

# YUE YANG

<https://yueyang1996.github.io> | [yueyang2333@gmail.com](mailto:yueyang2333@gmail.com) | (+1) 215-452-1517

## EDUCATION

---

Ph.D. in Computer and Information Science <b>University of Pennsylvania</b> Thesis: Language Priors for Visual Intelligence	2020 - 2025
M.S. in Robotics <b>University of Pennsylvania</b>	2018 - 2020
B.E. in Mechanical Engineering <b>Zhejiang University</b>	2014 - 2018

## PROFESSIONAL EXPERIENCE

---

<b>Allen Institute for AI</b> , Seattle, WA Research Scientist	08/2025 - Present
<b>University of Pennsylvania</b> , Philadelphia, PA Research Assistant	09/2020 - 08/2025
<b>Tencent AI Lab</b> , Bellevue, WA Research Scientist Intern	05/2022 - 08/2022

## RESEARCH INTERESTS

---

My research lies at the intersection of Natural Language Processing and Computer Vision. I aim to apply knowledge priors from large language models to build multimodal intelligence that can **interact** with both the **digital** [1] [II] and **physical** worlds [4].

## 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**, **SAC Highlights**)
- [2] Matt Deitke\*, Christopher Clark\*, Sangho Lee, Rohun Tripathi, **Yue Yang**, Jae Sung Park, et al. (51 authors in total) [Molmo and PixMo: Open Weights and Open Data for State-of-the-Art Multimodal Models](#). (**CVPR 2025**, **Best Paper Honorable Mention**)
- [3] **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**)
- [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. Christopher Clark, Jieyu Zhang, Zixian Ma, Jae Sung Park, Mohammadreza Salehi, Rohun Tripathi, Sangho Lee, Zhongzheng Ren, Chris Dongjoo Kim, Yinuo Yang, Vincent Shao, **Yue Yang**, Weikai Huang, Ziqi Gao, Taira Anderson, Jianrui Zhang, Jitesh Jain, George Stoica, Winson Han, Ali Farhadi, Ranjay Krishna. [Open Weights and Data for Vision-Language Models with Video Understanding and Grounding](#). (**CVPR**, 2026)
2. 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](#). (Findings of **EMNLP**, 2025)
3. Zhantao Yang, Ruili Feng, Keyu Yan, Huangji Wang, Zhicai Wang, Shangwen Zhu, Han Zhang, Jie Xiao, Pingyu Wu, Kai Zhu, Jixuan Chen, Chen-Wei Xie, **Yue Yang**, Hongyang Zhang, Yu Liu, Fan Cheng. [BACON: Improving Clarity of Image Captions via Bag-of-Concept Graphs](#). (**CVPR** 2025)
4. 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)
5. 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)
6. Runsheng Huang, Liam Dugan, **Yue Yang**, Chris Callison-Burch. [MiRAGeNews: Multi-modal Realistic AI-Generated News Detection](#). (Findings of **EMNLP** 2024)
7. 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)
8. 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)
9. 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)
10. **Yue Yang**, Wenlin Yao, Hongming Zhang, Xiaoyang Wang, Dong Yu, Jianshu Chen. [Z-LaVi: Zero-Shot Language Solver Fueled by Visual Imagination](#). (**EMNLP** 2022)
11. **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)
12. Shuyan Zhou\*, Li Zhang\*, **Yue Yang**, Qing Lyu, Graham Neubig, Chris Callison-Burch. [Show Me More Details: Discovering Event Hierarchies from WikiHow](#). (**ACL** 2022)
13. **Yue Yang**, Artemis Panagopoulou, Qing Lyu, Li Zhang, Mark Yatskar, Chris Callison-Burch. [Visual Goal-Step Inference using wikiHow](#). (**EMNLP** 2021)

## PREPRINTS & WORKSHOP PAPERS

---

- [1] Christopher Clark\*, **Yue Yang**\*, Jae Sung Park\*, Zixian Ma, Jieyu Zhang, Mohammadreza Salehi, Rohun Tripathi, Sangho Lee, Ranjay Krishna. [Molmo-Point: Better Pointing for VLMs with Grounding Tokens](#). (arxiv, 2026)

- [II] Tanmay Gupta\*, Piper Wolters\*, Zixian Ma\*, Peter Sushko\*, Rock Yuren Pang\*, Diego Llanes\*, **Yue Yang\***, Taira Anderson, Boyuan Zheng, Zhongzheng Ren, Harsh Trivedi, Taylor Blanton, Caleb Ouellette, Winson Han, Ali Farhadi, Ranjay Krishna. [MolmoWeb: Open Multimodal Action Policy for the Open Web.](#) (arxiv 2026).
- [III] Jianrui Zhang, **Yue Yang**, Rohun Tripathi, Winson Han, Ranjay Krishna, Christopher Clark, Yong Jae Lee, Sangho Lee. [Unified Spatial-Temporal Token Scoring for Efficient Video-Language Models.](#) (arxiv, 2026)
- [IV] Zixuan Bian\*, Ruohan Ren\*, **Yue Yang**, Chris Callison-Burch. [HOLODECK 2.0: Vision-Language-Guided 3D World Generation with Editing.](#) (arxiv, 2025)
- [V] 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)
- [VI] **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)

## TEACHING

---

### Teaching Assistant

CIS-521 Artificial Intelligence	2019 - 2022
CIS-530 Computational Linguistics	2021

## 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

---

<a href="#">VentureBeat &amp; Penn Engineering &amp; MarkTechPost</a>	2025
<b>AI Vision, Reinvented: The Power of Synthetic Data</b>	
<a href="#">Penn Engineering &amp; Penn Today &amp; Medical News Bulletin &amp; Medical Xpress</a>	2024
<b>Training Medical AI with Knowledge, Not Shortcuts.</b>	
<a href="#">TechCrunch &amp; WIRED &amp; MIT Technology &amp; VentureBeat</a>	2024
<b>Ai2's Molmo shows open source can meet, and beat, closed multimodal models.</b>	
<a href="#">Penn Engineering &amp; Tech Xplore &amp; Tech Times &amp; Tech Briefs</a>	2024
<b>Penn Engineers Recreate Star Trek's Holodeck Using ChatGPT and Video Game Assets.</b>	

## AWARDS

---

SAC Highlights Award, ACL	2025
AWS-ASSET Fellowship, Amazon	2025
Best Paper Honorable Mention Award, CVPR	2025
Outstanding Intern of the Year Award, Ai2	2023
Outstanding Teaching Award, University of Pennsylvania	2020