Precise and Generalized Robustness Certification for Neural Networks

Yuanyuan Yuan, *The Hong Kong University of Science and Technology and ETH Zurich;* Shuai Wang, *The Hong Kong University of Science and Technology;* Zhendong Su, *ETH Zurich*

https://www.usenix.org/conference/usenixsecurity23/presentation/yuan-yuanyuan-certification

This artifact appendix is included in the Artifact Appendices to the Proceedings of the 32nd USENIX Security Symposium and appends to the paper of the same name that appears in the Proceedings of the 32nd USENIX Security Symposium.

August 9–11, 2023 • Anaheim, CA, USA

978-1-939133-37-3

USENIX'23 Artifact Appendix: Precise and Generalized Robustness Certification for Neural Networks

[1,2]Yuanyuan Yuan, [1]Shuai Wang, and [2]Zhendong Su
[1]The Hong Kong University of Science and Technology, [2]ETH Zurich
yyuanaq@cse.ust.hk, shuaiw@cse.ust.hk, zhendong.su@inf.ethz.ch

A Artifact Appendix

A.1 Abstract

We provide code and data of our paper in this artifact. Our artifact is publicly available at `https://github.com/Yuanyuan-Yuan/GCert` with detailed documents. Using our tool, users can certify neural network robustness towards various semantic-level mutations.

A.2 Description & Requirements

A.2.1 Security, privacy, and ethical concerns

None

A.2.2 How to access

An archived copy of the initial version is available at: `https://zenodo.org/record/8062051`.

Our artifact is actively maintained at: `https://github.com/Yuanyuan-Yuan/GCert`.

A.2.3 Hardware dependencies

We do not have any particular requirements for the hardware. Our artifact may need GPUs to speed up the certification; we suggest evaluators having at least one GPU.

A.2.4 Software dependencies

Our tool is built based on Pytorch; evaluators need to first install Pytorch. See detailed instructions in our documents.

A.2.5 Benchmarks

None.

A.3 Set-up

A.3.1 Installation

Users only need to install Pytorch first. See details in our documents.

A.3.2 Basic Test

To test the basic functionality, evaluators can first run `cd experiments` to change the current directory. Then run `python augment_geometrical.py`. This script will start training a generative model with regulation proposed in our paper.

Detailed instructions are provided in our documents.

A.4 Version

Based on the LaTeX template for Artifact Evaluation V20220926. Submission, reviewing and badging methodology followed for the evaluation of this artifact can be found at `https://secartifacts.github.io/usenixsec2023/`.