# Boyang Zhang

E-mail: theostnc@bu.edu / Phone: +1 (650)-788-4328
Address: Allston, MA 02134 | **Homepage:** byzhang811.github.io


## Education Background

**Boston University** — **Boston, United States**
College of Engineering — 09/2025 – 01/2027 (Expected)
Master of Science in Electrical and Computer Engineering — GPA: 3.76/4.00

**Manchester Metropolitan University, Wuhan Campus** — **Wuhan, China**
School of Computer Science and Information Engineering — 09/2021 – 06/2025
Bachelor of Science in Software Engineering — GPA: 2.87/4.00

**Hubei University** — **Wuhan, China**
Manchester Metropolitan Joint Institute — 09/2021 – 06/2025
Bachelor of Engineering in Software Engineering

**Summer Course Program – University of California, Berkeley** — **Berkeley, United States**
Undergraduate Non-Degree Program — GPA: 3.50/4.00 — 07/2023 – 08/2023

## Research Interests

(1) Reliable Vision-Language-Action Learning (2) Data-centric Learning (3) Multimodal Visual Understanding

## Publications

**Boyang Zhang**, Lianlei Shan. *MPCoT: Reward-Guided Multi-Path Latent Reasoning for Test-Time Scalable Vision-Language-Action.* arXiv Preprint. 06/2026. Submitted to IEEE Conference on Robotics and Automation (ICRA 2027) arXiv:2606.06245.

Yu Xia, **Boyang Zhang**, Zhihong Pan, Kang Gao, Hanbo Yang, Ruoyu Chen, Kang Yang. *Deployment-Oriented Benchmark for Automated UAV-Based Facade Defect Detection.* Automation in Construction. Published. DOI 10.1016/j.autcon.2026.107280

**Boyang Zhang**, Kang Gao, Jianan Gu, Li Ai and Yuang Geng. *Improving Stage-Transition Detection in Long-Term Guided-Wave Monitoring through Autoencoder Regularization.* Structural Durability & Health Monitoring. 06/2026. Published. DOI 10.32604/sdhm.2026.083281.

**Boyang Zhang**. *RTDS: A Robust Two-Stage Tongue Diagnosis System with Swin-Hybrid Architecture.* Signal, Image and Video Processing. 04/2026. Published. DOI 10.1007/s11760-026-05316-3.

Yuang Geng, Zhongzheng Ren Zhang, Jianan Gu, **Boyang Zhang**, Kang Yang and Ivan Ruchkin. *Unsupervised Vision-Action Misbehavior Detection for Self-Evolving Training Data Cleaning in Autonomous Driving.* IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2026). Accepted. Manuscript draft.

## Manuscripts Under Review

Samu Yang†, **Boyang Zhang**†, Lianlei Shan. ActRight: Counterfactual Training for Causally-Grounded Vision-Language-Action Models. Submitted to ICLR 2027. Manuscript draft. († Equal contribution.)

**Boyang Zhang**, Junhua Zhou, Zhenchen Wei, Zheyuan Huai, Weidai Xia. *LaCETrack: Language-Guided Counterfactual Evidence Learning against Shortcut Cues in RGB-T Tracking.* Submitted to ICASSP 2027. Manuscript draft.

**Boyang Zhang**, Zheyuan Huai, Kang Gao, Yuang Geng, Huiyu Chen and Kang Yang. *Re-Querying the Detector: Counterfactual Response Verification for Anomaly Detection under Pose Variation.* Submitted to ICASSP 2027. Manuscript draft.

**Boyang Zhang**, Jordana Muroff and Janusz Konrad. *Assessment of Room Clutter via Image-Derived Cue Fusion.* Under review at Image and Vision Computing. Manuscript draft.

## Research Experience

**Research Assistant**, Tsinghua University, Beijing, China, Hybrid — 05/2025 – Present

**MPCoT: Reward-Guided Multi-Path Latent Reasoning for Test-Time Scalable Vision-Language-Action**

Project Lead, Advisor: Prof. Lianlei Shan, Tsinghua University

- Proposed a latent multi-path reasoning module for VLA policies, enabling test-time scaling through reasoning depth and hypothesis width while preserving the original action interface.
- Designed recurrent latent refinement and confidence-aware aggregation to combine multiple latent hypotheses.
- Introduced a training-only reward-guided path preference objective using action consistency, progress, and success signals; validated consistent gains on LIBERO and CALVIN under matched protocols. (*Submitted to ICRA 2027)*

**ActRight: Counterfactual Training for Causally-Grounded Vision-Language-Action Models**

Research Assistant, Advisor: Prof. Lianlei Shan, Tsinghua University

- Contributed to the core conceptual design of a counterfactual causal training framework for VLA models, emphasizing structured V→L→A information flow and language-grounded action prediction.
- Provided method-level guidance on counterfactual alignment objectives and adversarial intervention design to reduce spurious visual reliance and improve instruction sensitivity. *(Submitted to ICLR 2027)*

**Research Assistant**, Central South University, Changsha, China — 04/2026 – Present

**LaCETrack: Language-Guided Counterfactual Evidence Learning against Shortcut Cues in RGB-T Tracking**

Research Assistant, Advisors: Prof. Weidai Xia, Central South University, Hybrid

- Led the end-to-end development of LaCETrack, a language-guided counterfactual training framework using identity-preserving RGB-T feature conflicts and joint spatial-semantic ranking to reduce shortcut reliance.
- Improved a strong DINOv2-based comparator by 0.7 MPR and 0.7 MSR on RGBT234, and by 0.6 PR, 0.7 NPR, and 1.1 SR on LasHeR, with no additional inference-time overhead. *(Submitted to ICASSP 2027)*

**Research Assistant**, SmartDATA Lab & TEA Lab, University of Florida, Hybrid 01/2026 – Present

**Deployment-Oriented Benchmark for Automated UAV-Based Facade Defect Detection**

Researcher, Advisors: Prof. Kang Yang & Prof. Ruoyu Chen, University of Florida & Jiangsu University of Science and Technology.

- Led the technical and experimental design of a route-level UAV facade-defect benchmark, translating practical inspection requirements into a unified detector-evaluation and annotation-budget analysis framework.
- Developed an automated benchmarking pipeline for training, validating, and evaluating across six detector architectures.
- Designed the cross-family, qualitative, and failure-case analyses, deriving deployment-oriented recommendations for detector selection and labeled-data planning. *(**Published** in Automation in Construction)*

**Unsupervised Vision-Action Misbehavior Detection for Self-Evolving Training Data Cleaning in Autonomous Driving**

Research Assistant, Advisors: Prof. Kang Yang & Prof. Ivan Ruchkin, University of Florida

- Implemented key components of the dual-encoder vision-action autoencoder, including latent reference loss and PCC-based anomaly scoring for separating normal and misbehavior image-action pairs.
- Participated in evaluation of the self-evolving data-cleaning pipeline, which reduced the anomaly ratio from 20% to below 5% and improved controller CTE from 0.120 m to 0.053 m over four iterations. *(**Accepted** by IROS 2026)*

**Re-Querying the Detector: Counterfactual Response Verification for Anomaly Detection under Pose Variation**

Project Lead, Advisors: Prof. Kang Yang & Dr. Huiyu Chen, University of Florida & Southwest Jiaotong University

- Developed CRV, a training-free verification framework for robotic anomaly detection, combining normal-reference image synthesis with re-querying of frozen detectors to improve detection and localization.
- Improved PatchCore's image AUROC from 83.31% to 89.75% and pixel AP from 29.37% to 46.16% on RAD.
- Evaluated five detector backends on RAD and verified additional gains on MAD-Sim and PCAD; controlled ablations isolated a 6.58-percentage-point pixel AP improvement from re-querying. *(Submitted to ICASSP 2027)*

**Improving Stage-Transition Detection in Long-Term Guided-Wave Monitoring Through Autoencoder Regularization**

Project Lead, Advisor: Prof. Kang Yang, University of Florida

- Conducted a controlled empirical study comparing dropout, weight decay, and sparsity regularization under one fixed autoencoder pipeline, isolating the effect of each method on unsupervised adjacent stage-transition detection.
- Improved detected transitions from 2/12 (vanilla AE) to up to 7/12 (sparsity), with best mean AUROC reaching 0.9288. *(**Published** in Structural Durability & Health Monitoring)*

**Researcher**, Visual Information Processing (VIP) Lab, Boston University, Boston, United States, On-site 09/2025 – Present

**Assessment of Room Clutter via Image-Derived Cue Fusion**

Advisors: Prof. Janusz Konrad & Prof. Jordana Muroff, Boston University

- Developed an image-derived cue-fusion framework for hoarding-clutter assessment on HINDER-2025, combining RGB images with structural, semantic, and volumetric cues derived without additional manual annotations.
- Strengthened the RGB-only ViT baseline via partial transformer fine-tuning and DINOv2 initialization.
- Achieved 63.42% CCR and 96.27% CCR-1 with RGB+Depth fusion and DINOv2, improving over the prior ViT benchmark by +9.96 and +3.03 percentage points. *(Under Review at Image and Vision Computing)*

**Independent Research Project**, Hubei University 01/2025 – 05/2025

**RTDS: A Robust two-stage tongue diagnosis system with U-Net++ segmentation and swin-hybrid classification**

- Constructed a dataset of 2,100 in situ tongue images.
- Engineered a segmentation module using U-Net++ to isolate the Region of Interest.
- Designed the Swin-Hybrid architecture to capture both local textural details and global semantic context to mitigate class ambiguity. *(**Published** in Signal, Image and Video Processing)*

## Professional Experience

**DingTalk, Alibaba Cloud** **Hangzhou, China**

Algorithm Engineering Intern, Research Department 06/2024 – 09/2024

- Engineered the SAM (ViT-B) inference pipeline for an interactive education platform.
- Developed an adaptive preprocessing module for unconstrained inputs, improving robustness on low-light, noisy.

**iSoftStone Technology Service Co., Ltd** **Wuhan, China**

Machine Learning Intern, Algorithm Department 01/2024 – 06/2024

- Led development of a box-office prediction framework using XGBoost and LightGBM, building the end-to-end pipeline from multi-source data cleaning to baseline evaluation.
- Diagnosed covariate shift across years and redesigned the feature system with relative change rates and exponential decay to stabilize generalization for new release schedules.

## Honors, Awards & Intellectual Property

Boston University ECE MS Research Award 03/2026
First Prize - National College Student AI Technology Competition 01/2025
Hubei University Scholarship for Academic Excellence 07/2024
First Prize - National College Student Innovation and Entrepreneurship Ability Competition 06/2024

## Skills

**Programming & Tools**: Python, C/C++, SQL, Java, MATLAB, Shell Scripting, LaTeX.
**Languages**: Chinese (Native), English (Advanced), Spanish (Elementary)