Demo: Invisible Adversarial Stripes against Traffic Sign Recognition in Autonomous Driving

Dongfang Guo
Nanyang Technological University
Singapore
dongfang.guo@ntu.edu.sg

Yuting Wu
Nanyang Technological University
Singapore
yuting.wu@ntu.edu.sg

Yimin Dai
Nanyang Technological University
Singapore
yimin006@e.ntu.edu.sg

Pengfei Zhou
University of Pittsburgh
Pittsburgh, USA
pengfeizhou@pitt.edu

Xin Lou
Singapore Institute of Technology
Singapore
lou.xin@singaporetech.edu.sg

Rui Tan
Nanyang Technological University
Singapore
tanrui@ntu.edu.sg

Abstract

Camera-based computer vision is crucial for autonomous vehicle perception. We demonstrate *GhostStripe* [5], an attack system that uses light-emitting diodes and exploits the camera's rolling shutter effect to generate adversarial stripes that are invisible to humans while misleading traffic sign recognition. To maintain stable attack effectiveness, *GhostStripe* controls the timing of the modulated light emission, adapting to both the camera's framing operation and the movement of the victim vehicle. Evaluated on real testbeds, *GhostStripe* can stably spoof traffic sign recognition results for up to 97% of frames to a wrong class when the victim vehicle passes the road section.

CCS Concepts

• **Computer systems organization** → **Embedded and cyber-physical systems**; • **Security and privacy** → **Systems security**; **Side-channel analysis and countermeasures**.

Keywords

Autonomous vehicle, CMOS camera sensor, rolling shutter effect, adversarial attack

ACM Reference Format:
Dongfang Guo, Yuting Wu, Yimin Dai, Pengfei Zhou, Xin Lou, and Rui Tan. 2024. Demo: Invisible Adversarial Stripes against Traffic Sign Recognition in Autonomous Driving. In *The 22nd ACM Conference on Embedded Networked Sensor Systems (SENSYS '24), November 4–7, 2024, Hangzhou, China.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3666025.3699401

1 Introduction

Camera-based perception is crucial for autonomous vehicles, making its reliability essential for safety. Recent work on adversarial examples [3] has highlighted vulnerabilities in these systems. To explore these risks, we demonstrate a physically deployable and stealthy optical adversarial-example attack that exploits the camera's rolling shutter effect to fool the traffic sign recognition in autonomous vehicles.

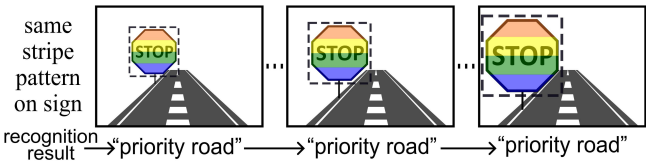

Figure 1: Stable attack effectiveness as the victim vehicle moves forward.

Complementary metal oxide semiconductor (CMOS) sensors are widely adopted in automotive cameras [1, 2]. They expose and read out the pixel values on a row-wise basis, typically from top to bottom. However, CMOS cameras exhibit the rolling shutter effect (RSE) [4]. Specifically, as each row of the CMOS sensor is exposed at slightly different times, rapid changes in input light can cause image distortion through varied color shades across scanlines. Recent studies [6–8] have shown the security implication of RSE, i.e., attackers can control the input light to create colored stripes on the captured image to mislead the computer vision interpretation. However, while previous studies have implemented basic RSE attacks on single frames in controlled environments, they fall short of achieving stable attack results over a sequence of frames [5].

GhostStripe aims to achieve stable attack results which render clearer security implications in the autonomous driving context. First, it deploys a LED near a traffic sign, projecting controlled flickering light onto the sign. As the flickering frequency exceeds human eye's perception limit, it remains invisible, making the LED appear benign. Meanwhile, the RSE-induced colored stripes captured by the camera mislead the traffic sign recognition. Second, for misleading the autonomous driving program to make erroneous decisions unconsciously, the traffic sign recognition results should be wrong and same across sufficient consecutive frames. Without such stability, anomaly detectors could trigger a fail-safe mechanism, reducing the attack's effectiveness. As the vehicle moves, the position and size of the cropout containing the sign in the camera's field of view (FoV) change , requiring the attack to adapt to both camera operations and vehicle movement to stably overlay the stripes, as envisaged in Figure. 1. To achieve this, *GhostStripe* controls the LED flickering according to the real-time sensing results of the victim

SENSYS '24, November 4–7, 2024, Hangzhou, China

ACM ISBN 979-8-4007-0697-4/24/11
https://doi.org/10.1145/3666025.3699401

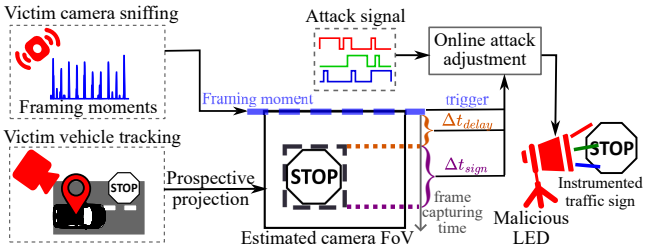

Figure 2: Overview of GhostStripe.

camera's operation and location to maintain stationary adversarial stripes on the traffic sign cropout in consecutive frames.

2 System Overview

Figure 2 overviews the *GhostStripe* with the following modules:

Attack signal generation. During the offline attack preparation phase, the attacker optimizes an time-modulated LED flickering signal corresponding to an adversarial designed colored stripes pattern for the minimum traffic sign size in the FoV that can be detected.

Victim camera sniffing. We observed that a camera's internal operations cause variations in current draw and the resulting magnetic emanations, with time-domain spikes indicating the camera's framing moments (i.e., when a frame's top scanline starts exposure) [5]. By triggering the attack replay at the framing moments, the attack signal can achieve phase synchronization with the victim camera. To detect these spikes, we build a sniffer by integrating a YHDC SCT-006 current transducer with a 330 Ω resistor to sample the voltage changes with an Arduino Due. The sniffer uses a threshold to detect the time-domain spikes.

Victim vehicle tracking. We use a LightWare SF30/C LiDAR rangefinder on the roadside to track the victim vehicle's position in real time. By tracking the real-time position of the victim vehicle and knowing the fixed location of the traffic sign, the attacker can calculate the relative position of the sign to the vehicle. Using this information, along with the camera's specifications (i.e., focal length, sensor size, image resolution, and on-vehicle position) and the traffic sign's dimensions, the traffic sign's vertical position and size in the camera's FoV can be derived by prospective projection. The vertical position and size indicate the time Δt_{delay} between the framing moment and the exposure of the sign's top scanline, and Δt_{sign} between the sign's top and bottom scanlines, respectively.

Online attack adjustment & light emission. During the attack, both the vehicle tracker and framing sniffer continuously transmit data to the LED controller built upon an Arduino Due. Upon receiving reports from either the vehicle tracker (via USB) or the sniffer (via nRF24L01+ transceivers at 2.4 GHz), the LED controller updates the parameters for attack signal control. Specifically, it triggers the attack replay at each framing moment, delays it by Δt_{delay}, and scales the duration to Δt_{sign}, to induce the designed adversarial stripes on the captured traffic sign in each frame. The replay of the attack signal is implemented by pulse-width modulation (PWM) for the LED's power supply with the Arduino. We integrate Marktech XM-L RGB LEDs to emit the attack light. To increase attack light intensity, we customize three buck converters, one for each color channel, to form an LED driver. Each converter uses the PWM signal from the Arduino to regulate high input voltage from a DC power supply, driving the LEDs to emit the attack light.

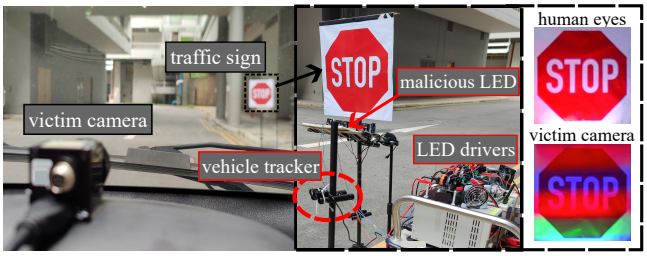

Figure 3: Outdoor testbed. The adversarial stripes can be captured by the victim camera but remain physically invisible to human eyes.

3 Results and Demonstration Description

Outdoor testbed results. We mount Leopard Imaging AR023ZWDR (default main camera in Baidu Apollo [2]) as the victim camera on a real vehicle, which is driven toward the instrumented traffic sign. Figure 3 shows the testbed. *GhostStripe* achieves up to 97% attack success rate in stably spoofing the traffic sign classification results towards a semantically conflicting class (e.g., misrecognizing "stop" as "priority road").

Live demonstration. In our demonstration, we will present a scaled-down tabletop version of the whole system to fit the provided demo space. One or more tables will simulate the road, with smaller models of the traffic sign and victim vehicle. During the demo, the audience will observe how *GhostStripe* remains invisible to the naked eye while inducing adversarial stripes in the victim camera FoV. These stripes will stably appear on the traffic sign in the camera's varying FoV across consecutive frames as the vehicle moves, and consistently spoof the traffic sign recognition. Additionally, a video demo of our outdoor real-road experiment will be displayed on a separate screen.

Acknowledgments

This research/project is supported by the National Research Foundation Singapore and DSO National Laboratories under the AI Singapore Programme (AISG Award No: AISG2-GC-2023-006).

References

[1] 2020. *Teardown: Teslaś hardware retrofits for model 3.* https://www.eetasia.com/teslas-hardware-retrofits-for-model-3/
[2] 2023. *Apollo hardware development platform.* https://developer.apollo.auto/platform/hardware.html
[3] Nicholas Carlini. 2023. *A complete list of all (arXiv) adversarial example papers.* https://nicholas.carlini.com/writing/2019/all-adversarial-example-papers.html.
[4] GETCAMERAS. 2020. *Rolling versus global shutter.* https://www.get-cameras.com/FAQ-ROLLING-VS-GLOBAL-SHUTTER
[5] Dongfang Guo, Yuting Wu, Yimin Dai, Pengfei Zhou, Xin Lou, and Rui Tan. 2024. Invisible Optical Adversarial Stripes on Traffic Sign against Autonomous Vehicles. In *ACM MobiSys.* 534–546.
[6] Sebastian Köhler, Giulio Lovisotto, Simon Birnbach, Richard Baker, and Ivan Martinovic. 2021. They see me rollin': Inherent vulnerability of the rolling shutter in cmos image sensors. In *ACSAC.* 399–413.
[7] Athena Sayles, Ashish Hooda, Mohit Gupta, Rahul Chatterjee, and Earlence Fernandes. 2021. Invisible perturbations: Physical adversarial examples exploiting the rolling shutter effect. In *IEEE/CVF CVPR.* 14666–14675.
[8] Chen Yan, Zhijian Xu, Zhanyuan Yin, Stefan Mangard, Xiaoyu Ji, Wenyuan Xu, Kaifa Zhao, Yajin Zhou, Ting Wang, Guofei Gu, et al. 2022. Rolling colors: Adversarial laser exploits against traffic light recognition. In *USENIX Security.* 1957–1974.