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 – Present
Major: Computer Science (Multimodal AI)

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

EXPERIENCE

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

- 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

EngageAI — Mar 2023 – Present
Researcher

- Video Processing and Understanding

ACL Year-Round Mentorship — Oct 2021 – Present
Organizer

- Event Organizer and Logo Designer

MetaAI Research Collaboration — Feb 2022 – Mar 2023
Researcher

- Video/Moment Retrieval, Segmentation, and Captioning
- CVPR 2023 Research Paper

Capital One — Jun 2022 – Aug 2022
Machine Learning Intern

- Document Search
- Vector Database Development

Éclat Engineering — Dec 2021 – Jan 2022
Software Developer Intern

- Systems Development

PUBLICATIONS

* Indicates equal contribution

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

2. Hyounghun Kim*, Abhay Zala*, 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 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"

HONORS AND AWARDS

Outstanding Undergraduate Researcher Award, Honorable Mention 2022
Computing Research Association (CRA)

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

RELEVANT COURSEWORK

Java Programming, C Programming, Assembly Programming and Hardware Design, Data Structures, Models of Language and Computation, 2D Computer Graphics, Algorithms and Analysis, Modern Web Programming, Data Science, Programming Language Concepts, Compilers, Computational Photography, Machine Learning, Robotics

Linear Algebra, Calculus (I – III), Mechanics and Relativity, Discrete Mathematics, Probability, Data Models and Inference

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
- 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
- ACL Year-Round Mentorship Logo Designer