NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

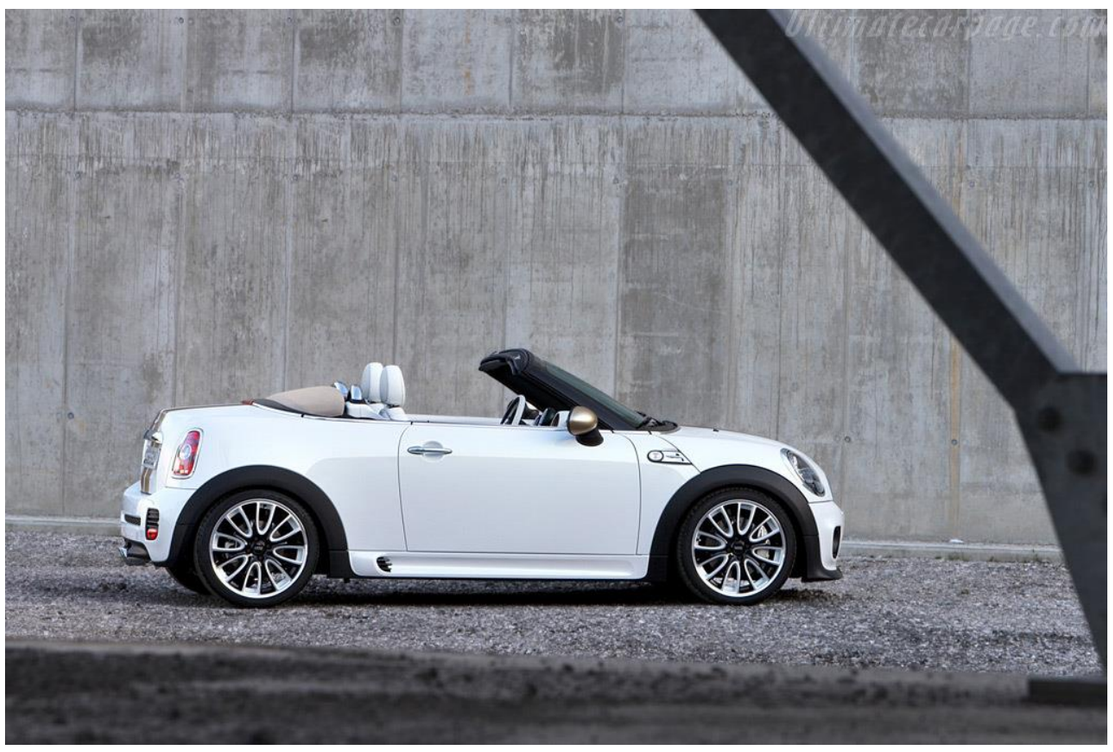
School of Computer Science and Engineering
College of Engineering

Adversarial Patch Detection

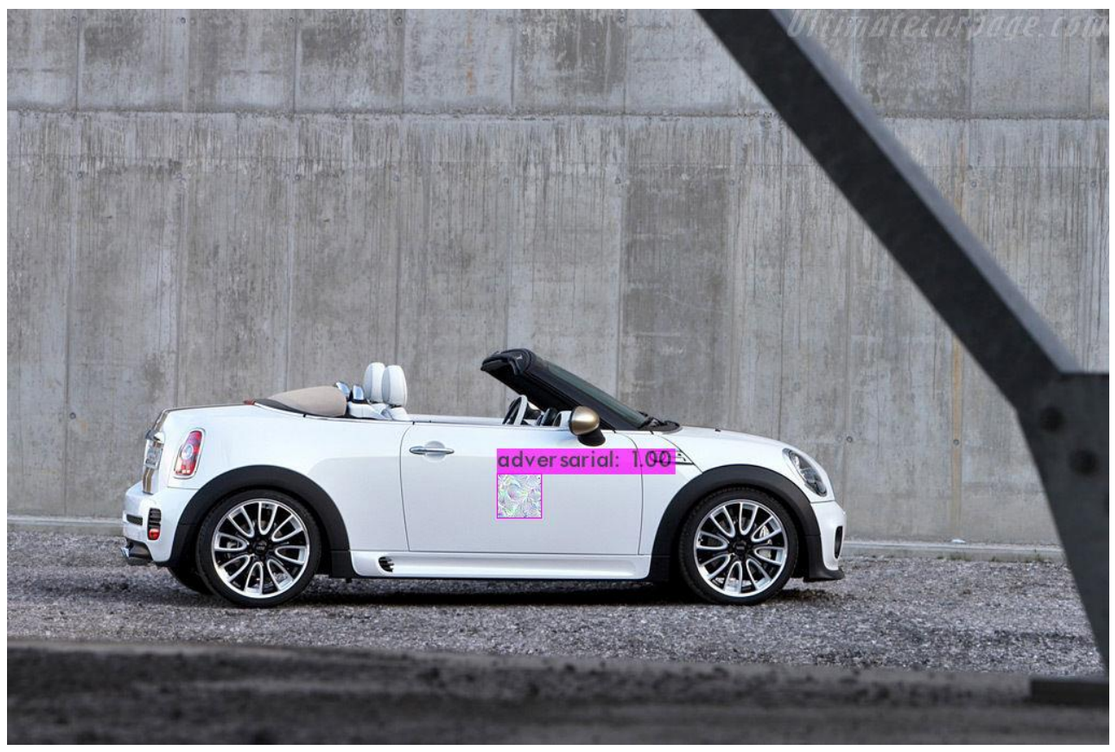
Defence against attacks on digital twinning

Student: Yeong Ler Yuen Joash Supervisor: Asst Prof Zhao Jun

Applying and Detecting Google Adversarial Patch



Original Image



Attacked Image with YOLO annotations

Project Objectives:

This paper aims to develop a YOLOv4 detector to find adversarial patches in images captured by cameras employed on the Internet of Vehicles (IoV) and an edge orchestrator to offload the detection task to systems with stronger computing capabilities and to address the optimization problem of minimising the trade-off between network latency and detection accuracy. As such, adversarial patches could be found in real time, with high precision, when moving through a complex environment like the Metaverse.

Results:

The Proximal Policy Optimization (PPO) algorithm was capable of achieving the highest global reward and accuracy while maintaining low latency, thanks to its robust structure.