Lehrstuhl für Sicherheit in der Informatik
Prof. Dr. Claudia Eckert

Fraunhofer AISEC

Student Assistant (Hilfswissenschaftler) & Research Internship (Forschungspraxis)

Evaluation of Adversarial Example Detection Systems

Motivation and Task Description

In recent years, deep neural networks (DNNs) have achieved remarkable results and even showed super-human capabilities in a broad range of domains. Often, this includes security-sensitive environments. For example, in today's driver-assistance systems and future autonomous cars, DNNs are used to automatically detect and identify street signs. In this scenario, misclassifications can lead to severe incidents threatening the passengers' lives.

Unfortunately, DNNs are vulnerable to so-called adversarial examples. Such inputs differ only slightly from their original counterparts, yet provoke missclassifications of the DNNs. To generate adversarial examples, attackers need to add only slight changes to the input, which makes the attack process simple and easily accessible. As the required perturbations are often human imperceptible, the detection of such attacks is a challenging task.

In this work, different adversarial examples detection approaches should be evaluated. For this purpose, state-of-the-art detection tools should be summarized. Then, based on currently used methods and our concepts, an end-to-end framework is implemented and tested. Finally, a comparison with current detection systems is performed and analyzed.

Requirements

- Sophisticated programming skills in Python
- First experience in Deep Learning libraries
- Interest in Deep Learning and Security
- Ability to work self-directed and systematically
- Motivation and self-organization

Contact

Jan-Philipp Schulze
Telefon: +49 89 322-9986-195
E-Mail: jan-philipp.schulze@aisec.fraunhofer.de

Philip Sperl
Telefon: +49 89 322-9986-141
E-Mail: philip.sperl@aisec.fraunhofer.de

Please attach your CV and your current transcript of records to your application.

Fraunhofer Research Institution for Applied and Integrated Security (AISEC)
Cognitive Security Technologies
Lichtenbergstraße 11, 85748 Garching (near Munich), Germany
https://www.aisec.fraunhofer.de Ausschreibungsdatum: 26. Oktober 2020