

Kanishk Jain

📱 mobile | @ email | 🔗 linkedin | 🐙 github | 🌐 webpage

EDUCATION

International Institute of Information Technology

MS by Research in CSE; **GPA: 9.33/10**

Hyderabad, India

Aug 2021 – Nov 2022

International Institute of Information Technology

B.Tech (Honors) in ECE; **GPA: 6.73/10**

Hyderabad, India

Aug 2013 – Jun 2017

RESEARCH INTERESTS

I am interested in the following research topics: learning from multiple data modalities, language understanding in autonomous systems during navigation, explainable deep learning, multi-object tracking, improving robustness to domain shifts and adversarial attacks, learning in low-data regimes, and ensemble learning.

RESEARCH PUBLICATIONS

Test-Time Amendment with a Coarse Classifier for Fine-Grained Classification

Kanishk Jain, Shyamgopal Karthik, Vineet Gandhi

Under Review

- A post-hoc correction approach for reducing mistake severity in neural networks by leveraging coarse-grained predictions at test-time to improve the performance of fine-grained classification.

Ground then Navigate: Language-guided Navigation in Dynamic Scenes

Kanishk Jain*, Varun Chhangani*, Amogh Tiwari, K Madhava Krishna, Vineet Gandhi

ICRA 2023

- A novel visual-grounding based approach to language-guided navigation in dynamic outdoor environment.

Bringing Generalization to Deep Multi-View Pedestrian Detection

Jeet Vora, Swetanjali Dutta, **Kanishk Jain**, Shyamgopal Karthik, Vineet Gandhi

WACV-W 2023

- Designed a comprehensive evaluation framework along with a novel dataset to assess the generalization capabilities of existing MVD methods.

Comprehensive Multi-Modal Interactions for Referring Image Segmentation

Kanishk Jain, Vineet Gandhi

ACL Findings 2022

- Proposed a novel architecture for Referring Image Segmentation which captures comprehensive interactions between visual and linguistic modalities in a synchronous manner with effective multi-hierarchy aggregation.

Grounding Linguistic Commands to Navigable Regions

Kanishk Jain*, Nivedita Rufus*, Unni Krishnan*, Vineet Gandhi, K Madhava Krishna

IROS 2021

- Introduced the novel task of Referring Navigable Regions (RNR), i.e., grounding regions of interest on road for navigation based on the linguistic command.

WORK EXPERIENCE

Optra Sports

Co-Founder

Dec 2022 – Present

- Developed top-view player tracking toolkit for the game of cricket.
- Collaborating with Start Sports (a Disney sports broadcast company) and AE Live (a graphics company for sports broadcasting) for live deployment in future cricket tournaments.

CVIT, IIIT Hyderabad

Research Assistant

Sep 2019 – Present

- Worked under guidance of Dr. Vineet Gandhi on Visual Grounding and its application in different multi-modal problem setting.
- Collaborated with Dr. K. Madhava Krishna on Language-guided Navigation leading to a publication at IROS-2021
- Part of the super winner team of Qualcomm Innovation Fellowship (QIF) 2020.
- An analytics tool for CSGO games providing relevant insights into winning strategies.

Turvo

Software Engineer

Jul 2017 – Aug 2019

- Integrated Xero Accounting Platform with Turvo platform using Pub-Sub messaging pattern for handling different accounting scenarios.
- Added capability of Batch Payment Processing to allow users to schedule and process multiple payments at once.
- OCR over Document Images using active learning based template detection for extracting information from unstructured documents.
- Implemented New Feature Notification Modal for users.

SELECTED PROJECTS

Top-View Player Tracking: Player tracking solution in the Bird's Eye View, deployed live in 2022 Asia Cup.

Stereo SLAM: Generate 3D point clouds using stereo images and use 2D-3D correspondences to estimate motion/pose using iterative Perspective-from-n-Points (PnP) algorithm.

Pose Graph Optimization: Used Levenberg–Marquardt algorithm to optimize for robot's poses by applying Odometry and Loop Closure constraints for 1D & 2D SLAM.

Neural Nearest Neighbor Networks: Implemented the NeurIPS paper, "Neural Nearest Neighbor Networks" as part of course project.

Unity Game for Amblyopia (*Hons. Project*): Developed a Unity Game for diagnosis of Amblyopia. The game is played using eye gaze movements captured using an eye tracker.

Text to Braille Converter: A learning tool for people with no vision. Converts a given text to braille and audio.

Neuro Rehab Systems: A rehabilitation tool which aims to aid recovery from a nervous system injury and minimize any functional alterations resulting from it.

TECHNICAL STRENGTHS

Languages: Python, Java, C++, C#, Node JS, Javascript

Frameworks: PyTorch, Keras, Tensorflow, OpenCV, scikit-learn, scikit-image

Tools: CARLA, Unity 3D, Matlab, Spring Boot, Maven, REST

Databases: MySQL, Mongo DB, ElasticSearch, Apache Solr, Redis

RELEVANT COURSES

ML/AI Courses: Statistical Methods in AI, Computer Vision, Mobile Robotics, Topics in Optimization Methods, Topics in ML, Cognitive Science and AI

Core Science: Computer Programming, Operating Systems and Algorithms, Data Structures

Other Courses: Digital Image Processing, Digital Signal Processing, Linear Algebra, Probability and Random Processes, Discrete Mathematics

ACHIEVEMENTS

Qualcomm Innovation Fellowship: Led the super winner team of Qualcomm Innovation Fellowship (QIF) 2020 India.

JEE Mains: Ranked in National Top 0.2% (amongst 1,200,000 candidates) in JEE Mains.

JEE Advanced: Secured 4539 rank in JEE Advanced among 150,000 candidates in JEE Advanced.

R&D Showcase: Presented the Amblyopia Game at college's annual R&D showcase.