Tairan Chen

Beijing, China | +86-17614416037 | terrencechentr@gmail.com
Homepage: https://thechentr.github.io/

EDUCATION

Jilin University
Jilin, China

Bachelor in Computer Science and Technology — *Sept. 2020 - Jun. 2024*

Overall GPA : 86.01 / 100

Related Coursework: C Programming (95), Compiler Principles (94), Computer Organization (90), Digital Logic Circuits (90), Microcomputer Systems (89), Discrete Mathematics (89), Data Structures (87), Operating System(86)

Tsinghua University
Beijing, China

Research Assistant, *Tsinghua SAIL Group* — *Dec. 2023 - Present*

Team leaders: Prof. Jun Zhu and Associate Prof. Hang Su

Research Focus: Computer Vision, Natural Language Processing , Adversarial Machine Learning

RESEARCH EXPERIENCE

Prompt-Guided Environmentally Consistent Adversarial Patch
Tsinghua University

Research Assistant — *May. 2024 - Nov. 2024*

Advisor: Associate Prof. **Hang Su** in the Department of Computer Science and Technology at Tsinghua University

Background: Vision-based systems are increasingly threatened by adversarial patches attacks in the physical world. Traditional adversarial patches lack environmental consistency, making them easily detectable by humans. We proposed PG-ECAP, a novel method for generating adversarial patches that seamlessly blend into their environments while maintaining effective attack performance.

- Developed an adversarial patch generation framework leveraging **text-to-image diffusion model** guided by environmental prompts, ensuring patches visually blend into real-world scenes.
- Introduced **Prompt Alignment Loss** and **Latent Space Alignment Loss** to maintain both the adversarial effectiveness and environmental consistency of generated patches.
- Achieved an average reduction of object detection models' mAP to **27.09** (from **66.96**) on the **INRIA dataset** and a **94.41%** attack success rate (ASR) in **physical-world experiments** across various settings, significantly surpassing existing methods.

Achievement: Published on **arXiv** and currently under review for **CVPR 2025**. For more details, please refer to the following link: https://arxiv.org/abs/2411.10498

Active Perception Defense Against 3D Adversarial Patches
Tsinghua University

Research Assistant — *Apr. 2024 - Present*

Advisor: Associate Prof. **Hang Su** in the Department of Computer Science and Technology at Tsinghua University

Background: Critical visual models (e.g., identity verification, autonomous driving) are highly vulnerable to 3D adversarial attacks. Traditional defenses struggle with dynamic conditions. We propose an active defense framework that improves robustness via adaptive exploration in 3D adversarial settings.

- Implemented a **differentiable rendering environment based on EG3D** to support consistent multi-view image generation, enabling more accurate 3D representation.
- Developed an active perception defense system that integrates an **RL-based policy module** and a **perception module** in a feedback loop, enhancing environmental exploration and robustness.
- Achieved up to a **95% reduction** in attack success rate (ASR) against various adversarial attacks, surpassing passive defenses and demonstrating effectiveness in 3D adversarial settings.

Achievement: Source code is available at: https://github.com/thechentr/EAD-YOLOv5

Red Teaming Multimodal Large Language Model Security Challenge **Tsinghua University**

Research Assistant *Jun. 2024 - Jul. 2024*

Advisor: Associate Prof. **Hang Su** in the Department of Computer Science and Technology at Tsinghua University

Background: Visual Language Models (VLMs) are vulnerable to jailbreak attacks caused by crafted inputs. Existing attack methods lack generalization and diversity. We proposed an adversarial sample generation method that significantly improves success rates in black-box scenarios.

- Developed an **adversarial samples generation method** initiated from the visual modality to enhance attack effectiveness on VLMs.
- Integrated the **COCO dataset** with **Stable Diffusion** to improve attack stealth and diversity.
- Achieved a **26.5% increase** in attack success rate (ASR) through optimized image generation and prompt tuning.

Achievement: Our team secured **the Second Place** in the Red Teaming Multimodal Large Language Model Security Challenge hosted by **CCF**.

PROJECT EXPERIENCE

Automated C++ Code Generator for Syntax Analysis **Jilin University**

Undergraduate Researcher *Mar. 2023 - May. 2023*

Advisor: Associate Prof. **Huaxiao Liu** in the Department of Computer Science and Technology at Jilin University

Background: Syntax analysis is crucial in compiler construction, but building efficient parsers is complex and time-consuming. To address this, we developed an automated C++ code generator for syntax analysis that converts grammar rules into highly optimized, executable code.

- Built an **automated C++ code generator for syntax analysis**, greatly reducing manual coding effort.
- Added support for **recursive descent parsing** and **LL(1) parsing**, improving code reusability and efficiency.
- Developed an **automatic prediction set calculation module**, easing integration with various grammar rules.

Achievement: Source code is available at: https://github.com/thechentr/SNL-compiler

AWARDS, SCHOLARSHIPS & LEADERSHIP

- Outstanding Undergraduate Graduate (**Top 5%**) *2024*
- Academic Scholarship (**Top 10%**) *2021, 2022, 2023*
- **Top 8.38%** of CCF Certified Software Professional in C++ in China *2021*
- Provincial **Second Prize** in the National Undergraduate Mathematical Contest in Modeling *2022*
- Outstanding Student Leader (**Top 3%**) *2021, 2022, 2023*
- Cultural and Sports Activities Scholarship (**Top 10%**) *2022, 2024*
- Social Work Scholarship (**Top 10%**) *2023*
- **President** of the Student Union, School of Computer Science and Technology *Sept. 2022–Sept. 2023*
- **Monitor** of Class 31 of 2020, School of Computer Science and Technology *Sept. 2020–Jul. 2024*

TECHNICAL PROFICIENCIES

Programming Languages: C/C++, Python(Pytorch), R, Java, etc.

Deep learning models: YOLO, EG3D, Stable Diffusion, GPT-3, etc.

Other Tools: Git, Docker, Latex, etc.