

Tong Wang

Multimodal Deep Learning · Representation Learning

Beijing, China • tovi.myo@gmail.com • tongwang.info • WeChat: gnotoge

Education

- M.S. University of Southern California**, M.S. in Computer Science – Los Angeles, CA May 2019 – Dec 2020
- GPA: 3.587/4.0, Viterbi School of Engineering
- B.S. University of California, San Diego**, B.S. in Mathematics-Computer Science – San Diego, CA Sept 2016 – Dec 2018
- GPA: 3.814/4.0, Minor in Speculative Design
 - Transferred from UC Santa Cruz (2014–2016), 4-year Dean’s Scholarship (\$20,000)

Experience

- Meitu Inc.**, Senior Computer Vision Researcher, MT Lab – Beijing, China Dec 2023 – present
- Scene text editing: model achieves industry-leading performance, generating millions in monthly revenue
 - Video editing: system outperforms Pika and Kling in quality and cost
 - 3 papers accepted at CVPR 2025 and ICML 2026; awarded MT Lab 2025 Most Innovative Technology Award (1/~100 projects)
 - Rated A (highest) in 2024 annual performance review
- Xianyuan Electronics**, AIGC Algorithm Engineer, 3D Modeling – Beijing, China Mar 2023 – Nov 2023
- Led feasibility research on industrial-grade text/image-to-3D mesh generation
 - Implemented full fine-tuning and PEFT strategies for text-to-image products; filed 1 patent
- Protoqiang Information Technology**, Algorithm R&D Engineer, Research – Beijing, China Apr 2021 – Mar 2023
- Built unified lip-reading/speech/multimodal recognition model achieving SOTA on CMLR (WER: lip 19.06%, speech 1.27%, multimodal 1.17%)
 - Developed TTS system (FastPitch + HiFi-GAN) with few-shot voice cloning; filed 3 patents
 - Rated S (highest) in 2021 annual performance review
- University of Southern California**, Research Intern, Vision & Image Lab – Los Angeles, CA June 2020 – Nov 2020
- Led research on domain-adaptive multi-view semantic segmentation
 - Co-authored DenseGAP, accepted at ICPR 2022 (Oral)

Publications

- MiVE: Multiscale Vision-Language Features for Reference-Guided Video Editing** May 2026
Tong Wang, Meng Zou, Chengjing Wu, Xiaochao Qu, Luoqi Liu, Xiaolin Hu, Ting Liu
arxiv.org/abs/2605.14664 (ICML 2026)
- Self-Prompting Diffusion Transformer for Open-Vocabulary Scene Text Editing via In-Context Learning** May 2026
Hongxi Li, **Tong Wang**, Chengjing Wu, Tianbao Liu, Jiangtao Yao, Xiaochao Qu, Xinxiao Wu, Luoqi Liu, Ting Liu
arxiv.org/abs/2605.15523 (ICML 2026)
- GlyphMastero: A Glyph Encoder for High-Fidelity Scene Text Editing** June 2025
Tong Wang, Ting Liu, Xiaochao Qu, Chengjing Wu, Luoqi Liu, Xiaolin Hu
openaccess.thecvf.com/content/CVPR2025/html/Wang_GlyphMastero_A_Glyph_Encoder_for_High-Fidelity_Scene_Text_Editing_CVPR_2025_paper.html (CVPR 2025)
- DenseGAP: Graph-Structured Dense Correspondence Learning with Anchor Points** Aug 2022
Zhengfei Kuang, Jiaman Li, Mingming He, **Tong Wang**, Yajie Zhao
10.1109/ICPR56361.2022.9956472 (ICPR 2022 (Oral))

Patents

合成语音评估方法、装置、设备及存储介质, CN114493232B, 第一发明人 (2022)
一种新视角图像生成方法、装置、设备及可读存储介质, CN117456031A, 第一发明人 (2024)
一种语音数据获取方法、装置、电子设备和存储介质, CN114822494A, 第一发明人 (2022)
处理图像中文本的方法、装置、可读存储介质和程序品, CN120339462A, 第二发明人 (2025)
语音克隆模型生成方法、装置及电子设备, CN115831088A, 第二发明人 (2023)
语音合成方法以及装置、存储介质、电子装置, CN115240631A, 第三发明人 (2022)

Honors & Awards

MT Lab 2025 Most Innovative Technology Award (1/~100 projects)
MT Lab 2024 Annual Performance Rating A (Rank 1)
MT Lab 2024 Newcomer Award
Protoqiang 2021 Annual Performance Rating S (Rank 1)
Protoqiang 2022 Technical Innovation Award
UCSD Dean's Scholarship (\$20,000 over 4 years)