

Utsav Nandi

linkedin.com/in/utsavnandi

Email : utsav.nandi@gmail.com

Mobile : +1-857-379-4804

EDUCATION

- **Northeastern University** Boston, MA
Master of Science in Computer Science Jan. 2023 – May. 2025

EXPERIENCE

- **ABB Robotics** San Jose, CA, USA
Computer Vision AI Engineer Intern Jan 2026 – Present
 - **Item picking pipeline:** Added user configurable image resolution and 3D Zivid camera support for the inference pipeline of ABB Robotics item picking solution. The computed 6D grasp location remains within tolerance range across multiple supported cameras and image resolutions.
 - **VLA Implementation:** Established a reusable VLA stack for a 7-DoF ABB YuMi, combining RAPID/ABB EGM robot control with a Python-LeRobot interface to support real-robot data collection, policy fine-tuning, and real-time chunked inference of fine-tuned models on hardware.
- **rStream Recycling** Somerville, MA, USA
Computer Vision Engineer Co-op Jan 2024 – Aug 2024
 - **Recyclable object detection and recognition:** Fine-tuned very large CLIP-like vision encoders using LoRA, integrated with open-vocab detection and SAM to achieve 67%+ recall for recyclable classification and localization on a highly imbalanced and diverse dataset; cut annotation time by 60% via semi-supervised labelling.
 - **Multimodal LLM Evaluation:** Benchmarked zero- and few-shot performance of LLMs like GPT4V, Gemini Pro, Claude Sonnet, etc. for image classification and localization on real world recycling dataset; evaluated tradeoffs against fine-tuned custom models.
 - **Annotation Data Store:** Designed and deployed a robust pipeline and store for ingesting and processing 100K+ annotation data from diverse sources into a PostgreSQL/pgvector database, creating the foundational image data catalog for all computer vision model development.
- **Siemens** Bengaluru, India
Senior Software Engineer - Machine Learning Jul 2018 – Dec 2022
 - **Perception System for Smart City:** Finetuned and deployed SOTA detection models achieving 0.74 mAP (\uparrow 37% vs. baseline) for real-time dust detection on 10K+ drone images; trained with PyTorch Distributed, quantized via ONNX Runtime, and deployed to a 2-node Kubernetes cluster processing 10 K images/day.
 - **Pose Estimation and Tracking:** Developed perception and tracking system combining vision + LIDAR achieving 97% accuracy for automated operator compliance system; reduced GSK facility compliance violations by 40%, securing contract extension through cross-team collaboration.
 - **NextWork HR Platform:** Built REST backend APIs with IAM serving 10K+ employees using Spring Boot and PostgreSQL; achieved sub-100ms response times and 99.5% uptime, deployed with AWS EKS. Project successfully serving global Siemens HR department.
 - **OCR System for Manufacturing:** Developed an industrial OCR system using a CRAFT + ResNet ensemble and OpenCV for embossed text reading, automating a critical compliance task and reduced manual effort; engineered pre- and post-processing for edge detection, achieving 97% accuracy. Deployed model via FastAPI.

PUBLICATION

- X. Bai, H. Liang, B. Galoaa, **U. Nandi**, et al. “MoReGen: Multi-Agent Motion Reasoning Engine for Code-based Text-to-Video Synthesis” to appear at **CVPR 2026**.
- B. Galoaa, X. Bai, **U. Nandi**, et al. “UniTrack: Differentiable Graph Representation Learning for Multi-Object Tracking” to appear at **ICLR 2026**.
- B. Galoaa, X. Bai, S. Moezzi, **U. Nandi**, et al. “Look Around and Pay Attention: Multi-camera Point Tracking Reimagined with Transformers,” in the proceedings of **3DV 2026**, (**oral, best paper award**).

TECH STACK

Languages: Python, C++, SQL **Platforms/Tools:** PyTorch, Jax, OpenCV, NumPy, ROS2, HuggingFace, MLOps, AWS, Azure, Docker