

YONGTAO GE

(+86) 188-5208-3041 · yongtao.ge@adelaide.edu.au · <https://yongtaoge.github.io/>

EDUCATION

- The University of Adelaide**, Adelaide, Australia Apr. 2021 - present
– Ph.D. student at Australian Institute for Machine Learning (AIML)
– *Advisor: Prof. Chunhua Shen*
- Southeast University**, Nanjing, China Sept. 2017 - Jun. 2020
– M.Sc. School of Automation
- Nanjing University of Science and Technology**, Nanjing, China Sept. 2013 - Jun. 2017
– B. Eng. School of Automation.

RESEARCH INTERESTS

My research centers on multimodal foundation models, exploring their applications in advanced understanding, generation, and reasoning. I've previously worked at the intersection of 3D computer vision and generative modeling, with key contributions in photorealistic human digitization and scene understanding. I am particularly interested in:

- Developing multimodal large language models (LLMs) for advanced understanding, reasoning, and generation, applying them as AI agents across various downstream tasks.
- Developing and applying generative models (especially large image-to-video models) for creating high-fidelity digital avatars, and digital content.
- Creating robust representations for 3D human reconstruction and animation, human-human, and human-scene interaction.
- Exploring how these technologies can be applied to downstream applications in the gaming and film industries.

This work builds on my experience in 2D/3D human pose estimation, 3D human pose and shape recovery, and 3D geometric reconstruction from visual data.

RESEARCH EXPERIENCE

- Part-time Research Scientist at SpreeAI (*Mentor: Dr. Minh P. Vo*) Jul. 2024 -
Topic: 3D Human Measurement for Virtual Try-on
- Research Intern at Ant Group (*Advisor: Dr. Xuan Wang*) Jul. 2023 - Dec. 2023
Topic: 3D Human Generation with Pretrained Latent Diffusion Models
- Research Intern at Meituan (*Advisor: Dr. Zhi Tian*) Jul. 2022 - Apr. 2023
Topic: 3D BEV Object Detection from Multi-view Images
- Research Intern at Tencent (*Advisor: Dr. Gang Yu*) Dec. 2021 - Apr. 2022
Topic: 3D Human Pose and Shape Estimation
- Research Intern at Alibaba (*Advisor: Dr. Zhibin Wang and Dr. Qiang Zhou*) Sept. 2020 - Nov. 2021
Topic: Data-Efficient Object Detection
- Research Intern at SenseTime (*Advisor: Dr. Ruimao Zhang*) Mar. 2020 - May. 2020
Topic: Semantic Face Manipulation with conditional GANs
- Research Assitant at The University of Hong Kong (*Advisor: Prof. Ping Luo*) Aug. 2019 - Jan. 2020
Topic: Symbolic Mathematical Question Reasoning with Transformers

PUBLICATIONS AND PATENTS

- **Yongtao Ge**, Kangyang Xie, Guangkai Xu, Hao Chen, Chunhua Shen. Generative Video Matting. (Siggraph 2025, [code link](#))
- Kangyang Xie, **Yongtao Ge**, Yanlong Sun, Hao Chen, Chunhua Shen. MAP: Matting Any Portraits with Synthetic Data and Depth Priors. (Under Review for ICCV 2025)
- Songyang Zhang*, **Yongtao Ge***, Jinyuan Tian*, Guangkai Xu, Hao Chen, Chen Lv, Chunhua Shen. PO-MATO: Marrying Pointmap Matching with Temporal Motions for Dynamic 3D Reconstruction. (Under Review for ICCV 2025, [code link](#))

- **Yongtao Ge**, Guangkai Xu, Zheng Huang, Yanlong Sun, Libo Sun, Hao Chen, Chunhua Shen. GeoBench: Benchmarking and Analyzing Monocular Geometry Estimation Models. (Under Review for IJCV, [code link](#))
- **Yongtao Ge**, Wenjia Wang, Yongfan Chen, Lei Yang, Hao Chen, Chunhua Shen. 3D Human Reconstruction in the Wild with Synthetic Data Using Generative Models. (Under review for TPAMI [project link](#), [huggingface demo](#))
- Guangkai Xu*, **Yongtao Ge***, Mingyu Liu, Chengxiang Fan, Zhiyue Zhao, Hao Chen, Chunhua Shen. Diffusion Models Trained with Large Data Are Transferable Visual Models. (ICLR 2025 [code link](#), [huggingface demo](#))
- Wenjia Wang*, **Yongtao Ge***, Haiyi Mei, Zhongang Cai, Qingping Sun, Chunhua Shen, Yanjun Wang, Lei Yang, Taku Komura. Zolly: Zoom Focal Length Correctly for Perspective-Distorted Human Mesh Reconstruction. (ICCV 2023 [Oral Presentation](#), [arxiv link](#), full marks of review score)
- **Yongtao Ge**, Qiang Zhou, Xinlong Wang, Chunhua Shen, Zhibin Wang, Hao Li. Point-Teaching: Weakly Semi-Supervised Object Detection with Point Annotations. (AAAI 2023, [arxiv link](#))
- Weian Mao, **Yongtao Ge**, Chunhua Shen, Zhi Tian, Xinlong Wang, Anton van den Hengel, Zhibin Wang. Poseur: Direct Human Pose Regression with Transformers. (Co-first author. ECCV 2022, [github link](#))

MISCELLANEOUS

- **Programming Languages and Libraries:** C/C++, Python, MATLAB, PyTorch.
- **Languages:** English, Mandarin (native).

ACADEMIC SERVICES

- **Computer Vision:** Reviewer for ICCV 2023/2025, CVPR 2024, ECCV 2024
- **Machine Learning:** Reviewer for NeurIPS 2024/2025, ICLR 2025, ICML 2025, NIPS 2025.