

Tsu-Jui (Ray) Fu

✉ tsujiifu@gmail.com
📄 tsujiifu.github.io

About Me

I am a Ph.D. candidate at UCSB. My research lies in vision+language and text-guided visual editing. I am also interested in language grounding and information extraction. My goal is to bridge the gap between multiple modalities via the AI system.

Education

- 2019–2024 **Ph.D. at University of California, Santa Barbara**, *Computer Science*, CA, USA.
Advisor: William Yang Wang
- 2014–2018 **B.S. at National Tsing Hua University**, *Computer Science*, Hsinchu, Taiwan.
Advisor: Shan-Hung Wu, Chun-Yi Lee, and Hwann-Tzong Chen

Experience

- 2019–now **UCSB NLP**, *Research Assistant*, CA, USA.
Advisor: William Yang Wang
- Summer 2023 **Apple AI/ML**, *Research Intern*, WA, USA.
Advisor: Zhe Gan and Yinfei Yang
- Summer 2022 **Meta AI**, *Research Intern*, CA, USA.
Advisor: Licheng Yu and Sean Bell
- Summer 2021 **Microsoft Azure AI**, *Research Intern*, WA, USA.
Advisor: Linjie Li, Zhe Gan, and Lijuan Wang
- Summer 2020 **Microsoft Research**, *Research Intern*, WA, USA.
Advisor: Yale Song and Daniel McDuff
- Summer 2019 **Preferred Networks**, *Research Intern*, Tokyo, Japan.
Advisor: Yuta Tsuboi and Jason Naradowsky
- 2018–2019 **Academia Sinica CKIP Lab**, *Research Assistant*, Taipei, Taiwan.
Advisor: Wei-Yun Ma

Publication

- ICLR'24 **Guiding Instruction-based Image Editing via Multimodal Large Language Models.**
Tsu-Jui Fu, Wenze Hu, Xianzhi Du, William Yang Wang, Yinfei Yang, and Zhe Gan
International Conference on Learning Representations (ICLR), 2024 (**Spotlight**)
- AAAI'24 **VELMA: Verbalization Embodiment of LLM Agents for Vision and Language Navigation in Street View.**
Raphael Schumann, Wanrong Zhu, Weixi Feng, **Tsu-Jui Fu**, Stefan Riezler, and William Yang Wang
Association for the Advancement of Artificial Intelligence (AAAI), 2024
- arXiv'23 **Discriminative Diffusion Models as Few-shot Vision and Language Learners.**
Xuehai He, Weixi Feng, **Tsu-Jui Fu**, Varun Jampani, Arjun Akula, Pradyumna Narayana, Sugato Basu, William Yang Wang, and Xin Eric Wang
arXiv:2305.10722
- EMNLP'23 **Text-guided 3D Human Generation from 2D Collections.**
Tsu-Jui Fu, Wenhan Xiong, Yixin Nie, Jingyu Liu, Barlas Oğuz, and William Yang Wang
Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023 (Findings)
- EMNLP'23 **EDIS: Entity-Driven Image Search over Multimodal Web Content.**
Siqi Liu*, Weixi Feng*, **Tsu-Jui Fu**, Wenhua Chen, and William Yang Wang
Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023 (Long)

- EMNLP'23 **Collaborative Generative AI: Integrating GPT-k for Efficient Editing in Text-to-Image Generation.**
Wanrong Zhu, Xinyi Wang, Yujie Lu, **Tsu-Jui Fu**, Xin Eric Wang, Miguel Eckstein, and William Yang Wang
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2023 (Short)
- NeurIPS'23 **Photoswap: Personalized Subject Swapping in Images.**
Jing Gu, Yilin Wang, Nanxuan Zhao, **Tsu-Jui Fu**, Wei Xiong, Qing Liu, Zhifei Zhang, He Zhang, Jianming Zhang, HyunJoon Jung, and Xin Eric Wang
Conference on Neural Information Processing Systems (**NeurIPS**), 2023
- NeurIPS'23 **LayoutGPT: Compositional Visual Planning and Generation with Large Language Models.**
Weixi Feng*, Wanrong Zhu*, **Tsu-Jui Fu**, Varun Jampani, Arjun Akula, Xuehai He, Sugato Basu, Xin Eric Wang, and William Yang Wang
Conference on Neural Information Processing Systems (**NeurIPS**), 2023
- CVPR'23 **Tell Me What Happened: Unifying Text-guided Video Completion via Multimodal Masked Video Generation.**
Tsu-Jui Fu, Licheng Yu, Ning Zhang, Cheng-Yang Fu, Jong-Chyi Su, William Yang Wang, and Sean Bell
Conference on Computer Vision and Pattern Recognition (**CVPR**), 2023
- CVPR'23 **An Empirical Study of End-to-End Video-Language Transformers with Masked Visual Modeling.**
Tsu-Jui Fu*, Linjie Li*, Zhe Gan, Kevin Lin, William Yang Wang, Lijuan Wang, and Zicheng Liu
Conference on Computer Vision and Pattern Recognition (**CVPR**), 2023
- ICLR'23 **Training-Free Structured Diffusion Guidance for Compositional Text-to-Image Synthesis.**
Weixi Feng, Xuehai He, **Tsu-Jui Fu**, Varun Jampani, Arjun Akula, Pradyumna Narayana, Sugato Basu, Xin Eric Wang, and William Yang Wang
International Conference on Learning Representations (**ICLR**), 2023
- EMNLP'22 **ULN: Towards Underspecified Vision-and-Language Navigation.**
Weixi Feng, **Tsu-Jui Fu**, Yujie Lu, and William Yang Wang
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2022 (Long)
- EMNLP'22 **CPL: Counterfactual Prompt Learning for Vision and Language Models.**
Xuehai He, Diji Yang, Weixi Feng, **Tsu-Jui Fu**, Arjun Akula, Varun Jampani, Pradyumna Narayana, Sugato Basu, William Yang Wang, and Xin Eric Wang
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2022 (Long)
- ECCV'22 **Language-Driven Artistic Style Transfer.**
Tsu-Jui Fu, Xin Eric Wang, and William Yang Wang
European Conference on Computer Vision (**ECCV**), 2022
- CVPR'22 **M³L: Language-based Video Editing via Multi-Modal Multi-Level Transformers.**
Tsu-Jui Fu, Xin Eric Wang, Scott Grafton, Miguel Eckstein, and William Yang Wang
Conference on Computer Vision and Pattern Recognition (**CVPR**), 2022
- AAAI'22 **DOC2PPT: Automatic Presentation Slides Generation from Scientific Documents.**
Tsu-Jui Fu, William Yang Wang, Daniel McDuff, and Yale Song
Association for the Advancement of Artificial Intelligence (**AAAI**), 2022
- arXiv'21 **VIOLET: End-to-End Video-Language Transformers with Masked Visual-token Modeling.**
Tsu-Jui Fu, Linjie Li, Zhe Gan, Kevin Lin, William Yang Wang, Lijuan Wang, and Zicheng Liu
arXiv:2111.12681
- ACL'21 **H-FND: Hierarchical False-Negative Denoising for Robust Distantly-Supervised Relation Extraction.**
Jhih-Wei Chen*, **Tsu-Jui Fu***, Chen-Kang Lee, and Wei-Yun Ma
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2021 (Findings)
- NAACL'21 **Semi-Supervised Policy Initialization for Playing Games with Language Hints.**
Tsu-Jui Fu and William Yang Wang
North American Chapter of the Association for Computational Linguistics (**NAACL**), 2021 (Short)

- EACL'21 **L2C: Describing Visual Differences Needs Semantic Understanding of Individuals.**
An Yang, Xin Wang, **Tsu-Jui Fu**, and William Yang Wang
European Chapter of the Association for Computational Linguistics (EACL), 2021 (Short)
- EACL'21 **Multimodal Style Transfer Learning for Outdoor Vision-and-Language Navigation.**
Wanrong Zhu, Xin Wang, **Tsu-Jui Fu**, An Yan, Pradyumna Narayana, Kazoo Sone, Sugato Basu, and William Yang Wang
European Chapter of the Association for Computational Linguistics (EACL), 2021 (Long)
- EMNLP'20 **SSCR: Iterative Language-Based Image Editing via Self-Supervised Counterfactual Reasoning.**
Tsu-Jui Fu, Xin Eric Wang, Scott Grafton, Miguel Eckstein, and William Yang Wang
Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020 (Oral)
- ECCV'20 **Counterfactual Vision-and-Language Navigation via Adversarial Path Sampler.**
Tsu-Jui Fu, Xin Eric Wang, Matthew Peterson, Scott Grafton, Miguel Eckstein, and William Yang Wang
European Conference on Computer Vision (ECCV), 2020 (Spotlight)
- AAAI'20 **Why Attention? Analyzing and Remediating BiLSTM Deficiency in Modeling Cross-Context for NER.**
Peng-Hsuan Li, **Tsu-Jui Fu**, and Wei-Yun Ma
Association for the Advancement of Artificial Intelligence (AAAI), 2020 (Oral)
- NeurIPS'19 **Learning from Observation-Only Demonstration for Task-Oriented Language Grounding via Self-Examination.**
Tsu-Jui Fu, Yuta Tsuboi, Sosuke Kobayashi, and Yuta Kikuchi
Conference on Neural Information Processing Systems (NeurIPS), 2019 (ViGIL workshop)
- ICCD'19 **A Distributed Scheme for Accelerating Semantic Video Segmentation on An Embedded Cluster.**
Hsuan-Kung Yang*, **Tsu-Jui Fu***, Kuan-Wei Ho, Po-Han Chiang and Chun-Yi Lee
International Conference on Computer Design (ICCD), 2019 (Oral)
- CoRL'19 **Adversarial Active Exploration for Inverse Dynamics Model Learning.**
Zhang-Wei Hong, **Tsu-Jui Fu**, Tzu-Yun Shann, Yi-Hsiang Chang, and Chun-Yi Lee
Conference on Robot Learning (CoRL), 2019 (Oral)
- ACL'19 **GraphRel: Modeling Text as Relational Graphs for Joint Entity and Relation Extraction.**
Tsu-Jui Fu, Peng-Hsuan Li, and Wei-Yun Ma
Annual Meeting of the Association for Computational Linguistics (ACL), 2019 (Long)
- WACV'19 **Attentive and Adversarial Learning for Video Summarization.**
Tsu-Jui Fu, Shao-Heng Tai, and Hwann-Tzong Chen
Winter Conference on Applications of Computer Vision (WACV), 2019 (Oral)
- ACCV'18 **Region-Semantics Preserving Image Synthesis.**
Kang-Jun Liu, **Tsu-Jui Fu**, and Shan-Hung Wu
Asian Conference on Computer Vision (ACCV), 2018
- NeurIPS'18 **Diversity-Driven Exploration Strategy for Deep Reinforcement Learning.**
Zhang-Wei Hong, Tzu-Yun Shann, Shih-Yang Su, Yi-Hsiang Chang, **Tsu-Jui Fu**, and Chun-Yi Lee
Conference on Neural Information Processing Systems (NeurIPS), 2018
- EMNLP'18 **Speed Reading: Learning to Read ForBackward via Shuttle.**
Tsu-Jui Fu and Wei-Yun Ma
Conference on Empirical Methods in Natural Language Processing (EMNLP), 2018 (Long)
- ECCV'18 **Visual Relationship Prediction via Label Clustering and Incorporation of Depth Information.**
Hsuan-Kung Yang, An-Chieh Cheng*, Kuan-Wei Ho*, **Tsu-Jui Fu**, and Chun-Yi Lee
European Conference on Computer Vision (ECCV), 2018 (PIC workshop)
- CVPR'18 **Dynamic Video Segmentation Network.**
Yu-Syuan Xu, **Tsu-Jui Fu***, Hsuan-Kung Yang*, and Chun-Yi Lee
Conference on Computer Vision and Pattern Recognition (CVPR), 2018

Organizer

- 2024 **3rd Workshop on Advances in Language and Vision Research (ALVR)**.
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2024
- 2021 **2nd Workshop on Advances in Language and Vision Research (ALVR)**.
North American Chapter of the Association for Computational Linguistics (**NAACL**), 2021

Award

- Spring 2023 **Second Place**, *Alexa Prize SimBot Challenge*, Amazon, United States.
- Spring 2019 **Third Place**, *Formosa Grand Challenge*, Ministry of Science and Technology (**MOST**), Taiwan.
- Summer 2018 **Second Place**, *Person In Context Challenge*, ECCV 2018.

Teaching

- Winter 2022 **Teaching Assistant**, *CS190I Introduction to Natural Language Processing*, UC Santa Barbara.
- Summer 2018 **Teaching Assistant**, *Vision and Learning Summer School*, National Tsing Hua University.
- Spring 2018 **Teaching Assistant**, *Nvidia Deep Learning Workshop*, National Central University.
- Fall 2017 **Teaching Assistant**, *CS565600 Deep Learning*, National Tsing Hua University.
- Fall 2017 **Teaching Assistant**, *CS210401 Hardware Design and Lab*, National Tsing Hua University.