

Yash Jain

Email | [Linkedin](#) | [Webpage](#) | [Google Scholar](#)

EMPLOYMENT

Microsoft Office AI

ML Scientist II in Office AI Science Team

June 2023 – Present

EDUCATION

Georgia Institute of Technology

Masters in Computer Science (Thesis advisor: Prof. Zolt Kira)

USA

2021 – 2023

Indian Institute of Technology Bombay

Bachelors of Technology in Computer Science (Thesis advisor: Prof. Soumen Chakrabarti)

India

2017 – 2021

RECENT PUBLICATIONS

* equal contribution

Local Prompt Optimization [\[PDF\]](#)

[Yash Jain](#), Vishal Chowdhary

[\[Oral Presentation\]](#) NAACL (Main Conference) 2025

PEEKABOO: Interactive Video Generation via Masked-Diffusion [\[PDF\]](#) [\[Code\]](#)

[Yash Jain](#)*, Anshul Nasery*, Vibhav Vineet, Harkirat Behl

[\[Invited Talk\]](#) IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2024

DAMEX: Dataset-aware Mixture-of-Experts for visual understanding of mixture-of-datasets [\[PDF\]](#) [\[Code\]](#)

[Yash Jain](#), Harkirat Behl, Zolt Kira, Vibhav Vineet

Advances in Neural Information Processing Systems (NeurIPS) 2023

Multi-Stage Multi-Modal Pre-Training for Automatic Speech Recognition [\[PDF\]](#)

[Yash Jain](#), D. Chan, P. Dheram, A. Khare, O. Shonibare, V. Ravichandran, Shalini Ghosh

Joint Int. Conf. on Computational Linguistics, Language Resources and Evaluation (LREC-COLING) 2024

RESEARCH EXPERIENCE

ML Scientist II at Microsoft

Summer 2023 - Present

- Training a next generation of instruct-tuned LLM based on Discrete Diffusion modeling.
- Lead scientist for shipping voice-based Copilot for next Microsoft release.
- Created an automatic prompt optimization method that improves production prompt across Microsoft internally.
- Generate synthetic data for PowerPoint and trained SLMs to replace GPT-4 in production.

Applied Scientist Intern at Amazon Alexa [\[paper\]](#)

Fall 2022

Guide: Shalini Ghosh

Amazon Alexa AI, US

- Topic: Large-scale AI training, Multi-modal learning (Speech and Vision), Speech Recognition
- Led the development and implementation of a novel ML algorithm that improves speech recognition accuracy by 38.45% compared to existing state-of-the-art, using videos as training data. Scaled training on 4000 GPUs.

Graduate Thesis at Georgia Tech [\[thesis\]](#)

Spring 2023

Guide: Vibhav Vineet, Zolt Kira

Microsoft Research & Georgia Tech

- Topic: Mixture-of-Experts, Object-Detector, Representation Learning
- Proposed MoE as an alternate ensembling strategy for mixing datasets in Object-Detection task.

Applied Scientist Intern at Microsoft [\[paper\]](#)

Summer 2022

Guide: Vibhav Vineet, Michael Bentley

Microsoft, US

U.S. Patent approved

- Topic: Multi-modal learning (Text and Vision), Self-Supervised learning
- Developed a novel pipeline of image difference captioning task for PowerPoint slide data by generating a synthetic dataset in a self-supervised manner, benefiting 4.4 million users in the US.

Research Scientist at Nokia Bell Labs [\[paper\]](#) [\[poster\]](#)

Summer 2021

Guide: Akhil Mathur

Nokia Bell Labs, UK

- Topic: Sensor (IMU) data training, Contrastive Learning, Self-Supervised learning
- Developed a collaborative ML algorithm that can utilize data from multiple wearable devices and improve activity detection by 7.9% F-1 score, potentially improving fitness and wellness monitoring of smartwatches

AWARDS

- **Undergraduate Research Award** for outstanding Bachelors Thesis at IITB 2021
- Recipient of **Dhirubai Ambani Foundation scholarship** for pursuing Masters studies at Georgia Tech 2021
- **All India Rank 29** in **JEE-Advanced** from 1.2 million students appearing for the engineering entrance exam 2017
- **Gold Medalist** in theory and **Silver Medalist** in practicals representing India at 11th International Junior Science Olympiad held in Mendoza, Argentina 2014

OTHER PUBLICATIONS

* equal contribution

Test-time Prompt Refinement for Text-to-Image models [\[PDF\]](#)

[Yash Jain](#) and others

Under submission at a conference 2025

GeoMeter: Probing Depth and Height Perception of Large Visual-Language Models [\[PDF\]](#)

Shehreen Azad, [Yash Jain](#), Rishit Garg, Yogesh S Rawat, Vibhav Vineet

Under submission at IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**) 2025

RiTtA: Modeling Event Relations in Text-to-Audio Generation [\[PDF\]](#)

Yuhang He, [Yash Jain](#), Xubo Liu, Andrew Markham, Vibhav Vineet

Under submission at ACL (2025)

PLUM: Improving Inference Efficiency By Leveraging Repetition-Sparsity Trade-Off [\[PDF\]](#)

Sachit Kuhar, [Yash Jain](#), Alexey Tumanov

Transactions on Machine Learning Research (**TMLR**) 2025

On the Utility of Virtual On-body Acceleration Data for Fine-grained Human Activity Recognition [\[PDF\]](#)

Zikang Leng, [Yash Jain](#), Hyeokhyen Kwon, Thomas Ploetz

ACM International Symposium on Wearable Computers (**ISWC**) 2023

Collossl: Collaborative self-supervised learning for human activity recognition [\[PDF\]](#) [\[Code\]](#)

[Yash Jain](#)*, Chi Ian Tang*, Chulhong Min, Fahim Kawsar, Akhil Mathur

ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (**UbiComp**) 2022

On the Effectiveness of Virtual IMU Data for Eating Detection with Wrist Sensors [\[PDF\]](#)

[Yash Jain](#), Hyeokhyen Kwon, Thomas Ploetz

ACM International Symposium on Wearable Computers (**ISWC**) 2022

Integrating transductive and inductive embeddings improves link prediction accuracy [\[PDF\]](#)

[Yash Jain](#)*, Chitrang Gupta*, Abir De, Soumen Chakrabarti

ACM International Conference on Information & Knowledge Management (**CIKM**) 2022

Group Supervised Learning: Extending Self-Supervised Learning to Multi-Device Settings [\[PDF\]](#)

[Yash Jain](#)*, Chi Ian Tang*, Chulhong Min, Fahim Kawsar, Akhil Mathur

Workshop on Self-Supervised Learning for Reasoning and Perception at **ICML** 2021

Rfid tattoo: A wireless platform for speech recognition [\[PDF\]](#)

Jingxian Wang, C. Pan, H. Jin, V. Singh, [Yash Jain](#), Jason I Hong, Carmel Majidi, Swarun Kumar

[\[Best Paper Award\]](#) ACM Interactive, Mobile, Wearable and Ubiquitous Technologies (**UbiComp**) 2020

TEACHING EXPERIENCE

Graduate Teaching Assistant

Georgia Tech

Primary Instructor: Prof. Zsolt Kira (GaTech)

Jan'23-May'23

- Course: CS4803/7643 Deep Learning; My Rating: 4.9/5.

Primary Instructor: Prof. Gerandy Brito (GaTech)

Jan'22-May'22

- Course: CS6515 Graduate Algorithms; My Rating: 4.8/5.