



清华大学
Tsinghua University

赵俊博

183-4308-7267 | zhaojb21@mails.tsinghua.edu.cn



基本信息

性别: 男

年龄: 24 周岁

出生年月: 1999.04

现住址: 北京市海淀区

GitHub: <https://github.com/A-LinCui>

预计毕业时间: 2024.07

教育经历

清华大学

2021.08 – 至今

电子与通信工程 硕士

清华大学

2017.08 – 2021.06

工程物理 本科

研究经历

清华大学电子工程系

2019.08-至今

纳米集成电路实验室——高能效计算组 (NICS-EFC)

研究方向

- 神经架构搜索 (Neural Architecture Search, NAS)
- 对抗鲁棒性 (Adversarial Robustness)

研究成果

- Junbo Zhao***, Xuefei Ning*, Enshu Liu, Binxin Ru, Zixuan Zhou, Tianchen Zhao, Chen Chen, Jiajin Zhang, Qingmin Liao, Yu Wang, "**Dynamic Ensemble of Low-fidelity Experts: Mitigating NAS "Cold-Start"**", in **AAAI Conference on Artificial Intelligence (AAAI) Oral**, 2023.

基于预测器的神经架构搜索存在严重的“冷启动”问题，即需要大量的架构-真实性能数据才能获得一个有效的预测器。本工作提出利用低保真度信息以缓解预测器训练对数据量的需求，并提出动态集成预测框架以更好地融合不同类型的低保真信息提供的有益信息，在有限的数量下大幅提升了预测器的预测能力。

- Xuefei Ning*, Zixuan Zhou*, **Junbo Zhao**, Tianchen Zhao, Yiping Deng, Changcheng Tang, Shuang Liang, Huazhong Yang, Yu Wang, "**TA-GATES: An Encoding Scheme for Neural Network Architectures**", in **Neural Information Processing Systems (NeurIPS) Spotlight**, 2022.

既有基于图的神经架构编码方案（如GATES）仅对神经网络的前向推导过程进行了建模，而神经网络的性能与反向训练过程也息息相关。本工作提出了一种新的基于图的架构编码方案 TA-GATES，模拟架构的训练过程进行编码。实验验证该编码方案可以有效提高预测器在有限数据量下的预测能力，从而提高搜索得到的架构的性能。

- Xuefei Ning*, **Junbo Zhao***, Wenshuo Li, Tianchen Zhao, Yin Zheng, Huazhong Yang, Yu Wang, "**Discovering Robust Convolutional Architecture at Targeted Capacity: A Multi-Shot Approach**", in <https://arxiv.org/abs/2012.11835>.

神经网络的对抗鲁棒性与神经网络架构息息相关。此外，既有研究还发现提高一个架构拓扑的模型容量（比如拓宽宽度）可以带来一致的鲁棒性提高。这揭示了架构设计中鲁棒性与高效性间明显的权衡关系。考虑到实际应用场景中的

资源限制，本工作考虑使用单次评估 NAS 方法找到给定资源限制下的最优鲁棒架构，然而，不同架构拓扑在超网络中的容量不同，使得超网中容量较大的架构拓扑获得了不公平的评估优势。本工作提出训练多个容量不同的超网络，利用一个架构拓扑在不同超网络中的容量与性能插值得到其在目标容量下的性能作为评估结果，从而进行搜索。

项目经历

NICS-EFC 实验室

神经架构搜索开源框架 `aw_nas`¹

2019.09 – 至今

- 负责该框架的日常开发与维护
- 基于该框架为实验室的研究及工程项目提供支持

低比特神经网络在存算一体推理芯片中的验证研究项目（华为委托研究）

2022.09 – 至今

负责项目第二期的神经架构搜索研究，面向超分辨率、目标检测任务，联合量化设计高效容错的神经网络模型。目前，项目第二期已经验收通过，正在进行第三期“面向容错架构设计的架构性能预测器”的专利研究。

神经网络结构优化工具 `nx_nas`

2023.04 – 至今

针对 `aw_nas` 在设计上对用户不甚友好的缺点，该框架旨在提供用户友好的神经网络结构优化工具，便于快速搭建架构搜索、剪枝等结构优化系统。主要负责并推动该框架的开发工作。目前，该工具已经完成了最核心的搜索空间定义以及架构导出 (derive) 开发，并开发了基本的搜索策略（随机搜索、进化搜索）、评估策略（独立评估）等。

AI量化压缩与AI软硬件架构设计（OPPO 委托开发）

2022.09 – 2023.02

负责项目第二期面向超分辨率任务的神经架构搜索软件开发。具体开发：1) 基于UNET模型（包含卷积类型、卷积核尺寸、上采样方式等决策的）搜索空间；2) 可微分搜索、基于强化学习以及进化算法的搜索策略；3) 独立评估的评估策略。

北京超星未来科技有限公司（实习）

2022.08-2022.09

神经网络剪枝框架 NOVA-SLIM

以校企合作的形式参与超星未来剪枝框架（工具）的开发。负责该框架的剪枝控制器开发。

无问芯穹科技有限公司（实习）

2022.07-至今

大模型国产芯片评测

测试大模型（GLM-6B、OPT-6.7B 等）在国产芯片上的实际性能，为公司正在进行的针对大模型的轻量化设计提供数据基础。目前已完成天数、海飞科等的测试工作，并正在参与统一芯片测试方案的设计与开发。

社工志愿实践经历

- 清华大学材料学院科协网宣部部员 2018.03 – 2019.01
- “兰州重器”暑期实践支队队员 2018.06
- 清华大学“五一”校园讲解志愿者 2019.05
- 建国七十周年国庆游行清华方阵队员 2019.10

相关技能

- 编程语言：（相对熟练）Python、（一般掌握）C/C++
- 开发工具 / 平台：Git、Vim、Linux

¹ https://github.com/walkerning/aw_nas