Yilun (Alan) Zhou

yzhou851@gatech.edu | https://ylz1992.github.io/personalweb/

Education

Georgia Institute of Technology — Atlanta, GA
M.S. in Computer Science, GPA: 3.8/4 — *Expected Dec 2024*
Relevant Coursework: ML, Adv ML, DL, CV, Deep Generative Models, Vision Language Models, Graduate Algorithms, Info Security, Network Security, Computer Networks, Software Development Processes

University of Florida — Gainesville, FL
M.S. in Geographic Information Systems — *Jan 2017*

Beijing Jiaotong University — Beijing, China
B.S. in Railway Engineering/GIS — *Jul 2014*

Technical Skills

Languages: Python, Java, JavaScript, C, C#, C++, LaTex
Tools & Frameworks: PyTorch, Huggingface Libs, Scikit-Learn, Git, Docker, React, Node.js, MongoDB, Unity, Android Studio, AWS Sets
Certificates: AWS Solutions Architect Associate, Data Science Certificate (Renmin University of China)

Research & Work Experience

Research Assistant — Aug 2024 – Present
Efficient and Intelligent Computing Lab, Georgia Tech — *Advisor: Prof. Celine Lin, Dr. Zhongzhi Yu*

- Investigating the role of attention sinks during adversarial attacks in large language models
- Developing generalized adversarial attack strategies across various language and generative models using token-based and string-based methods.
- Collaborating on a paper under review focusing on adversarial attacks in multi-modal systems.

Unity Developer Intern - Remote — Jun 2024 – Aug 2024
Plutonic — *Atlanta, GA*

- Implemented AWS workflows, including GameLift, Photon Fusion Kit, and S3, to support multiplayer experiences for a teletherapy platform.
- Developed OpenAI-based chatbot to enhance user interaction and improve the platform's engagement.

Research Assistant — Jan 2024 – Jun 2024
TReNDS Center, Georgia Tech — *Advisor: Dr. Sergey Plis*

- Developed deep learning models for diagnosing fMRI brain images, focusing on classification and segmentation tasks using models such as ResNet50 and U-Net.
- Improved diagnostic accuracy by 10% through model optimization and fine-tuning.

R&D Engineer — Jan 2017 – Dec 2022
China Architecture Design & Research Group — *Beijing, China*

- Led a team of 3 engineers in digital simulation and parametric modeling for infrastructures. Successfully designed and executed plans for over 10 buildings including the No.17 Bobsleigh track recognized by International Olympic Committee. Portfolio Highlights
- Took part into 10+ construction project; Achieved 8 structural design awards including 2 national constructional awards; 8 published articles and 1 published book(structural engineer);

Lab Technician — Dec 2016 – Aug 2017
Florida Department of Transportation — *Gainesville, FL*

- Data processing of geo parameters for deep foundation bridges in Florida by Matlab.
- Sensor data collection on Florida roads captures real-time information on traffic, road conditions for infrastructure planning.

Projects

Conditioned Denoising for Adversarial Defense Aug 2024 – Present

Generative Model, Robustness Analysis *Advisor: Prof. Zsolt Kira*

- Developed noise purification techniques to remove adversarial perturbations prior to image processing in Vision-Language Models, with the objective of mitigating misclassification.
- Leveraged multimodal information to enhance robustness against adversarial attacks in black-box settings.

Robustness of VLM Against Adversarial Concept Injection Aug 2024 – Present

VLM, Jailbreak *Advisor: Prof. Bo Dai*

- Led a project on adversarial attacks in Vision-Language Models (VLMs), using CLIP encoders to extract harmful concepts (e.g., nudity, violence) and injecting them into latent spaces to generate adversarial prompts.
- Optimized fixed-size prompts with gradient-based (PeZ) and Genetic Algorithms to jailbreak Text-to-Image (T2I) models like Flux, generate inappropriate images.
- Explored cross-modal vulnerabilities with ImageBind, extending attacks to generate inappropriate audio and depth, revealing weaknesses in multi-modal systems.
- Project Link

ML General Topics Jan 2023 – Jun 2023

Course-based Projects

- Conducted research on RL-based trading strategies for the Turkish Stock Market.
- Developed a model for housing price prediction in Georgia, USA, using supervised learning techniques.
- Project Link

Android App: Job Offer Comparison Aug 2023 – Dec 2023

Back-end Development

- Developed an Android 12 app to compare job offers with different benefit packages, integrating back-end development and UI design.
- GitHub Link

Command Line Utility: Txter Aug 2023 – Dec 2023

CLI Development

- Developed a Java-based command-line utility for text processing, featuring modular code design and robust testing using JUnit.
- GitHub Link

Computer Network Project Aug 2023 – Dec 2023

CN Class *Jan 2024 – Jun 2023*

- Network Setup based on EPF system. Involving configuring IP, OSPF and BGP for routers and hosts. Implemented VPN configurations for secure communications, CLI and RKPI for enhance the security and authenticity of the routing information exchanged between ASes.