

李炎静 (YANJING LI)

+86-15652145018 ◇ yanjingli98@gmail.com

性别: 女 籍贯: 山东烟台 出生年月: 1998.02.02 兴趣: 乐器 主页: [\[HomePage\]](#) [\[Scholar\]](#) [\[Github\]](#)
主研究方向: 神经网络模型量化、底层视觉、AIGC、3D 视觉、目标检测、剪枝、蒸馏

教育经历

-
- | | |
|----------------------------|-------------------|
| 北京航空航天大学 (985, 211) | 2020.09 - 2025.06 |
| 博士 (直博), 电子信息工程学院, 信号与信息处理 | |
| 导师: 曹先彬教授、张宝昌教授 | |
| 北京航空航天大学 (985, 211) | 2016.09 - 2020.06 |
| 学士, 高等理工学院, 电子信息 | |

实习经历

-
- | | |
|---|-------------------|
| 字节跳动-TikTok-AI 创新中心 | 2024.03 - 至今 |
| 大模型工程研发见习研究员
北京 | |
| · 开展大模型网络 GPU 推理加速相关工作的研究。 | |
| · 实习期间完成多模态大模型的 INT8 Weight-only 量化, 以及 FP8 和 INT4 量化, 实现性能损失在 1% 以下。 | |
| 上海人工智能实验室-通用视觉部 | 2022.10 - 2024.03 |
| 见习研究员
上海 | |
| · 开展神经网络压缩与加速相关工作的研究。 | |
| · 实习一年半时间里, 已以第一作者在 ECCV 2022 (Oral, 提出一种解耦二值神经网络优化算法)、ECCV 2022 (提出一种针对 1-bit 检测器信息差异感知的蒸馏学习方法)、NeurIPS 2022 (首次提出针对 ViT 网络的低比特量化方法)、AAAI 2023 (Oral, 提出一种自稳定的二值神经网络训练策略)、CVPR 2023 (Highlight, 首次提出针对 DETR 网络的低比特量化及训练策略)、ICCV 2023 (解决 3D 检测器在蒸馏学习过程中的表征差异问题)、NeurIPS 2023 (首次提出扩散模型的训练感知量化方法)、AAAI 2024 (首次提出二值 ViT 模型) 发表相关文章共计 8 篇。 | |
| 商汤科技有限公司-基础视觉组 | 2022.01 - 2022.10 |
| 自动驾驶感知见习研究员
北京 | |
| · 开展目标检测网络量化落地部署的工作。 | |
| · 实习半年时间里, 已以第一作者在 NeurIPS (首次提出一种信息矫正的 Low-bit ViT 量化训练方法) 发表论文 1 篇, 投稿论文 2 篇。 | |
| · 参与商汤基础视觉组的多项模型硬件部署任务。 | |

项目经历

-
- | | |
|--|------------------|
| (1) 华为技术有限公司-2012 实验室-算法应用部 神经网络量化研究与应用 | 2022.12 - 2024.8 |
| 项目学生领队、算法设计 | |

- 项目简介：如分类、检测和超分辨率重建，如何在实时性与高精度之间达到平衡是工业界的重要问题。提升量化网络的效率并且在小型设备上实现实时的 INT 2 ~ 4 推理，仍是技术的挑战。
- 主要挑战：传统的量化文章通常在 ResNet、VGG 等基础网络结构上验证性能，并且通过结构修改等方法提升性能。而对于更加前沿的网络结构 (MobileNet、GhostNet、ViT、DETR 等)，相关的量化方法设计仍是空白。设计针对新兴网络结构等量化方法以及结构，是一个新的挑战与方向。
- 完成情况：针对 ViT 网络的量化中存在的信息失真问题，我们设计了一种信息矫正模块与分布引导蒸馏方法，对于逐层的注意力特征进行矫正，使其接近于全精度教师网络的分布，从而在 2-bit ~ 4-bit 的 ViT 模型上进行性能提升。例如，我们在 4-bit ViT-S 模型上达到了 80.9%，超过了全精度 ViT-S。详细技术细节可查看我们最新 NeurIPS 2022 论文：Q-ViT: Accurate and Fully Quantized Low-Bit Vision Transformer (论文链接：<https://arxiv.org/abs/2210.06707>)。针对 DETR 网络存在的 query 特征信息失真问题，我们设计了一种分布矫正蒸馏方法，包含一种分布对齐模块与前景感知的 query 匹配方法，对于逐层的 query 特征进行分布矫正。我们在 4-bit DETR 上获得了接近全精度 DETR 的性能 (39.4% vs. 42.0%)。相关成果已形成论文 Q-DETR: An Efficient Low-Bit Quantized Detection Transformer (论文链接：<https://arxiv.org/abs/2304.00253>)，录用于 CVPR 2023 (highlight)。

代表性论文 (†: 共同一作)

二值神经网络

- (1) Bi-ViT: Pushing the Limit of Vision Transformer Quantization** 2024
Yanjing Li[†], Sheng Xu[†], Mingbao Lin, Xianbin Cao[✉], Chuanjian Liu, Xiao Sun, Baochang Zhang 二值 ViT
· AAAI Conference on Artificial Intelligence (AAAI)
- (2) DCP-NAS: Discrepant Child-Parent Neural Architecture Search for 1-Bit CNNs** 2023
Yanjing Li[†], Sheng Xu[†], Xianbin Cao, Li'an Zhuo, Baochang Zhang[✉], Tian Wang, Guodong Guo 二值神经网络、NAS
· International Journal of Computer Vision (IJCV)
- (3) Recurrent Bilinear Optimzation for Binary Neural Networks** 2022
Sheng Xu[†], Yanjing Li[†], Tiancheng Wang, Teli Ma, Baochang Zhang[✉], Peng Gao, Yu Qiao, Jinhu Lü, Guodong Guo 二值神经网络、循环双线性优化
· European Conference on Computer Vision (ECCV, Oral)
- (4) Learning 1-Bit Tiny Object Detector with Discriminative Feature Refinement** 2024
Sheng Xu[†], Mingze Wang[†], Yanjing Li[†], Mingbao Lin, Baochang Zhang[✉], David Doermann, Xiao Sun 1-bit 小目标检测、判别特征细化
· International Conference on Machine Learning (ICML)
- (5) IDa-Det: An Information Discrepancy-aware Distillation for 1-Bit Detectors** 2022
Sheng Xu[†], Yanjing Li[†], Bohan Zeng[†], Baochang Zhang[✉], Xianbin Cao, Peng Gao, Jinhu Lü 1-bit 检测器、信息差异感知蒸馏
· European Conference on Computer Vision (ECCV)

(6) Resilient Binary Neural Network

2023

Sheng Xu[†], Yanjing Li[†], Teli Ma[†], Mingbao Lin, Hao Dong, Baochang Zhang[✉],
Peng Gao, Jinhu Lü

二值神经网络、弹性 *scaling factor*

- AAAI Conference on Artificial Intelligence (AAAI, Oral)

低比特网络量化

(7) Q-ViT: Accurate and Fully Quantized Low-Bit Vision Transformer

2022

Yanjing Li[†], Sheng Xu[†], Baochang Zhang, Xianbin Cao[✉], Peng Gao,
Guodong Guo

Low-bit ViT 网络

- Conference on Neural Information Processing Systems (NeurIPS)

(8) Q-DM: An Efficient Low-Bit Quantized Diffusion Model

2023

Yanjing Li[†], Sheng Xu[†], Xianbin Cao[✉], Xiao Sun[✉], Baochang Zhang

Low-Bit Diffusion

- Conference on Neural Information Processing System (NeurIPS)

(9) Q-DETR: An Efficient Low-Bit Quantized Detection Transformer

2023

Sheng Xu[†], Yanjing Li[†], Mingbao Lin, Peng Gao, Guodong Guo, Jinhu Lü, Baochang Zhang[✉]

Low-bit

DETR

- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR, Highlight)

(10) Efficient Low-Bit Quantization with Adaptive Scales for Multi-Task Co-Training

2025

Boyu Liu, Haoyu Huang, Linlin Yang[✉], Yanjing Li[✉], Guodong Guo, Xianbin Cao, Baochang Zhang

Low-bit DETR

- International Conference on Learning Representations (ICLR)

3D 视觉

(11) Representation Disparity-aware Distillation for 3D Object Detection

2023

Yanjing Li[†], Sheng Xu[†], Mingbao Lin, Jihao Yin, Baochang Zhang,
Xianbin Cao[✉]

3D 目标检测、表征差异蒸馏

- IEEE International Conference on Computer Vision (ICCV)

(12) POEM: 1-Bit Point-wise Operations based on Expectation-Maximization for Efficient Point Cloud Processing

2021

Sheng Xu[†], Yanjing Li[†], Junhe Zhao, Baochang Zhang[✉],
Guodong Guo

1-bit 点云网络、EM 算法聚类

- British Machine Vision Conference (BMVC)

神经网络剪枝

(13) Filter pruning via expectation-maximization

2022

Sheng Xu[†], Yanjing Li[†], Linlin Yang, Baochang Zhang[✉], Dianmin Sun,
Kexin Liu

神经网络剪枝、EM 算法聚类

- Neural Computing and Applications (NCA)

扩散模型应用

(14) Implicit Diffusion Models for Continuous Super-Resolution

2023

Sicheng Gao, Xuhui Liu, Bohan Zeng, Sheng Xu, Yanjing Li, Xiaoyan Luo, Jianzhuang Liu, Xiantong Zhen, Baochang Zhang [✉]

扩散模型、任意倍率超分辨率重建

- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)

在审论文 (†: 共同一作)

二值神经网络

(1) Associative Recurrent Bilinear Optimization for Domain-Generalized Binary Neural Networks

2023

Sheng Xu [†], *Yanjing Li* [†], *Jinhu Lü*, *Chuanjian Liu*, *Baochang Zhang* [✉]

网络量化, 底层视觉

- International Journal of Computer Vision (IJCV, Major revise)

网络量化

(2) Learning Accurate Low-Bit Quantization towards Efficient Computational Imaging

2023

Sheng Xu [†], *Yanjing Li* [†], *Jinhu Lü*, *Chuanjian Liu*, *Baochang Zhang* [✉]

网络量化, 底层视觉

- International Journal of Computer Vision (IJCV, Major Revision)

公开专利

(1) 基于去相关二值网络的协同目标识别方法、系统及装置

吕金虎, 王滨, 徐昇, 张宝昌, 李炎静, 张峰, 王星

- CN Patent CN202210023260.5

(2) 图像处理方法、装置和存储介质

刘传建, 韩凯, 王云鹤, 李炎静, 张宝昌

- CN Patent CN117649586A

(3) 微小目标检测模型训练方法、装置及电子设备

张宝昌, 吕金虎, 王铭泽, 徐昇, 李炎静, 王田

- CN Patent CN118506154A

(4) 一种数据处理方法及其装置

刘传建, 韩凯, 张宝昌, 徐昇, 李炎静, 王云鹤

- CN Patent CN118506154A

(5) 一种基于特征信息差异的模型蒸馏方法及装置

张宝昌, 曾博涵, 徐昇, 李炎静

- CN Patent CN118506154A

学术活动

会议审稿人

- IEEE/CVF Computer Vision and Pattern Recognition (CVPR), 2022–2024
- IEEE/CVF International Conference on Computer Vision (ICCV), 2023
- European Conference on Computer Vision (ECCV), 2022/2024
- International Conference on Machine Learning (ICML), 2022–2024
- Conference on Neural Information Processing Systems (NeurIPS), 2022–2023
- International Conference on Learning Representations (ICLR), 2024
- AAAI Conference on Artificial Intelligence, 2024-2025

期刊审稿人

- IEEE Transactions on Neural Networks and Learning System (TNNLS)
- Neural Computing and Applications (NCA)

荣誉获奖

首批国家自然科学基金青年项目（博士研究生）（2023）

国家奖学金（2024）

北京航空航天大学一等学业奖学金（2022，2023）

北京航空航天大学优秀团员（2023）

北京航空航天大学优秀毕业生（2020）