SCIENCE · PASSION · TECHNOLOGY

TU Graz

Open Thesis / Project

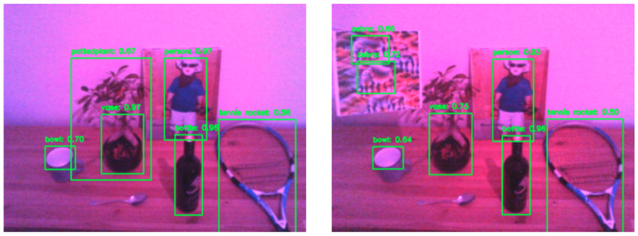
Real-world Challenges for Adversarial Patches in Object Detection

Embedded Learning and Sensing Systems Group

Motivation

Adversarial attacks make small, malicious changes to inputs that deceive ML models into making incorrect predictions. They might look normal or acceptable to human observers, but still manage to fool a ML system – e.g., a sticker in the observed scene in computer vision domain. These attacks are usually designed by employing digital transformation techniques designed to account for environmental variables. However, real-world factors – how light interacts with different shapes and materials – add extra complexity. Our project will study how these real-world conditions affect the performance of adversarial patches using a physically based rendering engine.

Interested? Please contact us for more details!



(a) Physically changed hue **(b) As in (a) + patch = not effective**

Figure 1: Patch efficiency in different light conditions [1].

Requirements / Skills

- Programming skills in Python and C++; interest in efficient code writing;
- Prior experience with deep learning frameworks (preferably PyTorch).

Goals and Tasks

In this project you will investigate the real-world performance of adversarial patches by systematically evaluating how diverse physical conditions – lighting-material interactions – affect their ability to deceive machine learning models. The project includes the following tasks:

- Literature review on adversarial learning and defense mechanisms;
- Model digital adversarial patches in real-world scenarios to examine how actual physical conditions influence their effectiveness; compare your method to existing methods;
- Summarize the results in a written report and prepare an oral presentation.

Target Group

Students in ICE and Computer Science.

Thesis Type

Master Project / Master Thesis.

References

[1] Jakob Shack et al. “Breaking the Illusion: Real-world Challenges for Adversarial Patches in Object Detection”. In: *Proceedings of 1st EMERGE Workshop on EWSN*. 2024. URL: https://arxiv.org/abs/2410.19863.

Contact Persons

- MSc Katarina Petrović (katarina.milenkovic@pro2future.at)
- Dr. Olga Saukh (saukh@tugraz.at)

Institute for Technical Informatics
Embedded Learning and Sensing Systems Group