Yue Yang

https://yueyang1996.github.io | yueyang1@seas.upenn.edu | (+1) 215-452-1517

Education

Ph.D. in Computer and Information Science — Aug 2020 - May 2025
University of Pennsylvania
Advisor: Chris Callison-Burch, Mark Yatskar

M.S. in Robotics — Aug 2018 - May 2020
University of Pennsylvania

B.E. in Mechanical Engineering — Aug 2014 - May 2018
Zhejiang University

Research Interests

My research lies at the intersection area of Natural Language Processing and Computer Vision. I aim to apply knowledge priors of large language models to build visual intelligence with better **interpretability** [5], **robustness** [2] and **data efficiency** [3][4][1].

Selected Works

[1] **Yue Yang***, Ajay Patel*, Matt Deitke, Tanmay Gupta, Luca Weihs, Andrew Head, Mark Yatskar, Chris Callison-Burch, Ranjay Krishna, Aniruddha Kembhavi, Christopher Clark. Scaling Text-Rich Image Understanding via Code-Guided Synthetic Multimodal Data Generation. (**ACL** 2025)

[2] **Yue Yang**, Mona Gandhi, Yufei Wang, Yifan Wu, Michael S. Yao, James C. Gee, Chris Callison-Burch, Mark Yatskar. A Textbook Remedy for Domain Shifts: Knowledge Priors for Medical Image Analysis. (**NeurIPS** 2024, **spotlight**)

[3] Matt Deitke*, Christopher Clark*, Sangho Lee, Rohun Tripathi, **Yue Yang**, Jae Sung Park, Mohammadreza Salehi, et al. (51 authors in total) Molmo and PixMo: Open Weights and Open Data for State-of-the-Art Multimodal Models. (**CVPR** 2025, **oral**)

[4] **Yue Yang***, Fan-Yun Sun*, Luca Weihs*, Eli Vanderbilt, Alvaro Herrasti, Winson Han, Jiajun Wu, Nick Haber, Ranjay Krishna, Lingjie Liu, Chris Callison-Burch, Mark Yatskar, Aniruddha Kembhavi, Christopher Clark. Holodeck: Language Guided Generation of 3D Embodied AI Environments. (**CVPR** 2024)

[5] **Yue Yang**, Artemis Panagopoulou, Shenghao Zhou, Daniel Jin, Chris Callison-Burch, Mark Yatskar. Language in a Bottle: Language Model Guided Concept Bottlenecks for Interpretable Image Classification. (**CVPR** 2023)

Publications

1. Zhantao Yang, Ruili Feng, Keyu Yan, Huangji Wang, Zhicai Wang, Shangwen Zhu, Han Zhang, Jie Xiao, Pingyu Wu, Kai Zhu, JixuanChen, Chen-Wei Xie, **Yue Yang**, Hongyang Zhang, Yu Liu, Fan Cheng. BACON: Improving Clarity of Image Captions via Bag-of-Concept Graphs. (**CVPR** 2025)

2. Long Le, Jason Xie, William Liang, Hung-Ju Wang, **Yue Yang**, Yecheng Jason Ma, Kyle Vedder, Arjun Krishna, Dinesh Jayaraman, Eric Eaton. Articulate-Anything: Automatic Modeling of Articulated Objects via a Vision-Language Foundation Model. (**ICLR** 2025)

3. Yifan Wu, Yang Liu, **Yue Yang**, Michael S. Yao, Wenli Yang, Xuehui Shi, Lihong Yang, Dongjun Li, Yueming Liu, James C. Gee, Xuan Yang, Wen-bin Wei, Shi Gu. A Concept-

based Interpretable Model for the Diagnosis of Choroid Neoplasias using Multimodal Data. (**Nature Communications**, 2025)

4. Runsheng Huang, Liam Dugan, **Yue Yang**, Chris Callison-Burch. MiRAGeNews: Multimodal Realistic AI-Generated News Detection. (Findings of **EMNLP** 2024)
5. Yiming Huang, Weilin Wan, **Yue Yang**, Chris Callison-Burch, Mark Yatskar, Lingjie Liu. CoMo: Controllable Motion Generation through Language Guided Pose Code Editing. (**ECCV** 2024)
6. Tuhin Chakrabarty, Arkady Saakyan, Olivia Winn, Artemis Panagopoulou, **Yue Yang**, Marianna Apidianaki, Smaranda Muresan. Large Language Models and Diffusion Models Co-Create Visual Metaphors. (Findings of **ACL** 2023)
7. Li Zhang, Hainiu Xu, **Yue Yang**, Shuyan Zhou, Weiqiu You, Manni Arora, Chris Callison-Burch. Causal Reasoning About Entities and Events in Procedural Texts. (Findings of **EACL** 2023)
8. **Yue Yang**, Wenlin Yao, Hongming Zhang, Xiaoyang Wang, Dong Yu, Jianshu Chen. Z-LaVI: Zero-Shot Language Solver Fueled by Visual Imagination. (**EMNLP** 2022)
9. **Yue Yang***, Artemis Panagopoulou*, Marianna Apidianaki, Mark Yatskar and Chris Callison-Burch. Visualizing the Obvious: A Concreteness-based Ensemble Model for Noun Property Prediction. (Findings of **EMNLP** 2022)
10. Shuyan Zhou*, Li Zhang*, **Yue Yang**, Qing Lyu, Graham Neubig, Chris Callison-Burch. Show Me More Details: Discovering Event Hierarchies from WikiHow. (**ACL** 2022)
11. **Yue Yang**, Artemis Panagopoulou, Qing Lyu, Li Zhang, Mark Yatskar, Chris Callison-Burch. Visual Goal-Step Inference using wikiHow. (**EMNLP** 2021)

Preprints & Workshop Papers

[I] Zhaowei Wang, Hongming Zhang, Tianqing Fang, Ye Tian, **Yue Yang**, Kaixin Ma, Xiaoman Pan, Yangqiu Song, Dong Yu. DivScene: Evaluating Large Vision Language Models for Object Navigation with Open-Vocabulary Targets in Diverse Scenes. (In submission, 2025)

[II] Josh Magnus Ludan, Qing Lyu, **Yue Yang**, Liam Dugan, Mark Yatskar, Chris Callison-Burch. Interpretable-by-Design Text Classification with Iteratively Generated Concept Bottleneck. (arxiv, 2023)

[III] **Yue Yang**, Joongwon Kim, Artemis Panagopoulou, Mark Yatskar, Chris Callison-Burch. Induce, Edit, Retrieve: Language Grounded Multimodal Schema for Instructional Video Retrieval. (O-DRUM Workshop at CVPR 2022)

Professional Experience

PRIOR @ Allen Institute for AI, Seattle, WA — Aug 2025 - Present
Research Scientist
Research Intern (Summer 2023, 2024)
Ai2 Outstanding Intern of the Year Award (2023)

Tencent AI Lab, Bellevue, WA — Summer 2022
Research Scientist Intern

Teaching

Teaching Assistant

CIS-521 Artificial Intelligence — Fall 2019 - Fall 2022
CIS-530 Computational Linguistics — Spring 2021
Outstanding Teaching Award, University of Pennsylvania (2020)

Academic Services

Paper Review

Computer Vision: CVPR, ECCV, ICCV, SIGGRAPH Asia.
Natural Language Processing: ACL, EMNLP, NAACL, EACL, COLM.
Machine Learning: NeurIPS, ICLR, ICML, TMLR.

Press Coverages

MarkTechPost — Feb 2025
CoSyn: An AI Framework that Leverages the Coding Capabilities of Text-only Large Language Models (LLMs) to Automatically Create Synthetic Text-Rich Multimodal Data

Penn Engineering & Penn Today & Medical News Bulletin & Medical Xpress — Oct 2024
Training Medical AI with Knowledge, Not Shortcuts.

TechCrunch & WIRED & MIT Technology & VentureBeat — Sep 2024
Ai2's Molmo shows open source can meet, and beat, closed multimodal models.

Penn Engineering & Tech Xplore & Tech Times & Tech Briefs — Apr 2024
Penn Engineers Recreate Star Trek's Holodeck Using ChatGPT and Video Game Assets.