通过对深度学习模型可解释性的基础原理、评测指标、优化方法的研究，指导解决当前深度学习中存在的“黑箱”问题带来的应用风险和道德伦理问题，提升跨场景模型鲁棒性和可靠性，杜绝算法歧视，促进深度学习基础理论研究。

【研究内容】

1. 可解释与可视化机理研究：研究深度学习语义概念抽象机理，一方面可以尝试对特征进行可视化，深入对深度学习表示原理的进一步探索；另一方面可以尝试构建本身具备可解释性的模型；
2. 可解释性与模型小型化方法研究：从可解释性角度评估模型能力，研究可解释的模型小型化方法，提高在指定场景领域内的小模型鲁棒性
3. 模型安全性评价研究：结合模型可解释性研究模型在高法律风险应用（医疗、自动驾驶等）中的模型安全性评价方法，提升模型生产中的风险规避能力。

课题七：基于 SRAM 的存算器件研究

【研究背景】

传统冯诺依曼架构计算和存储分离的特征，使得运算器和存储器之间的信息交换速度、功耗成为影响系统性能的主要因素。计算和存储一体化的存内计算形式可有效解决上述问题。其中，基于成熟工艺的 SRAM 的全数字存内计算是一种可行的方式，也有利于快速应用。

【研究内容】

1. 在现有较成熟的 coms 工艺平台，通过改变 foundry 标准 SRAM+逻辑门的边界，设计实现全新的带计算的 SRAM 的器件，定义新型的 SRAM_X 含特定的计算逻辑，目标是提升整体计算的效能，在相同工艺平台下，将效能提升达到 2 倍以上；
2. 技术难点：从晶体管级设计带计算的 SRAM 即 SRAM_X，现实自定义 SRAM_X 进行验证获得预期的效能收益。

课题八：大数据隐私计算加密技术

【研究背景】

大数据所蕴含的价值一方面切实地促进了产业的发展，另一方面也不可避免地给数据安全带来了新的挑战。特别是当大数据需要进行开放使用，需要对多方数据进行融合分析时，个人、组织的隐私数据保护越发重要，保证数据可用不可见，保护计算过程中多方数据的隐私安全成为刚性需求。针对上述问题，隐私计算是解决这一类问题的方向，其中同态加密是关键的技术点，其性能和应用过程带来的精度损失是当前业界的主要难题。

【研究内容】

同态加密方案与性能提速：

1. 能对加法同态、乘法同态、全同态等多种同态加密算法使用进行加速，同时安全级别达到业内领先；
2. 能够保证运算结果的完整性和准确性；
3. 数据规模支持亿级别及以上，可考虑在硬件层和算法软件层两层进行性能优化，性能较现有主流方案提高 10 倍；
4. 方案可在联邦学习算法中应用，且性能损耗不超过一个数量级。