Jingyang Zhang

zhjy227@gmail.com • https://zjysteven.github.io/

Summary

Jingyang is a Ph.D. candidate at Duke ECE with 5+ years of experience in designing advanced and robust training algorithms for **machine learning-based vision models**. His research spans **adversarial robustness**, **out-of-distribution detection**, and using **synthetic data** for better model evaluation and training. He combines **1)** outstanding research capabilities, with publications in top-tier ML conferences, and **2)** strong engineering skills, demonstrated through open-source implementations of ML models and algorithms.

Education

- **Duke University (Durham, NC)** **Aug 2019 - Apr 2024**
  Ph.D. student, Dept. of Electrical and Computer Engineering *GPA: 3.96/4.0*
- **Tsinghua University (Beijing, China)** **Sep 2015 - Jul 2019**
  B.Eng., Dept. of Electronic Engineering

Selected Projects

- **Adversarially Robust Ensemble Generation**
  - Proposed DVERGE, a novel ensemble training methodology for Deep Neural Networks (DNNs) that diversifies the learnt features of sub-models. With little degradation in clean accuracy, DVERGE was once the state-of-the-art ensemble-based defense against black-box transfer attacks.
  - *Supported by DARPA QED-RML program and was accepted by NeurIPS'20 (oral).* [Paper][Code]
- **Fine-Grained Out-of-Distribution Detection**
  - Proposed MixOE, a new DNN training algorithm that leads to 4%-13% improvement in true negative rate in large-scale, fine-grained OOD detection.
  - *Supported by AFRL and was accepted by WACV'23.* [Paper][Code]
- **Large-Scale Benchmark for Out-of-Distribution Detection**
  - Built OpenOOD v1.5, a large-scale, enhanced benchmark and test platform for OOD detection in the context of image classification. OpenOOD comprehensively evaluated existing methodologies and identified remaining challenges and future directions for the field.
  - A well-recognized project that receives *600+* stars; accepted by *NeurIPS'23 DistShift workshop (oral)*. [Paper][Code][Leaderboard]
- **Generating Natural Adversarial Examples with Stable Diffusion**
  - Developed an optimization technique that perturbs the conditional token embedding of Stable Diffusion to generate natural adversarial examples that deceive image classifiers.
  - *Accepted by ICLR'24 Tiny Paper track.* [Paper][Code]

Internship Experience

- **Machine Learning Research Intern @ Bosch Center for AI** *Jun 2022 - Dec 2022*
  - Was developing a "universal" adversarial defense that is robust to both ℓ_p (digital) and patch (physical) adversarial attacks against images. Demonstrated the effectiveness and potential of the defense through extensive experiments, which resulted in a patent.
- **Machine Learning Intern @ Tesla** *Jun 2023 - Sep 2023*
  - Implemented and adapted state-of-the-art deep learning models for trajectory prediction. Showed the efficacy of this method over baselines with proof-of-concept experiments in different scenarios.

Technical Skills

- Programming Languages: **Python**, C++, Matlab. Deep Learning Frameworks: **PyTorch**, TensorFlow.