Zi-Yi Dou

z_dou@apple.com | [Homepage](#) | [Google Scholar](#) | [GitHub](#)

EDUCATION

University of California, Los Angeles
Aug. 2020 – Jan. 2025

Ph.D. in Computer Science
- Advisor: Nanyun Peng
- Funded by Amazon PhD Fellowship, JPMC PhD Fellowship

Carnegie Mellon University
Aug. 2018 – Aug. 2020

M.S. in Language Technologies
- Advisor: Graham Neubig
- GPA: 4.01/4.33

Nanjing University
Sep. 2014 – Jul. 2018

B.S. in Computer Science, Minor in Statistics
- National Elite Program, Kuang Yaming Honors School
- Exchange student at UC Berkeley, worked with Prof. Joseph E. Gonzalez (Fall 2017)
- GPA: 91.12/100.0, Ranking: 1/18

EXPERIENCE

xAI
Nov. 2025 – Present

Member of Technical Staff
- post-training at the Macrohard team

Apple
Jan. 2025 – Nov. 2025

Research Scientist at Apple AI/ML
- Managers: Zhe Gan, Yinfei Yang
- Apple foundation models: multimodal posttraining; researching multimodal reasoning and agents

Meta
Jun. 2024 – Oct. 2024

Research Intern at FAIR Perception
- Peers: Pengchuan Zhang, Kate Saenko, Shoubhik Debnath, Didac Suris Coll-Vinent, Nicolas Carion
- Interned at the Segment Anything (SAM) team
- Worked on visually grounded reasoning for Llama

Meta
Jun. 2023 – Dec. 2023

Research Intern at FAIR Multimodal Foundations
- Peers: Fu-Jen Chu, Kris Kitani, Xitong Yang, Tushar Nagarajan, Huiyu Wang, Jing Huang
- Worked on egocentric video-language pretraining

Microsoft
May 2021 – Sep. 2022

Research Intern at Project Florence-Vision and Language
- Mentor: Zhe Gan
- Worked on vision-language pretraining and its applications in object detection/segmentation

PUBLICATIONS

- AFM Team, *Apple Intelligence Foundation Language Models*, 2025.

- Zhen Yang*, **Zi-Yi Dou***, Feng Di*, Forrest Huang, Anh Nguyen, Keen You, Omar Attia, Yuhao Yang, Michael Feng, Haotian Zhang, Ram Ramrakhya, Chao Jia, Jeffrey Nichols, Alexander Toshev, Yinfei Yang, Zhe Gan, *Ferret-UI Lite: Lessons from Building Small On-Device GUI Agents*, 2025.

- Yuhao Yang, Zhen Yang, **Zi-Yi Dou**, Anh Tuan Nguyen, Omar Attia, Andrew Szot, Keen You, Michael Feng, Ram Ramrakhya, Alexander Toshev, Chao Huang, Yinfei Yang, Zhe Gan, *UltraCUA: Scaling Computer Use Agent through GUI and Programmatic Control*, 2025.

- Yanghao Li, Rui Qian, Bowen Pan, Haotian Zhang, Haoshuo Huang, Bowen Zhang, Jialing Tong, Haoxuan You, Xianzhi Du, Zhe Gan, Hyunjik Kim, Chao Jia, Zhenbang Wang, Yinfei Yang, **Zi-Yi Dou**, Wenze Hu, Chang Gao, Dongxu Li, Philipp Dufter, Zirui Wang, Guoli Yin, Zhengdong Zhang, Chen Chen, Yang Zhao, Ruoming Pang, Zhifeng Chen, *MANZANO: A Simple and Scalable Unified Multimodal Model with a Hybrid Vision Tokenizer*, 2025.

- **Zi-Yi Dou**, Xitong Yang, Tushar Nagarajan, Huiyu Wang, Jing Huang, Nanyun Peng, Kris M. Kitani, and Fu-Jen Chu, *"Unlocking Exocentric Data for Egocentric Video Representation Learning"*, Preprint 2024.

- Wenbo Hu, **Zi-Yi Dou**, Liunian Harold Li, Amita Kamath, Nanyun Peng, and Kai-Wei Chang, *"MQT-LLaVA: Matryoshka Query Transformer for Large Vision-Language Models"*, NeurIPS 2024.

- **Zi-Yi Dou**, Cheng-Fu Yang, Xueqing Wu, Kai-Wei Chang, Nanyun Peng, *"Re-ReST: Reflection-Reinforced Self-Training for Language Agents"*, EMNLP 2024.

- Liunian Harold Li*, **Zi-Yi Dou***, Nanyun Peng, and Kai-Wei Chang, *"DesCo: Learning Object Recognition with Rich Language Descriptions"*, NeurIPS 2023. CVPR OmniLabel 2023 Challenge Winner (1st place).

- Xueyan Zou*, **Zi-Yi Dou***, Jianwei Yang*, Zhe Gan, Linjie Li, Chunyuan Li, Xiyang Dai, Harkirat Behl, Jianfeng Wang, Lu Yuan, Nanyun Peng, Lijuan Wang, Yong Jae Lee, and Jianfeng Gao, *"Generalized Decoding for Pixel, Image, and Language"*, CVPR 2023.

- **Zi-Yi Dou***, Aishwarya Kamath*, Zhe Gan*, Pengchuan Zhang, Jianfeng Wang, Linjie Li, Zicheng Liu, Ce Liu, Yann LeCun, Nanyun Peng, Jianfeng Gao and Lijuan Wang, *"Coarse-to-Fine Vision-Language Pre-training with Fusion in the Backbone"*, NeurIPS 2022.

- **Zi-Yi Dou**, Yichong Xu, Zhe Gan, Jianfeng Wang, Shuohang Wang, Lijuan Wang, Chenguang Zhu, Nanyun Peng, Zicheng Liu and Michael Zeng, *"An Empirical Study of Training End-to-End Vision-and-Language Transformers"*, CVPR 2022.

- **Zi-Yi Dou** and Nanyun Peng, *"FOAM: A Follower-aware Speaker Model for Vision-and-Language Navigation"*, NAACL 2022.

- **Zi-Yi Dou** and Nanyun Peng, *"Improving Pre-trained Vision-and-Language Embeddings for Phrase Grounding"*, EMNLP 2021.

- **Zi-Yi Dou** and Graham Neubig, *"Word Alignment by Fine-tuning Embeddings on Parallel Corpora"*, EACL 2021.

- Yixin Liu, **Zi-Yi Dou** and Pengfei Liu, *"RefSum: Refactoring Neural Summarization"*, NAACL 2021.

- **Zi-Yi Dou**, Pengfei Liu, Hiroaki Hayashi, Zhengbao Jiang and Graham Neubig, *"GSum: A General Framework for Guided Neural Abstractive Summarization"*, NAACL 2021.

- **Zi-Yi Dou**, Antonios Anastasopoulos and Graham Neubig, *"Dynamic Data Selection and Weighting for Iterative Back-Translation"*, EMNLP 2020.

- **Zi-Yi Dou**, Sachin Kumar and Yulia Tsvetkov, *"A Deep Reinforced Model for Zero-Shot Cross-Lingual Summarization with Bilingual Semantic Similarity Rewards"*, ACL 2020 WNGT.

- **Zi-Yi Dou**, Junjie Hu, Antonios Anastasopoulos and Graham Neubig, *"Unsupervised Domain Adaptation for Neural Machine Translation with Domain-Aware Feature Embeddings"*, EMNLP 2019.

- **Zi-Yi Dou**, Keyi Yu and Antonios Anastasopoulos, *"Investigating Meta-Learning Algorithms for Low-Resource Natural Language Understanding Tasks"*, EMNLP 2019.

- Graham Neubig, **Zi-Yi Dou**, Junjie Hu, Paul Michel, Danish Pruthi, Xinyi Wang and John Wieting, *"compare-mt: A Tool for Holistic Comparison of Language Generation Systems"*, NAACL 2019 Demo. Best Demo Runner-up.

- **Zi-Yi Dou**, Zhaopeng Tu, Xing Wang, Longyue Wang, Shuming Shi and Tong Zhang, *"Dynamic Layer Aggregation for Neural Machine Translation with Routing-by-Agreement"*, AAAI 2019.

- **Zi-Yi Dou**, Zhaopeng Tu, Xing Wang, Shuming Shi and Tong Zhang, *"Exploiting Deep Representations for Neural Machine Translation"*, EMNLP 2018.

- Xin Wang, Fisher Yu, **Zi-Yi Dou**, Trevor Darrel and Joseph E. Gonzalez, *"SkipNet: Learning Dynamic Routing in Convolutional Networks"*, ECCV 2018.

- **Zi-Yi Dou**, *"Capturing User and Product Information for Document Level Sentiment Analysis with Deep Memory Network"*, EMNLP 2017.

Awards

JPMC PhD Fellowship	July 2024
UCLA Outstanding Graduate Student Research Award	May 2023
NeurIPS Scholar Award	Oct. 2022, Oct. 2023
Amazon PhD Fellowship	Sep. 2022
NAACL Best System Demonstration Runner-up	Jun. 2019
AAAI Scholar Award	Jan. 2019
Nanjing University Outstanding Graduate Award	Jul. 2018
China National Scholarship	Sep. 2017

Service

Reviewer for ACL, EMNLP, NAACL, ACL Rolling Review, COLM, ICML, NeurIPS, ICLR, CVPR, ICCV, ECCV, AAAI

Area Chair for ACL, EMNLP, ACL Rolling Review