

EDUCATION

University College London (UCL)

London, UK

Master of Engineering in Robotics and Artificial Intelligence

Sep 2023 - Jun 2027 (Expected)

- ✧ **Degree Type:** Integrated Undergraduate and Master's Program | **Average Score: 79.5%**
- ✧ **Key Modules:** Reinforcement Learning for Robotics and Artificial Intelligence (86.3%), Introduction to Machine Learning for Robotics and Artificial Intelligence (78.16%), Object-Oriented Programming for Robotics and Artificial Intelligence (87.85%), Introduction to Programming for Robotics and Artificial Intelligence (80.75%)
- ✧ **Technical Skills:** Python, PyTorch, C, C++, MATLAB, Simulink, ROS, ROS 2, Linux, STM32 HAL, SolidWorks, Fusion 360, Docker

PUBLICATIONS

- [1] **Bai, N.**, et al. (2026). An Event-Triggered Robust Reinforcement Learning Scheme for Nonlinear Systems with Unknown Disturbances. *The 45th Chinese Control Conference (CCC 2026)*. (Accepted)
- [2] Chan, C. P., **Bai, N.**, et al. (2026). Reinforcement Learning Based Optimal Control: A Survey of Adaptive Dynamic Programming for Manipulators and Wheeled Mobile Robots. *Advanced Mechatronics*. (Accepted)
- [3] **Bai, N.**, et al. (2026). SignVLA: Real-Time Sign Language-Guided Robotic Manipulation via Attention LSTM and Vision-Language-Action Models. *The 31st International Conference on Automation and Computing*. (Accepted)
- [4] Yang, L., Zhu, H., Zhong, Z., **Bai, N.**, et al. (2026). Post-Training for Reasoning LLMs with Reinforcement Learning: A Stability-Efficiency Perspective. *The 31st International Conference on Automation and Computing*. (Accepted)
- [5] Zhu, L., **Bai, N.**, et al. (2026). Zero-shot Decomposed Retrieval Enhancement for Visually Rich Document via Broad Search and Deep Reasoning. (Under review)
- [6] Yang, L., Zhu, H., Zhong, Z., Tan, X., **Bai, N.**, et al. (2026). Post-Training for Reasoning Large Language Models with Reinforcement Learning: A Survey. *Unmanned System*. (Under review)

RESEARCH EXPERIENCES

Vision-Language-Action Model Control by Sign Language

London, UK

Supervised by Dr. Zezhi Tang | Dept of Computer Science, University College London

Sep 2025 - Jun 2026

- ✧ Deployed the NVIDIA GR00T framework onto a robotic manipulator, collected teleoperation demonstration data, and fine-tuned a Vision-Language-Action (VLA) model that completed a number of manipulation tasks.
- ✧ Utilized MediaPipe Hands to build a sign language feature extraction pipeline which was able to model and understand gestures and finger movements through the generation of absolute/relative 3D landmarks, normalized finger direction vectors, and temporal velocity features.
- ✧ Converted real-time sign language and leveraged an LLM to semantically complete it into natural language commands that could be used as inputs for a VLA model.
- ✧ First authored the paper, **SignVLA: Real-Time Sign Language-Guided Robotic Manipulation via Attention LSTM and Vision-Language-Action Models**, currently accepted by the 31st International Conference on Automation and Computing (ICAC 2026) and is pending publication.

A Survey on ADP Control for Manipulators and Wheeled Mobile Robots

London, UK

Supervised by Dr. Zezhi Tang | Dept of Computer Science, University College London

Sep 2025 - Apr 2026

- ✧ Reviewed the literature on Adaptive Dynamic Programming (ADP) with a special focus on application for robotics and how platform-specific physical constraints shape controller architectures and stability analysis.
- ✧ Compared typical ADP structures, including dual-network actor-critic, single-critic, and decentralized formulations with a focus on evaluating practical trade-offs between high-dimensional scalability, real-time computational overhead, and hardware deployment limitations.
- ✧ Summarized the core bottlenecks in the real-world deployment of ADP, particularly the discrepancies simulation results and what actually happens in real life applications, and authored a comprehensive survey paper, **Reinforcement Learning Based Optimal Control: A Survey of Adaptive Dynamic Programming for Manipulators and Wheeled Mobile Robots**, currently accepted by *Advanced Mechatronics* and is pending publication.

Post-Training for Reasoning LLMs with Reinforcement Learning: A Stability–Efficiency Perspective

London, UK

Supervised by Dr. Zezhi Tang | Dept of Computer Science, University College London

Sep 2025 - Jan 2026

- ✧ Developed a unified algorithm design framework to address the complexity of algorithm components in LLM RL post-training, and proposed a “Unified Design Map” that categorizes post-training methods via signal source and density, optimization regime and trust-region control, and structural granularity and agentic scope.
- ✧ Conducted a decomposition down to the component level and formulated deployment strategies, and built an auditable “method-by-motif matrix” for direct comparison of underlying design choices to select the best one. Analyzed stability-efficiency trade-offs across algorithm families, and formulated practical, deployment-oriented approaches and repair strategies grounded in measurable monitoring signals for specific trigger scenarios.
- ✧ Established standardized diagnostic and evaluation workflows for reasoning-model RL training, introducing compact monitoring metrics including degenerate-group rate, ratio-tail statistics, entropy trends, and success@budget, and proposed compute- and length-controlled evaluation protocols to avoid misleading gains caused by increased generation length or sampling budgets.

Reinforcement Learning for Nonlinear Control Systems

London, UK

Supervised by Dr. Zezhi Tang | Dept of Computer Science, University College London

Jun 2025 - Dec 2025

- ✧ Created an approach to help uncertain nonlinear systems avoid wasting computational resources through the integration of an Event-Triggered Mechanism (ETM), an Extended State Observer (ESO), and Adaptive Dynamic Programming (ADP).
- ✧ Designed an event-triggered learning scheme in parallel with an ESO-based disturbance estimation module to achieve real-time compensation for unknown disturbances. Also used an Actor-Critic architecture to approximate the optimal control policy without requiring prior knowledge of the system model.
- ✧ Demonstrated through simulation experiments that the proposed framework reduces redundant updates and computational costs by approximately 73% compared to traditional methods without any performance loss.
- ✧ First authored the research paper titled, *An Event-Triggered Robust Reinforcement Learning Scheme for Nonlinear Systems with Unknown Disturbances*, currently accepted by the 45th Chinese Control Conference (CCC 2026) and is pending publication.

Visual RAG Systems

Chengdu, China

Supervised by Dr. Yulong Ji | School of Aeronautics and Astronautics, Sichuan University

Sep 2024 - Jun 2025

- ✧ Designed a method that was zero-shot, plug and play and which required no retraining of the backbone model. The method was seamlessly integrated with existing base retrieval models by dynamically allocating retrieval quotas via an adaptive Top-K strategy.
- ✧ Decomposed a complex user query into three complementary semantic dimensions. Specifically; context, key evidence, and visual cues which were used to overcome the limitations of single vector retrieval, which often fails to fully represent multi-constraint or reasoning intensive queries.
- ✧ Experimental results demonstrate that the proposed method achieves a significant 15.4% improvement in Recall@1, validating the effectiveness of the multimodal query decomposition strategy.
- ✧ Authored the research paper titled, *Zero-shot Decomposed Retrieval Enhancement for Visually Rich Document via Broad Search and Deep Reasoning*, which is currently under review.

ACADEMIC PROJECTS

BabyBench Infant Self-Touch Reinforcement Learning Project

Feb 2025 - Mar 2025

- ✧ Built a self-touch learning environment for infant body exploration based on BabyBench and Mujoco, applying PPO to learn early-stage self-touch behaviors through reinforcement learning.
- ✧ Designed reward mechanisms and optimized training parameters to improve motion stability and exploration efficiency.

Emotion-Aware AI Intelligent Photo Album

Feb 2025 - Mar 2025

- ✧ Fine-tuned a Qwen-based dialogue model for empathetic and humorous conversation, leveraging YOLO for multi-person face detection, ResNet50 for facial expression classification, and Gradio for frontend interaction interface development.
- ✧ Responsible for fine-tuning the ResNet50 classifier and developing the Gradio-based frontend interface.