ABHAY ZALA

aszala.com | Google Scholar | GitHub | aszala@cs.unc.edu | zala.abhay@gmail.com

EDUCATION

MS University of North Carolina at Chapel Hill — May 2022 – May 2024
Major: Computer Science (Multimodal AI)

BS University of North Carolina at Chapel Hill — Aug 2019 – May 2022
Major: Computer Science

Outstanding Undergraduate Researcher Award**,** Honorable Mention
Computing Research Association (CRA)

EXPERIENCE

University of North Carolina at Chapel Hill — Sep 2019 – Present
Research Assistant, Supervised by Prof. Mohit Bansal

- Published several first and second author research papers and presented at prestigious conferences such as NeurIPS, CVPR, ICCV, NAACL, AAAI, and EMNLP
- Expertise in text-to-image generation and interpretable evaluation, text-to-video generation, visual programming with large language models, visual commonsense reasoning, vision-language navigation, image/video captioning, 3D pose understanding, PyTorch deep learning framework
- Serving as an organizer for the ACL Year-Round Mentorship program
- Mentor to undergraduate research students

EngageAI — Mar 2023 – Present
Researcher

- Leading a team of academic researchers on video processing and understanding for classroom settings
- Presented as an invited speaker to large audience on the topics of research in practice, deep learning, and large language models

Meta AI Research Collaboration — Feb 2022 – Mar 2023
Primary Researcher

- Lead project development on the video/moment retrieval, segmentation, and captioning in collaboration with researchers from Meta AI
- Published and presented a research paper at CVPR 2023

Capital One — Jun 2022 – Aug 2022
Machine Learning Intern

- Developed internal document retrieval application
- Demonstrated expertise in the document search, vector database development, and feature extraction (e.g., TF-IDF)

PUBLICATIONS

* Indicates equal contribution

9. Abhay Zala*, Jaemin Cho*, Han Lin, Jaehong Yoon, Mohit Bansal. **EnvGen: Generating and Adapting Environments via LLMs for Training Embodied Agents**. [preprint]

8. Abhay Zala, Han Lin, Jaemin Cho, Mohit Bansal. **DiagrammerGPT: Generating Open-Domain, Open-Platform Diagrams via LLM Planning**. [preprint]

7. Han Lin, Abhay Zala, Jaemin Cho, Mohit Bansal. **VideoDirectorGPT: Consistent Multi-scene Video Generation via LLM-Guided Planning**. [preprint]

6. Jaemin Cho, Abhay Zala, Mohit Bansal. **Visual Programming for Text-to-Image Generation and Evaluation**. Proceedings of NeurIPS 2023 [pdf]

5. Abhay Zala*, Jaemin Cho*, Satwik Kottur, Xilun Chen, Barlas Oğuz, Yasher Mehdad, Mohit Bansal. **Hierarchical Video-Moment Retrieval and Step-Captioning**. Proceedings of CVPR 2023 [pdf]

4. Jaemin Cho, Abhay Zala, Mohit Bansal. **DALL-Eval: Probing the Reasoning Skills and Social Biases of Text-to-Image Generative Models**. Proceedings of ICCV 2023 [pdf]

3. Abhay Zala*, Hyounghun Kim*, Mohit Bansal. **CoSIm: Commonsense Reasoning for Counterfactual Scene Imagination**. Proceedings of NAACL 2022 [pdf]

2. Abhay Zala*, Hyounghun Kim*, Graham Burri, Mohit Bansal. **FixMyPose: Pose Correctional Captioning and Retrieval**. Proceedings of AAAI 2021 [pdf]

1. Hyounghun Kim, Abhay Zala, Graham Burri, Hao Tan, Mohit Bansal. **ArraMon: A Joint Navigation-Assembly Instruction Interpretation Task in Dynamic Environments**. Findings of EMNLP 2020 [pdf]

TALKS, PRESENTATIONS, AND WORKSHOPS

2023 NeurIPS Paper Presentation, "Visual Programming for Text-to-Image Generation and Evaluation"

2023 EngageAI Panel, "Research in Practice"

2023 EngageAI Talk, "Deep Learning and Large Language Models"

2023 CVPR Paper Presentation, "Hierarchical Video-Moment Retrieval and Step-Captioning"

2023 EngageAI Research Presentation, "Video Moment-Retrieval and Moment-Captioning on Classroom Videos"

2022 NAACL Paper Presentation, "CoSIm: Commonsense Reasoning for Counterfactual Scene Imagination"

2021 AAAI Paper Presentation, "FixMyPose: Pose Correctional Captioning and Retrieval"

2020 SpLU Workshop, "ArraMon: A Joint Navigation-Assembly Instruction Interpretation Task in Dynamic Environments"

SKILLS

Programming Languages: Python, Java, C#, C/C++, Git, JavaScript, OpenGL, Svelte, NodeJS, LaTeX, HTML/CSS, PHP, SQL, MATLAB

Deep Learning Frameworks: PyTorch

Platforms: Unity Engine, Amazon Mechanical Turk, Docker, Kubernetes, GitHub, Amazon Web Services, Google Cloud, IBM Cloud, Linux, Windows, MacOS, Google Firebase Adobe Suite, Autodesk Suite, Microsoft Office Suite, Overleaf

Applications: Machine Learning, Natural Language Processing, Computer Vision, Robotics, Dataset Creation, Simulator Development, Website Development, Software Development, Database Management, VR Development, Graphics Rendering

OTHER

- Red team member for OpenAI's DALL-E 2 [information]
- Site Developer/Maintainer of nlp.cs.unc.edu, murgelab.cs.unc.edu, perfect-type.com
- ACL Year-Round Mentorship Logo Designer
- Developed Social Media Platform for Networking
- Developed Online Peer Tutoring Service with AI Assistance
- Developed TV Show and Movie Showcase Website
- Developed VR Healthcare Training System
- Developed AI assistant for presentations
- Winner of several hackathons