Shanghua Gao

Postdoctoral Researcher at Harvard University

(+1) 3397772888 | shanghuagao@gmail.com | http://shgao.site | gasvn | Google scholar | 15 first author papers, 7000+ citations.

Summary

He is currently a Postdoctoral Researcher at Harvard University, where his work centers on developing AI agentic models, multi-modal generative AI models, and universal representation learning methods, with the goal of building AI scientists for scientific discovery. His research emphasizes general techniques, including agentic AI, generative modeling, self-supervised learning, and the design of foundation models. His research has broad applications across fields such as AI for science (health and biomedical area), computer vision, multi-modal models, and time series analysis. He has published 14 first-author papers in top-tier journals and conferences, including **Cell (1), TPAMI (4)**, and leading conferences such as **CVPR and NeurIPS (6)**, obtaining over **7,000 citations**. His projects have gained over 8,000 GitHub stars, successfully transitioning into real-world products.

Education & Research Experience

Harvard University, Zitnik Lab — *Boston, USA*

Postdoctoral Researcher (AI for Science and Agentic AI, Collaborator: Marinka Zitnik) — *Sept. 2023 - Now*

- Stanford's List of **World's Top 2% Scientists**, 2023-2025.
- 1 first-author paper in Cell, 1 first-author paper in NeurIPS.
- Propose the **1st** idea of biomedical discovery with AI agents (**first author on Cell**).
- ToolUniverse: an open-source ecosystem under one AI-tool interaction protocol that unifies tools, data, and workflows to build AI scientists capable of collaborative scientific discovery across diverse models and domains (**Under review at Nature Biotechnology**).
- TxAgent: First general-purpose therapeutic AI agent with automatic multi-agent multi-step reasoning and connected to massive number of tools (ToolUniverse) for biomedical applications (**preparing a Nature submission**).
- CureBench (`https://curebench.ai/`): organize a **NeurIPS competition** benchmarking AI reasoning for therapeutic decision-making at scale, with more than 1,400 entrants.
- Time-series foundation model for healthcare analysis: the **1st work** for unifying predictive and generative tasks across various data sources.
- Conduct protein structure editing with conditional generative models.

Sea AI Lab, Shuicheng Yan's Team — *Singapore*

Research Intern (Foundation model representation learning, Collaborator: Shuicheng Yan) — *Mar. 2022 - Aug. 2023*

- Generative models: 1) Propose a Mask Diffusion Transformer for image generation, boosting DiT training speed by **10x** and achieving new SOTA. 2) highly flexible image editing.
- Efficient self-supervised learning for vision models, accelerated MAE training **16x** (1600→100 epochs) with improved performance.
- **Most starred GitHub repository in SEA AI Lab** (3.4K Github stars).

Nankai University, Media Computing Lab — *Tianjin, China*

Ph.D (Computer vision, Advisor: Ming-Ming Cheng) — *Aug. 2018 - Jun. 2023*

- **7,000+ Google scholar citations.**
- **10** first-author papers (4 in top machine learning journal TPAMI; 5 in top conferences like CVPR, 1 oral presentation).
- Research shipped to **2** products (National Disaster Reduction Center of China, Infervision) and led to **4** patents.
- **Early Ph.D. completion** (graduated one year ahead of schedule), 2023.
- **China's 100 most influential international academic papers award**, 2022.
- **Most influential paper** award, Jittor Developer Conference, 2021.
- CVPR outstanding reviewer, 2021.
- Recipient of scholarships from SK Artificial Intelligence, Tianjin Student Innovation, and Nankai University, among others, 2018-2023.
- Develop the **1st work** of large-scale, fully unsupervised semantic segmentation method across 1,000+ classes.
- Build the **1st multi-scale backbone** model Res2Net with wide applications (**3700+ citations**), shipped to multiple products and win **Grand Prize in the China International Industry Fair**.

Xidian University — *Xi'an, China*

Undergraduate for communication engineering — *Aug. 2014 - Jun. 2018*

- **Top 10 Distinguished Graduates**, Xidian University, 2018.
- National College Students' Innovation and Entrepreneurship Training Program Outstanding Project (top 0.5%), 2017.
- 2 first-author papers on OAM.

Selected Publications

Highlight

- 28 co-author papers (15 first author, 1 corresponding author), 1 Cell, 5 TPAMI, 1 oral, 1 most influential paper. **7,000+ google scholar citations**.
- 3,700+ google scholar citations for first-author paper Res2Net (TPAMI).
- **All first-author papers are open-sourced**, receiving **8,000 Github stars**: 1 repo > 3k, 2 repo > 1k, 6 repo > 500 stars.
- 3.4k Github stars for the first-author paper EditAnything (ACM MM).
- Biomedical AI agent (Cell): **1st work** that proposes a AI agent for biomedical discovery.
- MDT (ICCV): **1st mask diffusion transformer** for image generation, new SoTA, **10x faster** than DiT in training.
- UniTS (Neurips): **1st** unified multi-task multi-domain **time series foundation model**.
- LUSS (TPAMI): **1st** large-scale, fully **unsupervised semantic segmentation** method across 1,000+ classes.
- TEC (Science China): accelerates masked autoencoder (MAE) training by **16x** (1600→100 epochs).
- RF-Next (TPAMI): achieved the **1st place** in MegCup RAW image denoising.
- Deployed SOD100K to the National Disaster Reduction Center of China.
- Integrated Res2Net into Infervision's CT system.

AI for science: Agentic AI

Advancing agentic AI to build autonomous scientists capable of reasoning across tools and domains, with foundational contributions including TxAgent, ToolUniverse, and CURE-Bench.

Advancing agentic AI to build autonomous scientists capable of reasoning across tools and domains, with foundational contributions including TxAgent, ToolUniverse, and CURE-Bench.

Gao, S., Fang, A., Huang, Y., Giunchiglia, V., Noori, A., Schwarz, J. R., Ektefaie, Y., Kondic, J., and Zitnik, M. (2024). "Empowering biomedical discovery with ai agents". In: Cell.

Gao, S., Zhu, R., Kong, Z., Noori, A., Su, X., Ginder, C., Tsiligkaridis, T. and Zitnik, M., "TxAgent: An AI agent for therapeutic reasoning across a universe of tools". Preparing submission to Nature.

Gao, S., Zhu, R., Kong, Z., Aldogom, S., Huang, Y., Sui, P., Tsiligkaridis, T. and Zitnik, M., "Democratizing AI scientists using ToolUniverse." Under review at Nature Biotechnology.

Gao, S., Zhu, R., Kong, Z., Su, X., Ginder, C., Aldogom, S., Das, I., Evans, T., Tsiligkaridis, T. and Zitnik, M., "CURE-Bench: Competition on Reasoning Models for Drug Decision-Making in Precision Therapeutics". In NeurIPS 2025, competition.

Su, X., Messica, S., Huang, Y., Johnson, R., Fesser, L., **Gao, S.**, Sahneh, F. and Zitnik, M., (2025). Multimodal Medical Code Tokenizer. International Conference on Machine Learning (ICML).

Su, X., Wang, Y., **Gao, S.**, Liu, X., Giunchiglia, V., Clevert, D.-A., and Zitnik, M. (2024). "KGARevion: an AI agent for knowledge-intensive biomedical QA". In: International Conference on Learning Representations (ICLR).

AI for science: Foundation models

Advancing foundation models that learn unified representations across time-series, vision, and biomedical data to enable generalizable scientific representation learning.

Gao, S., Koker, T., Queen, O., Hartvigsen, T., Tsiligkaridis, T., and Zitnik, M. (2024). "UniTS: a Unified Multi-task Time Series Model". In: Conference on Neural Information Processing Systems (NeuriPS).

Wu, Y.-H., **Gao, S.**, Mei, J., Xu, J., Fan, D.-P., Zhang, R.-G., and Cheng, M.-M. (2021). "JCS: An Explainable COVID-19 Diagnosis System by Joint Classification and Segmentation". In: IEEE Transactions on Image Processing (TIP).

Su, H., Wang, W., Du, Z., Peng, Z., **Gao, S.**, Cheng, M.M. and Yang, J., 2021. "Improved protein structure prediction using a new multi-scale network and homologous templates". In Advanced Science.

Generative models & Multimodal models

Developming generative and multimodal models that unify predictive and creative intelligence, enabling cross-modal understanding and controllable generation across vision and language.

Gao, S., Zhou, P., Cheng, M.-M., and Yan, S. (2023). "Masked Diffusion Transformer is a Strong Image Synthesizer". In: IEEE/CVF International Conference on Computer Vision (ICCV).

Gao, S., Lin, Z., Xie, X., Zhou, P., Cheng, M.-M., and Yan, S. (2023). "Edit Anything by Segmenting Anything". In: ACM Multimedia.

Wu, G., Zhang, S., Shi, R., **Gao, S.**, Chen, Z., Wang, L., ... & Li, X. (2025). "Representation Entanglement for Generation: Training Diffusion Transformers Is Much Easier Than You Think." In: Conference on Neural Information Processing Systems (NeurIPS). **Oral**

Zhong, S., **Gao, S.**, Huang, Z., Wen, W., Zitnik, M., and Zhou, P. (2024). "MoExtend: Tuning new experts for modality and task extension". In: Association for Computational Linguistics (ACL) SRW

Huang, Z., Zhong, S., Zhou, P., **Gao, S.**, Zitnik, M., Lin, L. (2025). A Causality-aware Paradigm for Evaluating Creativity of Multimodal Large Language Models. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI).

Zhong, S., Huang, Z., **Gao, S.**, Wen, W., Lin, L., Zitnik, M., and Zhou, P. (2024). "Let's Think Outside the Box: Exploring Leap-of-Thought in Large Language Models with Creative Humor Generation". In: IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR).

Kolbeinsson, A., O'Brien, K., Huang, T., **Gao, S.**, Liu, S., Schwarz, J.R., Vaidya, A., Mahmood, F., Zitnik, M., Chen, T. and Hartvigsen, T., 2024. "Composable interventions for language models". In: International Conference on Learning Representations (ICLR).

Foundation model representation learning

Developing universal self-supervised learning frameworks that discover semantics and structure from large-scale unlabeled data, laying the foundation for scalable and generalizable scientific models.

Gao, S., Li, Z.-Y., Yang, M.-H., Cheng, M.-M., Han, J., and Torr, P. (2022). "Large-scale Unsupervised Semantic Segmentation". In: IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE TPAMI).

Gao, S., Zhou, P., Cheng, M.-M., and Yan, S. (2022). "Towards Sustainable Self-supervised Learning". In: SCIENCE CHINA Information Sciences.

Li, Z.-Y., **Gao, S.**, and Cheng, M.-M. (2023). "SERE: Exploring Feature Self-relation for Self-supervised Transformer". (**Corresponding author**) In: IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE TPAMI).

Zhao, K., **Gao, S.**, Wang, W., and Cheng, M.-M. (2019). "Optimizing the F-measure for Threshold-free Salient Object Detection". In: IEEE/CVF International Conference on Computer Vision (ICCV).

Efficient network architecture

Designing efficient and scalable neural architectures, such as Res2Net and RF-Next, that redefine multi-scale representation learning and enable high-performance, resource-efficient inference.

Gao, S., Cheng, M.-M., Zhao, K., Zhang, X.-Y., Yang, M.-H., and Torr, P. (2021). "Res2Net: A New Multi-scale Back- bone Architecture". In: IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE TPAMI).

Gao, S., Li, Z.-Y., Han, Q., Cheng, M.-M., and Wang, L. (2022). "RF-Next: Efficient Receptive Field Search for Convo- lutional Neural Networks". In: IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE TPAMI).

Gao, S., Han, Q., Li, D., Cheng, M.-M., and Peng, P. (2021). "Representative Batch Normalization with Feature Calibration". In: IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR). **Oral**.

Cheng*, M.-M., **Gao*, S.**, Borji, A., Tan, Y.-Q., Lin, Z., and Wang, M. **Joint first author**. (2021). "A Highly Efficient Model to Study the Semantics of Salient Object Detection". In: IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE TPAMI).

Gao, S., Han, Q., Li, Z.-Y., Peng, P., Wang, L., and Cheng, M.-M. (2021). "Global2Local: Efficient Structure Search for Video Action Segmentation". In: IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR).

Gao, S., Tan, Y.-Q., Cheng, M.-M., Lu, C., Chen, Y., and Yan, S. (2020). "Highly Efficient Salient Object Detection with 100K Parameters". In: European Conference on Computer Vision (ECCV).

Gu, Y.-C., **Gao, S.**, Cao, X.-S., Du, P., Lu, S.-P., and Cheng, M.-M. (2021). "iNAS: Integral NAS for Device-Aware Salient Object Detection". In: IEEE/CVF International Conference on Computer Vision (ICCV).

Tan, Y.-Q., **Gao, S.**, Li, X.-Y., Cheng, M.-M., and Ren, B. (2020). "VecRoad: Point-based Iterative Graph Exploration for Road Graphs Extraction". In: IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR).

Fan, D.-P., Cheng, M.-M., Liu, J.-J., **Gao, S.**, Hou, Q., and Borji, A. (2018). "Salient objects in clutter: Bringing salient object detection to the foreground". In: European Conference on Computer Vision (ECCV).

Interdisciplinary collaborations

Collaborating with leading researchers in visualization and biomedicine to extend general-purpose agentic AI and foundation models into new domains, including biomedical and genomics data visualization.

Lange, D., Sui, P., **Gao, S.**, Zitnik, M., & Gehlenborg, N. (2025). "DQVis Dataset: Natural Language to Biomedical Visualization". In NeurIPS.

Nguyen, H. N., L'Yi, S., Smits, T. C., **Gao, S.**, Zitnik, M., & Gehlenborg, N. (2025). "Multimodal retrieval of genomics data visualizations".

Lange, D., **Gao, S.**, Sui, P., Money, A., Misner, P., Zitnik, M., & Gehlenborg, N. (2025). "YAC: Bridging Natural Language and Interactive Visual Exploration with Generative AI for Biomedical Data Discovery". arXiv preprint arXiv:2509.19182.

Lange, D., **Gao, S.**, Sui, P., Money, A., Misner, P., Zitnik, M., & Gehlenborg, N. (2025). "A Generative AI System for Biomedical Data Discovery with Grammar-Based Visualizations". arXiv preprint arXiv:2509.16454.

Professional Activities

Organize the NeurIPS 2025 competition
CURE-Bench: Competition on Reasoning Models for Drug Decision-Making in Precision Therapeutics — *May.-Dec. 2025*

Invited Talk at AstraZeneca
Invited talk with title "Empowering Scientific Discovery with AI Scientists". — *Nov. 2025*

Invited Talk at Roche
Invited talk with title "TxAgent: An AI Agent for Therapeutic Reasoning Across a Universe of Tools". — *May. 2025*

Invited Talk at Chan Zuckerberg Initiative Rare As One Network
Invited talk with title "Therapeutic reasoning across a universe of tools". — *Apr. 2025*

Invited Talk at Transformative Computational Biology Investigators Meeting
Invited talk with title "The future of drug repurposing with AI, today: Foundation and agentic AI". — *Dec. 2024*

Invited Talk at Nankai University
Invited talk with title "Generative AI for Science". — *Dec. 2024*

Invited Talk at Huawei
Invited talk with title "Unified multi-task time series foundation model". — *Jun. 2024*

Invited Talk at Princeton University
Invited talk with title "Self-adaptive representation learning". — *Mar. 2023*

Invited Talk at IEE Signal Processing Society
Invited talk with title "Towards Designing an Explainable COVID-19 Diagnosis System Registration". — *Aug. 2022*

Peer Review

Program committee member & Reviewer

- Program committee member: AAAI, IJCAI
- Journal reviewer: Cell Reports Medicine, TPAMI, TIP, TNNLS.
- Conference reviewer: CVPR, ICCV, ECCV, NeurIPS, ICLR, ICML.