

Education

- 2024-Present **PhD CSE at The Ohio State University, GPA:4/4.**
PhD in Computer Science
◦ *Coursework:* Data Visualization, Advanced NLP.
- 2022-2024 **MS CIS at University of Pennsylvania, GPA: 3.90/4.**
Masters in Computer and Information Science
◦ *Coursework:* Artificial Intelligence, Natural Language Processing, Machine Learning, Analysis of Algorithms, Database Management Systems, Computer Vision, Network Systems.
- 2018-2022 **B.Tech CS at Veermata Jijabai Technological Institute, GPA: 9/10.**
Bachelors in Computer Science
◦ *Coursework:* Operating System, Machine Learning, Database Management Systems, Data Interpretation and Analysis, Web Technologies, Software Engineering, Cloud Computing, Wireless Networks, Network Security.

Publications and Presentations

- NeurIPS 2024 **A Textbook Remedy for Domain Shifts: Knowledge Priors for Medical Image Analysis,** Yue Yang, **Mona Gandhi**, Yufei Wang, Yifan Wu, Michael S. Yao, Chris Callison-Burch, James C. Gee, Mark Yatskar.
The Annual Conference on Neural Information Processing Systems, 2024, **spotlight**
Received the **Best Paper Award** at AIM-FM 2024 Workshop
- CVPR 2023 **CREPE: Can Vision-Language Foundation Models Reason Compositionally?,** Zixian Ma*, Jerry Hong*, Mustafa Omer Gul*, **Mona Gandhi**, Irena Gao, Ranjay Krishna.
IEEE Conference on Computer Vision and Pattern Recognition, 2023, **highlight**
- CVPR 2022 **Measuring Compositional Consistency for Video Question Answering,** **Mona Gandhi***, Mustafa Omer Gul*, Eva Prakash, Madeleine Grunde-McLaughlin, Ranjay Krishna, Maneesh Agrawala.
IEEE Conference on Computer Vision and Pattern Recognition, 2022

Selected Research Experience

- 2025-Present **Design Evaluation and Editing, Adobe Research.**
◦ *Mentor:* Dr. Sayan Nag
◦ *Domain:* Computer Vision and Natural Language Processing
◦ *Work:* Developed a design-aware coherence evaluation framework by creating large-scale poster datasets and training multimodal embedding models. Built an iterative editing pipeline using generative models with beam-search-based reversible edits and automated design quality metrics.
- 2024-Present **Skill Composition Evaluation, The Ohio State University.**
◦ *Mentor:* Prof. Srinivasan Parthasarathy
◦ *Domain:* Computer Vision and Natural Language Processing
◦ *Work:* Researched skill composition in multimodal reasoning, decomposing tasks into sub-skills (object, attribute, relation identification) and evaluating how different skill combinations affect performance on challenging negatives.
- 2023-2024 **Interpretable Radiology, University of Pennsylvania.**
◦ *Mentor:* Prof. Mark Yatskar
◦ *Domain:* Computer Vision and Natural Language Processing
◦ *Work:* Building an interpretable concept bottleneck model inspired by LaBo for the medical domain, specifically chest X-rays, for assisting radiologists in analysing them with natural language explanations in the form of concepts supporting the final model output.

2021-2022 **Question Decomposition**, *Stanford University*.

- *Mentors*: Prof. Maneesh Agrawala, Prof. Ranjay Krishna
- *Domain*: Computer Vision and Natural Language Processing
- *Publication*: IEEE CVPR 2022
- *Work*: Developed a question decomposition engine that deconstructs a compositional question from AGQA into a subquestion hierarchy, a directed acyclic graph. Evaluated the performance of HCRN and HME models using consistency metrics on subquestion hierarchies for the AGQA dataset and learnt that easier subquestions are incorrect while complex questions are answered correctly.

Peer Review Service

KDD 2025, ACL 2025

Selected Teaching Experience

Graduate Teaching Assistant

2024 **CIS 4000/4010: Senior Design Project.**

2024 Spring: instructed by Prof. Boon Thau Loo

2023 Fall: instructed by Prof. Boon Thau Loo

- Mentoring and managing five teams of undergraduate students for their senior design project, helping them plan their weekly goals and giving feedback and suggestions.

2023 **CIS 5300: Natural Language Processing.**

2023 Fall: instructed by Prof. Mark Yatskar

Technical Skills

Languages Python, C++, C, Java, SQL

DL Tools Pytorch, Tensorflow, OpenCV, Hugging Face

Web-Dev HTML, CSS, Javascript, PHP, React JS, NodeJS, Oracle DB, Mongo DB

Others Linux, Visual Studio, Git

Achievements

2024 Received the Outstanding Teaching Award at University of Pennsylvania.

2018 1st rank among girls and 10th rank overall in MHT-CET (state-level entrance exam) among 280 thousand students.

2008-2020 Pursued Visharath Pratham (highest level of degree in Indian Classical Dance) in Bharatnatyam (Indian Classical Dance) and performed Araangetram (first stage performance).

Selected Projects

2024 **Examining the Reversal Curse on GPT models.**

- *Technologies*: LLMs
- *Description*: The Reversal Curse paper ([link](#)) highlights a simple task that these models fail at. If the model has seen "A is B", it is not guaranteed that the model can generalize "B is A" - this is coined as Reversal Curse in the paper. In addition to replicating the results from the paper, we investigate the model on a verification task, where the model is asked a yes-no question. The model struggles to respond to these questions and even contradicts itself within the same response.

2023 **AutoArt, Neural Style Transfer.**

- *Technologies*: Computer Vision, Deep Learning
- *Course*: Computer Vision, CIS 5810
- *Description*: Implemented two novel methods to improve our outputs for neural style transfer: (i) fine-tuning the model as a classification for a particular style, and (ii) flattening the layers to allow us the convenience of adding style and content loss inside the blocks. Concluded that mobilenetv2 with flattening with a fine-tuned model gave the best visual results.

2022 **Could this be any better?**, *Multimodal Sarcasm Detector*.

- *Domain*: Computer Vision and Natural Language Processing
- *Course*: Natural Language Processing, CIS 5300
- *Description*: Implemented a multimodal sarcasm detector using video, audio, and text features from the MUSTARD dataset. Trained and analysed the performance of LSTMs with different types of attention. Concluded that the best-performing model learns a bias towards labelling data as sarcastic, but it does very well in detecting non-sarcastic data.