Vardaan Pahuja

pahuja.9@osu.edu | vardaanpahuja.github.io

Summary

Final-year Ph.D. researcher working on **agentic and multimodal AI for real-world workflows**: combining **vision-language modeling with structured retrieval and graph-based reasoning**. Experienced in **end-to-end pipelines** (data, training/fine-tuning, evaluation) and **computer-use agents**.

Education

The Ohio State University Aug 2019 – Present
Ph.D. in Computer Science — GPA: 4.0/4.0

Université de Montréal Sept 2017 – Jul 2019
M.Sc. in Computer Science — GPA: 4.3/4.3

IIT Kharagpur Jul 2012 – Jul 2016
B.Tech (Hons.), E&ECE + CSE Minor — GPA: 9.61/10 (Rank 1)

- Awarded **Institute Silver Medal 2016** for best academic performance in department at graduation.
- Awarded Nilanjan Ganguly Memorial Award for **Best Bachelor's Thesis** in department.

Qualifications

- **Agents & LLMs:** LLM agents and Multimodal Foundation Models; experience building end-to-end pipelines (data, training/fine-tuning, and evaluation) for computer-use agents; Retrieval-augmented Generation.
- **Multimodal Representation Learning:** Large-scale pretraining and representation learning for data from heterogeneous sources including unstructured text, KGs, images, and videos.
- **ML/NLP Foundations:** Machine Learning, Representation Learning, Knowledge Graphs, KG Reasoning, Dialog Systems.
- **Frameworks/Infra:** PyTorch, Hugging Face, DeepSpeed, AWS, Linux, Git.
- **CS Foundations:** Algorithms, Data Mining, Computational Linguistics, Object Oriented System Design.
- **Programming:** Python, C/C++, Java.

Experience

Research Intern May 2024 – Aug 2024
Microsoft Research *Redmond, WA*

• Led development of an exploration-driven, multi-agent web trajectory synthesis pipeline to generate a dataset of 94K trajectories with diverse, grounded tasks and verified multi-step demonstrations on live websites.

• Drove cost-efficient synthesis at $0.28 per successful trajectory, and scaled generation to 50 hours with 60 parallel processes, enabling large-scale dataset creation without human annotation bottlenecks.

• Fine-tuned small multimodal models into end-to-end web agents and demonstrated strong gains on realistic web agent benchmarks (**ACL Findings'25**).

Ph.D. Student Researcher May 2022 – Aug 2022
Google Research *Mountain View, CA*

• Proposed diversified joint vision–language tokenization to learn disentangled cross-modal tokens (encouraging non-overlapping token attention maps), improving robustness and generalization for VQA/VideoQA.

• Integrated the diversity objective into an iterative co-tokenization architecture for video–text representation learning, yielding more informative and less redundant token sets for downstream Transformer reasoning (**T4V@CVPR'23**).

Neuro-Symbolic AI Intern May 2021 – Aug 2021
Bosch Center for AI *Pittsburgh, PA*

• Implemented and adapted transformer + pointer-network architectures for QA-based information extraction from long documents.

Research Software Engineer July 2016 – July 2017
IBM Research *Bengaluru, India*

• Created a large-scale conversational QA dataset over Wikidata, built via a semi-automatic pipeline with 200K dialogs / 1.6M turns featuring logical, quantitative, and comparative reasoning over KG subgraphs (**AAAI'18**.)

• Proposed Corr-BiRNN, a joint learning framework for correlated sequence labeling (punctuation + capitalization) for ASR transcripts, improving both tasks without relying on prosodic features (**Interspeech'17**.)

Publications

Automatic Image-Level Morphological Trait Annotation for Organismal Images. V. Pahuja, S. Stevens, A. East, S. Record, Y. Su. **ICLR 2026**.

Explorer: Scaling Exploration-driven Web Trajectory Synthesis for Multimodal Web Agents. V. Pahuja*, Y. Lu*, C. Rosset, B. Gou, A. Mitra, S. Whitehead, Y. Su, A. Awadallah. **ACL 2025 (Findings)**.

Reviving the Context: Camera Trap Species Classification as Link Prediction on Multimodal Knowledge Graphs. V. Pahuja, W. Luo, Y. Gu, C. Tu, H. Chen, T. Berger-Wolf, C. Stewart, S. Gao, W. Chao, Y. Su. **CIKM 2024**.

A Retrieve-and-Read Framework for Knowledge Graph Link Prediction. V. Pahuja, B. Wang, H. Latapie, J. Srinivasa, Y. Su. **CIKM 2023**.

Diversifying Joint Vision-Language Tokenization Learning. V. Pahuja, AJ Piergiovanni, A. Angelova. **Transformers for Vision Workshop, CVPR 2023**.

A Systematic Investigation of KB-Text Embedding Alignment at Scale. V. Pahuja, Y. Gu, W. Chen, M. Bahrami, L. Liu, W. Chen, Y. Su. **ACL 2021**.

Complex Sequential Question Answering: Towards Learning to Converse Over Linked Question Answer Pairs with a Knowledge Graph. A. Saha*, V. Pahuja*, M. Khapra, K. Sankaranarayanan, S. Chandar. **AAAI 2018**.

Joint Learning of Correlated Sequence Labeling Tasks Using Bidirectional Recurrent Neural Networks. V. Pahuja*, A. Laha*, S. Mirkin, V. Raykar, L. Kotlerman, G. Lev. **Interspeech 2017**.

Knowledge Base Question Answering: A Semantic Parsing Perspective. Yu Gu, V. Pahuja, Gong Cheng, Yu Su. **AKBC 2022**.

Holistic Transfer: Towards Non-Disruptive Fine-Tuning with Partial Target Data. C. Tu, H. Chen, J. Zhong, Z. Mai, V. Pahuja, T. Berger-Wolf, S. Gao, C. Stewart, Y. Su, W. Chao. **NeurIPS 2023**.

Fine-Tuning is Fine, if Calibrated. Z. Mai*, A. Chowdhury*, P. Zhang*, C. Tu, H. Chen, V. Pahuja, T. Berger-Wolf, S. Gao, C. Stewart, Y. Su, W. Chao. **NeurIPS 2024**.

Structure Learning for Neural Module Networks. V. Pahuja, J. Fu, S. Chandar, C. J. Pal. **LANTERN Workshop, EMNLP 2019**.

Learning Sparse Mixture of Experts for Visual Question Answering. V. Pahuja, J. Fu, C. J. Pal. **Visual Question Answering and Dialog Workshop, CVPR 2019**.

Tooling framework for instantiating natural language querying system. M. Jammi, J. Sen, A. Mittal, S. Verma, V. Pahuja, R. Ananthanarayanan, P. Lohia, H. Karanam, D. Saha, K. Sankaranarayanan. **VLDB Endowment 2018**.

Towards a Robust and Generalizable Embodied Agent. C. Song*, J. Wu*, J. Byun, Z. Xu, V. Pahuja, G. Bajaj, S. Stevens, Z. Chen, Y. Su. **Proceedings of Alexa Prize SimBot Challenge 2023**.

Awards

Honorable Mention Award for Poster, OSU CSE Graduate Student Research Poster Exhibition (2024).

Prof. J.C. Ghosh Memorial Prize, IIT Kharagpur — best academic performance (VI semester) (2015).

Class of 1970 Alumni (US) Association Prize, IIT Kharagpur — best academic performance in the institute (IV semester) (2014).

IIT Kharagpur Alumni (California Chapter) Award, IIT Kharagpur — best academic performance in the institute (IV semester) (2014).

Service

Area Chair: ARR/EMNLP 2025.

Program Committee: ICLR'26, AAAI'26, NeurIPS'25, AAAI'25, ICCV'25, ACL'25, NAACL'25, CVPR'25, WACV'25, COLM'24, CVPR'24, EMNLP'23, ACL'23, NAACL'22, Transactions on Big Data'24.