

Yue Su

+86 18062089249 • s3702681@gmail.com • selen-suyue.github.io
in selen-su-008867334 • selen-suyue • He/Him

Education

Xidian University

XiAn, China

Undergraduate in School of Artificial Intelligence

2022–2026

- **GPA:** 3.70/4.00 **Score:** 85.00/100.00 **Ranking:** 6.95%(Specialized), 12.88%(All)
- **Core Courses:** Algorithm (99), Data Structure (98), Circuit (96), Pattern Recognition (94), Advanced Mathematics (94), Engineering (93), Program (93), Machine Learning (92), Optimization (90), Probability (90), Numerical Methods (90).

Research Interests

My research interests and the deep learning paradigm I aim to shape primarily focus on the following aspects:

- **Reasoning-based Learning:** Inferring knowledge through interaction with the real world, such as deducing object motions, contacts, mechanical interactions, affordances, etc., from observations. By leveraging the ability to reason about the real world, robots or agents can acquire enhanced generalization capabilities, potentially leading to the development of active learning abilities in the future.
- **Generative Modeling:** The aim is to use generative methods to model objects governed by physical laws, including feature-level, semantic-level, and scene-level representations, with the ultimate goal of building a world model.

Research Experience

- **Machine Vision and Intelligence Group** **Shanghai Jiao Tong University**
Research Assistant advised by Prof. Cewu Lu, Prof. Lixin Yang
Pursuing research in robot learning with particular attention to perception and inference about the real world. July 2024 - Now
- **Key Laboratory of Cooperative Intelligent Systems** **Xidian University**
Research Assistant advised by Prof. Maoguo Gong, Prof. Hao Li
Contributed to research on adversarial attacks against computer vision systems. September 2023 - July 2024

Publications

- **Dense Policy: Bidirectional Autoregressive Learning of Actions**
Yue Su,^{*} Xinyu Zhan^{*}, Hongjie Fang, Han Xue, Haoshu Fang, Yong-Lu Li, Cewu Lu, Lixin Yang[†]
Propose Dense Policy, A bidirectional robotic autoregressive policy, which infers trajectories by gradually expanding actions from sparse keyframes, has demonstrated capabilities exceeding diffusion-based policies.
[\[arXiv\]](#) [\[code\]](#) [\[website\]](#)
- **Motion Before Action: Diffusing Object Motion as Manipulation Condition**
Yue Su,^{*} Xinyu Zhan^{*}, Hongjie Fang, Yong-Lu Li, Cewu Lu, Lixin Yang[†]
Propose MBA, a novel module that employs two cascaded diffusion processes for robot action generation under object motion guidance. Designed as a plug-and-play component, MBA can be flexibly integrated into existing robotic manipulation policies with diffusion action heads.
[\[arXiv\]](#) [\[code\]](#) [\[website\]](#)
- **Generative Adversarial Patches for Physical Attacks on Cross-Modal Pedestrian Re-Identification**
Yue Su, Hao Li[†], Maoguo Gong[†]
A generative adversarial attack on VI-ReID models perturbs modality-invariant features, creating patches that expose sota vulnerabilities and highlight the need for enhanced feature extraction.
[\[arXiv\]](#)
- **AdvDisplay: Adversarial Display Assembled by Thermoelectric Cooler for Fooling Thermal Infrared Detectors**
Hao Li[†], Fanggao Wan, Yue Su, Yue Wu, Mingyang Zhang, Maoguo Gong[†]
Historically, infrared adversarial attacks were single-use and inflexible. Using TEC, we implemented attacks adaptable to various scenarios, causing pedestrian detection models to misjudge.
Accepted at AAAI 2025.

Projects

- **MetaPalace: Let you in a meta world of The Palace Museum**

We've done what the Old Palace official website couldn't: offering 3D artifact views with single-view reconstruction and an interactive LLM-powered tour guider using RAG technology.

[\[website\]](#) [\[front-end code\]](#) [\[back-end code\]](#)

- **U-pre: U-Net is an excellent learner for time series forecasting**

Time series forecasting is suited for U-Net's architecture due to its consistent input-output distributions and strong mathematical alignment. Combining U-Net with Bert-Encoder improved performance by incorporating both local and global attention.

[\[code\]](#) [\[report-cn\]](#)

- **AgentCrossTalk: Perform a Crosstalk between two LLM agents**

This project uses the Google Gemini to create a simple chatbot application simulating two crosstalk performers performing based on user-provided topics.

[\[code\]](#) [\[website\]](#)

- **FGSM3D: Is the point cloud gradient perturbation attack feasible?**

We tried to extend FGSM to the 3D field and achieved significant success within a certain gradient range, but the sampling method of 3D models tells us that things seem to be not that simple...

[\[code\]](#) [\[report-cn\]](#)

- **AcoFlow: Heuristic Search for Maximum Flow Problem**

The problem of finding the maximum flow lies in how to design better heuristic information to find the augmenting path. We boldly challenge this problem through the ant colony algorithm.

[\[code\]](#) [\[report-cn\]](#)

Awards

- **First Prize, Provincial Level, 2023 China Mathematical Contest in Modeling**
- **First Prize, Provincial Level, 2024 China Mathematical Contest in Modeling**
- **Second Prize, Northwestern, 2024 China Computer Design Contest**

Community Experience

- **Head of the Research Department of Microsoft Club, Xidian University.**