

# ABHAY ZALA

[aszala.com](http://aszala.com) | [Google Scholar](https://scholar.google.com/citations?user=aszala) | [GitHub](https://github.com/aszala) | [aszala@cs.unc.edu](mailto:aszala@cs.unc.edu) | [zala.abhay@gmail.com](mailto:zala.abhay@gmail.com)

## EDUCATION

---

<b>MS</b>	University of North Carolina at Chapel Hill Major: Computer Science (Multimodal AI)	May 2022 – Present
<b>BS</b>	University of North Carolina at Chapel Hill Major: Computer Science	Aug 2019 – May 2022

## EXPERIENCE

---

<b>University of North Carolina at Chapel Hill</b> Research Assistant, Supervised by Prof. Mohit Bansal <ul style="list-style-type: none"><li>Text-to-Image Model Exploitability and Interpretability</li><li>Visual Programming with Large Language Models</li><li>Visual Commonsense Reasoning</li><li>Vision-Language Navigation</li><li>Image/Video Captioning</li><li>3D Pose Understanding</li></ul>	Sep 2019 – Present
<b>EngageAI</b> Researcher <ul style="list-style-type: none"><li>Video Processing and Understanding</li></ul>	Mar 2023 – Present
<b>ACL Year-Round Mentorship</b> Organizer <ul style="list-style-type: none"><li>Event Organizer and Logo Designer</li></ul>	Oct 2021 – Present
<b>MetaAI Research Collaboration</b> Researcher <ul style="list-style-type: none"><li>Video/Moment Retrieval, Segmentation, and Captioning</li><li>CVPR 2023 Research Paper</li></ul>	Feb 2022 – Mar 2023
<b>Capital One</b> Machine Learning Intern <ul style="list-style-type: none"><li>Document Search</li><li>Vector Database Development</li></ul>	Jun 2022 – Aug 2022
<b>Éclat Engineering</b> Software Developer Intern <ul style="list-style-type: none"><li>Systems Development</li></ul>	Dec 2021 – Jan 2022

## PUBLICATIONS

---

\* Indicates equal contribution

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 AACL 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

---

**Panel**, “Research in Practice”  
EngageAI, 2023

**Talk**, “Deep Learning and Large Language Models”  
EngageAI, 2023

**Paper Presentation**, “Hierarchical Video-Moment Retrieval and Step-Captioning”  
CVPR, 2023.

**Research Presentation**, “Video Moment-Retrieval and Moment-Captioning on Classroom Videos”  
EngageAI SV, 2023.

**Paper Presentation**, “CoSim: Commonsense Reasoning for Counterfactual Scene Imagination”

NAACL, 2022.

**Paper Presentation**, “FixMyPose: Pose Correctional Captioning and Retrieval”  
AAAI, 2021.

**Workshop**, “ArraMon: A Joint Navigation-Assembly Instruction Interpretation Task in  
Dynamic Environments”  
SpLU, 2020.

---

## HONORS AND AWARDS

<b>Outstanding Undergraduate Researcher Award</b> , Honorable Mention Computing Research Association (CRA)	2022
---	------

---

## 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