Guobin Shen

https://floyedshen.github.io/ +86 13931425808
floyed_shen@outlook.com shenguobin2021@ia.ac.cn
Guobin Shen No.95 Zhongguancun East Road, Beijing, China

Research Interests

My research interests focus on biologically-inspired neural networks and their applications in cognitive science and artificial intelligence. I am particularly interested in integrating brain-inspired models with advanced AI systems, as well as exploring the safety and interpretability of large-scale models to address complex real-world challenges.

Education

2021 – 2026 **Ph.D., Institute of Automation, Chinese Academy of Sciences**
Bio-Inspired AI / LLM Alignment, Safety & Interpretability
National Scholarship (~1%) / CAS President's Scholarship (~1%)

2017 – 2021 **B.Eng., Communication Engineering Sun Yat-sen University**
GPA: 4.05 / 5.0, **Rank**: 1 / 85 (Overall)
National Scholarship (~2%) × **2**

Publications

During my doctoral studies, I have published over ten papers as first author in prestigious journals such as *PNAS* and *Cell Patterns*, as well as top-tier conferences including *NeurIPS*, *ICLR*, *CVPR*, *ICCV*, etc. Additionally, I have contributed as co-author to works published in *TPAMI*, *TEVC*, *Pattern Recognition*, *IJCAI*, *AAAI*, and other leading journals and conferences. To date, my work has received over 400 citations with an H-Index of 11.

LLM & AI Safety

1. Shen, Guobin, Zhao, Dongcheng, Feng, Linghao, He, Xiang, Wang, Jihang, Shen, Sicheng, Tong, Haibo, Dong, Yiting, Li, Jindong, Zheng, Xiang, and others. "PandaGuard: Systematic Evaluation of LLM Safety in the Era of Jailbreaking Attacks." 2025. `AI Safety` `Framework`

2. Shen, Guobin, Zhao, Dongcheng, Dong, Yiting, He, Xiang, and Zeng, Yi. "Jailbreak Antidote: Runtime Safety-Utility Balance via Sparse Representation Adjustment in Large Language Models." *Proceedings of the 13th International Conference on Learning Representations*, 2025. `AI Safety` `Interpretability` `ICLR`

3. Shen, Guobin, Zhao, Dongcheng, He, Xiang, Feng, Linghao, Dong, Yiting, Wang, Jihang, Zhang, Qian, and Zeng, Yi. "Neuro-Vision to Language: Image Reconstruction and Interaction via Non-invasive Brain Recordings." *Proceedings of the 38th Conference on Neural Information Processing Systems (NeurIPS 2024)*, 2024. `Multimodal LLM` `fMRI` `NeurIPS`

Publications (continued)

4. Shen, Guobin, Zhao, Dongcheng, Bao, Aorigele, He, Xiang, Dong, Yiting, and Zeng, Yi. "StressPrompt: Does Stress Impact Large Language Models and Human Performance Similarly?" *Proceedings of the 39th AAAI Conference on Artificial Intelligence (AAAI 2025)*, 2025. `LLM Analysis` `Cognitive Science` `AAAI`

5. Shen, Guobin, Zhao, Dongcheng, Dong, Yiting, Li, Yang, Li, Jindong, Sun, Kang, and Zeng, Yi. "Astrocyte-Enabled Advancements in Spiking Neural Networks for Large Language Modeling." *arXiv preprint arXiv:2312.07625*, 2023. `Astrocyte` `LLM Pre-training`

6. Wu, Ping, Shen, Guobin, Zhao, Dongcheng, Wang, Yuwei, Dong, Yiting, Shi, Yu, Lu, Enmeng, Zhao, Feifei, and Zeng, Yi. "CVC: A Large-Scale Chinese Value Rule Corpus for Value Alignment of Large Language Models." *arXiv preprint arXiv:2506.01495*, 2025. `Value Alignment` `Dataset`

7. Dong, Yiting, Shen, Guobin, Zhao, Dongcheng, He, Xiang, and Zeng, Yi. "Harnessing Task Overload for Scalable Jailbreak Attacks on Large Language Models." *arXiv preprint arXiv:2410.04190*, 2024. `Jailbreak` `LLM`

Spiking Neural Networks & Brain-Inspired AI

8. Shen, Guobin, Zhao, Dongcheng, Dong, Yiting, and Zeng, Yi. "Brain-Inspired Neural Circuit Evolution for Spiking Neural Networks." *Proceedings of the National Academy of Sciences*, vol. 120, no. 39, 2023, p. e2218173120. `Neuro-Evolution` `SNN` `PNAS`

9. Shen, Guobin, Zhao, Dongcheng, and Zeng, Yi. "Backpropagation with Biologically Plausible Spatiotemporal Adjustment for Training Deep Spiking Neural Networks." *Cell Patterns*, vol. 3, no. 6, 2022. `SNN` `Backpropagation` `Cell Patterns`

10. Shen, Guobin, Zhao, Dongcheng, Li, Tenglong, Li, Jindong, and Zeng, Yi. "Are Conventional SNNs Really Efficient? A Perspective from Network Quantization." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 27538-27547. `SNN` `Efficiency` `CVPR Hightlight`

11. Shen, Guobin, Zhao, Dongcheng, and Zeng, Yi. "Exploiting High-Performance Spiking Neural Networks with Efficient Spiking Patterns." *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2025. `SNN` `Efficiency` `TETCI`

12. Shen, Guobin, Zhao, Dongcheng, and Zeng, Yi. "Exploiting Nonlinear Dendritic Adaptive Computation in Training Deep Spiking Neural Networks." *Neural Networks*, vol. 170, 2024, pp. 190-201. `SNN` `Dendritic Dynamic` `Neural Networks`

13. Shen, Guobin, Zhao, Dongcheng, Shen, Sicheng, and Zeng, Yi. "Enhancing Spiking Transformers with Binary Attention Mechanisms." *The Second Tiny Papers Track at ICLR 2024*. `Transformer` `Binary Attention` `ICLR Tiny Paper`

14. Shen, Guobin, Zhao, Dongcheng, Dong, Yiting, Li, Yang, and Zeng, Yi. "Dive into the Power of Neuronal Heterogeneity." *arXiv preprint arXiv:2305.11484*, 2023. `Neuronal Heterogeneity` `SNN`

15. Shen, Guobin, Zhao, Dongcheng, Dong, Yiting, Li, Yang, Zhao, Feifei, and Zeng, Yi. "Learning the Plasticity: Plasticity-Driven Learning Framework in Spiking Neural Networks." *arXiv preprint arXiv:2308.12063*, 2023. `Plasticity` `Learning Framework`

Publications (continued)

16 Han, Bing, Zhao, Feifei, Zeng, Yi, and Shen, Guobin. "Developmental Plasticity-Inspired Adaptive Pruning for Deep Spiking and Artificial Neural Networks." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024. `Pruning` `Plasticity` `TPAMI`

17 Pan, Wenxuan, Zhao, Feifei, Shen, Guobin, Han, Bing, and Zeng, Yi. "Brain-Inspired Multi-Scale Evolutionary Neural Architecture Search for Deep Spiking Neural Networks." *IEEE Transactions on Evolutionary Computation*, 2024. `NAS` `Evolution` `TEVC`

18 Zhao, Dongcheng, Shen, Guobin, Dong, Yiting, Li, Yang, and Zeng, Yi. "Improving Stability and Performance of Spiking Neural Networks through Enhancing Temporal Consistency." *Pattern Recognition*, vol. 159, 2025, p. 111094. `SNN` `Stability` `Pattern Recognition`

19 Zeng, Yi, Zhao, Dongcheng, Zhao, Feifei, Shen, Guobin, Dong, Yiting, Lu, Enmeng, Zhang, Qian, Sun, Yinqian, Liang, Qian, Zhao, Yuxuan, and others. "BrainCog: A Spiking Neural Network Based, Brain-Inspired Cognitive Intelligence Engine for Brain-Inspired AI and Brain Simulation." *Patterns*, 2023, p. 100789. `Framework` `Brain-inspired` `Patterns`

20 Han, Bing, Zhao, Feifei, Zeng, Yi, Pan, Wenxuan, and Shen, Guobin. "Enhancing Efficient Continual Learning with Dynamic Structure Development of Spiking Neural Networks." *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI 2023)*, 2023. `Continual Learning` `SNN` `IJCAI`

21 Yu, Yonghao, Zhao, Dongcheng, Shen, Guobin, Dong, Yiting, and Zeng, Yi. "Brain-Inspired Stepwise Patch Merging for Vision Transformers." *IJCAI*, 2025. `Vision Transformer` `Brain-inspired` `IJCAI`

22 Shen, Sicheng, Zhao, Dongcheng, Shen, Guobin, and Zeng, Yi. "TIM: An Efficient Temporal Interaction Module for Spiking Transformer." *Proceedings of the 33rd International Joint Conference on Artificial Intelligence (IJCAI 2024)*, 2024. `Transformer` `Temporal` `IJCAI`

23 He, Xiang, Liu, Xiangxi, Li, Yang, Zhao, Dongcheng, Shen, Guobin, Kong, Qingqun, Yang, Xin, and Zeng, Yi. "CACE-Net: Co-guidance Attention and Contrastive Enhancement for Effective Audio-Visual Event Localization." *Proceedings of the 32nd ACM International Conference on Multimedia*, 2024, pp. 985-993. `Multimodal` `Event Localization` `MM`

24 He, Xiang, Zhao, Dongcheng, Li, Yang, Shen, Guobin, Kong, Qingqun, and Zeng, Yi. "An Efficient Knowledge Transfer Strategy for Spiking Neural Networks from Static to Event Domain." *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 1, 2024, pp. 512-520. `Knowledge Transfer` `Event Domain` `AAAI`

25 Feng, Linghao, Zhao, Dongcheng, Shen, Sicheng, Dong, Yiting, Shen, Guobin, and Zeng, Yi. "Time Cell Inspired Temporal Codebook in Spiking Neural Networks for Enhanced Image Generation." *arXiv preprint arXiv:2405.14474*, 2024. `Image Generation` `Time Cell`

Hardware Acceleration & System Optimization

26 Shen, Guobin, Li, Jindong, Li, Tenglong, Zhao, Dongcheng, and Zeng, Yi. "SpikePack: Enhanced Information Flow in Spiking Neural Networks with High Hardware Compatibility." *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2025. `SNN` `Hardware` `ICCV`

Publications (continued)

27 Li, Jindong, Shen, Guobin, Zhao, Dongcheng, Zhang, Qian, and Zeng, Yi. "Firefly v2: Advancing Hardware Support for High-Performance Spiking Neural Network with a Spatiotemporal FPGA Accelerator." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2024. Hardware FPGA TCAD

28 Li, Jindong, Shen, Guobin, Zhao, Dongcheng, Zhang, Qian, and Zeng, Yi. "Firefly: A High-Throughput Hardware Accelerator for Spiking Neural Networks with Efficient DSP and Memory Optimization." *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 31, no. 8, 2023, pp. 1178-1191. Hardware Accelerator TVLSI

29 Li, Jindong, Li, Tenglong, Shen, Guobin, Zhao, Dongcheng, Zhang, Qian, and Zeng, Yi. "Revealing Untapped DSP Optimization Potentials for FPGA-Based Systolic Matrix Engines." *2024 34th International Conference on Field-Programmable Logic and Applications (FPL)*, IEEE, 2024, pp. 197-203. FPGA Optimization FPL

30 Li, Tenglong, Li, Jindong, Shen, Guobin, Zhao, Dongcheng, Zhang, Qian, and Zeng, Yi. "FireFly-S: Exploiting Dual-Side Sparsity for Spiking Neural Networks Acceleration with Reconfigurable Spatial Architecture." *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2024. Acceleration Sparsity TCAS-I

Datasets & Data Augmentation

31 Shen, Guobin, Zhao, Dongcheng, and Zeng, Yi. "EventMix: An Efficient Data Augmentation Strategy for Event-Based Learning." *Information Sciences*, vol. 644, 2023, p. 119170. Event-based Augmentation Information Sciences

32 Dong, Yiting, He, Xiang, Shen, Guobin, Zhao, Dongcheng, Li, Yang, and Zeng, Yi. "EventZoom: A Progressive Approach to Event-Based Data Augmentation for Enhanced Neuromorphic Vision." *Proceedings of the 39th AAAI Conference on Artificial Intelligence (AAAI 2025)*, 2025. Event-based Neuromorphic AAAI

33 Dong, Yiting, Li, Yang, Zhao, Dongcheng, Shen, Guobin, and Zeng, Yi. "Bullying10K: A Large-Scale Neuromorphic Dataset Towards Privacy-Preserving Bullying Recognition." *Advances in Neural Information Processing Systems*, vol. 36, 2024. Dataset Neuromorphic NeurIPS

Project

PandaGuard: LLM Security Assessment Framework Stars 31 **Project Lead**

Designed and implemented a systematic LLM jailbreak attack security assessment framework, integrating multiple attack and defense algorithms

Built large-scale benchmark dataset PandaBench, providing multi-dimensional security evaluation metrics

BrainCog: Brain-Inspired Cognitive Intelligence Engine Stars 537 **Project Lead**

Led development of comprehensive spiking neural network framework supporting brain-inspired AI and brain simulation research

Implemented 50+ functional SNN algorithms covering cognitive functions including perceptual learning, decision-making, and knowledge representation

Project (continued)

- **aw_nas: Neural Architecture Search Framework** Stars 250 **Contributor**
  Contributed algorithm implementations including Once-for-all to the open-source NAS framework, along with related dataset construction

Academic Services

Serve as a reviewer for conferences including *NeurIPS*, *ICML*, *ICLR*, *CVPR*, *ICCV*, *ECCV*, *AAAI*, *MM*, *AISTATS*, and journals including IEEE Computational Intelligence Magazine, Neural Networks, and Neurocomputing.

Awards and Honors

Scholarships

2025.06 **CAS President's Scholarship** (~1%)
Academic honor from Chinese Academy of Sciences, recognizing doctoral students with outstanding academic achievements

2024.11 **National Scholarship for Doctoral Students** (~1%)
Awarded by Ministry of Education of China, recognizing doctoral students with excellence in academic research and comprehensive qualities

2020.11 **National Scholarship for Undergraduates** (~2%)
Awarded by Ministry of Education of China, recognizing undergraduates with outstanding academic performance and comprehensive excellence

2019.11 **National Scholarship for Undergraduates** (~2%)
Received the nation's highest academic honor for two consecutive years

2018.11 **Lin Bin-Liu Xiangdong Endowed Scholarship** (~1%)

Academic Honors

2023.11 **Cell Press Paper of the Year**
Recognizing outstanding academic achievements published in top-tier journals (first student author)

2022.11 **Cell Press Best Paper Award for Chinese Scientists**
Recognizing outstanding academic achievements published in top-tier journals (first author)

2021.06 **Outstanding Graduate & Outstanding Thesis, Sun Yat-sen University**

Outstanding Graduation Thesis, Sun Yat-sen University

Awards and Honors (continued)

Competition Awards

2019.09 **Runner-up, International Aerial Robotics Competition (Asia-Pacific Region)**
Premier international robotics competition, demonstrating engineering practice and teamwork capabilities

2019.08 **National Second Prize, National Undergraduate Electronics Design Contest**
Authoritative national competition in electronic design

2018.09 **Second Prize, National Undergraduate Biomedical Electronic Innovation Design Competition**

Other Honors

2025.03 **Model Student of Excellence**
Highest comprehensive honor at university level, recognizing exemplary students with outstanding character, academics, and overall development