A Software Tool for Securing Deep Learning against Adversarial Attacks for CAVs

Technology Transfer Activities

by

Pierluigi Pisu, Ph.D., Clemson University
Gurcan Comert, Ph.D., Benedict College
Negash Begashaw, Ph.D., Benedict College
Chunheng Zhao, Ph.D. Candidate, Clemson University
Kalpit Vadnerkar, Ph.D. candidate, Clemson University

Contact information
Pierluigi Pisu, Ph.D.
4 Research Drive, Greenville, SC 29607
Clemson University
Phone: (864) 283-7227; *E-mail:* pisup@clemson.edu

November 2024

Center for Connected Multimodal Mobility (C^2M^2)

200 Lowry Hall
Clemson, SC 29634

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by the Center for Connected Multimodal Mobility (C^2M^2) (Tier 1 University Transportation Center) Grant, which is headquartered at Clemson University, Clemson, South Carolina, USA, from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Non-exclusive rights are retained by the U.S. DOT.

ACKNOWLEDGMENT

This study is supported by the Center for Connected Multimodal Mobility (C^2M^2) (USDOT Tier 1 University Transportation Center) headquartered at Clemson University, Clemson, SC. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of C^2M^2, and the official policy or position of the USDOT/OST-R, or any State or other entity, and the U.S. Government assumes no liability for the contents or use thereof.

Table of Contents

TECHNOLOGY TRANSFER ACTIVITIES

1 Outputs

This proposal builds upon the PI's previously completed project "Securing Deep Learning against Adversarial Attacks for Connected and Automated Vehicles"[1]. In this prior project, we developed a robust image classification model with a tradeoff between accuracy and adversarial robustness. As a continuing work, in this project, we developed a graphical user interface (GUI) for automatically generating the robust image classification model given a baseline perception neural network, including options for training and evaluation under different publicly available image data sets. We built a user-friendly interface to help train the perception neural network to minimize the impact of adversarial attacks from perception outputs upon user preference requests. Users can select several popular adversarial attacks from an option menu and investigate their impact on the perception output from the panel interface.

1.1 Output #1

Dissemination of C^2M^2 research results was obtained through technical report as indicated below:

P. Pisu, G. Comert, N. Begashaw, C. Zhao and K. Vadnerkar, "A Software Tool for Securing Deep Learning against Adversarial Attacks for CAVs", Final Technical Report, C^2M^2, Dec. 2024.

Technical Reports: 1

1.2 Output #2

The following new methods and products resulted from this research:

We developed a software tool that would allow for easy-to-use interface features, seamless adapting to various perception neural networks and automatically capable of generating the adversary resilient outputs within the desired perception system. The implementation utilizes React with TypeScript for robust type safety and better development experience, along with a modular component architecture that ensures scalability and maintainability.

New Research Products: 1

2 Outcomes

The project allowed for collaboration of faculty from Clemson University and Benedict College and support training of students at graduate and undergraduate level working on the project.

2.1 Outcome #1

Two graduate students (Clemson University) and two undergraduate students at Benedict College were involved in the project.

[1] https://cecas.clemson.edu/C2M2/project-reports/

2.2 Outcome #2

We leverage Clemson Palmetto Computing platform for the use of cluster computing and high-performance computing. The incorporation of the proposed deep neural network algorithm and existing Cluster computing platforms is conducted in this project.

By incorporating an intuitive graphical user interface (GUI), we believe that even those who might not be well-versed in the intricacies of deep learning can comfortably navigate the platform. This focus on usability is crucial as it broadens the potential user base, allowing researchers from various backgrounds to efficiently generate neural networks that are resilient against adversarial attacks.

3 Impacts

This project focuses on the development of new software tools for making DNNs more resilient to adversarial attacks with a particular focus on the development of object detectors utilized in the perception module of CAVs. The results of this project will have a broad applicability not only to the transportation sector but also to many other engineering fields such as autonomous driving, biometric identification, speech and face recognition, intelligent transportation systems, and robotics (vision SLAM).

3.1 Impact #1

Currently, our software tool can train robust image classification models with the selection of various datasets, model architectures, and training parameters. The GUI can train and test intrinsically robust neural networks to make a variety of adversarial attacks less effective, which means obtaining correct recognition results even in the presence of adversarial attacks.