# Jialin Wu

Email: jialinwu@google.com
Phone: +1 (737) 781-1180
Linkedin: https://www.linkedin.com/in/jialin-wu-a50135175/
Homepage: https://jialinwu.netlify.app

## WORK EXPERIENCE

**Senior Research Scientist** Nov 2024 - Now
Google DeepMind

**Research Scientist** May 2024 - Nov 2024
Google DeepMind

**Research Scientist** Sep 2022 - May 2024
Google Research

## EDUCATION

**Ph.D. in Computer Science Department** September 2017 - August 2022
The University of Texas at Austin, TX, USA
Advisor: Dr. Raymond J. Mooney

**BEng. in the Department of Automation** September 2013 - July 2017
Tsinghua University, Beijing, China
Advisor: Dr. Xiangyang Ji

## RESEARCH INTERESTS

Recently, I am interested in enhancing the capabilities of image generation models on info-seeking queries. Some research questions I am exploring include (1) integrating search signals during the pre/post-training phases as well as during inference for image generation, and (2) enhancing the factual accuracy of images produced in response to info-seeking queries.

## RESEARCH PUBLICATIONS

Yue Zhao, Long Zhao, Xingyi Zhou, **Jialin Wu**, Chun-Te Chu, Hui Miao, Florian Schroff, Hartwig Adam, Ting Liu, Boqing Gong, Philipp Krähenbühl, Liangzhe Yuan "Distilling vision-language models on millions of videos" In *CVPR* 2024.

**Jialin Wu**, Xia Hu, Yaqing Wang, Bo Pang, Radu Soricut. "Omni-SMoLA:Boosting Generalist Multimodal Models with Soft Mixture of Low-rank Experts" In *CVPR* 2024. **(Highlight)**

Xi Chen, Josip Djolonga, Piotr Padlewski, Basil Mustafa, Soravit Changpinyo, **Jialin Wu**, et. al., "PaLI-X:On Scaling up a Multilingual Vision and Language Model" In *CVPR* 2024.

Nan Ding, Tomer Levinboim, **Jialin Wu**, Sebastian Goodman, Radu Soricut. "CausalLM Is Not Optimal for In-Context Learning" in *ICLR* 2024.

Brianna Zitkovich, Tianhe Yu, Sichun Xu, Peng Xu, Ted Xiao, Fei Xia, **Jialin Wu**, et. al., "RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Controlling" In *CoRL*, 2023.

**Jialin Wu**, Raymond J. Mooney. "Entity-Focused Dense Passage Retrieval for Outside-Knowledge Visual Question Answering." In *EMNLP*, 2022.

**Jialin Wu**, Jiasen Lu, Ashish Sabharwal, Roozbeh Mottaghi. "Multi-Modal Answer Validation for Knowledge-Based VQA." In *AAAI*, 2022. **(Oral, Acceptance rate 15%)**

**Jialin Wu**, Liyan Chen, Raymond J. Mooney. "Improving VQA and its Explanations by Comparing Competing Explanations." In *AAAI Explainable Agency in Artificial Intelligence Workshop*, 2021. **(Oral)**

Jungjun Kim, Hanbin Ko, **Jialin Wu** "CoNAN: A Complementary Neighboring-based Attention Network for Referring Expression Generation". In *COLING*, 2020.

**Jialin Wu** and Raymond J. Mooney. "Self-Critical Reasoning for Robust Visual Question Answering." In *NeurIPS*, 2019 **(Spotlight, 2.97%)**

**Jialin Wu** and Raymond J. Mooney. "Hidden State Guidance: Improving Image Captioning using An Image Conditioned Autoencoder." In *NeurIPS Vigil Workshop*, 2019.

**Jialin Wu**, Zeyuan Hu, Raymond J. Mooney. "Generating Question Relevant Captions to Aid Visual Question Answering" In *ACL*, 2019. **(Oral)**

**Jialin Wu** and Raymond J. Mooney. "Faithful Multimodal Explanation for Visual Question Answering." In *ACL BlackboxNLP Workshop*, 2019. **(Oral)**

**Jialin Wu**$^{\star}$, Dai Li$^{\star}$, Yu Yang$^{\star}$, Chandrajit Bajaj and Xiangyang Ji. "Dynamic Filtering with Large Sampling Field for ConvNets." In *ECCV*, 2018. **(Poster)**

**Jialin Wu**, Yu Yang, He Jiang, Yi Li, Guijin Wang, Xiangyang Ji. "Action Recognition and Localization with Instance FCNN" In *RCAR*, 2018.

## TECHNICAL REPORTS

**Gemini Team**. "Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities" In arXiv preprint arXiv:2507.06261 (2025).

**Jialin Wu** and Raymond J. Mooney. "Breaking Down Questions for Outside-Knowledge Visual Question Answering." In OpenReview, 2021.

**Jialin Wu**, Soumyajit Gupta, and Chandrajit Bajaj. "Higher Order Mutual Information Approximation for Feature Selection." In arXiv preprint arXiv:1612.00554 (2016).

**Jialin Wu**, Gu Wang, Wukui Yang, Xiangyang Ji. "Action Recognition with Joint Attention on Multi-Level Deep Features." In arXiv preprint arXiv:1607.02556 (2016).

## INTERNS

**AI2 Ph.D. Research Intern** May 2020 - Aug 2020

PRIOR Team, AI2 Seattle.

**Google Ph.D. Research Intern** May 2019 - Aug 2019

Research Machine Intelligence Team, Google NYC.

**Research Assistant** May 2018 - present

Machine Learning Research Group , UT Austin.

**Research Assistant** May 2016 - Sep 2016

Computational Visualization Center, UT Austin, USA

**Research Assistant** Jan 2016 - July 2017

Broadband Network and Digital Media Lab, Tsinghua University.

## PROFESSIONAL SERVICES

Program Committee Member: IJCAI 2019 Workshop on Explainable Artificial Intelligence (XAI)
Conference Reviewer: CVPR (2019, 2020), BMVC(2020), ECCV (2020), ICCV (2019), WACV (2022)
AAAI (2020, 2021,2023), NeurIPS (2020, 2022, 2023), EMNLP (2021), ICLR (2022, 2023), ICML (2023)
Organizer: UT Explainable Artificial Intelligence (XAI) Reading Group

## TEACHING EXPERIENCE

CS371R: Information Retrieval and Web Search (Fall 2021)
CS303E: Elements of Computers and Programming (Spring 2018)
CS394N: Neural Networks (Fall 2017)

## TECHNICAL SKILLS

**Programming Languages**: Python, C/C++, CUDA
**Deep Learning Frameworks**: Jax, Caffe, Caffe2, PyTorch, TensorFlow