Kanishk Jain

mobile | email | linkedin | github | webpage

Education

Université de Montréal — Montréal, Canada
PhD in CSE; ***GPA: 4.3*** — *Sep 2023 – present*

International Institute of Information Technology — Hyderabad, India
MS by Research in CSE; ***GPA: 9.33/10*** — *Aug 2021 – Dec 2022*

International Institute of Information Technology — Hyderabad, India
B.Tech (Honors) in ECE; ***GPA: 6.73/10*** — *Aug 2013 – Jun 2017*

Research Publications

Test-Time Amendment with a Coarse Classifier for Fine-Grained Classification — Paper
Kanishk Jain*, Shyamgopal Karthik, Vineet Gandhi* — *NeurIPS 2023*

- Introduced a post-hoc correction strategy that leverages coarse-grained predictions at test-time to reduce mistake severity and improve the accuracy of fine-grained classification

Instance-Level Semantic Maps for Vision Language Navigation — Paper
Laksh Nanwani, Anmol Agarwal, ***Kanishk Jain****, et al.* — *ROMAN 2023*

- Designed instance-level semantic maps enabling robust language-directed navigation across diverse environments.

Ground then Navigate: Language-guided Navigation in Dynamic Scenes — Paper
Kanishk Jain**, Varun Chhangani*, Amogh Tiwari, K Madhava Krishna, Vineet Gandhi* — *ICRA 2023*

- Developed a novel visual-grounding approach for language-guided navigation in dynamic outdoor environments

Bringing Generalization to Deep Multi-View Pedestrian Detection — Paper
Jeet Vora, Swetanjal Dutta, ***Kanishk Jain****, Shyamgopal Karthik, Vineet Gandhi* — *WACV-W 2023*

- Formulated an innovative evaluation framework and dataset for benchmarking generalization in multi-view pedestrian detection systems.

Comprehensive Multi-Modal Interactions for Referring Image Segmentation — Paper
Kanishk Jain*, Vineet Gandhi* — *ACL Findings 2022*

- Proposed a novel architecture for Referring Image Segmentation, leveraging synchronous multi-modal interactions and hierarchical aggregation to enhance performance.

Grounding Linguistic Commands to Navigable Regions — Paper
Kanishk Jain**, Nivedita Rufus*, Unni Krishnan*, Vineet Gandhi, K Madhava Krishna* — *IROS 2021*

- Introduced the task of Referring Navigable Regions (RNR), which involves grounding navigable road regions based on linguistic commands to facilitate language-guided autonomous navigation.

Work Experience

Université de Montréal | Mila
PhD Student — *Sep 2023 – present*

- Developing methods to automatically discover failure modes in Vision Language Models using Large Language Models in adversarial attack setups.
- Analyzing state space models to improve the efficiency of text-to-image generation.
- Investigating cultural understanding in Vision Language Models across diverse cultural contexts.

CVIT, IIIT Hyderabad
Research Engineer — *Mar 2023 – Aug 2023*

- Conducted research on reducing mistake severity in fine-grained classification using Hierarchical Ensembles, demonstrating significant improvements in both fully supervised and semi-supervised settings.
- Developed and deployed a player tracking solution in Bird's Eye View for live use during the 2022 Asia Cup.

CVIT, IIIT Hyderabad
Research Assistant — *Sep 2019 – Dec 2022*

- Conducted research on language-guided autonomous navigation, focusing on explicit grounding of navigable regions and the integration of visual feedback.
- Developed multi-modal interaction techniques for referring image segmentation, enhancing the accuracy and efficiency of visual-linguistic model interactions.
- Developed an analytics tool for CSGO games, providing 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.

Selected Projects

Top-View Player Tracking: Developed a player tracking solution in Bird's Eye View, deployed live during the 2022 Asia Cup Cricket Tournament.

Stereo SLAM: Generated 3D point clouds using stereo images and estimated motion/pose through 2D-3D correspondences with the iterative Perspective-n-Point (PnP) algorithm.

Pose Graph Optimization: Applied the Levenberg–Marquardt algorithm to optimize robot poses using Odometry and Loop Closure constraints for 1D and 2D SLAM.

Unity Game for Amblyopia: Developed a Unity game for diagnosing Amblyopia, controlled using eye gaze movements captured by an eye tracker.

Neuro Rehab Systems: Created a rehabilitation tool aimed at aiding recovery from nervous system injuries and minimizing functional alterations.

Technical Strengths

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

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

Tools: CARLA, TensorRT, Open3D, 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

Academic Service

Reviewer: **ACL 2024**, **ECCV 2024**, **WACV 2024**, **ICRA 2024**

Achievements

Morocco Solidarity Hackathon: Member of the winning team; developed a solution to predict human trafficking behavior from social media posts.

Qualcomm Innovation Fellowship: Led the winning team of the Qualcomm Innovation Fellowship (QIF) 2020 in India.

JEE Mains: Ranked in the top 0.2% nationally among 1,200,000 candidates in JEE Mains.

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

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

References

Dr. Vineet Gandhi, Dr. K. Madhava Krishna