Artifact Evaluated
NDSS
Available
Functional
Reproduced

The Heat is On: Understanding and Mitigating Vulnerabilities of Thermal Image Perception in Autonomous Systems

Sri Hrushikesh Varma Bhupathiraju[1], Shaoyuan Xie[2], Michael Clifford[3], Qi Alfred Chen[2], Takeshi Sugawara[4], and Sara Rampazzi[1]

1 UF UNIVERSITY of FLORIDA

2 UC Irvine | AS²Guard Autonomous & Smart Systems Guard Research Group

3 TOYOTA INFOTECH Envisioning Mobility

4 UEC 国立大学法人 電気通信大学 The University of Electro-Communications

Thermal Cameras in Autonomous Systems

Beyond Visible Light

Capture infrared radiation from surrounding objects

Build a heat map of the environment

Thermal Cameras in Autonomous Systems

Beyond Visible Light

Capture infrared radiation from surrounding objects

Build a heat map of the environment

Improve Visibility

Enhance visibility in low lighting conditions (nighttime)

Improve performance under severe weather conditions

Thermal Cameras in Autonomous Systems

Beyond Visible Light

Capture infrared radiation from surrounding objects

Build a heat map of the environment

Improve Visibility

Enhance visibility in low lighting conditions (nighttime)

Improve performance under severe weather conditions

Adoption into Real-World Autonomous Systems

Integrated by robotaxies like Zoox, Waymo, Nuro

Used in drone makers DJI & Skydio for reliable perception

Thermal Cameras in Autonomous Systems

Beyond Visible Light

Capture infrared radiation from surrounding objects

Build a heat map of the environment

Improve Visibility

Enhance visibility in low lighting conditions (nighttime)

Improve performance under severe weather conditions

Adoption into Real-World Autonomous Systems

Integrated by robotaxies like Zoox, Waymo, Nuro

Used in drone makers DJI & Skydio for reliable perception

Can thermal camera-based perception be used for obstacle detection?

What are the limitations of such technology under adversarial manipulation?

Previous Works

Security analysis in conventional RGB cameras

Lensing

Ghost Image
(Man et al., 2020)

Invisible Reflections
(Sato et al., 2024)

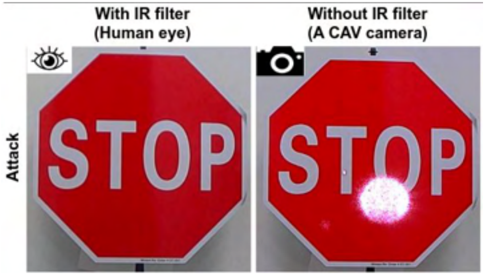


Image acquisition

Rolling Colors
(Yan et al., 2022)

They see me rollin
(Kohler et al., 2021)

Hardware

Poltergeist
et al., 2021)

Signal Injection
(Kohler et al., 2021)

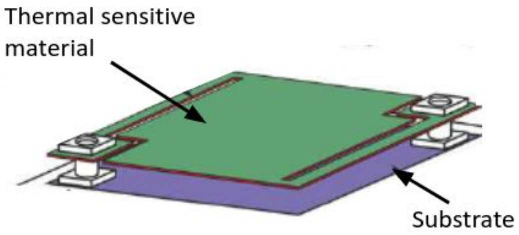
Thermal Camera Sensors

Previous works vulnerabilities in RGB cameras **do not** extend to thermal cameras

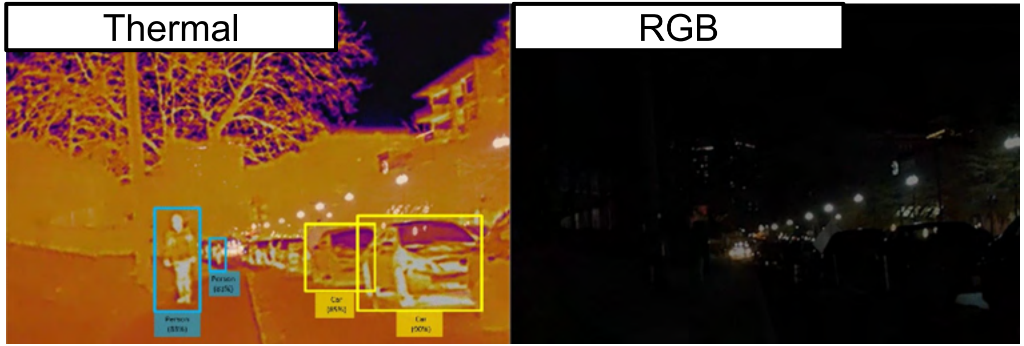
Thermal Camera Sensors

Previous works vulnerabilities in RGB cameras **do not** extend to thermal cameras

Thermal sensor



[Bieszczad et al., 2011]



[Gallagher et al., 2023]

Unique physical properties

- ➔ Long range infrared (8-12 μm)
- ➔ Arrays of radiation detectors (bolometers)

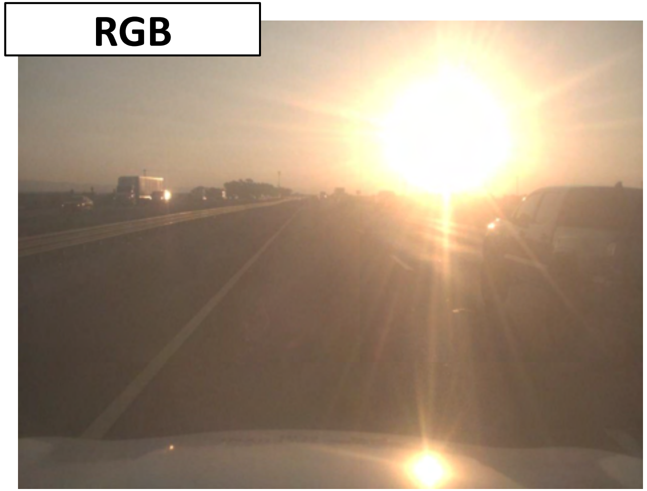
Unique sensor characteristics

- ➔ High dynamic range
- ➔ Periodic calibration
- ➔ Special lenses

Thermal Camera-based Perception

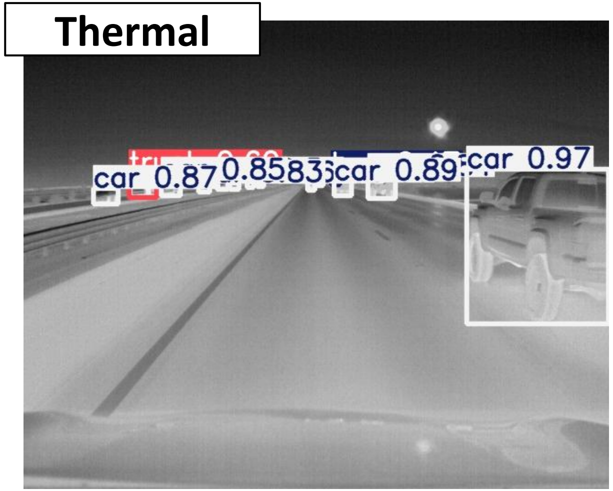
➔ Capture infrared radiation and measure relative temperature of the scene

RGB



Environment in physical domain

Thermal

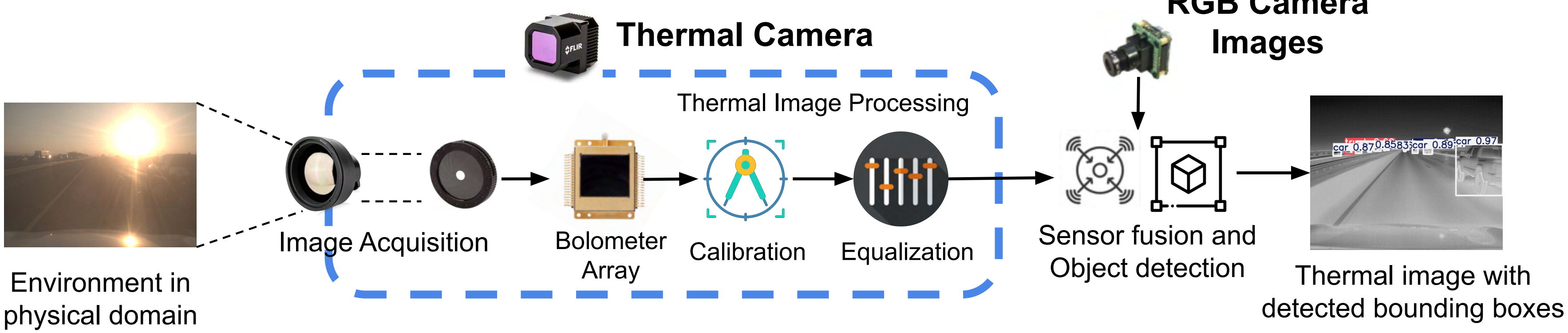


Thermal image with detected bounding boxes

Thermal Camera-based Perception

➔ Capture infrared radiation and measure relative temperature of the scene

Thermal image perception pipeline



Thermal Camera

RGB Camera Images

Thermal Image Processing

Environment in physical domain

Image Acquisition

Bolometer Array

Calibration

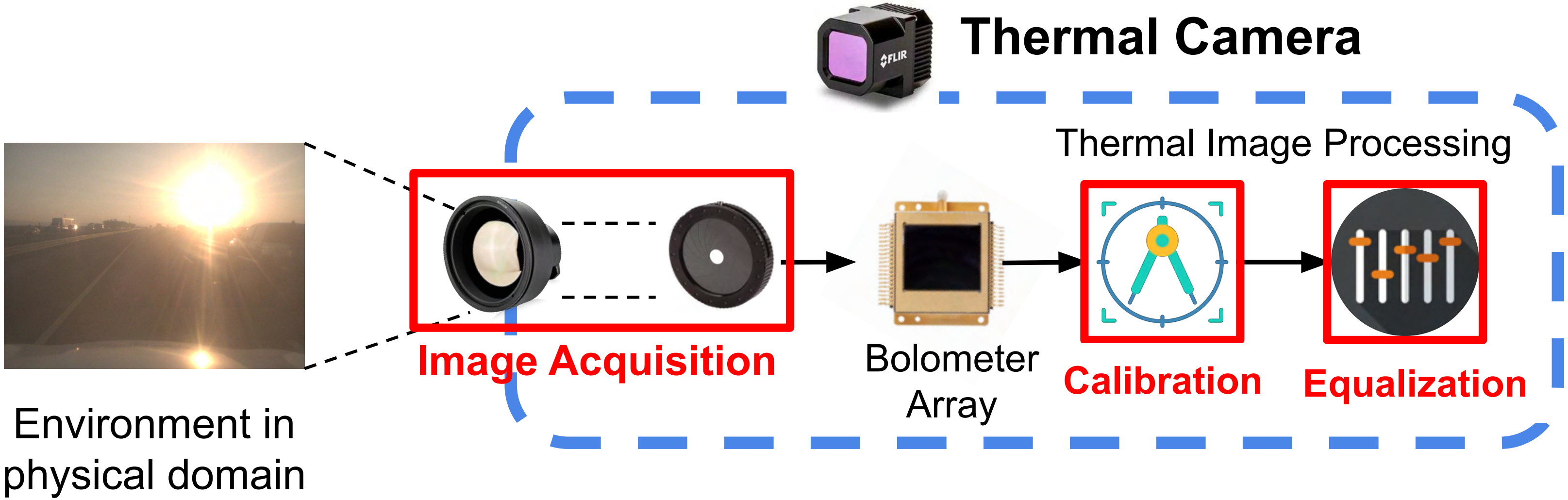
Equalization

Sensor fusion and Object detection

Thermal image with detected bounding boxes

➔ DNN object detectors + RGB image fusion for obstacle avoidance

Thermal Camera-based Perception



New vulnerabilities **unique** to thermal cameras and their signal processing

Threat Model

Goal of the adversary: Induce misdetection or detection of fake obstacles

Threat Model

Goal of the adversary: Induce misdetection or detection of fake obstacles

Generation of arbitrary shape ghost objects

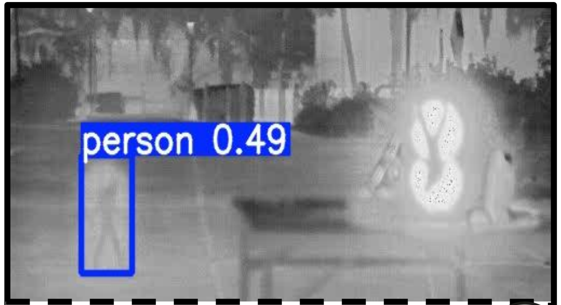


Image Acquisition

Threat Model

Goal of the adversary: Induce misdetection or detection of fake obstacles

Generation of arbitrary shape ghost objects

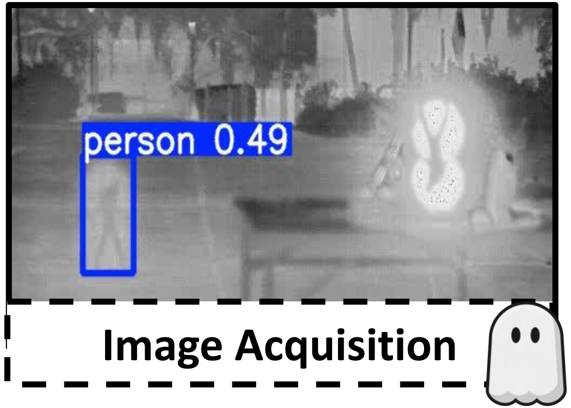
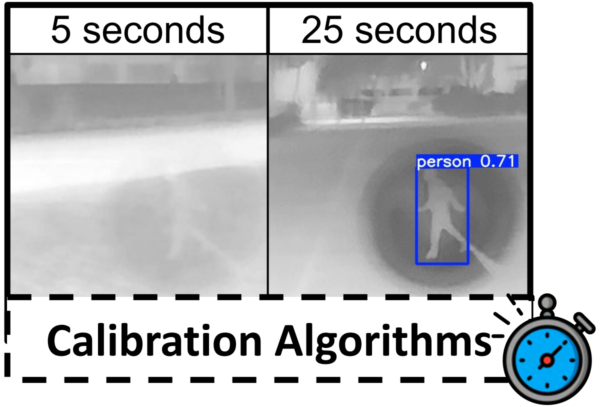


Image Acquisition

Generation of delayed fake obstacles



Calibration Algorithms

Threat Model

Goal of the adversary: Induce misdetection or detection of fake obstacles

Generation of arbitrary shape ghost objects

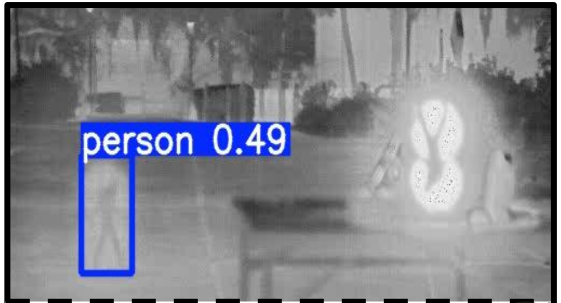
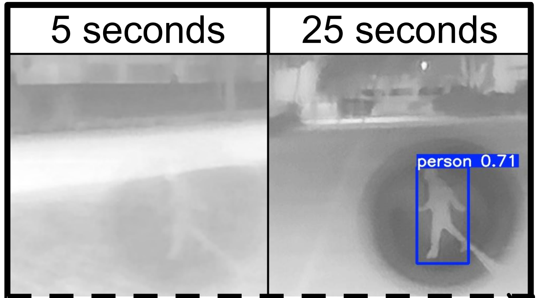


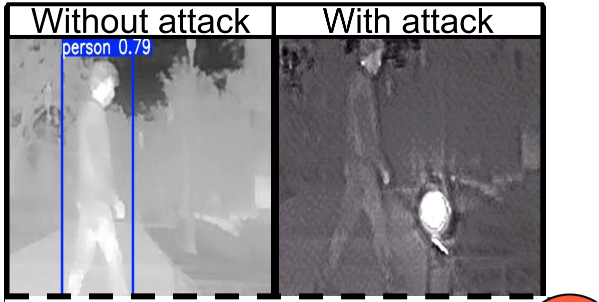
Image Acquisition

Generation of delayed fake obstacles



Calibration Algorithms

Misdetection of genuine obstacles



Equalization Algorithms

Threat Model

Goal of the adversary: Induce misdetection or detection of fake obstacles

Generation of arbitrary shape ghost objects

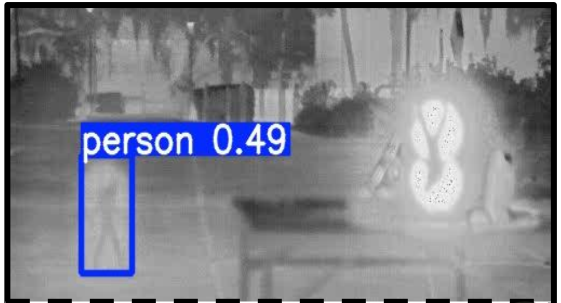
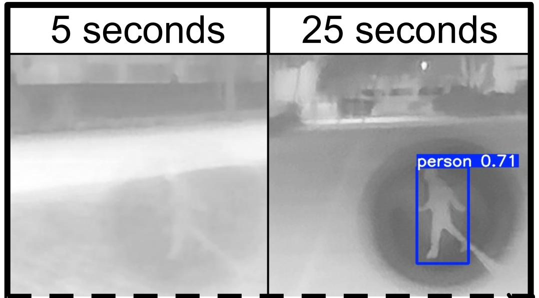


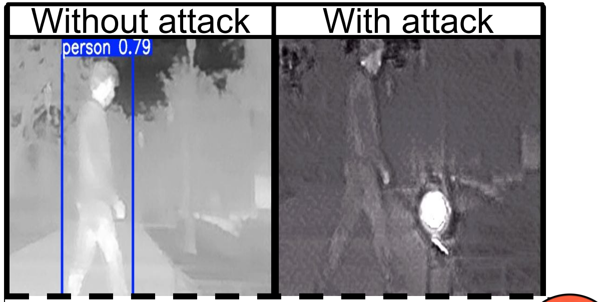
Image Acquisition

Generation of delayed fake obstacles



Calibration Algorithms

Misdetection of genuine obstacles



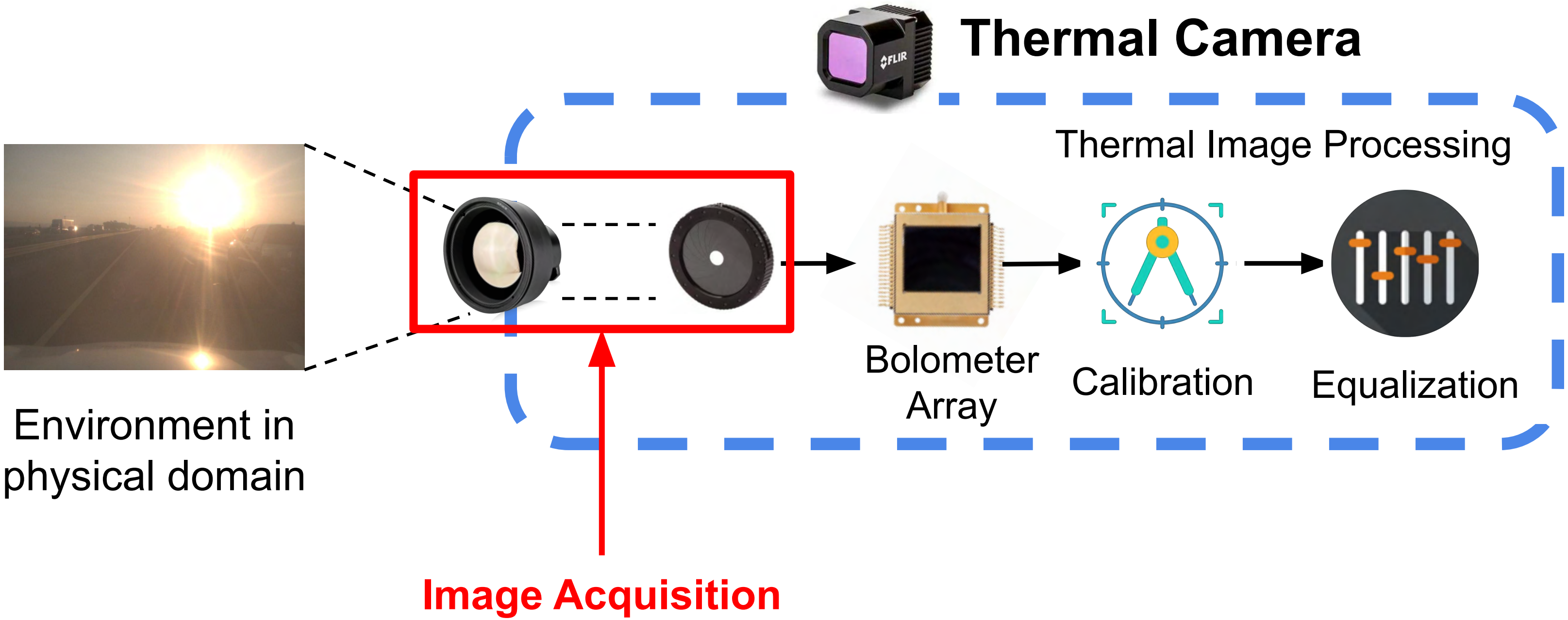
Equalization Algorithms

We assume

➔Knowledge about the thermal imaging pipeline used in the victim camera

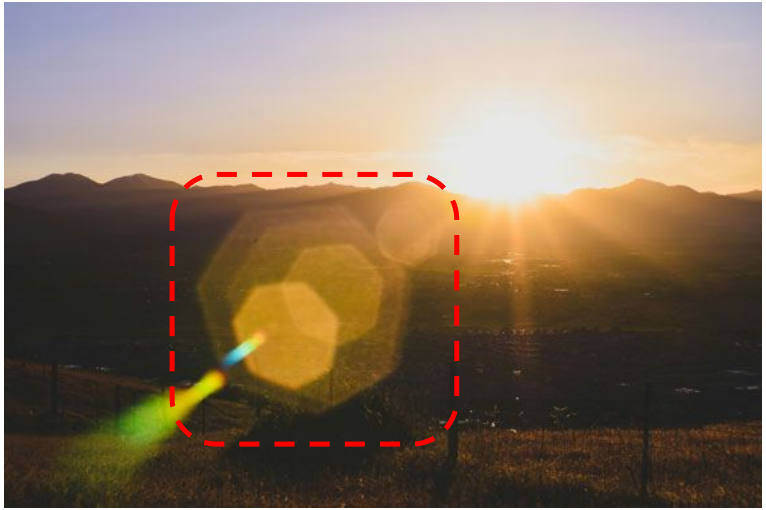
➔Control of the temperature, position, structure, and duration of the heat source

Vulnerability in Image Acquisition



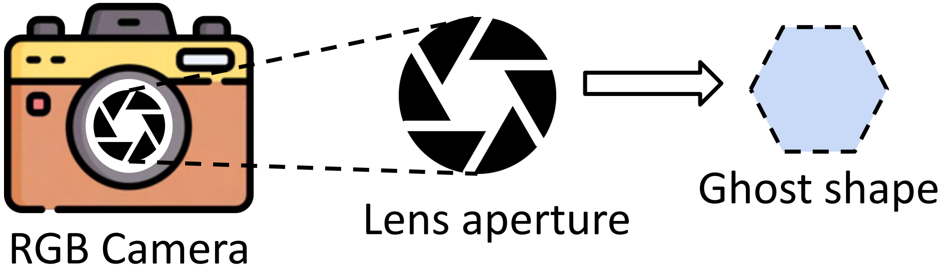
Vulnerability in Image Acquisition

Bright light reflects from the lens of regular RGB cameras, creating **Ghosts**



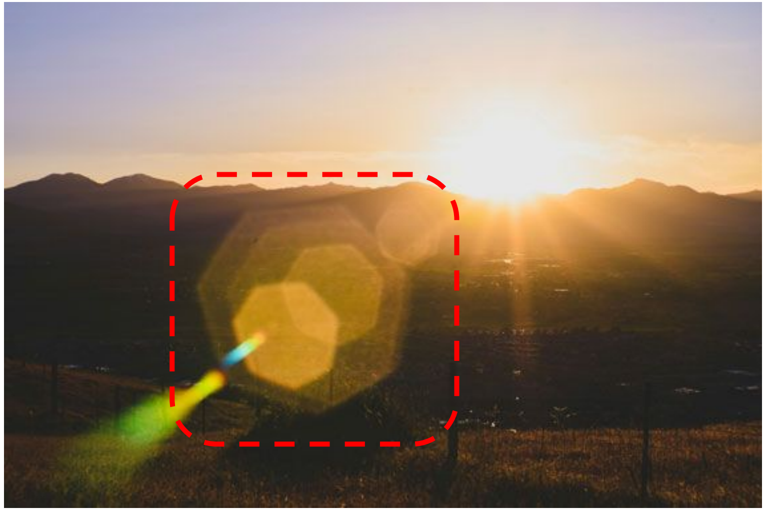
[Picturecorrect]

Ghosts take the shape of camera aperture



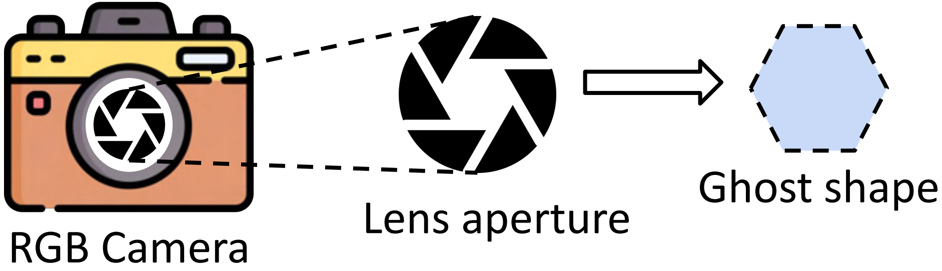
Vulnerability in Image Acquisition

Bright light reflects from the lens of regular RGB cameras, creating **Ghosts**



[Picturecorrect]

Ghosts take the shape of camera aperture



Ghosts are exploited in previous work [Man et al.]

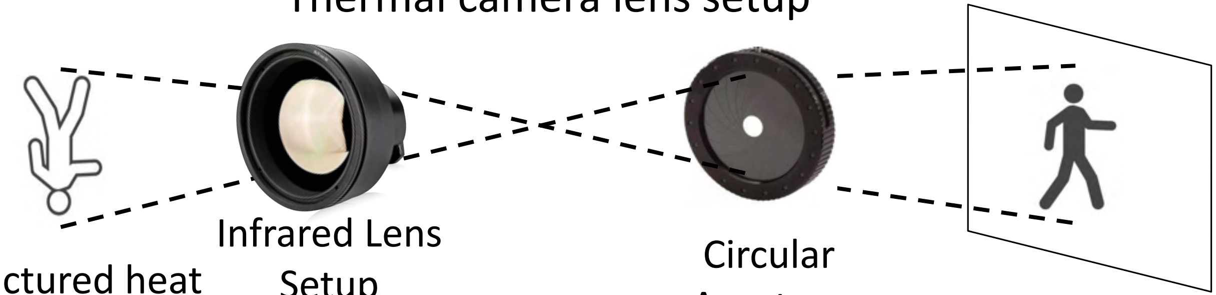
[Man et al.]

Artifacts limited to aperture shapes

Vulnerability in Image Acquisition

Circular apertures and **special materials** (germanium or zinc selenide)

Thermal camera lens setup



Structured heat source

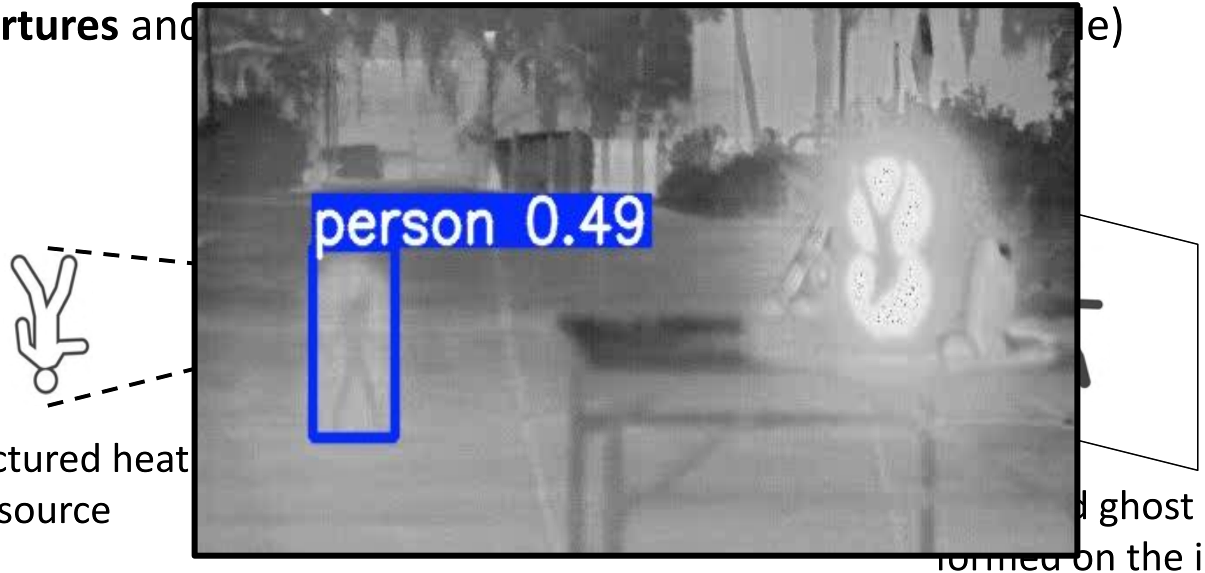
Infrared Lens Setup

Circular Aperture

Structured ghost artifact formed on the image

Vulnerability in Image Acquisition

Circular apertures and e)

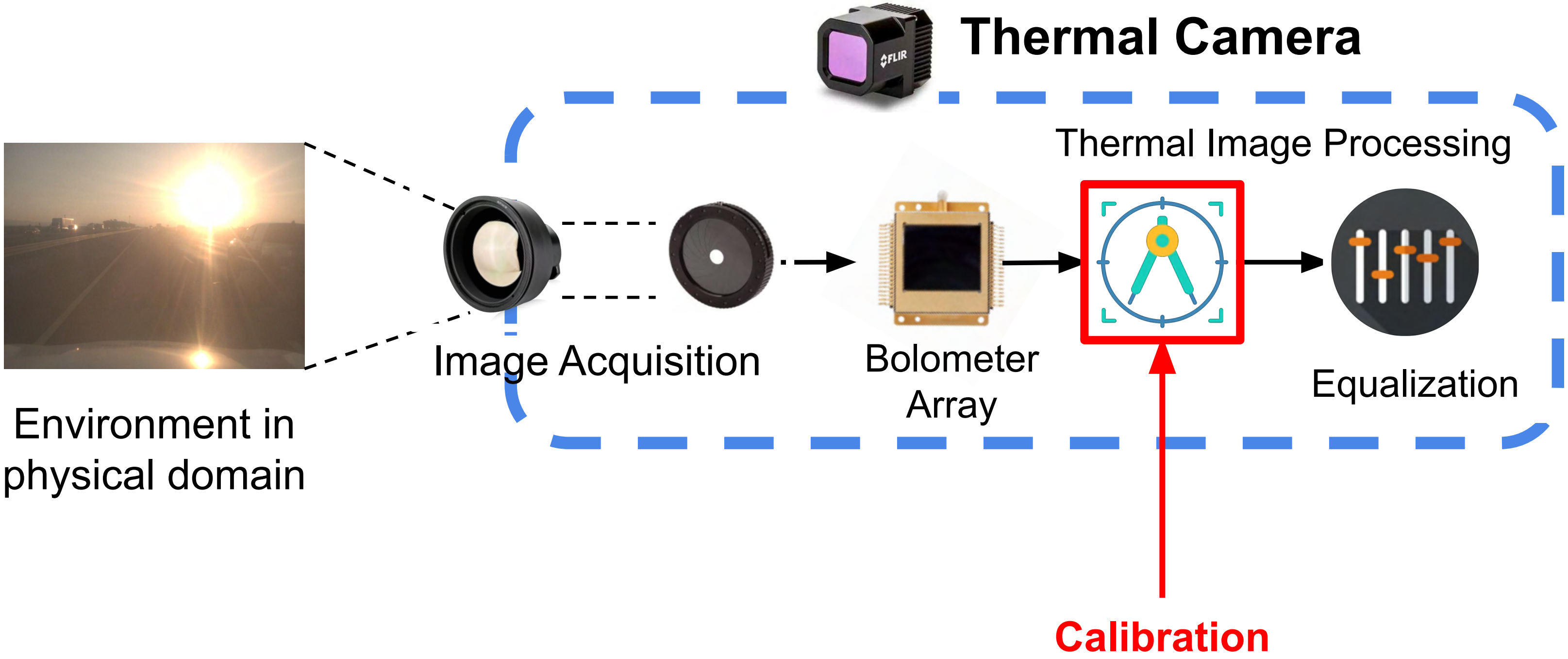


Structured heat source

d ghost artifact formed on the image

The lens setup preserve the structure of arbitrary shaped ghost artifacts = arbitrary shaped spoofed obstacle!

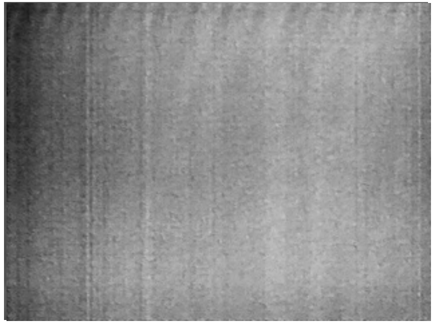
Vulnerability in Thermal Calibration



Vulnerability in Thermal Calibration

Thermal cameras are prone to **thermal noise** due to variations in sensor temperature

Image with thermal noise



Vulnerability in Thermal Calibration

Thermal cameras are prone to **thermal noise** due to variations in sensor temperature

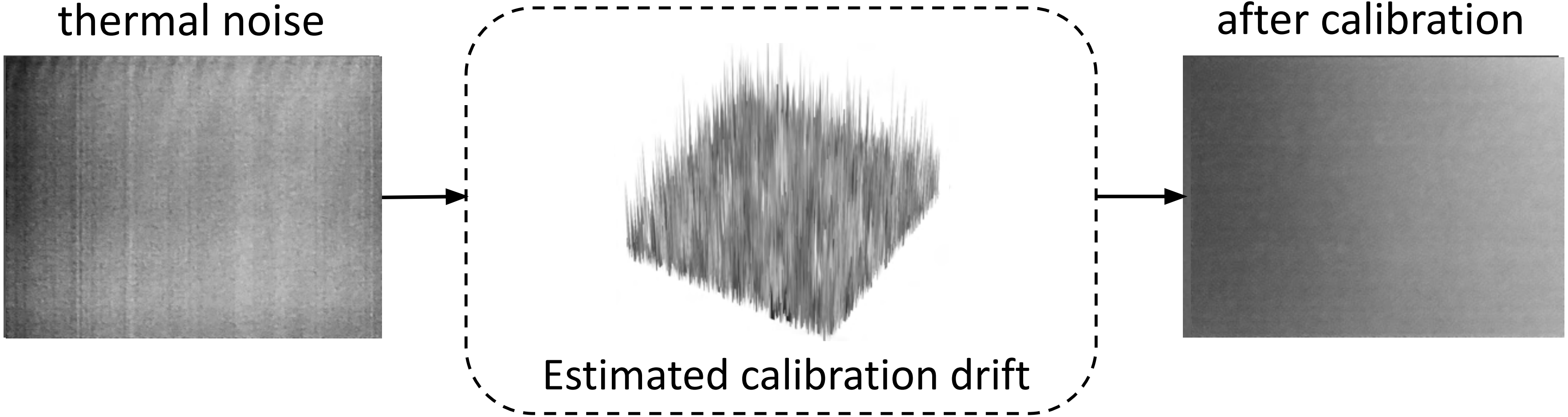


Image with thermal noise

Calibration Algorithm

Estimated calibration drift

Denoised image after calibration

Specialized **calibration algorithms** periodically denoise the images

Vulnerability in Thermal Calibration

Heat sources trigger **overcompensation** of the thermal noise

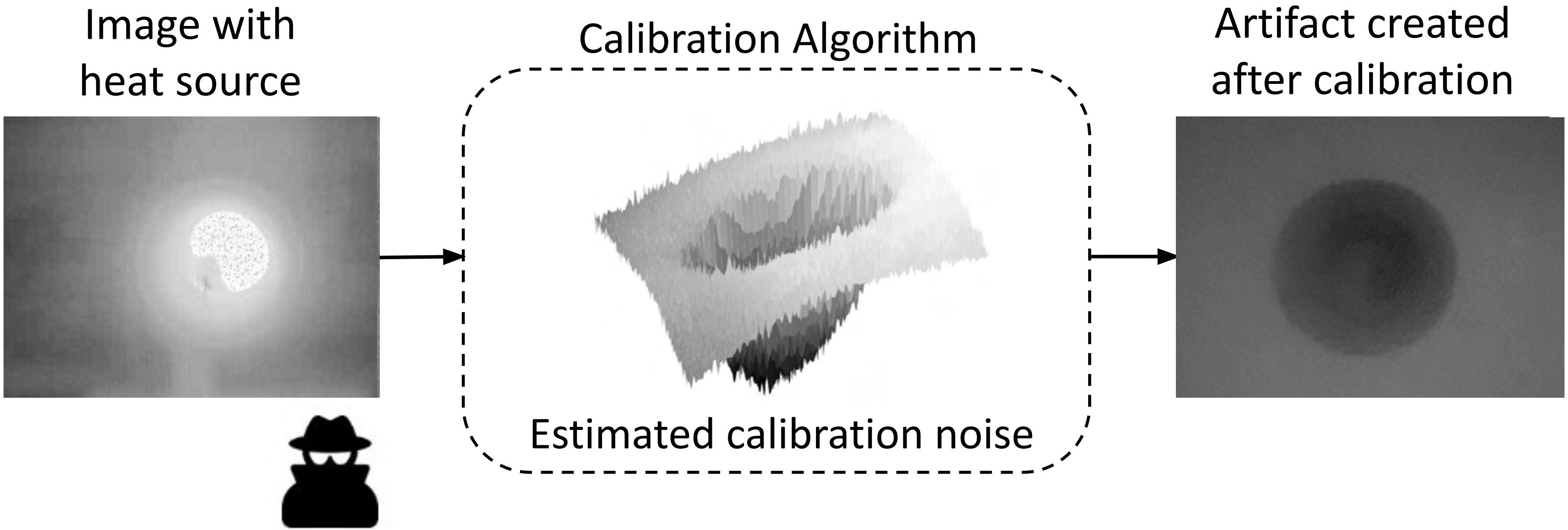


Image with heat source

Calibration Algorithm

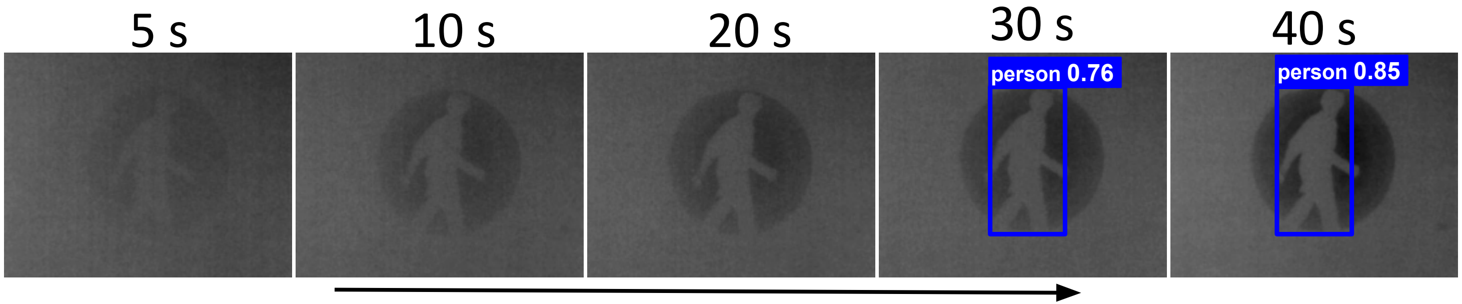
Estimated calibration noise

Artifact created after calibration

Characteristics of Calibration Artifact

Calibration artifacts contrast are amplified **even if the heat source is not present anymore**

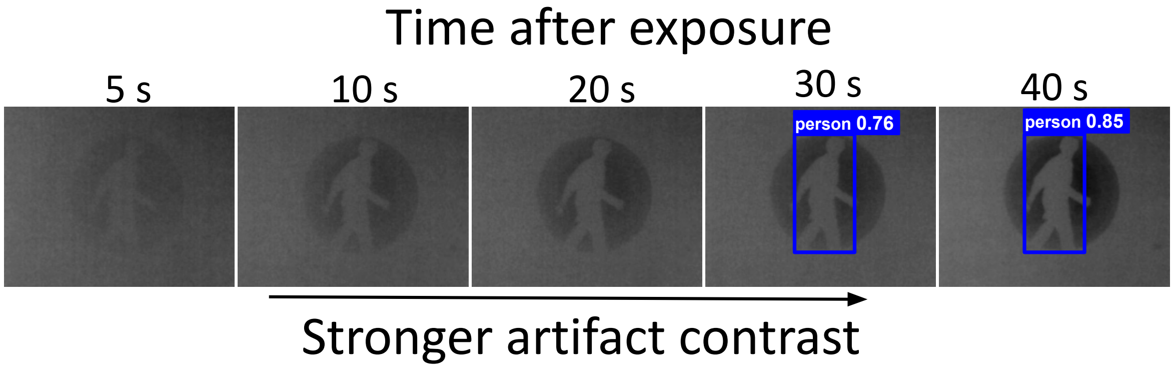
Time after exposure



Stronger artifact contrast

Characteristics of Calibration Artifact

Calibration artifacts contrast are amplified **even if the heat source is not present anymore**



Time after exposure

5 s

10 s

20 s

30 s

40 s

Stronger artifact contrast

The contrast of the artifact can be **controlled** by:

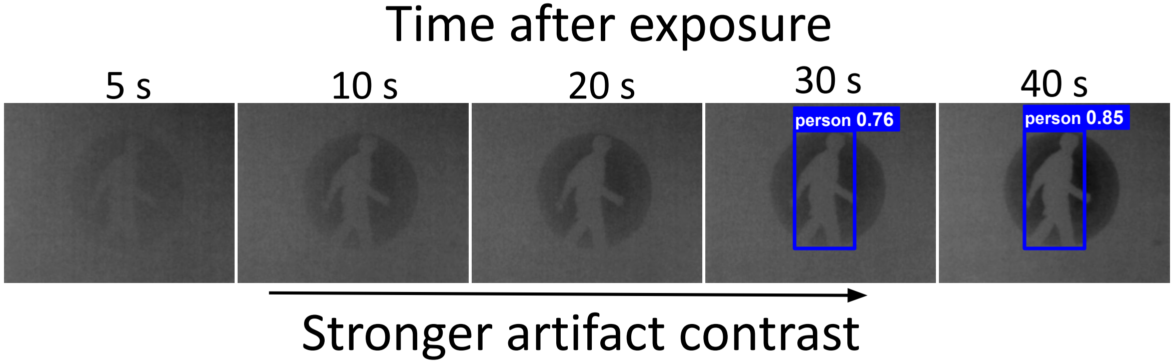
➔ **Exposure time** of the malicious heat source

➔ **Duration** of the attack

➔ **Temperature** of heat source (≅240°C)

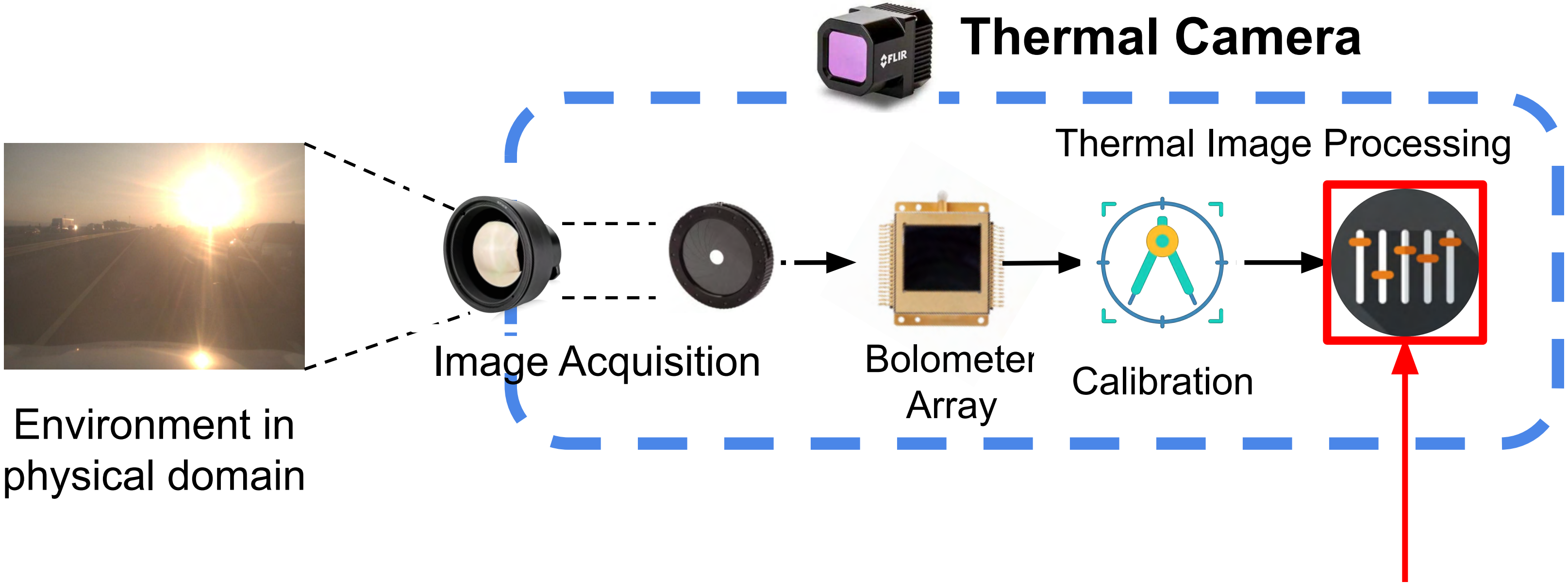
Characteristics of Calibration Artifact

Calibration artifacts contrast are amplified **even if the heat source is not present anymore**



Delayed artifacts which trigger obstacle detection after the attack is terminated!

Vulnerability in Image Equalization

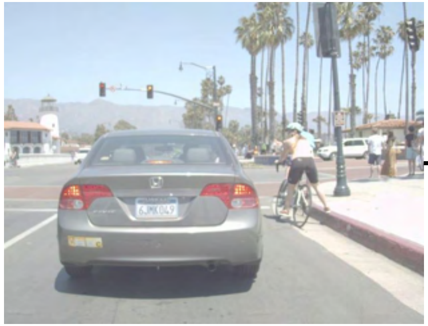


Equalization

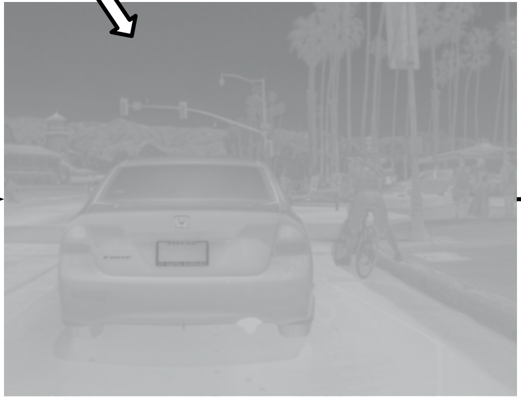
Vulnerability in Image Equalization

Narrow dynamic ranges produce low contrast images

Equalization algorithms redistribute the pixel intensities



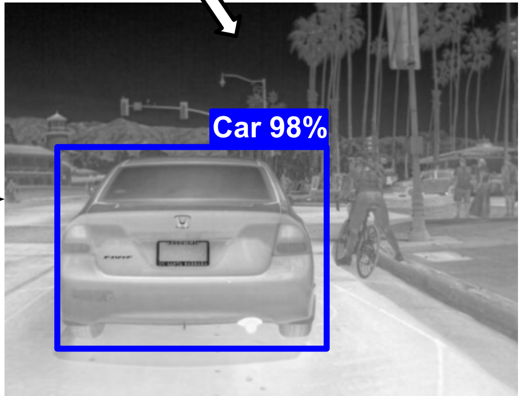
Real world scene



Thermal image before Equalization

Equalization Algorithm

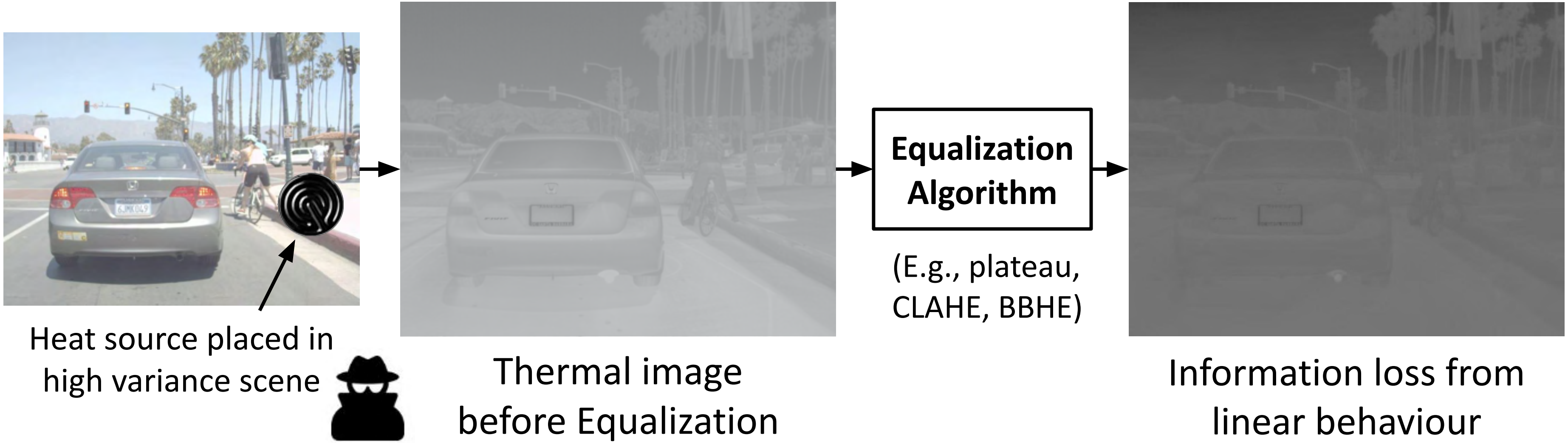
(E.g., plateau, CLAHE, BBHE)



Preserved dynamic range of the image

Vulnerability in Image Equalization

Scenes with **high thermal contrast** trigger **linear behaviour** and **information loss**



Heat source placed in high variance scene

Thermal image before Equalization

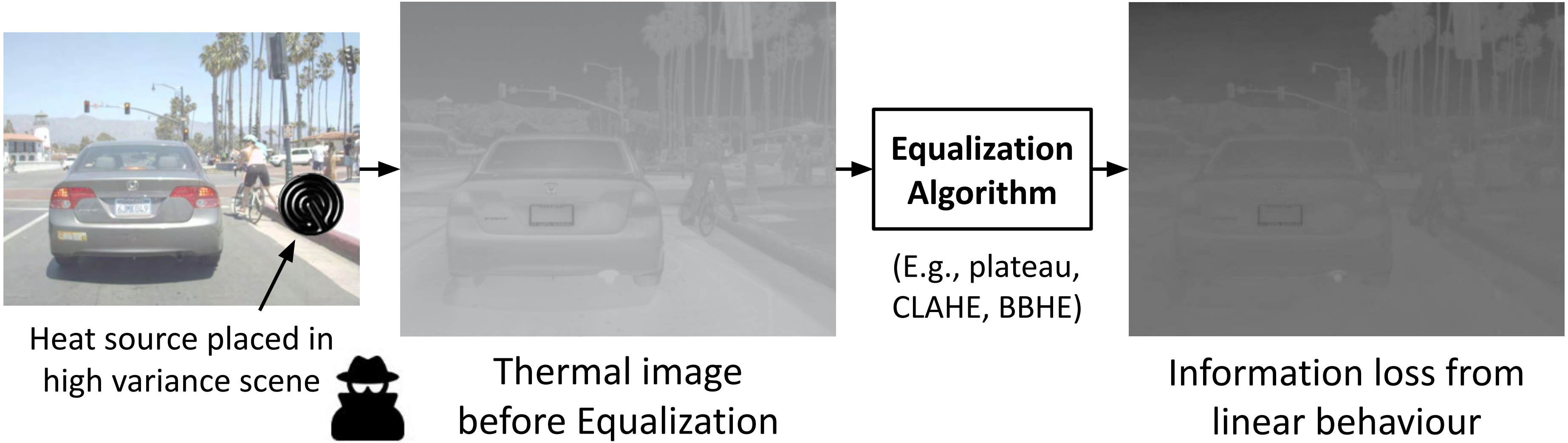
Equalization Algorithm

(E.g., plateau, CLAHE, BBHE)

Information loss from linear behaviour

Vulnerability in Image Equalization

Scenes with **high thermal contrast** trigger **linear behaviour** and **information loss**



Heat source placed in high variance scene

Thermal image before Equalization

Equalization Algorithm

(E.g., plateau, CLAHE, BBHE)

Information loss from linear behaviour

Heat sources to trigger linear behaviour and consequent misdetection!

Experimental Setup

Standard **reptile heating lamp** as a heat source (240°C)

- ➔ Cheap and accessible (< 20$) means for attack
- ➔ Heat radiation is invisible to human eye

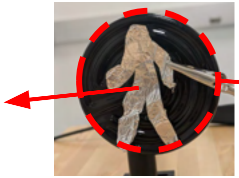
Experimental Setup

Standard **reptile heating lamp** as a heat source (240°C)

→ Cheap and accessible (< 20$) means for attack

→ Heat radiation is invisible to human eye

Structure heat source using aluminium foil to create **arbitrary shaped artifacts**



Structured aluminium foil

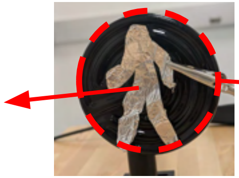
Heat source (reptile heating lamp)

Experimental Setup

Standard **reptile heating lamp** as a heat source (240°C)

- ➔ Cheap and accessible (< 20$) means for attack
- ➔ Heat radiation is invisible to human eye

Structure heat source using aluminium foil to create **arbitrary shaped artifacts**



Structured aluminium foil

Heat source (reptile heating lamp)

Target three commercial thermal cameras

Automotive

FLIR Boson

Robotics

InfiRay T2S

Drones

FPV XK-C130

Evaluation

- ➔ Characterize the pixel properties of the artifacts
- ➔ Synthesize the artifacts on the thermal images of **FLIR dataset**

Evaluation

- ➔ Characterize the pixel properties of the artifacts
- ➔ Synthesize the artifacts on the thermal images of **FLIR dataset**

Obstacle Detection

Models:

- ➔ YOLOv5
- ➔ YOLOv8
- ➔ FasterRCNN

Sensor Fusion (RGB + Thermal)

Models:

- ➔ DAMSDet
- ➔ FasterRCNN

Evaluation

- ➔ Characterize the pixel properties of the artifacts
- ➔ Synthesize the artifacts on the thermal images of **FLIR dataset**

Obstacle Detection

Models:
- ➔ YOLOv5
- ➔ YOLOv8
- ➔ FasterRCNN

Sensor Fusion (RGB + Thermal)

Models:
- ➔ DAMSDet
- ➔ FasterRCNN

- ➔ Evaluate artifacts with different pedestrian poses (no adversarial optimization)

Front face

Running

Crossing

Evaluation of Ghost Attack

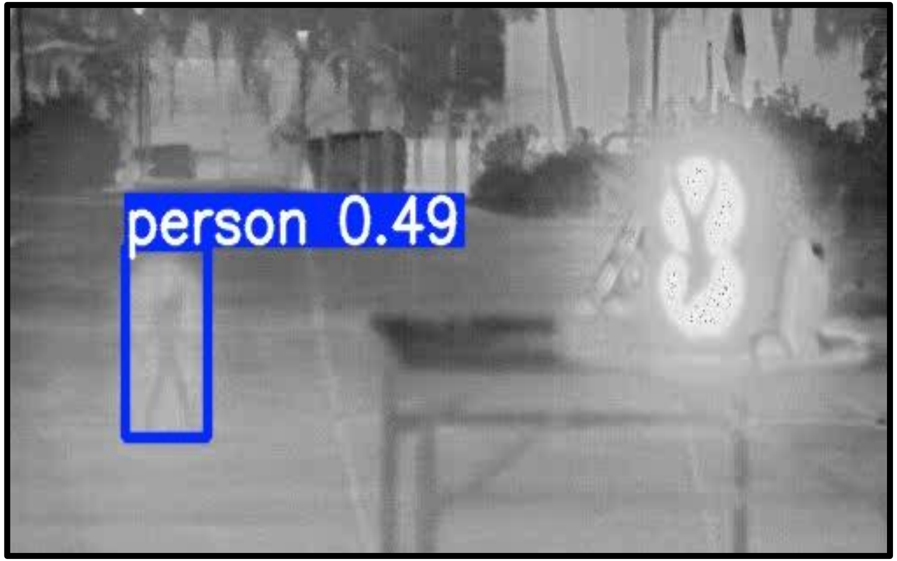
Ghost Attack

Metric: **Attack Success Rate**

Results:

→ Object Detection: > **98%**

→ Sensor Fusion: > **91%**



→ Heat source temperature of **80°C** is sufficient to achieve attack success rate **> 90%**

→ Attack success rate is high regardless of the artifact pose

Evaluation of Calibration Attack

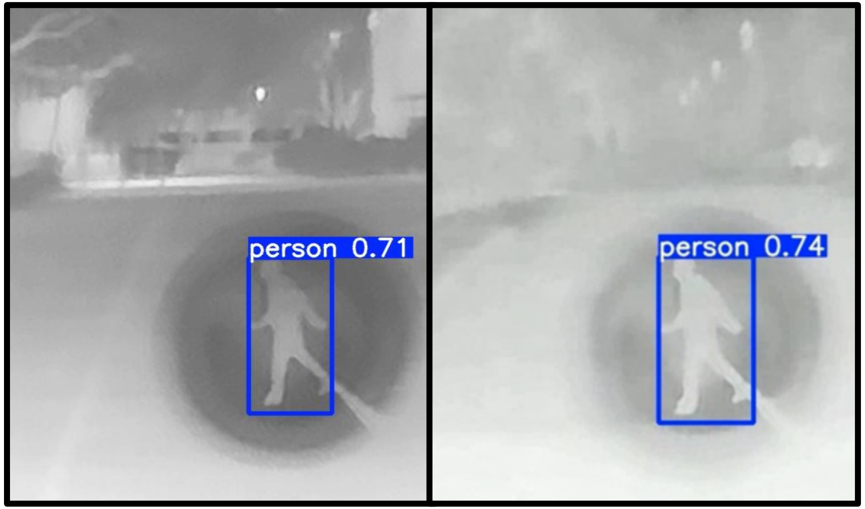
Calibration Attack

Metric: **Attack Success Rate**

Results:

➔ Object Detection: > **81%**

➔ Sensor Fusion: > **56%**



➔ A **30 second exposure** of heat source causes attack success rate **> 90%**

➔ Human poses with high structural features show higher attack success rates

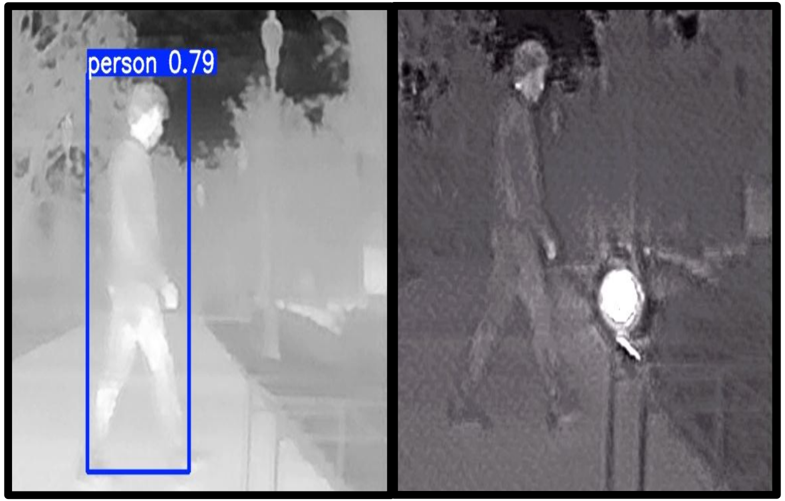
Evaluation of Equalization attack

Equalization Attack

Metric: **Mean Average Precision**

Results:

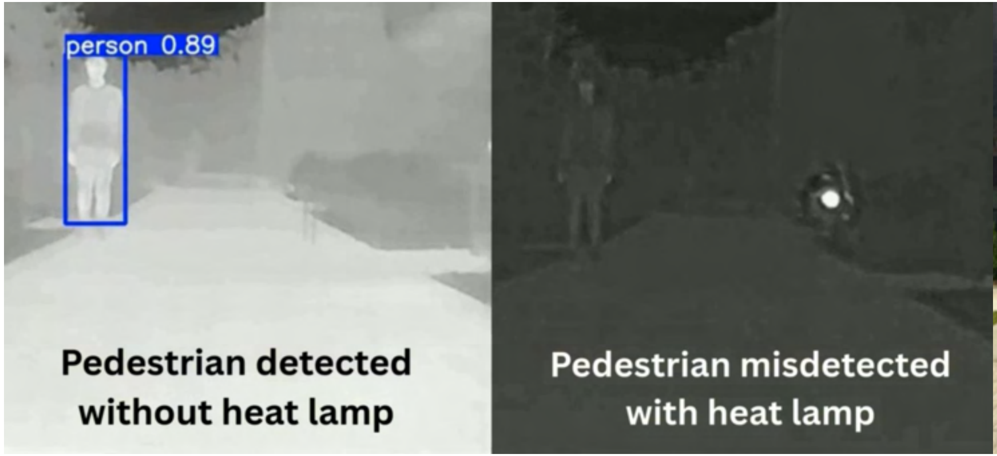
- ➔ Object Detection: < **0.09**
- ➔ Sensor Fusion: < **0.63**



- ➔ The model performance drop linearly with temperature
- ➔ Attack is stronger against pedestrian obstacles

Real World Driving Scenarios

Equalization Attack: Vehicle driving towards target pedestrian at **40 km/h** from 50 m away


person 0.89
Pedestrian detected without heat lamp
Pedestrian misdetected with heat lamp

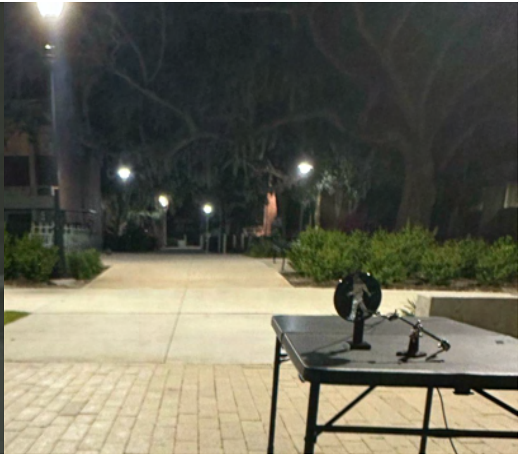
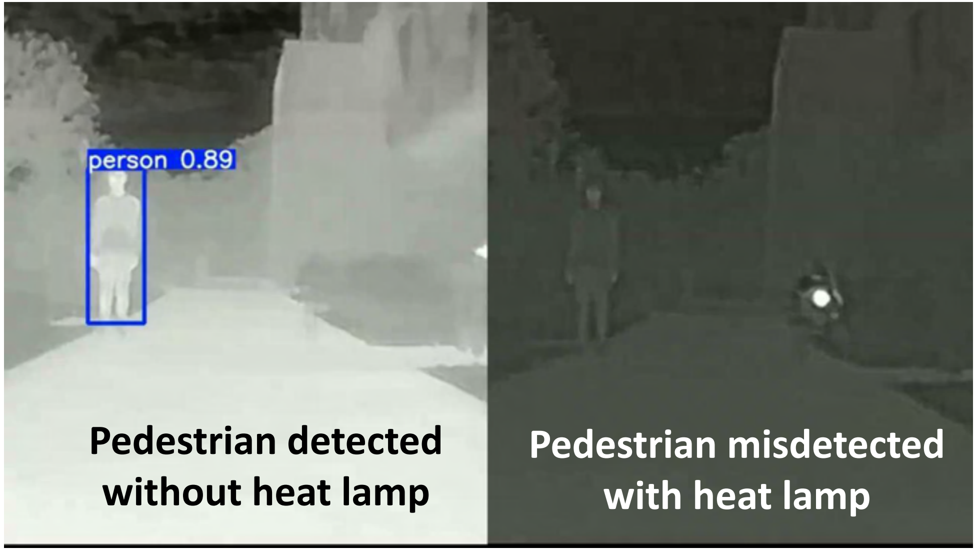
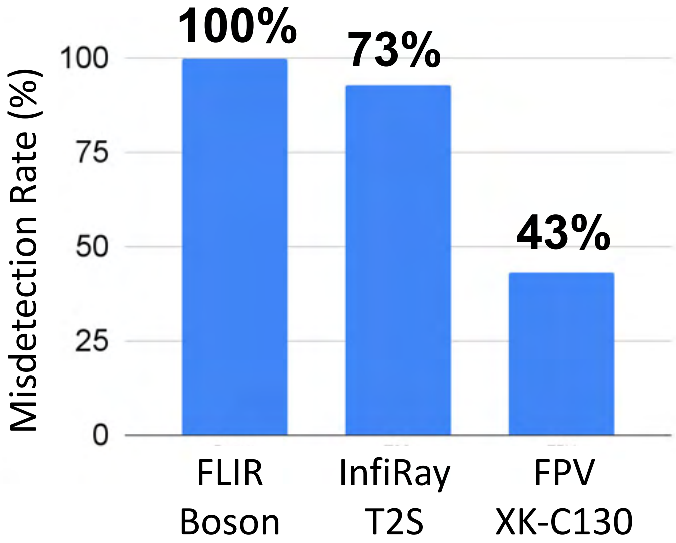


Image from RGB camera
with heat source ON

Real World Driving Scenarios

Equalization Attack: Vehicle driving towards target pedestrian at **40 km/h** from 50 m away

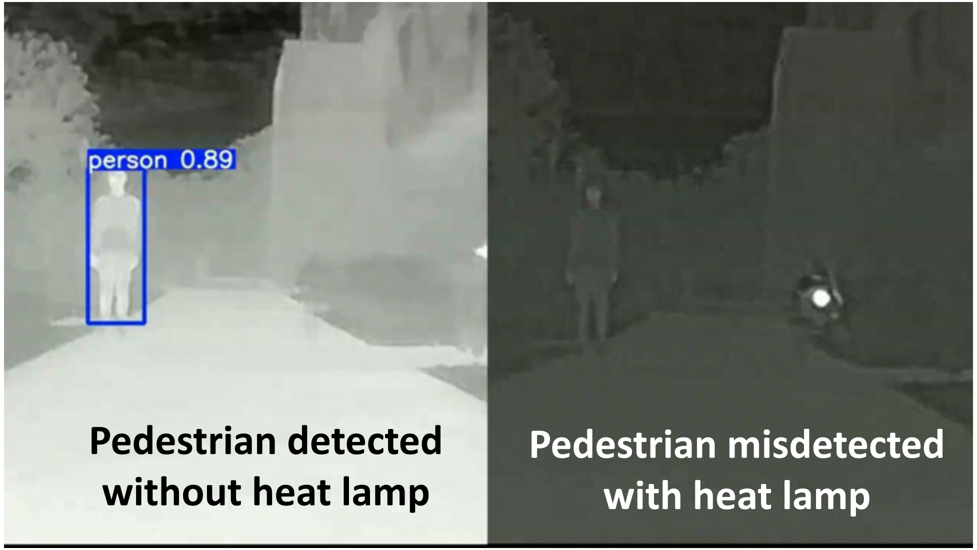

person 0.89
Pedestrian detected without heat lamp
Pedestrian misdetected with heat lamp



Attack is effective in realistic high speed scenarios

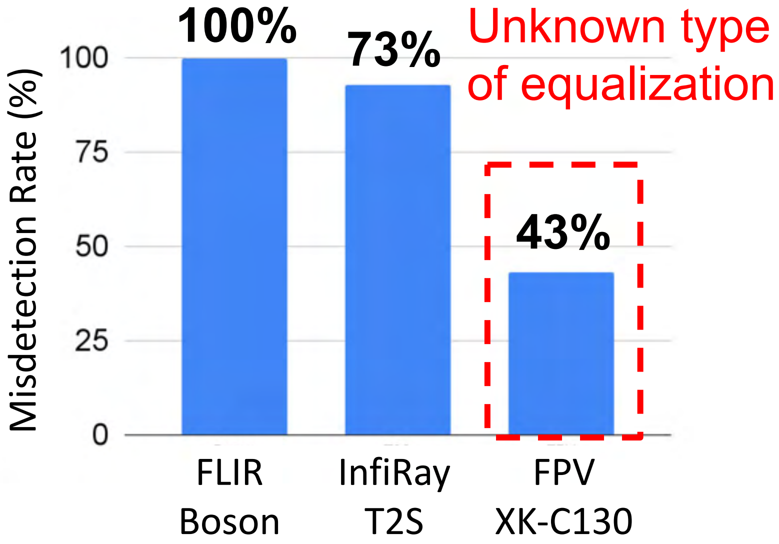
Real World Driving Scenarios

Equalization Attack: Vehicle driving towards target pedestrian at **40 km/h** from 50 m away



Pedestrian detected without heat lamp

Pedestrian misdetected with heat lamp



Unknown type of equalization

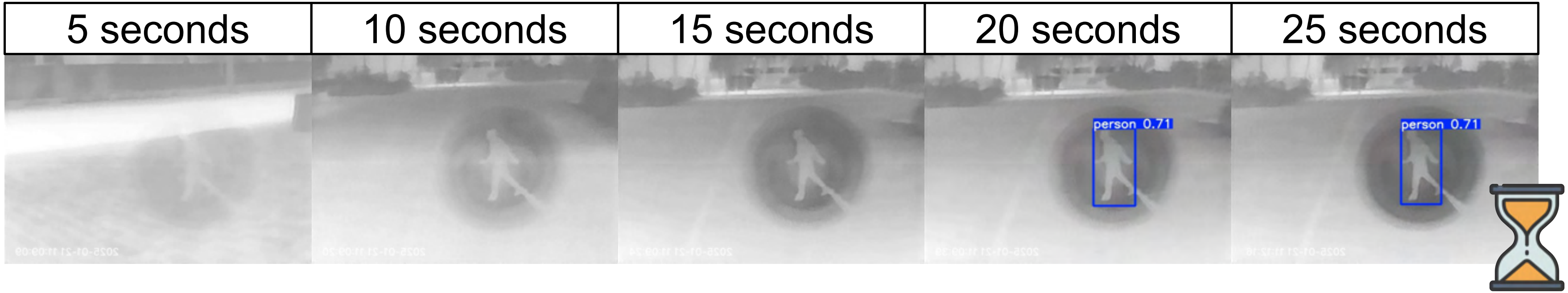
Real World Driving Scenarios

Calibration Attack: Vehicle drives at speeds of **40 km/h** after the attack

Real World Driving Scenarios

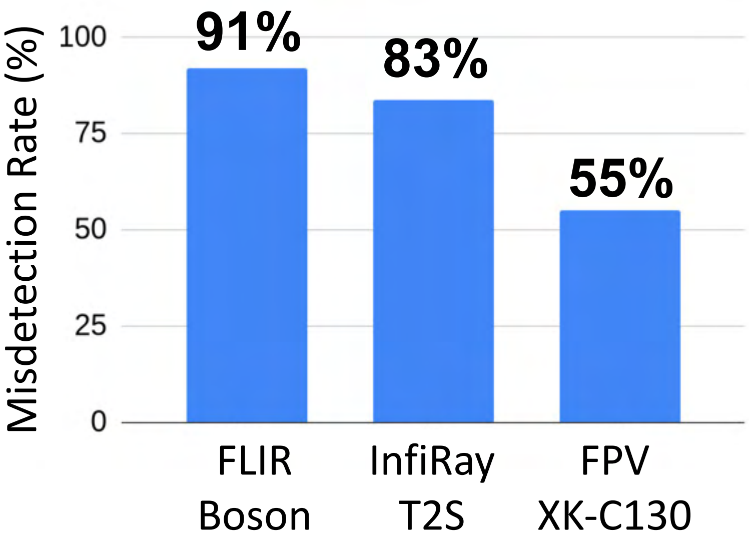
Calibration Attack: Vehicle drives at speeds of **40 km/h** after the attack

- Attack can trigger detection with **10-60 seconds** of exposure
- Attack induced detection from **10 seconds** up to **2 minutes** after exposure depending on the camera calibration period


5 seconds | 10 seconds | 15 seconds | 20 seconds | 25 seconds

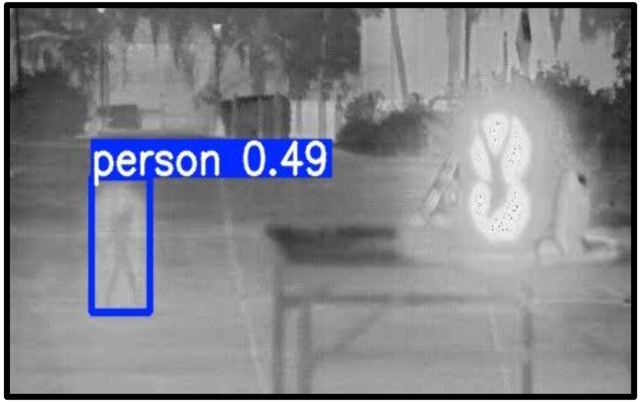
Real World Driving Scenarios

Calibration Attack: Vehicle drives at speeds of **40 km/h** after the attack



Real World Driving Scenarios

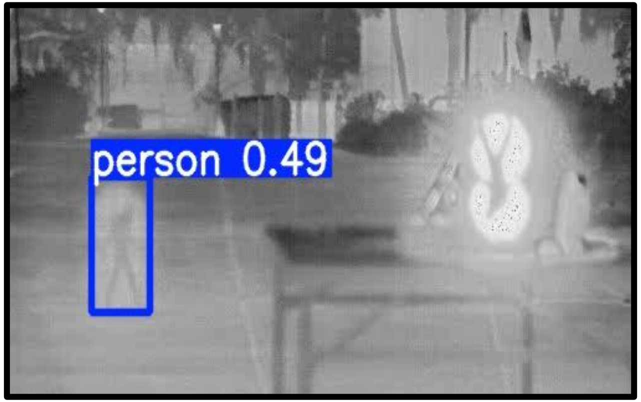
Ghost Attack: Vehicle approaches heat lamp at speed of **2.5 km/h** from 2.5 meters from the **FLIR Boson** camera



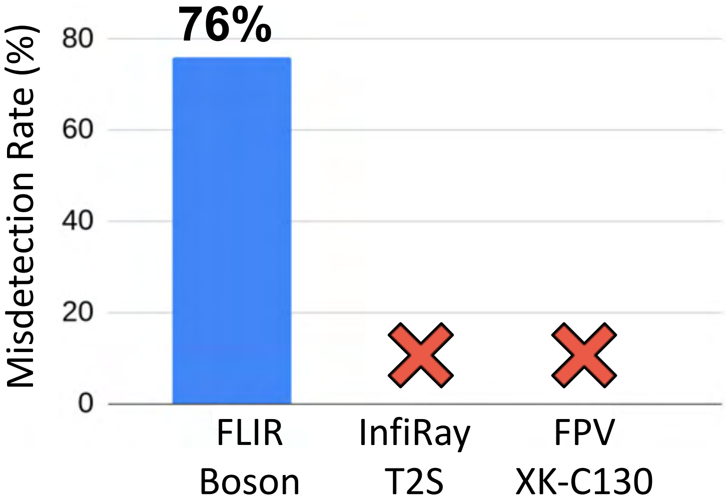
Creation of arbitrary shaped artifacts

Real World Driving Scenarios

Ghost Attack: Vehicle approaches heat lamp at speed of **2.5 km/h** from 2.5 meters from the **FLIR Boson** camera



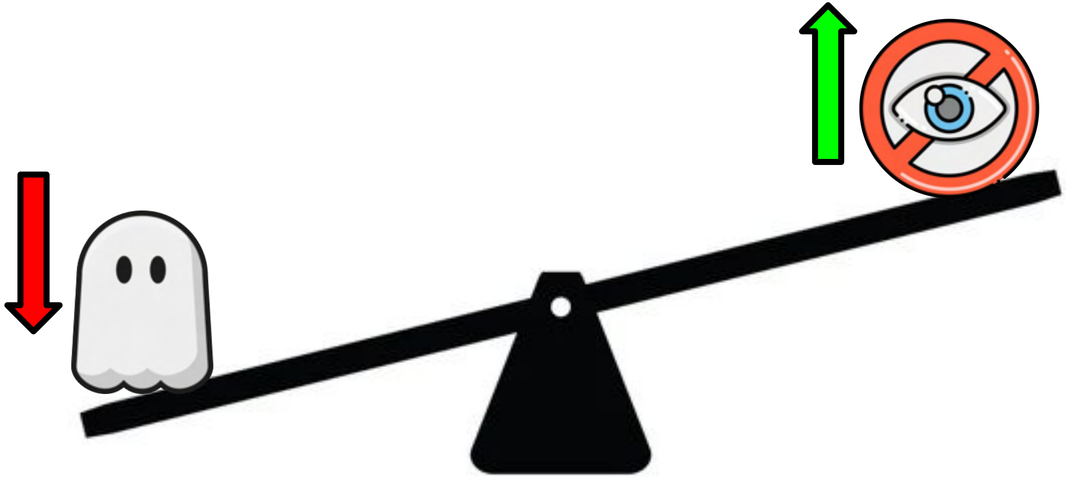
Creation of arbitrary shaped artifacts



Ghost artifacts in InfiRay T2S and FPV XK-C130 are **not strong enough** to induce detection

Interdependent Vulnerabilities

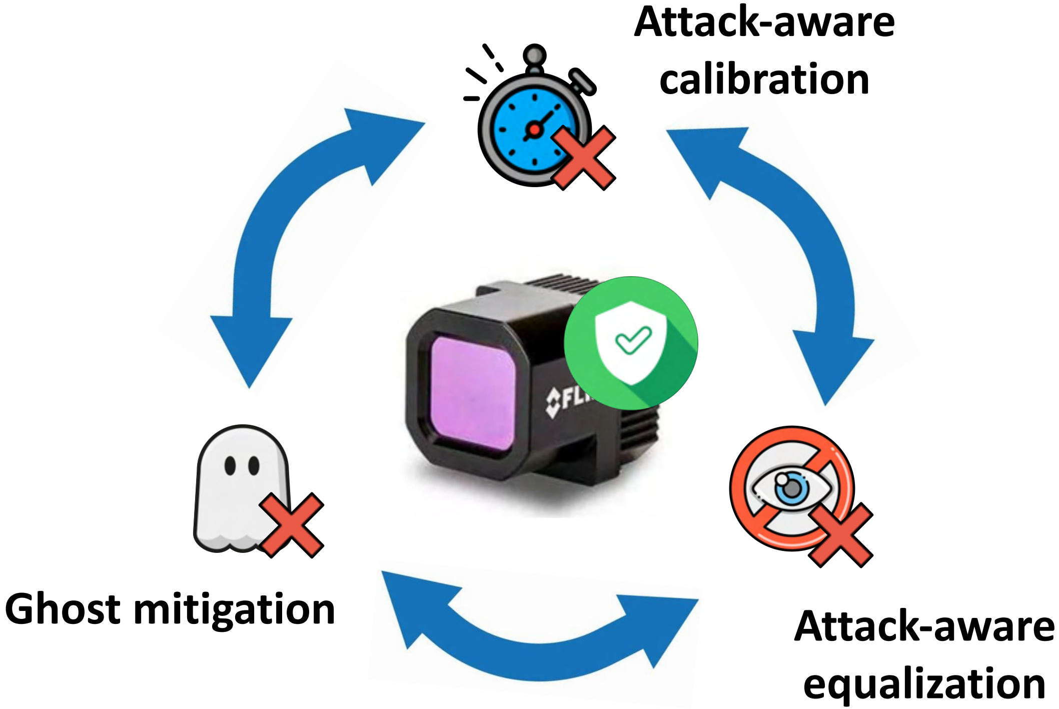
Inverse relation between Equalization attack and Ghost attack



- ➔ Drop in pixel intensity from equalization attack causes drop in ghost artifact intensity
- ➔ Calibration attack is proportional to the pixel intensity of equalization attack
- ➔ Individually addressing one vulnerability, in turn, strengthens the other

Attack-aware Signal Processing

Design signal processing algorithms that mitigate the attacks simultaneously

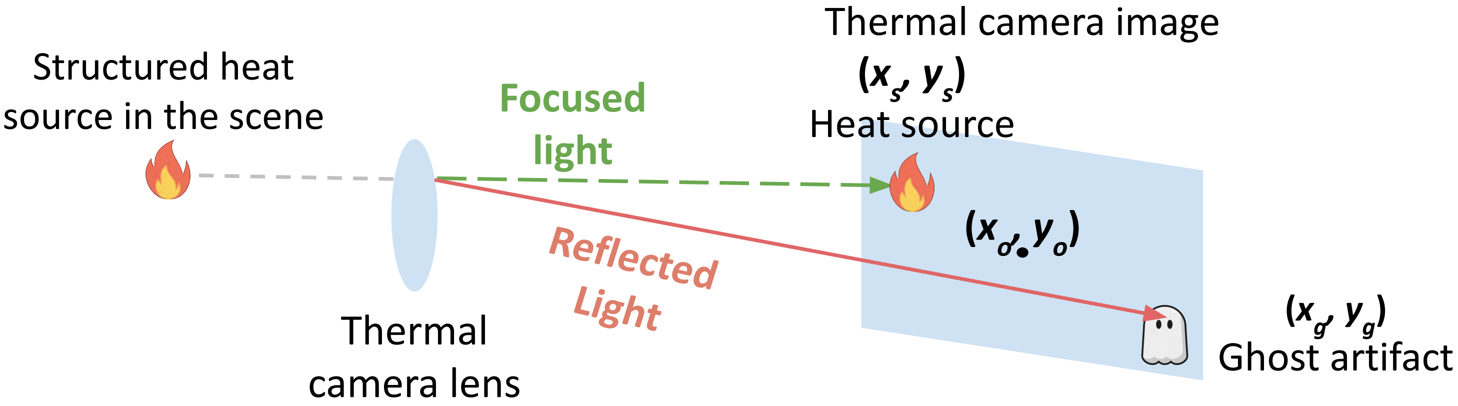


Ghost Mitigation

➔ Lens hoods used to suppress ghosts are not suitable for automotive applications

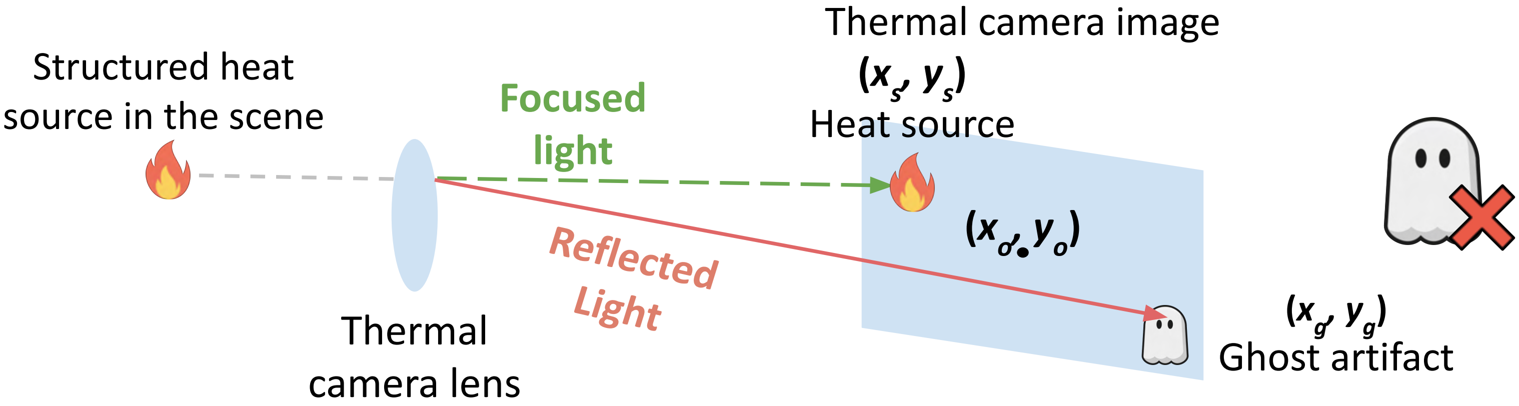
Ghost Mitigation

- ➔ Lens hoods used to suppress ghosts are not suitable for automotive applications
- ➔ We leverage the geometric properties of thermal lenses to locate ghosts



Ghost Mitigation

- ➔ Lens hoods used to suppress ghosts are not suitable for automotive applications
- ➔ We leverage the geometric properties of thermal lenses to locate ghosts



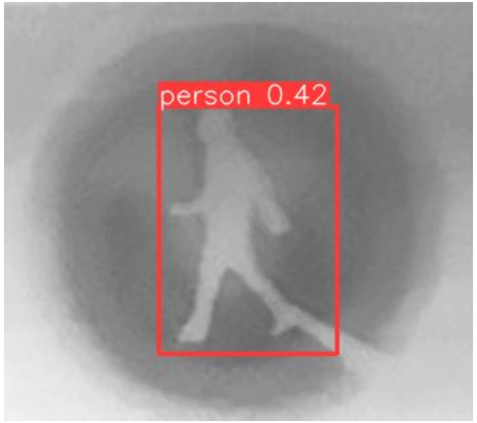
100% success rate in detecting heat sources (>80°C) and suppress the ghosts

Attack Aware Calibration

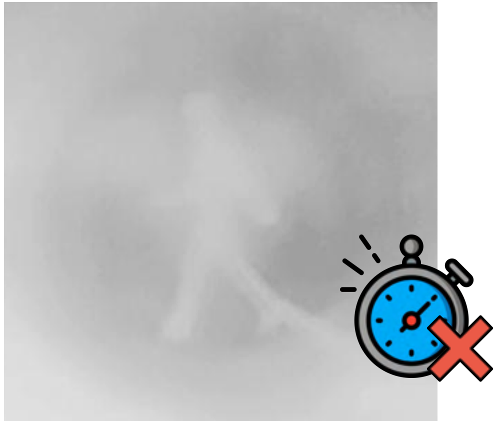
➔ Calibration artifacts are caused by overcompensation of thermal drift

Attack Aware Calibration

- ➔ Calibration artifacts are caused by overcompensation of thermal drift
- ➔ Our attack-aware algorithm limits the offset to a calculated threshold
- ➔ Thresholds calculated based on camera specific attack requirements



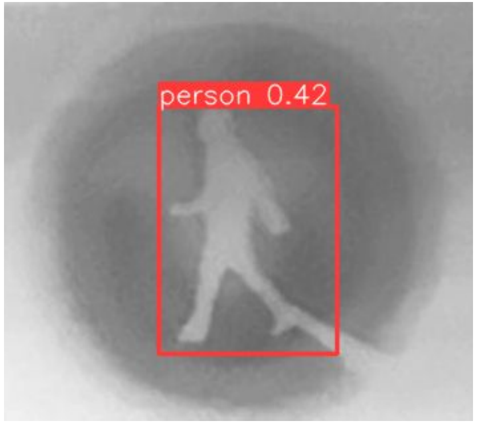
With Attack



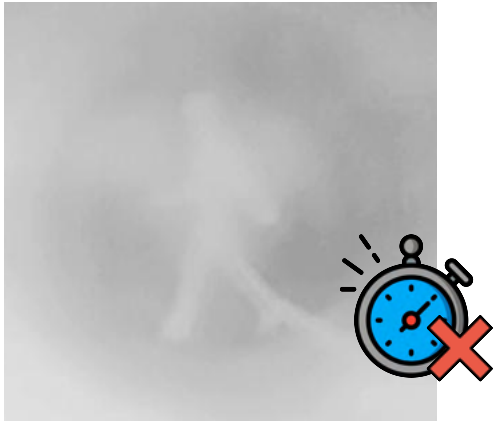
Our defense

Attack Aware Calibration

- ➔ Calibration artifacts are caused by overcompensation of thermal drift
- ➔ Our attack-aware algorithm limits the offset to a calculated threshold
- ➔ Thresholds calculated based on camera specific attack requirements



With Attack



Our defense

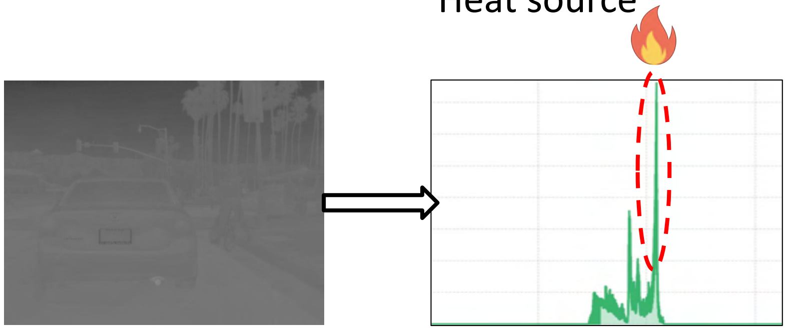
100% success rate in suppressing calibration artifacts

Obstacle detection accuracy remains steady at 92%, with **no performance degradation**

Attack Aware Equalization

➔ Heat sources triggering linear behaviour create spikes in image histograms

Heat source

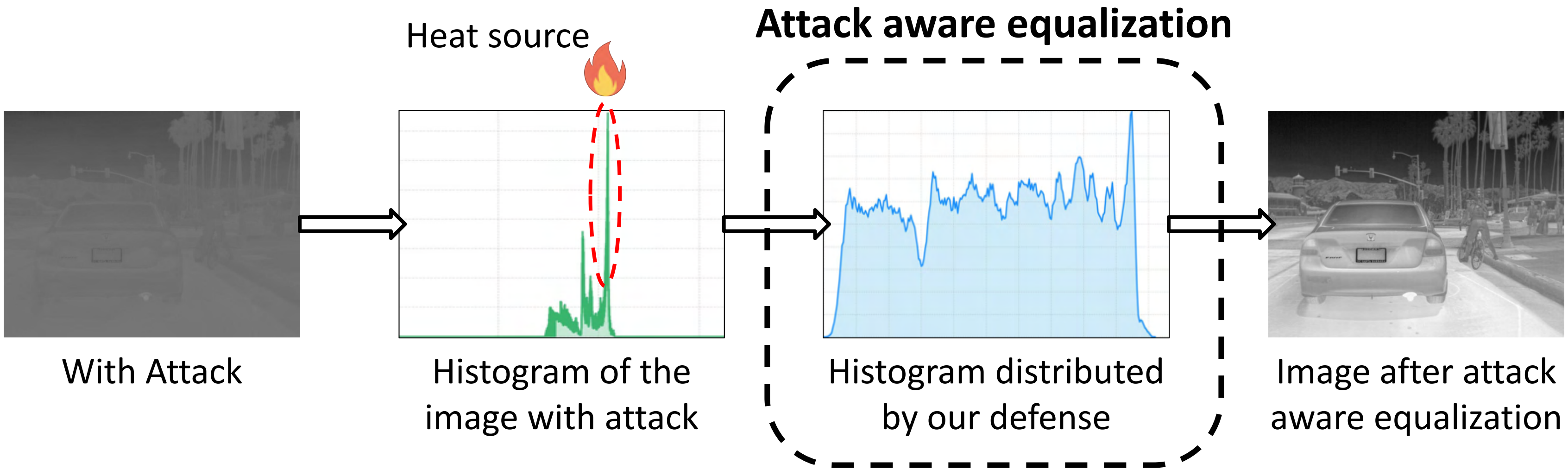


With Attack

Histogram of the image with attack

Attack Aware Equalization

- ➔ Heat sources triggering linear behaviour create spikes in image histograms
- ➔ Our methodology detects and excludes such spikes in histograms



Heat source

Attack aware equalization

With Attack

Histogram of the image with attack

Histogram distributed by our defense

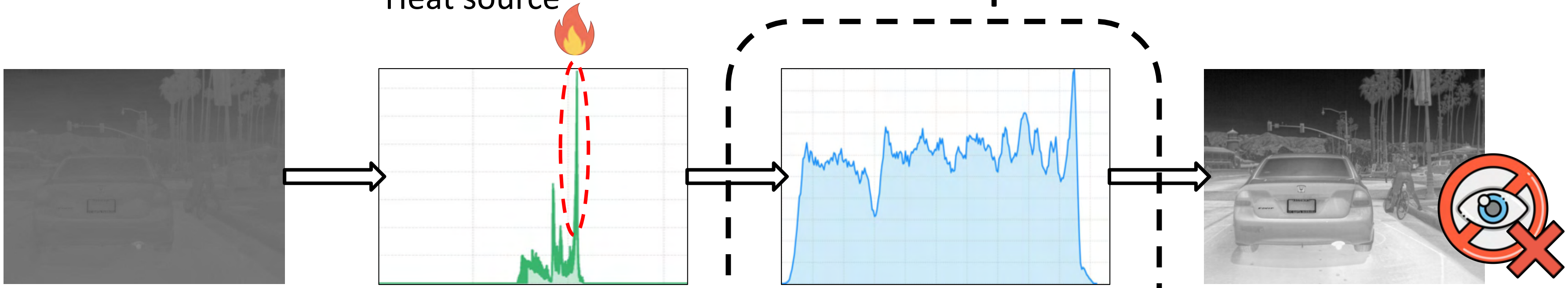
Image after attack aware equalization

Attack Aware Equalization

- ➔ Heat sources triggering linear behaviour create spikes in image histograms
- ➔ Our methodology detects and excludes such spikes in histograms

Attack aware equalization

Heat source



With Attack

Histogram of the image with attack

Histogram distributed by our defense

Image after attack aware equalization

Increase in pedestrian detection rate under attack from **0% to 96%**

Takeaways

Uncover three new vulnerabilities in thermal imaging pipelines, namely in **image acquisition**, **calibration** and **equalization algorithms**

Takeaways

Uncover three new vulnerabilities in thermal imaging pipelines, namely in **image acquisition**, **calibration** and **equalization algorithms**

Demonstrate the attacks on three commercial thermal cameras used in autonomous systems; **FLIR Boson, InfiRay T2S, FPV XK-C130**

Takeaways

Uncover three new vulnerabilities in thermal imaging pipelines, namely in **image acquisition**, **calibration** and **equalization algorithms**

Demonstrate the attacks on three commercial thermal cameras used in autonomous systems; **FLIR Boson, InfiRay T2S, FPV XK-C130**

Driving experiments with vehicle reaching **40 km/h** show **100%** misdetection and **91%** fake obstacle detection

Takeaways

Uncover three new vulnerabilities in thermal imaging pipelines, namely in **image acquisition**, **calibration** and **equalization algorithms**

Demonstrate the attacks on three commercial thermal cameras used in autonomous systems; **FLIR Boson, InfiRay T2S, FPV XK-C130**

Driving experiments with vehicle reaching **40 km/h** show **100%** misdetection and **91%** fake obstacle detection

Densign novel **attack aware signal processing** techniques that effectively mitigate the consequences of the attacks in real-time

Takeaways

Uncover three new vulnerabilities in thermal imaging pipelines, namely in **image acquisition**, **ca**

De … eras used in … **V XK-C130**

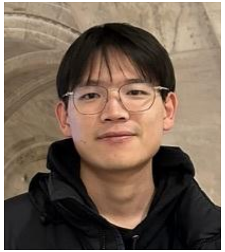
Disclosed vulnerabilities to corresponding vendors ✓

Driving experim … nd **91%** fake obsta

Densign novel **attack aware signal processing** techniques that effectively mitigate the consequences of the attacks in real-time

Thank You!

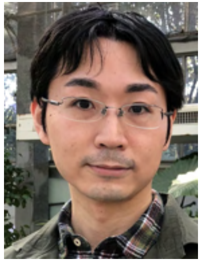
S. Hrushikesh Bhupathiraju



Shaoyuan Xie

Michael Clifford

Qi Alfred Chen



Takeshi Sugawara

Sara Rampazzi

✉ bhupathirajus@ufl.edu

https://sites.google.com/view/thermal-vuln-ad/

Supported by

TOYOTA InfoTech Envisioning Mobility

NSF

DEPARTMENT OF TRANSPORTATION UNITED STATES OF AMERICA

JSPS

SoCAL HUB