

Nelson Tian

lunshuo.tian@student.unsw.edu.au | (+61) 401-454-389 | Sydney, NSW 2020

EDUCATION

THE UNIVERSITY OF NEW SOUTH WALES, SYDNEY, AUSTRALIA

Master of Information Technology (Artificial Intelligence)

Sep 2025 - Sep 2026

- **Core Courses:** Computer Vision, Natural Language Processing, Advanced Machine Learning, Recommender Systems, etc.

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY, SHENZHEN, CHINA

Bachelor of Engineering, Computer Science and Technology

Aug 2021 - Jul 2025

- **Core Courses:** Machine Learning, Deep Learning, Intelligent Robotics, Computer Vision, Artificial Intelligence, etc.

SUMMARY

- **AI Perception Researcher & Open-Source Developer** currently pursuing a Master's degree at UNSW, specializing in 3D Computer Vision, Embodied AI, and Autonomous Navigation.
- **Authored** a paper on real-time 6D pose tracking submitted to **IROS 2026** (3rd Author), and successfully **integrated** core algorithm optimizations into the official **OpenCV** repository.
- Proficient in C++ and Python, with hands-on experience in deploying deep learning models (**VLMs, YOLO**) on edge devices under **restricted VRAM**, and building end-to-end robotics pipelines using ROS and hardware SDKs.

INTERNSHIP EXPERIENCE

Dobot | Shenzhen, China

Jan 2026 - Mar 2026

Perception Algorithm Engineer Intern

- **Engineered** an end-to-end voice-controlled grasping system for Atom robots, seamlessly integrating Whisper-based speech recognition, visual perception, and motion control modules.
- **Deployed** the **Qwen-7B** Vision-Language Model (VLM) locally under strict hardware constraints (**9GB VRAM** peak) to enable intelligent scene understanding.
- **Constructed** a real-time target detection and tracking pipeline utilizing **YOLO-Seg** and **SORT** algorithms, formulated robust grasping strategies via the Atom SDK.
- **Achieved** high real-time accuracy for natural language instruction execution; the project was officially adopted into the company's open-source repository (GitHub: Dobot-Arm/voice_txt_command).

Google Summer of Code (OpenCV) | Remote

Mar 2025 - Aug 2025

Open Source Contributor

- **Engineered** the Fractal ArUco marker detection feature for OpenCV's objdetect module from scratch, enabling accurate near-and-far marker recognition to expand OpenCV's applications in **drone** scenarios.
- **Constructed** a comprehensive benchmark dataset encompassing complex conditions (e.g., rotation, occlusion, noise) and established standardized evaluation metrics and pipelines.
- **Optimized** the original algorithm structure and patched existing vulnerabilities, boosting the keypoint matching capability by up to **28%** and significantly enhancing overall detection accuracy and stability.
- **Delivered** robust code branches with complete unit tests and performance benchmarks adhering to strict upstream standards, leading to successful approval and merging by the OpenCV community (PR: opencv/opencv#27934).

RESEARCH EXPERIENCE

Click-to-Model: Real-Time Interactive Object Modeling and Robust 6D Pose Tracking

Sep 2025 - Mar 2026

Core Developer & 3rd Author (Submitted to IROS 2026, Under Review)

- **Developed** an end-to-end interactive perception framework (CLM) for unseen objects to enable real-time robotic grasping and manipulation, eliminating the need for pre-scanned CAD models via single-click initialization.
- **Proposed** a joint shape-color driven particle size matching algorithm that reconstructs high-fidelity 3D models by fusing depth map information, significantly enhancing feature extraction efficiency and scale recovery consistency for model-free objects in complex interactive environments.
- **Achieved** superior performance on the YCB-V and self-built datasets, reaching an ADD(-S) accuracy of **0.96** for novel objects and a **97%** Relocalization Success Rate (RSR) while maintaining a stable tracking speed of **19 FPS**.

Explainable AI in PCGRL via Action-Perturbation Saliency Maps

Sep 2023 - Jul 2024

Core Developer

- **Pioneered** an action-perturbation saliency map approach for PCGRL, utilizing the gradient of the output probability distribution (measured by Euclidean distance) to quantify saliency and decode complex agent decision-making.
- **Conducted** a rigorous evaluation framework based on extensive human visual annotation, achieving a **90.5%** saliency match rate that significantly outperformed the reward-based baseline (**80.9%**).
- **Optimized** the underlying algorithms and code architecture, accelerating computation efficiency by **10x** compared to the initial prototype, enabling the real-time analysis of large-scale procedural map generation.
- **Bridged** the interpretability gap in reinforcement learning, providing actionable visual insights for developers to debug and understand AI-driven generative logic.

PROJECT EXPERIENCE

Autonomous Driving and Object Recognition based on Limo Robot

Jun 2025 - Jul 2025

- **Built** a GMapping SLAM and AMCL localization pipeline within the ROS framework, enabling autonomous mapping and navigation for a Limo robot in complex map environments.
- **Developed** path recording and Pure Pursuit cruising modules to achieve stable, end-to-end autonomous driving, path tracking, and speed control along pre-recorded trajectories.
- **Engineered** a traffic scene perception system by integrating deep learning (YOLO for traffic lights) and traditional visual algorithms (Canny edge detection, polygon fitting, and white area sampling for crosswalks).
- **Realized** real-time detection and decision-making for traffic scenarios, successfully supporting automated responses such as stopping at red lights, decelerating at yellow lights, and executing 1-second stops at crosswalks.

OpenCV ArUco Module Optimization and Evaluation

Sep 2024 - Jan 2025

- **Conducted** a systematic evaluation of OpenCV's ArUco, ArUco3, and ArUcoNano modules, designing a comprehensive dataset of 1.2k real-world samples covering 7 influencing factors like perspective distortion, marker density, and noise.
- **Quantified** algorithmic differences by comparing key metrics including detection rate, pose estimation accuracy, and computational efficiency, generating complete evaluation pipelines and visual reports.
- **Identified** conservative default parameters in ArUco3 and optimized them, boosting the detection rate by 35%-300% (significantly reducing missed detections for small-sized markers) while maintaining similar processing times.
- **Submitted** a Pull Request to optimize ArUco3 default parameters (PR: [opencv/opencv#28645](#)) and provided data-driven evaluations on ArUcoNano to guide future community development decisions.

SKILLS & EXPERTISE

Languages & Tools: C/C++, Python, Linux, Git.

Frameworks & Libraries: PyTorch, TensorFlow, Transformers, ONNX, OpenCV.

Computer Vision (2D & 3D): 6D Pose Estimation, Semantic Segmentation, Object Detection, Object Tracking.

Robotics, Simulation & Edge: ROS, CARLA, NVIDIA Isaac Sim, Edge Deployment.

Advanced Machine Learning: Reinforcement Learning (PPO, OpenAI Gym), Federated Learning, Cross-Modal Learning, Explainable AI (XAI), Benchmark Design & Evaluation.

ADDITIONAL INFORMATION

Language Proficiency: English (Professional Working Proficiency), Mandarin (Native).

Awards: Top 5 Global Proposal in ERNIE Hackathon -- Multimodal Track (2025).