Muhammad Raza

+1 226 698 7292 | iamraza1998@gmail.com | Waterloo, Canada
LinkedIn | GitHub | Google Scholar | Portfolio

Education

University of Waterloo — Waterloo, Canada
Doctor of Philosophy in Engineering (3D Computer Vision, Physical AI) — *Sept 2026 – July 2030*

Kumoh National Institute of Technology — Gumi-si, South Korea
Masters in IT Convergence Engineering; ***GPA: 4.33/4.50*** — *Feb 2023 – Feb 2025*

Karachi Institute of Economics and Technology (PAF-KIET) — Karachi, Pakistan
Bachelors in Computer Science; ***GPA: 3.34/4.00*** — *Aug 2018 – Sept 2022*

Skills

Programming: Python, C++

Deep Learning: PyTorch, TensorFlow, model training, fine-tuning, representation learning

Computer Vision & 3D Perception: 3D Gaussian Splatting, 3D Reconstruction, point cloud analysis/processing, feature matching/tracking, Physical AI, 2D/3D object detection, segmentation, object tracking, depth estimation,

Efficient AI: model compression, quantization, knowledge distillation, LoRA, memory-efficient inference

Multimodal AI: vision–language models (VLMs), vision–language–action models (VLAs), open-vocabulary perception

Deployment & Systems: TensorRT, ONNX, NVIDIA Jetson, Docker, mobile and edge inference

Development Tools: Git, Jira, Confluence, Slack

Experience

Neubility Inc. — Seoul, South Korea
R&D AI Robotics Engineer — *Oct 2025 – Oct 2026*

- **Multi-Camera 2D/3D Tracking & Pseudo-Labeling:** Developed a geometry-based tracking-by-detection pipeline for a 5-camera robotic platform, projecting 2D detections into a shared 3D world frame using calibrated depth and camera extrinsics. Generated **19K+ pseudo-labels across 646 global tracks** with zero ID switches on static objects, and incorporated appearance embeddings using temporal-consistency and triplet losses.
- **Custom DINOv3 2D/3D Perception Model:** Designed an FCOS-style anchor-free detector on a frozen DINOv3 backbone with stride-8 FPN features, cross-level attention, deformable convolutions, and auxiliary DETR supervision. Achieved **mAP@50 of 0.739** and **macro-F1 of 0.458** on the 16-class Neubie domain.
- **Shared-Backbone Multi-Task Perception:** Architected a unified DINOv3-based perception framework supporting 2D detection, 3D bounding-box lifting, monocular depth estimation, and semantic segmentation through independent task-specific heads sharing a common feature backbone.
- **3D Perception & Automated Dataset Generation:** Developed automated 3D labeling and perception pipelines combining open-vocabulary detection, monocular depth estimation, camera geometry, point-cloud reconstruction, and 3D scene understanding to reduce manual annotation requirements.
- **Edge Deployment & ROS2 Integration:** Engineered ONNX/TensorRT-compatible perception models, including INT8-optimized inference pipelines, and integrated real-time 2D/3D perception into the *Neubie* autonomous robot through ROS2.

3i Inc. — Seoul, South Korea
R&D AI Software Engineer — *Jan 2025 – Oct 2025*

- **PIVO Sports Tracking System:** Developed and deployed end-to-end AI modules for the PIVO sports tracking application, including object detection, single-object tracking, human pose estimation, and real-time inference optimization for consumer devices.

- **Occlusion-Robust Single-Object Tracking:** Designed a lightweight occlusion-handling framework combining fine-tuned object detectors with CSRT-based continuous tracking, enabling stable target localization under partial and full occlusions.
- **Distractor-Aware Memory (DAM):** Implemented a memory-based recovery mechanism for re-detection and identity preservation during occlusion and tracking failure, improving robustness against target switches and visual ambiguity.
- **3D Avatar and Model Optimization:** Contributed to 3D avatar generation pipelines, including 3D mesh simplification and compression for efficient real-time rendering and deployment.
- **Core Technologies:** Object detection, single-object tracking, occlusion handling, human pose estimation, memory-based tracking, model optimization, and real-time deployment.

WENS Lab Gumi-si, South Korea

Research Assistant – Deep Learning and Computer Vision *Feb 2023 – Present, Full-time*

- **Autonomous Person-Following:** Developed a real-time YOLOv8-based detection and tracking system for autonomous grocery carts.
- **UAV Vision for Smart Factories:** Built a lightweight YOLOv8n + ROS pipeline for real-time QR code detection and recognition.
- **Lightweight Vision Models:** Designed depthwise separable CNN architectures for efficient image classification and edge deployment.
- **UxV Perception & Mapping:** Developed smart-farming instance segmentation and collaborative indoor 3D mapping using VINS-Fusion and COVINS-G.
- **Industrial Computer Vision:** Implemented ROS-based detection systems for person, traffic light, and chemical leakage monitoring.

First-Author Publications

ASD Accepted August 2026

BMVC - 2026 Results

- Proposed **Anti-Shortcut Distillation via Temporal Negative Knowledge Transfer**

FSGM-Sum Submitted June 2026

Pattern Recognition Results

- Proposed **FSGM-Sum: Factorized Spatio-Temporal Gating Mixer for Efficient Video Summarization**

EdgeDAM Submitted August 2026

WACV 2027 - Results

- Proposed **EdgeDAM: Distractor-Aware Real-time Object Tracking for Mobile Devices.**

References

Chul Min Yeum

Professor,
University of Waterloo,
Waterloo, Canada.

Mail - cmyeum@uwaterloo.ca

Address - E2-2313 University Ave W,
N2L 3G1, Canada

Phone - (519) 888-4567

Omer Tariq

Research Scientist,
Neubility Inc.,
Seoul, South Korea.

Mail - omertariq2000@gmail.com

Dept. - Perception AI.

Phone - +82 (10) 2175-3294