

Mr. Honghao Fu

Email: honghao.fu@uq.edu.au | Tel: +86-18428043844

[\[Homepage\]](#) [\[Github\]](#) [\[Google Scholar\]](#)

Education background

The University of Queensland (UQ)

Brisbane, Australia

◆ *Doctor of Philosophy in Computer Science*

07/2025-Now

◆ Research Interests: Large Language Models, Agentic Systems, Multimodal Understanding

◆ Supervisor: Prof. Yujun Cai, Prof. Miao Xu and Prof. Yiwei Wang

Nanyang Technological University (NTU)

Singapore

◆ *Master of Science in Electronics*

08/2023-01/2025

◆ GPA: 4.13/5.00

◆ Research Interests: Generative AI, Computer Vision

◆ Supervisor: Prof. Bihan Wen and Dr. Yufei Wang

Southeast University (SEU)

Nanjing, China

◆ *Bachelor of Engineering in Measuring Control Technology & Instruments*

08/2019-06/2023

◆ GPA: 85.85/100

◆ Outstanding Dissertation Award (Top 1%)

Publications

1. **Fu H.**, Xu M., Wang Y., Zhang D., Liu J. & Cai Y. (2026). VideoStir: Understanding Long Videos via Spatio-Temporally Structured and Intent-Aware RAG. *ACL 2026 Main*. (Oral, Top 0.8%) [\[Paper\]](#)
2. **Fu H.**, Ouyang Y., Chang K., Wang Y., Huang Z. & Cai Y. (2026). ContextNav: Towards Agentic Multimodal In-Context Learning. *ICLR 2026*. [\[Paper\]](#)
3. Yang Y., Cheng W., Chen S., **Fu H.**, Zeng X., Cai Y., Yu G. & Ma X. (2026). OmniLottie: Generating Vector Animations via Parameterized Lottie Tokens. *CVPR 2026*. [\[Paper\]](#)
4. **Fu H.***, Ren J.*, Chai Q., Ye D., Cai Y. & Wang H. (2025). VistaWise: Building Cost-Effective Agent with Cross-Modal Knowledge Graph for Minecraft. *EMNLP 2025 Main*. [\[Paper\]](#)
5. **Fu H.**, Shen Z., Chin J. J., & Wang H. (2025). BrainVis: Exploring the Bridge between Brain and Visual Signals via Image Reconstruction. *ICASSP 2025*. [\[Paper\]](#)
6. **Fu H.**, Gu Y., Yan Y., Shen Y., Wu Y. & Sun L. (2025). SDR-GAIN: A High Real-Time Occluded Pedestrian Pose Completion Method for Autonomous Driving. *IEEE Transactions on Intelligent Transportation Systems*. (Q1, IF=8.4) [\[Paper\]](#)
7. Gu Y., Yan X., Qin H., Akhtar N., Yuan S., **Fu H.**, Yang S. & Mian A. (2025). HDTCNet: A hybrid-dimensional convolutional network for multivariate time series classification. *Pattern Recognition*. (Q1, IF=7.5) [\[Paper\]](#)
8. Liu S., Han Y., Xing P., Yin F., Wang R., Cheng W., Liao J., Wang Y., **Fu H.**, ... , Yu G. & Jiang D. (2025). Step1X-Edit: A Practical Framework for General Image Editing. *arXiv:2504.17761*. (Technical Report) [\[Paper\]](#)
9. Ren J., Zhang G., **Fu H.**, Wu P. & Wang H. (2025). WaMo: Wavelet-Enhanced Multi-Frequency Trajectory Analysis for Fine-Grained Text-Motion Retrieval. *arXiv:2508.03343*. (Under Review) [\[Paper\]](#)
10. **Fu H.***, Wang Y.*, Yang W., & Wen B. (2024). DP-IQA: Utilizing Diffusion Prior for Blind Image Quality Assessment in the Wild. arXiv preprint *arXiv:2405.19996*. (Under Review) [\[Paper\]](#)

11. **Fu H.***, Shen Y.*, Liu Y., Li J., & Zhang X. (2023). SGCN: A multi-order neighborhood feature fusion landform classification method based on superpixel and graph convolutional network. *International Journal of Applied Earth Observation and Geoinformation*. (Q1, IF=8.6) [[Paper](#)]

Patent

1. A Hidden Cable Positioning Detection and AR Visualization System, Patent No. CN114089271A, 2022. [[Link](#)]

Work Experience

StepFun

Shanghai, China

Research Intern

03/2025-06/2025

- ◆ Worked on generative AI. Trained text-to-image diffusion models tailored to business applications.
- ◆ Participated in the development of the diffusion-based image editing model Step1X-Edit and the VLM-based vector animation generation model LottieGen.

Hong Kong University of Science and Technology (Guangzhou)

Remote

Research Assistant

08/2023-01/2025

- ◆ LLM-based agents in video games.
- ◆ 2D/3D generative model and scene reconstruction.
- ◆ Multimodal tasks between biological signals and computer vision.

Sun Yat-sen University

Remote

Research Assistant

11/2022-06/2023

- ◆ Semantic segmentation of remote sensing images.
- ◆ The application of graph neural networks in computer vision.

Competition

SEU Robo Cup Robot Competition: Visual Group, Prize of Excellence

09/2020

SEU Smart Car Competition: Visual Group, Prize of Third Place

11-12/2019

Technical Skills

- ◆ **Programming skills:** Python (PyTorch), C/C++, HTML, Matlab and vibe coding tools.
- ◆ **Software skills:** Linux, Visual Studio, PyCharm and Interllij IDEA.
- ◆ **Service:** Reviewer of ICLR, NeurIPS, ICML, ACL, CVPR, ACM-MM, IJCNN, ICASSP, IEEE Transactions on Intelligent Transportation Systems (T-ITS), IEEE Internet of Things Journal (IoT), Journal of Selected Topics in Signal Processing (JSTSP), Pattern Recognition (PR), Neural Networks, Neurocomputing, etc.