Trung M. Bui, PhD

Robotics & Computer Vision Engineer — Perception, Localization & Edge Deployment

Morgan Hill, CA • +1 669-326-2460 • bmtrungvp@gmail.com
linkedin.com/in/trung-m-bui • github.com/mtbui2010 • **U.S. Green Card**

Summary

Computer Vision and Robotics Engineer with 7+ years shipping production perception systems and 11+ years total CV experience including PhD research. Specializes in **object detection, multi-object tracking, 6D pose estimation, and RGB-D depth perception**, with a strong track record of deploying reliable, high-throughput systems in unstructured real-world environments. Optimized real-time inference pipelines on **NVIDIA Jetson** (TensorRT, INT8, 30+ FPS, sub-200ms latency) and built end-to-end data pipelines processing large-scale image datasets from production deployments. Published in *IEEE TIP* (Q1) and *IEICE Transactions*. Experienced integrating classical geometric vision (feature matching, calibration, RANSAC) with modern deep learning for robust, field-deployable solutions.

Experience

Senior Computer Vision & Robotics Engineer *March 2019 – Present*

Korea Electronics Technology Institute (KETI), Seongnam, South Korea

- Led perception ML for a **production CV system** that improved success from 70% to **92% in cluttered real-world scenes**, processing hundreds of items per hour across multiple deployment sites; system in production 3+ years.
- Optimized inference for edge deployment with **TensorRT, ONNX, INT8 quantization** on NVIDIA Jetson AGX, sustaining **30+ FPS** with **50% latency reduction** versus FP32 baseline and sub-200ms end-to-end latency.
- Trained and shipped deep learning models for **object detection, segmentation, and 6D pose estimation** on RGB-D data; handled long-tail challenges via **active learning** and open-vocabulary detection.
- Designed a **vision transformer architecture** with fine-grained attention for precision detection and pose estimation in cluttered scenes, achieving **20% precision improvement over CNN baselines**. Published in *IEICE Transactions*, Nov 2025.
- Built end-to-end **training pipelines**: dataset curation (synthetic + real + production failures), augmentation strategies, hyperparameter tuning, evaluation harnesses, and continuous improvement loops.
- Architected a mobile manipulation system (**MARS**) with ROS2 framework, RGB-D cameras, and LLM-based task planning; integrated modular perception packages for detection, segmentation, grasp pose estimation, and VLM-based grounding. Published in *Ubiquitous Robots 2025*.
- Built **semi-automatic data annotation pipeline**: YOLOv8 initial detection → GroundingDINO open-vocabulary refinement → FastSAM segmentation → human verification → fine-tuning loop; processed 200K+ annotations.
- Led cross-functional collaboration with hardware, manufacturing, and deployment teams; mentored 3 junior engineers; contributed core CV components to national R&D projects with multi-million USD funding.

Selected Projects

Fine-Aware Vision Transformer for Precision Detection *Published, IEICE 2025*

- Designed a novel transformer architecture with fine-grained attention module for cluttered RGB-D scenes; standard ViT struggled with global pooling treating background as equally important as foreground.
- Achieved **20% precision improvement** over CNN baselines; trained on 200K+ annotations with custom augmentation; L1 loss on pose regression + BCE on quality scoring.
- Deployed on **NVIDIA Jetson** with TensorRT, sustained **30+ FPS** at production accuracy in real-world deployment.

Tracker Lab — Multi-Object Tracking from Scratch *2025*

github.com/mtbui2010/vision_tracking

- Implemented **SORT, DeepSORT, ByteTrack, and a custom tracker from scratch** — Kalman filter, Hungarian assignment, motion/appearance/IoU cost functions; no third-party tracking libraries.
- Detection with fine-tuned **YOLOv11** on MOT17/DanceTrack; ReID embeddings via **OSNet**; custom **MOTA / IDF1 / HOTA** metrics benchmarked against MOT17-val.
- Full-stack research platform with FastAPI + Next.js 14 Canvas UI, side-by-side tracker comparison, frame-by-frame stepping; deployed on Docker, Vercel, and RunPod serverless GPU workers.

GroundingDINO Annotation & Active Learning Pipeline *2024*

github.com/mtbui2010/groundingdino_tool

- Streamlit-based tool for semi-automatic image labeling and open-vocabulary detector fine-tuning on novel object categories.
- Pipeline: YOLOv8n → GroundingDINO open-vocabulary refinement → FastSAM segmentation → human verification → fine-tuning; closes long-tail label gaps without large pre-labeled datasets.
- Designed for **high-throughput field data**: iterative active learning loop continuously improves model accuracy as new deployment data arrives.

VLA Manipulation Policies on Franka Panda (Isaac Lab + MuJoCo) *2025*

github.com/mtbui2010/robot_sim_vla

- Built a vision-language-action (VLA) evaluation platform for a Franka Panda arm across Isaac Lab and MuJoCo with standardized observation/action spaces.
- Reached **98.75% success on the LIBERO benchmark** (80 episodes), edging the 97.1% paper baseline; built reproducible fine-tuning infra (4×A6000, 150K steps).

Single Image Dehazing Using Color Ellipsoid Prior *IEEE TIP 2018*

github.com/mtbui2010/Single-Image-Dehazing

- Novel algorithm for single-image dehazing under adverse outdoor visual conditions; PhD research basis.
- Published in *IEEE Transactions on Image Processing* (Q1 journal); outperformed existing methods on standard benchmarks.

Selected Publications

- **T. M. Bui**, J. Hwang, S. Jun, W. Kim, D. Shin. "A Fine-Aware Vision Transformer for Precision Grasp Pose Detection." *IEICE Transactions on Information and Systems*, Nov. 2025.
- **T. M. Bui**, W. Kim. "Single Image Dehazing Using Color Ellipsoid Prior." *IEEE Transactions on Image Processing*, Feb. 2018. *(Q1 journal)*
- **T. M. Bui**, Y. Kim, S. J. Moon, M. Cho, M. Seo, D. Shin. "Development of a Mobile Assistive Robot for Daily Living Support." *Ubiquitous Robots*, 2025.
- **T. M. Bui**, H. N. Tran, W. Kim, S. Kim. "Segmenting Dark Channel Prior in Single Image Dehazing." *IET Electronics Letters*, March 2014.

Education

Ph.D. in Computer Vision *2014 – 2019*

Kyung Hee University, South Korea

Thesis: *Single Image Dehazing Using Color Ellipsoid Prior* (basis of IEEE TIP 2018 publication)

M.Eng. in Computer Vision *2011 – 2014*

Kyung Hee University, South Korea

B.Eng. in Electrical and Electronics Engineering *2005 – 2010*

Ho Chi Minh City University of Technology, Vietnam — Excellent Engineer Training Program

Technical Skills

Computer Vision: object detection, semantic / instance segmentation, 6D pose estimation, multi-object tracking (SORT / DeepSORT / ByteTrack), stereo / RGB-D depth perception, point cloud processing, feature matching, visual localization
Edge Deployment: TensorRT optimization, INT8 quantization, **NVIDIA Jetson** (AGX/Orin), real-time inference (30+ FPS), sub-200ms latency, model quantization & pruning
Classical CV: OpenCV, camera calibration, RANSAC, ICP, feature matching, geometric vision, homography estimation
Deep Learning: vision transformers, CNN architectures, attention mechanisms, transfer learning, self-supervised learning, hyperparameter tuning
Production ML: large-scale training pipelines, dataset curation, active learning, augmentation, evaluation harnesses,

shadow deployment, canary rollout, drift detection
Robotics: ROS2 / ROS, MoveIt2, hand-eye calibration, mobile manipulation, grasp planning, sensor integration
Frameworks: PyTorch, TensorFlow, Hugging Face, ONNX, TensorRT, CUDA
Languages: Python (expert), C++, C
Infrastructure: Docker, Linux, Git, distributed training, FastAPI, Streamlit
Certifications: Visual Perception for Self-Driving Cars (Coursera, 2023); Improving Deep Neural Networks (DeepLearning.AI, 2021)