**Poltergeist: Acoustic Adversarial Machine Learning against Cameras and Computer Vision**

**Rahavee Prabakaran**

The paper talks about how computer vision systems are crucial to the safety of autonomous systems because these systems are a key to sense the environment for autonomous vehicles. Tragic consequences can happen when dealing with autonomous vehicles if computer vision systems are manipulated against attacker, for example the autonomous vehicle can fail to detect the pedestrian in front of it and keep moving ahead. The paper goes on to talk about how adversarial attacks exploit the vulnerability of ML algorithms which will be threat against computer vision systems. Thus, it is important to either focuses on altering physical objects such as stickers on stop sign or just adding perturbations on digital images. The camera systems in autonomous vehicles nowadays use image stabilizers to reduce image blur caused by motions. Inertial sensors like Accelerometer are vulnerable to resonant acoustic signals. Thus, the paper proposes Poltergeist where the universal sensor readings are manipulated by acoustic signals such that the images are blurred to be adversarial ones which can be misclassified by the detectors in a computer vision system. To investigate feasibility of poltergeist preliminary analysis is performed by simulating different blur patterns and fit the blurred images into the detector. There are challenges which arises although the results are promising. One challenge is to find an efficient and effective blurry pattern to achieve targeted goals and the second challenge is to quantify the impact of acoustic signals upon the blur pattern and blur levels. To tackle problem one, Bayesian optimizer is used. The paper then proposes the design for a poltergeist attack systems design. The three building are Blur Pattern Modeling, Attack parameter optimization and sensor output injection. To evaluate the performance of poltergeist, the paper studies the and compares simulation and real-world evaluations. To study the simulation evaluation BDD100K, KITT datasets were used and it was found that hiding attack achieves 100% success rate against all of the tested object detectors, which means Hiding attack can hide any object of interest such as person or traffic lights. For creating attacks, targeted attacks are easier than non-targeted attacks. For Altering attacks, targeted attacks are easier than non-targeted attacks. For studying Real-world attacks Ultrasonic Speaker is used on target Samsung S20 smartphone in a moving vehicle and the scenes are in city lane, city crossroad. The paper also compares simulation vs real-world for different attacks under different scenes. And it was found that all three attacks are feasible and Hiding attack has the best performance. The two are also compared under attacks and the ones simulated by the blur models. Countermeasures against Poltergeist includes MEMS Inertial Sensors Safeguarding, Image Stabilization Techniques, Object Detection Algorithms, and Sensor Fusion Techniques. Finally, the paper concludes new class of system-level vulnerabilities, AMpLe attacks, injecting physics into Adversarial Machine Learning and proposed Poltergeist attacks, acoustic adversarial machine learning against cameras and computer vision and evaluation showed high performance against 4 academic and 1 commercial object detectors

**Three strong points:**

1. Attacks under various scenes are studied well in the paper

2. The Poltergeist system is explained in detail

3.The datasets, object detectors, and object of interest is clearly mentioned for each simulation and real-world.

**Three weak points:**

1.The paper did not describe about introduction to Poltergeist clearly.

2. The paper did specify on what basis the countermeasures against Poltergeist are proposed.

3. The paper did not talk about existing work like Walnut in detail.