# Ghost-Probe: NLOS Pedestrian Rushing Detection with Monocular Camera for Automated Driving


## ABSTRACT

One of the most serious factors compromising driving safety is when people in drivers' non-line-of-sight areas rush out suddenly. Existing studies on non-line-of-sight imaging rely on expensive equipment or are limited to severe laboratory conditions (e.g., massive planar reflectors and controlled illumination), rendering these technologies inapplicable in complex driving scenarios. In this paper, we propose a non-line-of-sight moving obstacle detection system *Ghost-Probe*, which can provide an advanced driver assistance system (ADAS) with sufficient time to respond and stop safely. We design a shadow signal discriminator to assess the weak shadows created by a moving obstacle, such as pedestrians in the blind area, while simultaneously filtering out the impacts of other complicated illumination. Note that we merely use commercial monocular cameras and our system is robust to a wide range of lighting scenarios and planar reflectors. We evaluate the generalizability of our approach using the datasets collected in real-world driving scenarios with a variety of road surface and lighting circumstances. The results indicate that our system can detect the moving pedestrian in the non-line-of-sight area at a distance of 20 meters and offer the ADAS system advance warning to keep a safe distance.



## CCS CONCEPTS

• **Computer systems organization** → *Embedded software.*

## KEYWORDS

Non-line-of-sight, Shadow signal analyze, Driving safety, Detection system

**ACM Reference Format:**

Hao Yan[1], Feng Lin[1*], Jin Li[1], Meng Zhang[1], Zhisheng Yan[2], Jian Xiao[3], Kui Ren[1]. 2023. Ghost-Probe: NLOS Pedestrian Rushing Detection with Monocular Camera for Automated Driving. In *ACM Conference on Embedded Networked Sensor Systems (SenSys '23), November 12–17, 2023, Istanbul, Turkiye.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3625687.3625791



*Feng Lin is the corresponding author.



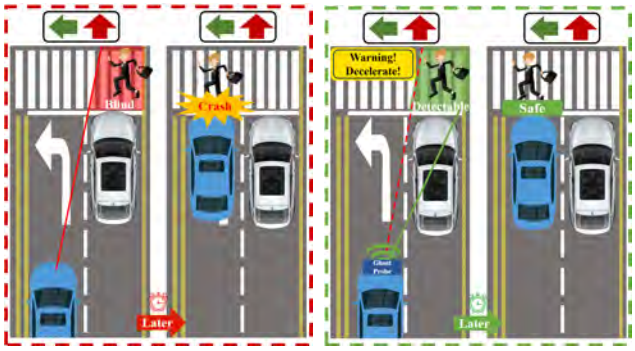


*SenSys '23, November 12–17, 2023, Istanbul, Turkiye*





ACM ISBN 979-8-4007-0414-7/23/11…$15.00
https://doi.org/10.1145/3625687.3625791


(a) Diagram of the Ghost rush (b) Scenario with Ghost-Probe

**Figure 1: Scenes at a multi-directional intersection. Scenes at a multi-directional intersection. White car stops temporarily while waiting for a red light and blue car turns left according to regulations. (a) A schematic of *Ghost rush* accident. A pedestrian rushes out of the blind region. The blue vehicle lacked reaction time and braking distance to cause a collision. (b) A schematic of *Ghost rush* accident with Ghost-Probe. With Ghost-Probe, pedestrians moving in the blind region can be detected from a distance, warning the driver or ADAS system in advance to reduce speed and avoid collisions.**

## 1 INTRODUCTION

Driving safety is one of the most critical aspects of vehicle safety configurations, with a wide range of research prospects. According to the World Health Organization (WHO) [40], over 1.3 million people die annually in road accidents, with nighttime driving being particularly hazardous. While just one-quarter of our driving at night, 50% of traffic deaths happen during this period [25], with the majority of them being *"Ghost rush"* [35], which are caused by pedestrians suddenly emerging from areas outside the driver's line of sight, as shown in Fig. 1 (a). *Ghost rush* incidents have become one of the greatest risks to road safety due to their suddenness and unpredictability. Compared to pedestrians darting out from behind trees or walls at the roadside, the risk is significantly elevated when they emerge abruptly from in front of stationary vehicles. This is primarily due to the nature of the blind spots they create. Blind spots formed by roadside obstructions generally exist at a greater distance, providing drivers with ample reaction time. Conversely, the blind spots created by stationary vehicles in adjacent lanes are larger and proximate, leaving drivers with considerably less time to react to emergent situations.

Because of the high risk of such mishaps, researchers have made great efforts to prevent such accidents. A new strategy on the internet of vehicles, vehicle to X (V2X) or vehicle to vehicle (V2V)

[49], has emerged to deal with them. Nonetheless, it necessitates extensive infrastructure development in both onboard and roadside sensors and complex communication protocols with huge amounts of data, which raises enormous financial and technological impediments, prohibiting it from being extensively deployed in a short period time [2]. A more pragmatic approach is non-line-of-sight (NLOS) imaging, which pays more attention to recovering occluded objects around the corner. Researchers utilized laser [5, 16] or coherent light sources [7, 20] to capture scatter patterns, including information about hidden objects from the reflected surface to reconstruct the original images. Nevertheless, ***the hardware is too expensive to deploy on a large scale and requires the assistance of a trained technician.*** In comparison to the active NLOS imaging technology, the passive imaging technique [32] just requires a commercial camera without an active light source, thereby decreasing the cost of the necessary gear. However, ***it requires strict elaborate laboratory illumination***, which may not be available in a realistic driving scenario with complex variables. Woodford et al. [42] investigate the use of a new class of reflectors to accomplish NLOS object detection by utilizing several curved reflectors to offer complete coverage of vital sites near intersections. But ***not all scenes happen to be arranged with planar reflectors***.

To advance *Ghost rush* prevention in real-world driving scenarios without the need for high-cost and specially installed equipment, we aim to leverage existing onboard monocular cameras and non-supplementary facilities to detect non-line-of-sight objects. According to the observations in Section 3.2, a moving object in the blind area creates a shadow signal in the line of sight, even if it is very faint (as seen in Fig. 2 (c)). We find that it's plausible to design a system capable of detecting NLOS moving pedestrians using these faint shadow signals. While constructing this system, we confronted the following critical challenges:

***Challenge 1:*** *Discerning subtle shadow changes from a distance for a fast-moving vehicle is a formidable task.* The light alterations triggered by swift vehicles considerably outpace the shadow modifications caused by pedestrians in motion. Consequently, the vehicle's movement can instigate substantial interference with the detection of faint shadows.

***Challenge 2:*** *The complexity of illumination and the specificity of the projection surface result in a low signal-to-noise ratio (SNR) for shadows.* The interplay of intricate urban illumination can easily diminish the shadows cast by pedestrians. Additionally, we aim to detect shadows projected onto the ground, not onto a plane perpendicular to the light source, which further reduces the SNR for shadows and intensifies the detection challenge.

This paper principally presents a system capable of detecting moving pedestrians beyond the line of sight in real-world driving scenarios with intricate lighting conditions. Operating solely on the commercially available ADAS monocular camera on board, it requires no additional large, planar, vertical reflectors. This system supports both drivers and prevalent ADAS systems, aiding in early decision-making to prevent collisions.

To tackle the challenge 1, we initially employ a dynamic buffer to capture and process images in real time. These images are then aligned to the first frame using a homography matrix, significantly reducing the computational complexity of subsequent operations. Leveraging 3D vehicle detection technology, we accurately identify the area encompassing the shadows cast by moving obstacles. We then devise an algorithm for region of interest (ROI) extraction, which effectively filters out the pronounced noise from faint shadows introduced by rapidly moving vehicles.

In response to Challenge 2, dealing with the low signal-to-noise ratio of faint shadows, we design a dynamic weak shadow discriminator. This method amplifies weak shadows and analyzes the dynamic changes of pixels utilizing a causal temporal filter and camera-specific dynamic thresholds. Ultimately, we discern the presence of moving obstacles outside the line-of-sight position by examining dynamic pixels. The effectiveness and applicability of our approach are validated through an extensive evaluation using real-world datasets, ensuring robust performance in actual scenarios. Our contribution to this work is three-fold:

- We explore the potential of leveraging subtle shadow signals to identify pedestrians moving beyond the field of view. Our findings suggest that these minor shadows preceding the vehicle can shed light on the presence of moving obstructions within the blind spots.
- We design *Ghost-Probe*, a system for implementing early warning of moving pedestrians in blind areas by identifying dynamic shadow signals with low SNR by detecting pixel changes. We validate our system by conducting extensive experiments with real-world datasets. The results indicate that it can detect moving pedestrians in the blind area at a distance of 20 m and alert ADAS systems.
- We expand the application scenario of non-visual range imaging to overcome the challenge of detecting moving targets in non-visual range states in complicated situations, as well as providing novel active defensive driving protection for the ADAS.

## 2 RELATED WORKS

### 2.1 None-line-of-sight Detection

Time-of-Flight-based (TOF) NLOS imaging [17, 19, 26, 33] has recently attracted substantial interest. Due to the method's susceptibility to interference from other unpredictable light sources, a tightly controlled laboratory environment is required. Felix et al. proposed a method to detect shadow signals on a moving platform [24] that uses AprilTag for image registration alignment, determines whether there are moving objects around corners by detecting shadow changes in front of the moving platform, and achieves convincing results in the ground and corridor light environments. However, in the actual nighttime driving situation, the signal-to-noise ratio of the shadow signal is greatly lowered due to the extremely complicated lighting conditions. Moreover, it is impractical to apply AprilTag on actual roads, thus the technology remains limited to the laboratory. Woodford et al. [42] investigated a method that involves using radar to gather information from blind spots by reflecting signals off surfaces around roadside corners (such as utility poles or building walls). However, this approach demands a specific signal emission angle and it is challenging to find suitable reflecting surfaces at the required positions and angles in realistic driving environments.

### 2.2 Moving Object Detection

Conventional target detection techniques, such as YOLO [3], faster-RCNN [29], and other detectors, merely identify the object category


**Figure 2: (a) Shadow projection schematic. (b) Shadow casted by pedestrian with a close up view. (c) A rear view of shadow casted by pedestrians.**

and flag it, without analyzing the object's motion. In advanced driving assistance systems (ADAS), moving obstacle detection and prediction [22] are crucial for driving decision systems because the next decision for emergency braking and high-precision navigation heavily depends on the state and location of moving objects on the road and makes decisions accordingly to avoid possible collisions. In the field of object movement detection, the method represented by optical flow (FlowNet [15]) detects the intensity of image pixels over time and infers the speed and direction of object movement. However, if the photographer's movement is too great, a stationary object can be misclassified as a moving one. MSFgNet [27] presents an end-to-end neural network for extracting moving objects from films by differentiating the foreground from the background. FisheyeMODNet [45] presents a CNN architecture that employs a lightweight encoder for the detection of moving objects in an autonomous driving environment. However, these methods can only identify items that are easily identifiable from the background, such as vehicles and people.

## 2.3 Shadow Handling

Shadow detection is a common research area in computer vision, and the majority of current shadow detection research focuses on its elimination [10, 18]. Detection methods have also gradually shifted from prior feature-based approaches [11, 13] to those based on deep learning [8, 14, 38, 48]. However, the strategy based on networks severely restricts the system's detection time. Current shadow identification approaches, on the other hand, concentrate on a single image and a single image's content, and the bulk of shadow detection scenarios occur during well-lit daylight [37, 38, 48], making it easier to detect shadow images with big deviations from the background and obvious edges. Nighttime shadow identification is currently difficult due to the lighting situation's complexity and the lack of distinction between the shadow foreground and backdrop. Our method requires only deep learning to aid in localization and no network involved in the detection process, which significantly enhances the detection rate.


**Figure 3: (a) The ROI of the camera. (b) The result of shadow analysis with still camera aroung dark corner. (c) The result of shadow analysis with moving camera around dark corner. (d) The result of shadow analysis with still camera around bright corner.**

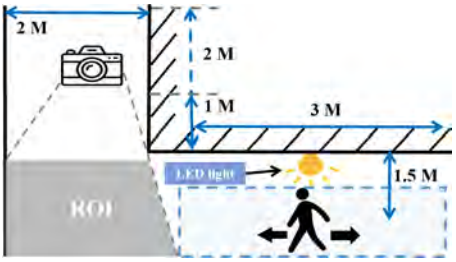



**Figure 4: Schematic diagram of the experimental setup.**

# 3 PRELIMINARY STUDY

In this section, we first analyze why existing shadow detection techniques cannot be used directly and then present a real-world experiment to verify the feasibility of NLOS pedestrian detection through its shadow signal.

## 3.1 Concept classification of shadows

The shadow status is strongly correlated with many factors, such as the relationship between the relative position of the light source and the projection surface, the distance between them as well as the type of light source, etc. As shown in Fig. 2 (a), the shadows occluded by people under the sun have distinct boundaries, which we refer to as *hard shadow*. The light source approximates parallel light and it can delineate the edge of the obstructing because the distance between the occluder and the projection surface is significantly shorter than that between the light source and the occluder [41]. Currently, a considerable study is devoted to the detection and eradication of such shadows [8, 11, 37, 38, 48] using their easily distinguishable hue and edges. However, there is another type of shadow that is more difficult to manipulate, and it is known as *soft shadow*. When the distance between the point light source and the occluder is less than or approximates the distance between the occluder and the projection surface, the shadow becomes divergent, which makes it hard to distinguish the occluder's edge. In real-world driving scenarios, the blind region often arises 1-3 meters in front of the vehicle, and the shadows caused by pedestrians in this area are often divergent and (Fig. 2 (b)), especially when viewed from a distance (Fig. 2 (c)). Due to their blurred edges and faint shadow signal, the prevailing shadow detection techniques are unable to detect such soft shadows.

## 3.2 Feasibility and challenges

To achieve the NLOS pedestrian detection, our intuitive idea is to detect the soft shadow cast by moving pedestrians in the ROI. We design an experiment to verify the feasibility and show the challenges of our solutions.

**Experiment setup.** We chose a dark corridor with a corner where we could control the illumination condition, as shown in Fig. 4. We switch off all lighting systems except for the vehicle-specific light on the wall, which is mounted 1.4 meters above the ground. One experimenter moved within an area of 3 meters in length and 1.5 meters from the wall where the lights were installed. We use the leopard camera (see Section 5.1) to record a 4-second video for each scenario while the experimenter moves in the same route.

**Scenarios.** We set up three different experimental scenarios to extensively verify the feasibility of the scheme. **(S1)** The camera

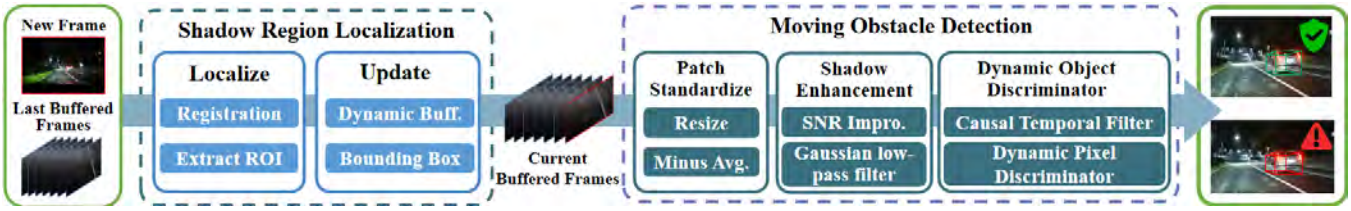



**Figure 5: An overview of a potential moving obstacle detection system *Ghost-Probe***

remains positioned 1 meter from the corner. The illumination facilities are kept off except for the light on the wall. **(S2)** The camera is moving behind the corner from a position of 3 meters at a speed of 0.5 m/s to 1 meter. The illumination facilities is kept off except for the wall light. **(S3)** The camera remains positioned 1 meter from the corner. Turn on the LED lighting equipment and maintain uniform light in the experimental area.

**Results.** We use inter frame difference method [23] to analyze the video captured by the camera in three different scenarios, and the results are shown in Fig. 3. When the camera is motionless and captures the video in a dark environment, the moving shadow in the ROI can be detected. Comparing Fig. 3 (b) and Fig. 3 (c), when capturing moving video, we find that the existing solution failed to detect any shadows because **the change of the surrounding environment is far greater than the motion of faint shadows**. Moreover, if we capture video in a well-lit environment, the shadows will be obliterated by other light. Consequently, **its signal power is further diminished**. Considering the complex illumination and moving scenes in a real-world driving scenario, detecting a faint shadow signal casted by pedestrians is a big challenge.

# 4 SYSTEM DESIGN

## 4.1 System Overview

In this work, we propose *Ghost-Probe*, a system that detects potentially hazardous moving pedestrians in the driver's blind region. As seen in Fig. 5, the *Ghost-Probe* system consists of two phases: shadow region localization and moving obstacle detection.

**Shadow Region Localization.** To meet challenge 1, we aim to extract ROI in real-time with the interference of complex scenes and design two phases: ROI localization and dynamic buffer update to deal with it. We employ 3D vehicle detection technology to locate the vehicle that creates a blind area in the driver's field of view. Upon completion of the 3D vehicle detection procedure, we can receive the 3D bounding box of the target vehicle, which allows us to gain the spatial mapping of the 3D object in the 2D frame. We can then analyze the actual road direction in the 2D image. Using the bounding box and orientation, a region of interest can be computed. To the best of our knowledge, this is the first work that uses the geometric relationship of 3D target detection results to locate the ROI. As 3D detection requires computer resources, we construct a dynamic image buffer to reduce future computational complexity, greatly enhance system efficiency and accomplish real-time processing. A buffer eventually is used to convey a succession of photos to the future phase.

**Moving Obstacle Detection.** To address challenge 2, we do a lot of research on how to detect potentially moving obstacles under influence of complex conditions. We designed a moving obstacle detection model consisting of the following steps: (1) patch standardization, (2) shadow enhancement, and (3) dynamic pixel detection. Patch standardization tries to increase the difference between frames while also filtering the influence of ambient light variations on the target signal. Then a shadow enhancement model is implemented to boost the SNR of the shadow signal. Finally, the dynamic pixel discriminator determines if there is a potential moving obstacle beyond the line-of-sight region.

## 4.2 Shadow Region Localization

*4.2.1 Image Registration.* Image registration refers to the process of converting several photographs into the same coordinate system. We use it to align two frames to reduce the complexity of interest region selection computation by a significant margin, as described in Section 4.2.3. This process is often split into three steps [50]:

(1) Feature detection. There are numerous methods for feature detection, including scale-invariant feature transform (SIFT) and speeded-up robust features (SURF). To satisfy our objectives, our system employs oriented FAST and rotated BRIEF (ORB) [31]. ORB algorithm mainly improves the real-time performance of mainstream algorithms such as SIFT and SURF. Our choice of the ORB algorithm is primarily motivated by two conditions:

- The ORB algorithm achieves a better balance between effectiveness and performance than others.
- The actual road situation, which will be used to test our system, offers a great deal of distinguishing feature points, which makes the work of the ORB algorithm easier.

(2) Feature matching. After this preliminary processing, we can obtain a sequence of picture key points and descriptors. We measure the similarity between two feature descriptors using the hamming distance. We locate similar features in the two photos, arrange them according to their similarity scores, and retain a tiny fraction of the initial matches. By utilizing these matched points, we can derive the H.

(3) Estimating the homography. The H matrix refers to the homography which transforms points of two planes with 8 Degrees of Freedom (DOF):

$$s\begin{bmatrix}x'\\y'\\1\end{bmatrix} = H\begin{bmatrix}x\\y\\1\end{bmatrix} = \begin{bmatrix}h_{11} & h_{12} & h_{13}\\h_{21} & h_{22} & h_{23}\\h_{31} & h_{32} & h_{33}\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}, \quad (1)$$

where $(x, y)$ represents the image in the original view and $(x', y')$ stands for the original image in the new perspective. Before estimating H, it must be noted that despite the ORB algorithm's best efforts, there will undoubtedly be a large number of wrong matches. Random Sample Consensus (RANSAC) [12] is an iterative approach to estimating parameters of a mathematical model from a set of observed data that contains outliers. We use it to construct an accurate H matrix based on a sequence of key point matches.

After obtaining an accurate homography matrix (H), we can align the newly acquired frame with the reference frame, thereby eliminating the need for the recalculation process outlined in the

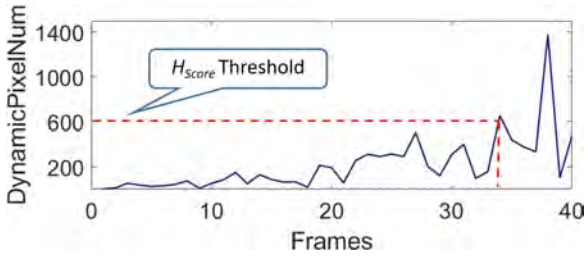



**Figure 6: Randomly select 40 consecutive frames to calculate the $H_{score}$ change curve.**

interest region selection (Section 4.2.3). It is worth noting that the movement of people within the frame does not impact the calculation of the H matrix, as the majority of feature points remain stable.

The *Ghost-Probe* detection method proposed in this research operates in real-time. Therefore, we construct a dynamic picture buffer to to balance performance and efficiency. It is assigned to pre-process newly added photographs and temporarily store the images to be detected. When the onboard cameras work, there is a series of photos awaiting processing. It is inefficient to process each frame since the content of two consecutive frames may not vary significantly. In this case, when a new image is loaded into the buffer, image registration is utilized to align it with the first frame. The buffer capacity is set to a variable value to support the update frequency of the image buffer at various vehicle speeds. A faster vehicle speed results in more change between frames, necessitating more frequent buffer updates, and vice versa. We define a threshold called $H_{score}$ to determine when the buffer should be updated.

The homography H matrix computes the $H_{score}$ value. Referring back to Section 4.2.1, we know that the 8-DOF H matrix may convert the points of two planes. Therefore, the H matrix represents the degree of transformation. When a new frame is about to be appended to the buffer, it will be aligned to the head of the sequence and the H matrix obtained from image registration will be saved. The $H_{score}$ can then be expressed using the following formula:

$$H_{score} = \sum_i \sum_j |H(i,j)| \,. \tag{2}$$

To demonstrate the validity of $H_{score}$, we randomly selected 40 consecutive frames from a video shot during typical vehicle operation and aligned each frame to the first image while computing the $H_{score}$. Fig. 6 depicts the transformation of each frame's $H_{score}$ to the first one. In the first 34 frames, H steadily increases as the gap between the viewpoints of the current frame and the first frame grows. Due to the viewpoint difference, the current frame is too different from the first frame after 34 frames, leading to a significant shift in key points and dramatic oscillations in H.

*4.2.2 Dynamic Image Buffer.* Through extensive experiments, we set a threshold: $H_{score} = 600$. A new image will be appended to the buffer if the current $H_{score} <= 600$, or the buffer will be reset otherwise. In this way, homography transformation accuracy can be guaranteed and the system does not affect operational efficiency by processing redundant images in the buffer.

*4.2.3 Interest Region Selection.* After picture registration, the most likely possible danger regions in each frame must be extracted for

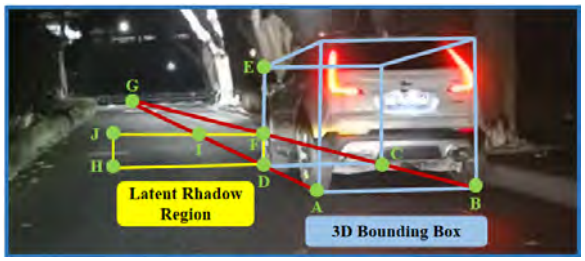



**Figure 7: The process of selecting potential shadow Region. The blue box is generated by 3D vehicle detection and the yellow one is our selection target area.**

targeted detection. This section explains how we identify interesting regions for the subsequent stage of detection. In addition, we devised a two-step method for progressively locating regions of interest, from big targets to small regions.

*3D Vehicle Detection.* Locating potentially hazardous sites to be detected directly from the entire content-rich image is computationally complex, and uneven selection of areas as the vehicle travels may directly lead to larger detection result mistakes. Therefore, we decide to start by locating the vehicle responsible for the blind area in our field of view. It will be more stable than determining the soft shadow region and effectively eliminating the complex influence in the driving scenario. The 3D vehicle detection approach proposed by *Wang et al.* [39] is used to collect information about the location and orientation of automobiles and roadways in the real world from photographs. It is based on FCOS and regresses the boundaries of the 2D frame to the six faces of the 3D frame. It's a simple yet efficient one-stage 3D object detection algorithm that uses only a monocular camera without any 2D-3D correspondence priors. There are amount of 3D vehicle detection techniques, but they are not the subject of this investigation. The result of the aforementioned 3D vehicle detection algorithm is a 3D bounding box of autos.

*Regional Orientation.* Using the above-described 3D vehicle detection, we can obtain 8 vertices of the 3D bounding box, as shown in Fig. 7. Next, we will utilize the Algorithm 1 to pick the latent shadow region. The workflow is introduced as follows:

First, we extend line $AD$, $BC$ to $G$. In the real world, our area of interest is on the ground, hence we choose point $A{\sim}D$ since the wheels in the actual world are all on the same plane (on the ground). In actuality, the points of lines $AD$ and $BC$ lie on the road plane. Then, let the line $BG$ intersect $ED$ at the point $F$. Line $ED$ is the vehicle's front edge, as well as the blind area delineation line. In other words, the region to the right of this line is where the prospective moving obstruction appears, and the region to the left is the latent zone of shadow signal produced by the vehicle headlights. Extend $CD$ to $H$ such that $DH$ equals $CD$. The line $CD$ represents the set distance to the front wheels in the real world. We employ it to limit the extent of the zone of interest and guarantee its stability. Finally, make $FI$ parallel to $CD$ and intersect $DG$ at $I$. Then extend $FI$ to point $J$ such that $JF$ is equal to $CD$ and connect point $J, H$. This above allows us to pick a stable latent shadow region with high signal-to-noise ratio (SNR) and obtain the coordinates of points $J, F, H, D$.


(a) (b)

**Figure 8: (a)Latent shadow region patch captured from 10 meters distance. (b)The patch after shadow enhancement.**


(a) (b)

**Figure 9: The result of dynamic obstacle detection. (a) The patch contains no moving obstacle beyond line of sight region. (b) The patch contains obstacle movement.**

Considering diverse driving characteristics and situations (e.g., Right-hand drive vehicles), vehicles causing blind spots may not always stop on the right side in real-world scenarios, such as at intersections. To address this issue, it's essential to establish the relative positioning of the vehicles. We assume that the vehicles are traveling in parallel directions on the same section of the road. In this case, we found experimentally that the 3D bounding box of the vehicle parked on the right side of the road is located in the right area of the 2D image, the opposite is true for vehicles on the left. According to this, when the vehicle is parked on the right side of the road and its 3D bounding box is in the right half of the frame, we will use Algorithm 1 to select the shadow area. On the other side, we mirror-flip all operations of Algorithm 1 to obtain the potential shadow region.

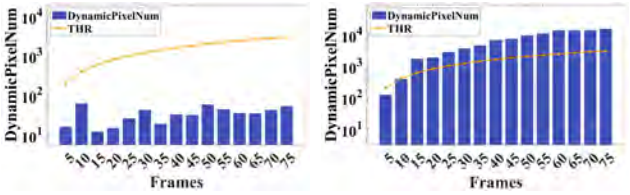
## 4.3 Moving Obstacle Detection

After obtaining a sequence of frames that have been processed by the preceding processes, we need to determine if there is a moving impediment in the area that might endanger driving safety. We propose an algorithm for detecting potential moving obstacles based on shadows projected by dangerous sources. We focus on the ROI selected in Section 4.2.3 where potential moving objects are physically capable of casting shadows. Due to its fragility, it may not be detected by drivers or standard ADAS detectors.

**Algorithm 1** Shadow region selection

1: Initialization: get points $A(x_A, y_A)$, $B(x_B, y_B)$, $C(x_C, y_C)$, $D(x_D, y_D)$, $E(x_E, y_E)$ from 3D vehicle detection.
2: The parameters of line $BC$: $a_1 \leftarrow y_C - y_B, b_1 \leftarrow x_B - x_C, c_1 \leftarrow x_C \times y_B - x_B \times y_C$.
3: The parameters of line $AD$: $a_2 \leftarrow y_D - y_A, b_2 \leftarrow x_A - x_D, c_2 \leftarrow x_D \times y_A - x_A \times y_D$.
4: $x_G \leftarrow (c_2 \times b_1 - c_1 \times b_2)/(a_1 \times b_2 - a_2 \times b_1), y_G \leftarrow (c_1 \times a_2 - c_2 \times a_1)/(a_1 \times b_2 - a_2 \times b_1)$.
5: $F \leftarrow BG \bigcap ED$.
6: Extend $CD \rightarrow H$ such that $DH = CD$.
7: Make the line $FI \ // \ CD$ and intersecting $DG$ at point $I$.
8: Extend $FI \rightarrow J$ such that $JF = CD$ and connect point $J, H$.

**Output:** Points $J, F, D, H$.




(a) (b)

**Figure 10: (a)The relationship between threshold and amount of dynamic pixels when there is no obstacle in the blind area. (b)The relationship between threshold and amount of dynamic pixels when there is a pedestrian in the blind area.**

*4.3.1 Region Standardization.* In every computation cycle, we can obtain a sequence of ROI patches from the picture buffer that is processed by the aforementioned procedures. To standardize image characteristics, we downsample each frame to a $100 \times 100$ patch using bilinear interpolation. Inspired by the widespread usage of mean image subtraction for a single image in the deep learning pre-process, we construct a global mean image during the dynamic buffer that is averaged for each channel of all patches individually and generates a new image.

$$\bar{V}(x, y, c) = \frac{1}{n} \sum_{i=1}^{n} V_i(x, y, c), \tag{3}$$

$$f_i = V_i - \bar{V}, (i = 1, 2, ..., n), \tag{4}$$

where $n$ stands for the frame number in the buffer and $c$ is the color channel for each patch. Regarding the effect of other light sources, such as street lights or opposing headlights, our insight is that their effect on luminance is global and that changes in the patch's global luminance have few effects on changes in the image's finer details. Therefore, we remove the global mean patch from each frame in the current sequence to reduce the impact of global brightness changes on local patches.

*4.3.2 Shadow Enhancement.* While eliminating the effect of global brightness, mean image subtraction significantly reduces the signal-to-noise ratio (SNR) of the shadow signal. Therefore, we must improve it to guarantee the validity of following detection. Before enhancing the shadow, a Gaussian low-pass filter (GLPF) is implemented. In a word, we compute

$$g_i = |\mathbb{G}(f_i, k, \sigma) \cdot \delta|, \tag{5}$$

where $\mathbb{G}$ is a linear blur filter of size $k$, which using isotropic Gaussian kernels with covariance matrix $diag(\sigma^2, \sigma^2)$. Because of the sensitivity of the weak shadow detection process to tiny perturbations in pixels, we use GLPF to filter out high-frequency signals while avoiding the "Ringing Effect". With the help of research on image blur[4], we set $\sigma$ depending on $k$, which is $\sigma = 0.3 * ((k-1) * 0.5 - 1) + 0.8$. At the same time, $k$ is set to 3 based on empirical observation. $\delta$ is an enhancement parameter that is computed with the help of the mean image of Eq. (4), that is

$$\delta = \sqrt{\frac{1}{p-1} \sum_{i=1}^{p} (f_i(x, y) - \bar{f}_i)^2}, \tag{6}$$

where $p$ is the pixel number of the patch and $p = x * y$. It seeks to raise SNR in order to increase the detectability of a shadow, even if the signal's origin is invisible to human sight.

**Table 1: Profile of the dataset**

| Properties | Condition Settings |
|---|---|
| **Light Condition** | (a) The dark road (b) The road with streetlights (c) Opposite direction lights |
| **Road Surface** | (a) Wet tarmac road (b) Dry tarmac road (c) Wet+Dry (d) Sprayed crosswalk (e) Epoxy-paint garage floor |
| **Obstacle Type** | (a) Adult walk (b) Child rush (c) Elderly stumble (d)Ride a bike |
| **Vehicle Speed** | (a) Uniform speed 10 km/h (b) Uniform speed 30 km/h |
| **Equipments** | (a) Leopard Image Camera (b) Huawei Mate 30 Pro cellphone (c) Canon camera |

*4.3.3 Dynamic Obstacle Discriminator.* According to shadow enhancement, obstructions in front of the vehicle cast a succession of patches with high SNR shadows. Throughout a series of patches in chronological order, the change of pixels at the same location on different patches has a temporal link. In other words, the state of a pixel at time $t$ depends on its state at time $t-1$. Then, a causal temporal filter is used to establish a connection between two moments of a pixel, resulting in a new sequence of patches.

$$C(x, y, t) = g(x, y, t) + \alpha \cdot g(x, y, t-1), \tag{7}$$

where the the filtered result $C(x, y, t)$ at time $t$ depends on the input patch at time $t$ and the output at the previous time instant $t-1$ multiplied by a parameter $\alpha$ which is set to 0.3 based on empirical observation. Then, each pixel is evaluated to see if it is a dynamic pixel caused by the moving obstacle. We are inspired by [43], in which a "dynamic threshold" is used to adjust to varied light situations in various traffic scenarios.

$$P_i(x, y) = \begin{cases} 0, |C(x, y, i) - C(x, y, i-1)| \leq \omega \cdot \sigma(C_{i-1}) \\ 1, |C(x, y, i) - C(x, y, i-1)| > \omega \cdot \sigma(C_{i-1}) \end{cases}, \tag{8}$$

where $\omega$ is an adjustable parameter that depends on the image noise and we set $\omega = 2$ for all our experiments. All pixels with a change bigger than the mean squared difference are assigned to 1, or "dynamic," while the rest are set to 0. To minimize noise and improve motion pixel performance, we compute

$$P_i \bullet K = [\cup \{(K)_z \mid (K)_z \cap P_i = \varnothing\}]^c \tag{9}$$

where $K$ is a morphological ellipse element with kernel size 2. $z$ means a set of dynamic pixels and $(K)_z$ stands for the translations of $K$, which are contained in $P_i$. After that, we sum up all dynamic pixels of the sequence and compare it with a *DynamicTHR* computed as below.

$$DynamicTHR = 100px \times 100px \times N \times NR, \tag{10}$$

where $N$ is the number of patches in the current sequence, and $100\times 100 \times N$ pixels are available to determine whether it is a dynamic (value=1) or static (value=0). $NR$ is a Noise Rate parameter that depends on the camera's resolution and noise level. The higher the noise levels, the higher the misclassification rates, hence we set $NR$ as a noise tolerance to improve the system's robustness. Ultimately, a potential moving obstruction is detected when dynamic pixel sum surpasses the camera-specific threshold.

# 5 EVALUATION

## 5.1 Equipment Setup

To statistically evaluate the performance of the system, we collect real-world datasets under a variety of scenarios with different equipment. We also hold an active institutional review board (IRB) approval for collecting action data from participants for driving safety research. All the evaluations tightly follow the rule of IRB regulation.

**Vehicles:** We collect data using two vehicles on a straight road with realistic urban night lighting conditions. *Vehicle A* is a Cadillac xt5 with a height of 1.68 meters and a width of 1.9 meters. Because of the perspective, it can completely obscure adult pedestrians (with a height of 1.7 meters) in front of the car so that oncoming traffic from behind is invisible to the front of the car and vice versa. *Vehicle A* remains at a critical stop with headlights on, acting as a sight barrier between pedestrians and *vehicle B. Vehicle B* is an XPeng P5 with a height of 1.53 meters and a width of 1.84 meters. It is coming from the rear of the adjacent lane as *Vehicle A* at a fixed speed with headlights on. We installed the camera in the middle-top position of the windshield, which represents the mounting height of most ADAS cameras. During the experiment, the experimental vehicle kept the low beam lights on to simulate the real driving situation.

**Dummy pedestrian:** Pedestrians are the main danger element in *Ghost rush* and our detection target. We loaded the dummy onto a remotely controllable moving platform to simulate a pedestrian moving in front of the vehicle. During the experiment, the vehicle will travel at a fixed speed, and the above settings can minimize the safety risk to the experimenters.

**Cameras:** Due to the remote location of the filming vehicles and the complex lighting conditions on the road, different filming qualities can produce large differences in the source data. We choose two types of cameras to ensure the diversity of the raw data and collect source frames.

**(1) Leopard Image LI-IMX490** is an autonomous driving camera module with 120° FOV lens, widely used by manufacturers such as Apollo and Nvidia. The framerate is set to 20 fps, using RGB888 ISP to process the raw data. The videos were saved as uncompressed AVI which has complete pixel and color information.

**(2) Canon M50** with the 18-150 mm lens: the framerate was set to 30 fps and isolation $1920 \times 1080$. The videos were finally saved as .mp4 format.

**Computers:** During data collection, we use a Lenovo laptop which has an R7-6800H core and 16GB RAM to control the Leopard Image camera. After that, a desktop computer which has an i7-10700 core, 32GB RAM, and a server with an Nvidia 3090 graphics card are used to process images.

## 5.2 Evaluation Scenario and Data Collection

According to the research about road safety [1], driving scenarios at night are specifically dangerous and drivers' inadequate surveillance is the main culprit that caused almost half of the crashes. As such, according to reports from the National Security Council(NSC)[25], we design a realistic driving scenario to support our following experiments, as shown in Fig. 11.

The *Vehicle A* temporarily stops at a city roadside with street lights and a pedestrian walks out from 1.5 meters in front of the


**Figure 11: (a) Rear view of *Vehicle A*. The pedestrian positioned ahead of *Vehicle A* remains invisible to the direct line of sight from *Vehicle B*. (b) Side view of the experimental scene. (c) Schematic diagram of the experimental scenario.**

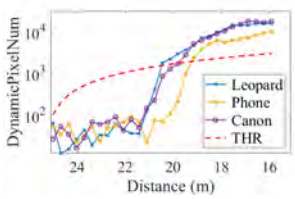



**Figure 12: Trend of dynamic pixels at different cameras.**

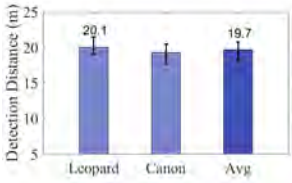



**Figure 13: The Detection distance with different camera.**

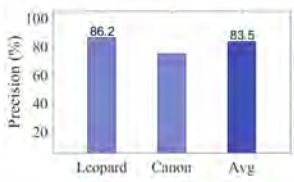



**Figure 14: The detection precision of different camera.**

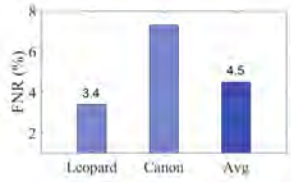



**Figure 15: The detection FNR of different camera.**

*Vehicle A*. *Vehicle B* is in the adjacent lane as *Vehicle A*, coming from 50 meters behind it at a speed of 10 km/h. In this process, the experimenter and *Vehicle B* are invisible to each other due to the obstruction of *Vehicle A*.

We use the equipments described in Section 5.1 and capture a series of videos in our evaluation scenarios with different conditions. As seen in Table 1, we collect our data in 5 different attributes and several settings are arranged in each of them. We set "the dark road + dry tarmac road + adult walk + uniform speed 10 km/h" as the baseline and select 10 scenario conditions through control variates. For each condition setting, we use our equipments for data acquisition, taking five videos of each device. Totally, we collect 100 videos of about 1.2 hours, which can be split into 90000 frames. Of these, 70% were adults walking, 15% were seniors hobbling, and 15% were children rushing. As our detection phase does not utilize deep learning, the entire dataset serves as the test set. It's worth noting that our data collection site was an open road. The frequent vehicular traffic during the data collection process enhanced the diversity and authenticity of our dataset.

## 5.3 Evaluation Metrics

We verify our novel approach in a real-world scenario expressed above and evaluate it through three aspects:

**(1) Detection distance**: at what distance can we detect potentially moving obstacles beyond our line of sight. In order to obtain more accurate results, we use frames of the video we captured to calculate the detection distance: $D_d = D - fr * n$, where $D$ is the distance between the start line and *Vehicle A* which is set 50 meters in our experiment scenario. The *Vehicle B* will go through the start line at a preset speed and keep it stably during our video. The $fr$ is the frame rate of the capture equipments and $n$ is the consumption of frames when the hidden obstacle is detected.

**(2) Detection efficiency**: The detection efficiency in our system is defined as the time gap from acquiring the frame to obtaining the result. Note that we calculate the detection time without the time to load the model because this part is already included in the vehicle initialization [46] in the real world.

**(3) Detection stability**: After the algorithm detects a potential hazard in a video frame, it is very important to be able to detect the hazard consistently and correctly in subsequent frames. It is defined as the detection stability which stands for the capability of the algorithm to maintain detection precision. We opt to use *Precision* and *False Positive Rate (FNR)* to quantify the stability of Ghost-Probe. *Precision* represents the probability that an unseen shadow signal casted by moving obstacle beyond line of sight can be correctly recognized, which can be calculated by: $Precision = F_{TP}/F_{all}$, where $F_{TP}$ refers to the quantity of frames that were correctly identified and labeled as positive. $F_{all}$ is the number of all testing frames after the first detected frame. *FNR* stands for the algorithm's capacity to accurately predict the positive sample and avoid injustice, which can be calculated by: $FNR = F_{err}/F_{N-all}$. $F_{err}$ is the number of false positive frames and $F_{N-all}$ is the total number of ground truth negative frames. The higher *Precision* and lower *FNR* indicates a more stable detection and higher confidence.

## 5.4 Overall Performance

*5.4.1 Detection distance.* To show the validity of the Ghost-Probe system, we calculate its detection distance which is the significant evaluate factor of ADAS. According to the research about braking distance [9], a car driving at 40km/h on an wet ordinary tarmac road (friction coefficient $\mu = 0.5$) has a braking distance of about 12.6 meters in an emergency, which includes the driver's reaction time and braking time. Higher coefficient of friction and slower speeds will result in shorter braking distances. Our system can detect potencial moving obstacle beyond line of sight through its faint shadow signal. We can observe in Fig. 12 that the dynamic pixels begin to rise 21 meters away from the hazard and surpass the threshold at approximately 20 meters. Specifically, the leopard camera, a commercial camera for autonomous driving, recognized a distance of more than 20 meters. Because it had the highest resolution and

**Table 2: Detection efficiency results of each phase**

| Frame Type | SRL | DBU | SSD | Totally |
|---|---|---|---|---|
| Normal Fr. | 2.467 ms | / | 4.467 ms | 6.934 ms |
| Buffer Update Fr. | 2.358 ms | 90ms | 7.065 ms | 99.423 ms |

**Table 3: Comparison result on moving object detection**

| Methods | Detection Distance |
|---|---|
| YOLO V4 [3] | 2.75 m |
| Faster-RCNN [29] | 3.15 m |
| BorderDet [28] | 2.9 m |
| **Ours** | **20.1 m** |

the least loss of pixel information without video compression. Although the presence of random errors lead some fluctuation in the detection distance (as shown in Fig. 13), their detection distances were all higher than the safe braking distance of 12.6 m, and the average detection distance reached 19.2 meters. These results above demonstrate that our approach establishes a safety margin that provides sufficient distance between the car and an impending obstruction in front, allowing sufficient time for the ADAS to react and stop safely.

*5.4.2 Detection efficiency.* Our detection process is delineated into three distinct stages: shadow region localization (SRL), dynamic buffer updates (DBU) and shadow signal detection (SSD) to evaluate the time consumption separately. Likewise, we divide the frame types into two: the *normal frame* and the *buffer update frame*.

When a frame is loaded into the buffer, we align it with the first frame and calculate the $H_{score}$ to determine whether it is necessary to update the buffer. If the buffer is unnecessarily updated, then the frame is a *normal frame* and will go directly to the subsequent process. Instead, the frame is a *buffer update frame*. As we can see in Table 2, shadow region localization and shadow signal detection during the whole process achieve better efficiency. The reason for the larger time consumption of the *buffer update frame* is that the dynamic buffer update phase requires running the 3D target detection module. We run python programs with a cluster server equipped with an Nvidia 3090 and the system processing efficiency was stable at 10 fps. This signifies that our algorithm can make judgments within 0.1 seconds, implying that at a vehicle speed of 40 km/h, it can make a decision within 1 meter, which adequately meets the needs of the driving environment. Comparatively, the refresh rate of commercial LiDAR systems typically stays within the range of 5-10 fps. With the enhancement of computational capabilities in vehicles, image processing modules have become commonly integrated into current vehicles, such as Tesla Autopilot. Commercial in-vehicle computing platforms often deeply optimize the detection algorithms used, thereby achieving detection speeds even faster than those possible with an Nvidia 3090 GPU.

*5.4.3 Detection stability.* To evaluate the detection stability of our algorithm, we conduct two sets of experiments using videos captured by different types of equipment.

**E1:** The frame where the dynamic pixel value in the region of interest exceeds the threshold is selected as the starting frame, and the detection precision is calculated during all subsequent frames until the pedestrian in the blind area appears in the line of sight. For each equipment, we randomly select 10 videos and calculate their *Precision* averages.

**E2:** We capture several vedios where there is no any pedestrians in front of the car. And we randomly select 10 videos and calculate their *FNR* averages.

As we can see in Fig. 14, the detection precision of Ghost-Probe using Leopard camera and Canon camera is maintained at 86% and 80%, respectively. Fig. 15 shows that the average detection *FNR* with different types of equipments is maintained at 4.5%. In autonomous driving, tolerating a modestly higher false positive rate (FPR) in detection algorithms is reasonable, given the emphasis on safety. A prudent approach can boost system performance and lower accident risks. For a binary classification issue, we observe that our algorithm performs strictly better than random no matter using each equipment.

**Table 4: Comparison experiments on shadow detection**

| Approach | Shadow Type | Detection Rate | Duration (ms/frame) |
|---|---|---|---|
| Gryka et al. [13] | Hard | 0 % | 5.185 |
| Finlayson et al. [11] | Hard | 1 % | 60 |
| Zheng et al. [48] | Hard | 2 % | 215 |
| Wang et al. [38] | Hard | 5 % | 487 |
| **Ours** | **Soft** | **98.5%** | **2.467** |

## 5.5 Comparison Experiments

*5.5.1 Shadow signal enhancement.* The method proposed in this paper detects potentially moving pedestrians by detecting changes in faint shadow signals. Actually, shadow detection has been extensively studied in the field of computer vision. So we conduct this saliency experiment using two types of recent approaches for shadow detection, one is for deep-learning based [38, 48] and the other is feature-based [11, 13]. We choose 200 images that contain shadows casted by pedestrians moving in front of the vehicle and send them to five detectors. The results are shown in Table 4.

Traditional feature-based shadow detection algorithms [11, 13] are incapable of detecting any shadows *because the faint shadow's edge is insufficiently distinct from the background.* Similarly, approaches based on deep learning [38, 48] can only detect a few random shadow instances and most of them are not generated by moving pedestrians. Shadow detection focuses mostly on shadows under parallel light (such as sunlight), whose hue and luminance are easily identifiable from the backdrop. As depicted in Fig. 10 (a), the complicated lighting and blurred edges in our scenario make it difficult to discern the shadows cast by nighttime pedestrians walking in front of vehicles from their surroundings. Our method has achieved state-of-the-art performance in terms of detection rate and detection time by focusing primarily on these faint shadows that can be difficult for the human eye to perceive.

*5.5.2 Potencial moving obstacle detection.* The Ghost-Probe we proposed aims to detect moving obstacles thus preventing *Ghost rush* described in Fig. 1. Researchers have proposed many excellent algorithms in the field of object detection, such as YOLO and R-CNN. These methods are also widely used to obstacle avoidance


**Figure 16: Three types of illumination situation: (a) Dark, (b) Opposite lights, (c) Road lights.**


**Figure 17: Five types of road surface: (a) wet tarmac road: after rained moderately, (b) dry tarmac road, (c) tarmac road not completely dry: two hours after moderate rain at 30 °C, (d) crosswalk, (e) garage floor with epoxy paint.**

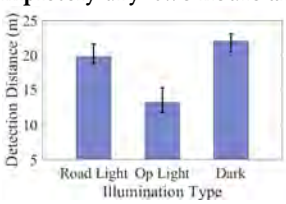



**Figure 18: Detection distance under different illumination conditions.**

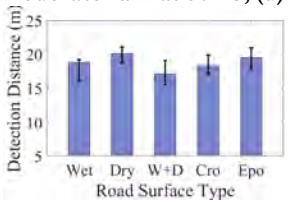



**Figure 19: Detection distance on different status of road surface.**

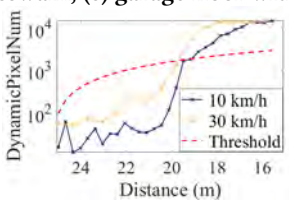



**Figure 20: Trend of dynamic pixel count at different speeds.**

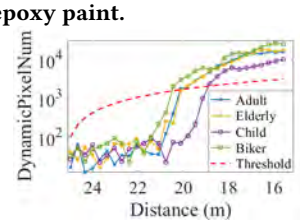



**Figure 21: Trend of dynamic pixel changes due to different obstacle types.**

and hazard warning [21, 34]. We choose three of the latest object detection algorithms with strong performance for our comparison experiments, as they represent the most widely used and commercially available detectors currently. We do not choose the NLOS object detection method because of its requirement of strict laboratory environmental control and the difficulty of applying it in real driving scenarios. We use 10 videos captured in the same scenario described in Section 5.2 and four methods work on them. The detection distance is calculated as a measure of the performance of the algorithm.

Due to limitations within their line of sight, the currently employed object detectors have poor detection distances, as shown in Table 3. When the moving pedestrian is detected, he actually has appeared in the driver's line of sight. At the same time, the head of the vehicle has traveled to the rear or even the middle of the vehicle that constitutes the blind area. With the help of the faint shadow signal process and dynamic pixel analysis, our system can detect the pedestrian moving in front of the vehicle at a greater distance (20.1 m). It will give more time to react thus reducing the probability of collision.

## 6 INCLUSIVENESS STUDY

Any perturbation such as complex light conditions will make the existing approaches unavailable. In this section, we consider several factors that could affect the performance of *Ghost-Probe*. For the following evaluation, we collect data under different conditions to verify the robustness of the *Ghost-Probe* we proposed.

### 6.1 Impact of Light Conditions

It is well known that driving at night not only has poor visibility conditions but is also disrupted by different lights. We evaluate our system in three distinct night-driving scenarios:

(1) **Dark.** We captured our video on a straight road without any lighting facilities except vehicles' headlights (Fig. 16 (a)). In this scenario, the shadow signals will not be perturbed by other illumination.

(2) **Opposite lights.** The data is gathered at an intersection, with two vehicles stationed in the opposite lane, their headlights activated (Fig. 16 (b)). In such circumstances, the ground shadow signals can be largely obliterated by the headlights of other vehicles, leading to a significant loss of information.

(3) **Road lights.** Data is gathered on a standard urban road, equipped with streetlights, to emulate nighttime driving conditions (Fig. 16 (c)). This represents a common nocturnal driving scenario where streetlights affect road surface shadows. The intensity of this impact is directly linked to the relative positioning of the streetlights and pedestrians. The closer they are, the more severe the annihilation of shadows by the lights, and vice versa. In our experiments, we opt for the most challenging setup, where the streetlight is situated directly above the pedestrians.

As seen in Fig. 18, our system has a detection distance of over 20 meters due to the fact that the shadow signals generated by obstacles can be captured more easily without the interference of other light. When the headlights of oncoming traffic cause serious interference in our detection area, we can still detect a moving pedestrian in front of the car through a faint shadow signal at a distance of 15 meters. In the scenario with street lighting, our system can detect NLOS hazards at a distance of 19 meters, even when the most demanding conditions were set.

### 6.2 Impact of Road Surface

We obtain the obstacle signal with the help of the road surface, which makes the road surface a significant factor in the performance of the approach. We collect real-world datasets on the dry


(a) (b)

**Figure 22: (a) A schematic of *Ghost rush* during daytime. (b) A schematic of *Ghost rush* during nighttime without headlights. As analyzed in Section 3.1, when the light source is distanced from the obstruction (the pedestrian) and directly faces the projection surface, it results in a hard shadow.**

tarmac road (Dry), the wet tarmac road (Wet), the tarmac road not completely dry (W+D), the garage floor with epoxy paint (Epo) and the spray-painted floor with crosswalks (Cro), as seen in Fig. 17. The result in Fig. 19 shows that, of the four different states of tarmac, the dry pavement reflects the shadow signals better and allows us to identify possibly moving pedestrians at a distance of 20.5 meters. Wet pavement poses a detection challenge due to reduced contrast between shadows and the road surface. Nonetheless, our detector managed to recognize shadows at a distance of 15.6 meters in these conditions. In contrast, crosswalks, which are more visually distinct, had a minimal impact on our system's detection performance. On epoxy-painted surfaces with their high reflectivity, our algorithm was able to capture shadow signals as far away as 20 meters. These findings substantiate that our system maintains considerable effectiveness on a range of common road surfaces and is capable of providing safety warnings in diverse scenarios.

## 6.3 Impact of Vehicle Speed

As introduced in Section 2.1, when the recording device and the detection target are moving at the same time, the light changes caused by fast-moving vehicles are far greater than the shadow alterations caused by moving pedestrians. Therefore, the vehicle's motion can cause strong interference on the transition of faint shadows. Considering *Ghost-Probe* applied on the vehicle in motion, we need to explore the effect of movement speed on the performance. We drove *Vehicle B* at 10 km/h and 30 km/h, and the other experimental settings were the same as in Section 5.2. A reasonable and legal speed for a vehicle to travel through an intersection is no more than 30km/h, which is why we chose this speed for our assessment. In addition, we need to consider the safety of people in the experiment, even in closed roads. All assessment programs must tightly follow the rule of IRB regulation. We can see the result in Fig. 20, our system can capture the faint shadow signal at 20 meters when 30 km/h and 19.2 meters when 10 km/h. This is attributable to the fact that as the vehicle's speed increases, the time interval between successive frames will be longer. Consequently, if there are shadow signal variations within the Region of Interest (ROI), the magnitude of these signal changes will be significantly greater than at lower speeds. It means that higher vehicle speeds correspond to longer detection distances, which is in line with our hopes. Because higher speeds require longer braking distances and longer detection distances give the driver and the ADAS plenty of time to react and make decisions.

## 6.4 Impact of Obstacle Types

During our solution on moving object detection beyond line-of-sight position, the signal casted by the obstacle is the indispensable element. Since different types of obstacles produce different signals, we invited different experimentalists to simulate kids, adults, the elderly and cyclists to validate the robustness of the system and the result is shown in Fig. 21. We used an adult of 1.7 meters in height and 1.2 m/s in walking speed as a baseline and it can be detected at a distance of 20 meters. The elderly had a slower movement speed (0.8 m/s) compared to adults but they are not far from the detection distance of adults. We observed that when a kid (1.5 m/s) walks around in the blind area, a longer detection distance is required. Because the child's height is much lower than an adult's (about 1.3 m) and his head is roughly flush with the headlights. When he moves, the area blocking the light is smaller, thus producing a fainter shadow signal and making it more difficult to detect. On the contrary, when a biker moves around in the blind area, a larger area of shadow can be created because of the larger size. So that we can distinguish a potential moving biker at a greater distance.

## 6.5 Daytime Scenario

As described in Section 1, *Ghost rush* accidents frequently occur during nighttime conditions when sufficient illumination is lacking. However, this does not imply that such accidents do not occur during the daytime, even if the probability is reduced due to clear visibility. Therefore, we have conducted evaluations of our system under well-lit daylight conditions, as seen in Fig. 22 (a).

We gathered data during daytime using the identical experimental configurations as outlined in Section 5.2. The outcomes demonstrate that our system is capable of detecting potential moving pedestrians at distances exceeding 30 meters, a significant improvement over the results observed in the nighttime scenario. This is primarily due to the single source of light (sunlight) in the well-lit daytime conditions, which reduces interference. Moreover, as mentioned in Section 3.1, given the substantial distance between the light source and obstructions, the light rays can be approximated as parallel, resulting in sharp, well-defined shadows, known as hard shadows. This makes it relatively straightforward for existing algorithms to detect them. Consequently, it becomes evident that detecting *Ghost rush* incidents during night, under complex lighting conditions using soft shadows, presents a substantial challenge.

## 6.6 Alternative Light Sources

We have proposed a novel method for detecting pedestrians suddenly emerging from in front of temporarily parked vehicles at night by identifying the soft shadows cast by the vehicle's headlights. However, the method we proposed is not confined to using the subtle shadows cast by vehicle headlights. When a pedestrian darts out from the front of a roadside parked vehicle with the headlights off, we can also opt to use the most prevalent light source in nighttime driving scenarios - streetlights, as seen in Fig. 22 (b).

We conducted an evaluation of our system under actual nighttime driving conditions using streetlights as the light source. The results demonstrated that our system could detect potential moving pedestrians within a distance exceeding 19 meters, virtually on par with the detection range when using vehicle headlights as the light source. While streetlights are positioned farther away from the

obstructions compared to headlights, their luminosity is higher and the illumination range is wider, resulting in more distinctly defined shadow boundaries.

# 7 DISCUSSION

## 7.1 Multi-vehicle Strategy

In an effort to maintain precision and control over the variables within our experiment, we decided to focus on using a single vehicle as the view obstruction for the initial phase of this study. However, it is important to acknowledge that real-world driving scenarios present a more complex environment, typically involving multiple vehicles. These blind areas, particularly those generated by vehicles in close proximity to us, pose a substantial risk.The inherent danger lies in the fact that the blind area created by the nearest vehicle allows for a smaller margin of error, demanding a shorter braking distance upon the detection of a potentially hazardous impediment such as a pedestrian unexpectedly entering the roadway. To tackle this, we have devised a detection strategy that prioritizes immediate risks. In scenarios where multiple vehicles are present, our 3D discriminator identifies the closest vehicle for detection. This strategy ensures that we are constantly alert to potential pedestrian threats in the blind area closest to us. This strategy, in essence, balances efficiency and safety, providing an effective solution to the problem of detecting pedestrians in a blind area in multi-vehicle scenarios. By focusing on the closest threat, we can respond promptly to possible immediate dangers, thereby enhancing overall safety.

## 7.2 Extreme Scenarios

Autonomous driving scenarios present a diverse range of complexities, which represent significant challenges for this field [47]. To ensure our system's robustness, we evaluated it across various lighting and pavement conditions. However, our limited experimental data made it difficult to assess performance in certain extreme weather and road surface conditions such as dense fog, heavy rain, sandstorms, or icy roads. These scenarios could influence the reflection and propagation of faint shadow signals, potentially affecting the system's performance. However, driving in such extreme conditions is inherently hazardous, and many Advanced Driver-Assistance Systems (ADAS) [30, 36] disallow their use in such environments. Nonetheless, Under the circumstance allowed by ADAS, our system effectively detects Non-Line-of-Sight (NLOS) moving pedestrians. Like other novel approaches, it cannot fully address all the corner case challenges on autonomous driving. As we continue to refine the system, we anticipate an improved ability to cope with a broader range of conditions.

## 7.3 Limitations

According to relevant research [1, 25], we've conducted extensive investigation into nighttime scenarios and also evaluated daytime situations. The results demonstrate robust performance of our approach across these different conditions. Our experiments have covered a wide spectrum of real-world driving scenarios. However, due to the 'long tail' problem inherent in autonomous vehicles, our algorithm has not yet been reliably tested in rainy conditions. This constraint arises from safety considerations during experiments, as the experimental data is collected from vehicles in operation on open roads. During rainy weather, due to slick road conditions and impaired visibility, it is not feasible to collect reliable data without compromising safety. Maciej et al. [13] give an exciting perspective on weak shadow capturing. They proposed that weak shadows can be segmented, and therefore edited, by learning a mapping function for image patches that generate shadow mattes. Wu et al. [44] create a new soft shadow detection approach and design a reasonable supervision strategy to alleviate the effect of annotation noises. On the other hand, various local governments have enacted laws requiring vehicles to keep headlights on during rainy weather [6] to ensure their own safety and the surrounding vehicles. We can use this light source for detecting moving objects in blind spots.

## 7.4 General Applicability

Indeed, the Ghost-Probe system's versatility extends beyond the scenarios presented in this study. As urban landscapes grow increasingly intricate, marked by a surge in overpasses, the presence of numerous supporting pillars introduces a host of blind spots in our driving environment. These blind spots, particularly prevalent at intersections, significantly escalate the safety risks in urban driving scenarios. To curb potential mishaps, a simple yet effective solution could involve installing basic lighting facilities on these pillars. Once these illumination aids are in place, vehicles fitted with the Ghost-Probe system could effectively detect pedestrians hidden within these blind spots, providing enough time for decision-making and evasive maneuvers. The merits of such infrastructural development are two-fold. Not only do these pillars create vision barriers, but they also contribute to poor illumination in their immediate vicinity, especially during nighttime. The introduction of additional lighting solutions would substantially alleviate this deficiency. Similarly, our system could be leveraged to mitigate collision risks in any driving scenario marred by blind spots.

# 8 CONCLUSION

*Ghost rush* refers to a specific type of accident where pedestrians unexpectedly emerge from blind areas beyond the driver's line of sight. The practical application of existing NLOS imaging methods has been hindered due to factors such as the expense of necessary equipment and stringent implementation conditions. Addressing this issue, we introduce *Ghost-Probe*, an innovative system meticulously designed to detect moving obstacles located in blind spots under complex lighting conditions. Remarkably, this system employs standard ADAS monocular camera technology, eliminating the need for any additional reflectors. To validate our system, we have thoroughly tested it in an array of real-world driving scenarios, with diverse lighting circumstances and varying road conditions. The promising results demonstrate the system's robust capacity to provide ample reaction time for both drivers and the existing Advanced Driver-Assistance Systems (ADAS), effectively mitigating the risk of accidents.

# REFERENCES


[1] National Highway Traffic Safety Administration(NHTSA). 2021. NHTSA promotes safe behaviors on our nation's roads. (2021). https://www.nhtsa.gov/road-safety

[2] automotive World. 2021. V2X is close. Here's what still needs to happen. (2021). https://www.automotiveworld.com/articles/v2x-is-close-heres-what-still-needs-to-happen/

[3] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. 2020. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934* (2020).

[4] Gary Bradski and Adrian Kaehler. 2008. *Learning OpenCV: Computer vision with the OpenCV library.* " O'Reilly Media, Inc.".

[5] Yanpeng Cao, Rui Liang, Jiangxin Yang, Yanlong Cao, Zewei He, Jian Chen, and Xin Li. 2022. Computational framework for steady-state NLOS localization under changing ambient illumination conditions. *Optics Express* 30, 2 (2022), 2438–2452.

[6] Tim Charlet. 2015. Headlight Use Laws for All 50 States. (2015). https://www.yourmechanic.com/article/headlight-use-laws-for-all-50-states

[7] Wenzheng Chen, Fangyin Wei, Kiriakos N Kutulakos, Szymon Rusinkiewicz, and Felix Heide. 2020. Learned feature embeddings for non-line-of-sight imaging and recognition. *ACM Transactions on Graphics (ToG)* 39, 6 (2020), 1–18.

[8] Zhihao Chen, Liang Wan, Lei Zhu, Jia Shen, Huazhu Fu, Wennan Liu, and Jing Qin. 2021. Triple-cooperative video shadow detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2715–2724.

[9] CSGNetwork. 2023. Brake Distance. (2023). http://www.csgnetwork.com/stopdistcalc.html

[10] Graham D Finlayson, Mark S Drew, and Cheng Lu. 2009. Entropy minimization for shadow removal. *International Journal of Computer Vision* 85, 1 (2009), 35–57.

[11] Graham D Finlayson, Steven D Hordley, Cheng Lu, and Mark S Drew. 2005. On the removal of shadows from images. *IEEE transactions on pattern analysis and machine intelligence* 28, 1 (2005), 59–68.

[12] Martin A. Fischler and Robert C. Bolles. 1981. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM* 24, 6 (jun 1981), 381–395. https://doi.org/10.1145/358669.358692

[13] Maciej Gryka, Michael Terry, and Gabriel J Brostow. 2015. Learning to remove soft shadows. *ACM Transactions on Graphics (TOG)* 34, 5 (2015), 1–15.

[14] Xiaowei Hu, Lei Zhu, Chi-Wing Fu, Jing Qin, and Pheng-Ann Heng. 2018. Direction-aware spatial context features for shadow detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 7454–7462.

[15] Eddy Ilg, Nikolaus Mayer, Tonmoy Saikia, Margret Keuper, Alexey Dosovitskiy, and Thomas Brox. 2017. Flownet 2.0: Evolution of optical flow estimation with deep networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2462–2470.

[16] Julian Iseringhausen and Matthias B Hullin. 2020. Non-line-of-sight reconstruction using efficient transient rendering. *ACM Transactions on Graphics (ToG)* 39, 1 (2020), 1–14.

[17] Achuta Kadambi, Hang Zhao, Boxin Shi, and Ramesh Raskar. 2016. Occluded imaging with time-of-flight sensors. *ACM Transactions on Graphics (ToG)* 35, 2 (2016), 1–12.

[18] Salman H Khan, Mohammed Bennamoun, Ferdous Sohel, and Roberto Togneri. 2015. Automatic shadow detection and removal from a single image. *IEEE transactions on pattern analysis and machine intelligence* 38, 3 (2015), 431–446.

[19] Martin Laurenzis, Andreas Velten, and Jonathan Klein. 2016. Dual-mode optical sensing: three-dimensional imaging and seeing around a corner. *Optical Engineering* 56, 3 (2016), 031202.

[20] Xin Lei, Liangyu He, Yixuan Tan, Ken Xingze Wang, Xinggang Wang, Yihan Du, Shanhui Fan, and Zongfu Yu. 2019. Direct object recognition without line-of-sight using optical coherence. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 11737–11746.

[21] Guan-Ting Lin, Vinay Malligere Shivanna, and Jiun-In Guo. 2020. A Deep-learning model with task-specific bounding box regressors and conditional back-propagation for moving object detection in ADAS applications. *Sensors* 20, 18 (2020), 5269.

[22] Zhenguang Liu, Shuang Wu, Shuyuan Jin, Qi Liu, Shijian Lu, Roger Zimmermann, and Li Cheng. 2019. Towards Natural and Accurate Future Motion Prediction of Humans and Animals. In *CVPR*. 10004–10012. https://doi.org/10.1109/CVPR.2019.01024

[23] Taichi Nakashima and Yoshito Yabuta. 2018. Object Detection by using Interframe Difference Algorithm. In *2018 12th France-Japan and 10th Europe-Asia Congress on Mechatronics*. 98–102. https://doi.org/10.1109/MECATRONICS.2018.8495743

[24] Felix Naser, Igor Gilitschenski, Guy Rosman, Alexander Amini, Fredo Durand, Antonio Torralba, Gregory W Wornell, William T Freeman, Sertac Karaman, and Daniela Rus. 2018. Shadowcam: Real-time detection of moving obstacles behind a corner for autonomous vehicles. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 560–567.

[25] The National Safety Council of America. 2018. The Most Dangerous Time to Drive. (2018). https://www.nsc.org/road/safety-topics/driving-at-night?

[26] Rohit Pandharkar, Andreas Velten, Andrew Bardagjy, Everett Lawson, Moungi Bawendi, and Ramesh Raskar. 2011. Estimating motion and size of moving non-line-of-sight objects in cluttered environments. In *CVPR 2011*. IEEE, 265–272.

[27] Prashant W Patil and Subrahmanyam Murala. 2018. MSFgNet: A novel compact end-to-end deep network for moving object detection. *IEEE Transactions on Intelligent Transportation Systems* 20, 11 (2018), 4066–4077.

[28] Han Qiu, Yuchen Ma, Zeming Li, Songtao Liu, and Jian Sun. 2020. Borderdet: Border feature for dense object detection. In *European Conference on Computer Vision*. Springer, 549–564.

[29] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems* 28 (2015).

[30] Chang-Gyun Roh, Jisoo Kim, and I-Jeong Im. 2020. Analysis of impact of rain conditions on ADAS. *Sensors* 20, 23 (2020), 6720.

[31] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. 2011. ORB: An efficient alternative to SIFT or SURF. In *2011 International Conference on Computer Vision*. 2564–2571. https://doi.org/10.1109/ICCV.2011.6126544

[32] Sheila W Seidel, John Murray-Bruce, Yanting Ma, Christopher Yu, William T Freeman, and Vivek K Goyal. 2020. Two-dimensional non-line-of-sight scene estimation from a single edge occluder. *IEEE Transactions on Computational Imaging* 7 (2020), 58–72.

[33] Dongeek Shin, Ahmed Kirmani, Vivek K Goyal, and Jeffrey H Shapiro. 2015. Photon-efficient computational 3-D and reflectivity imaging with single-photon detectors. *IEEE Transactions on Computational Imaging* 1, 2 (2015), 112–125.

[34] Le-Anh Tran, Truong-Dong Do, Dong-Chul Park, and My-Ha Le. 2021. Enhancement of Robustness in Object Detection Module for Advanced Driver Assistance Systems. In *2021 International Conference on System Science and Engineering (ICSSE)*. IEEE, 158–163.

[35] Arizona State University. 2023. Pedestrian Injuries and Fatalities. (2023). https://popcenter.asu.edu/content/pedestrian-injuries-fatalities-0

[36] Cadillac vehicles. 2016. super-cruise: the first hands free driver-assistance technology for compatible roads. (2016). https://www.gmc.com/connectivity-technology/super-cruise

[37] Tianyu Wang, Xiaowei Hu, Chi-Wing Fu, and Pheng-Ann Heng. 2021. Single-stage instance shadow detection with bidirectional relation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 1–11.

[38] Tianyu Wang, Xiaowei Hu, Qiong Wang, Pheng-Ann Heng, and Chi-Wing Fu. 2020. Instance shadow detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 1880–1889.

[39] Tai Wang, Xinge Zhu, Jiangmiao Pang, and Dahua Lin. 2021. FCOS3D: Fully Convolutional One-Stage Monocular 3D Object Detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*. 913–922.

[40] World Health Organization (WHO). 2016. Road safety. (2016). https://www.who.int/data/gho/data/themes/road-safety

[41] A. Woo, P. Poulin, and A. Fournier. 1990. A survey of shadow algorithms. *IEEE Computer Graphics and Applications* 10, 6 (1990), 13–32. https://doi.org/10.1109/38.62693

[42] Timothy Woodford, Xinyu Zhang, Eugene Chai, and Karthikeyan Sundaresan. 2022. Mosaic: leveraging diverse reflector geometries for omnidirectional around-corner automotive radar. In *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*. 155–167.

[43] Jing Wu, Xin Du, Yun-fang Zhu, and Wei-kang Gu. 2008. Adaptive fuzzy filter algorithm for real-time video denoising. In *2008 9th International Conference on Signal Processing*. 1287–1291. https://doi.org/10.1109/ICOSP.2008.4697367

[44] Xian-Tao Wu, Wen Wu, Lin-Lin Zhang, and Yi Wan. 2022. Don't worry about noisy labels in soft shadow detection. *The Visual Computer* (2022), 1–12.

[45] Marie Yahiaoui, Hazem Rashed, Letizia Mariotti, Ganesh Sistu, Ian Clancy, Lucie Yahiaoui, Varun Ravi Kumar, and Senthil Yogamani. 2019. Fisheyemodnet: Moving object detection on surround-view cameras for autonomous driving. *arXiv preprint arXiv:1908.11789* (2019).

[46] Michael Ying Yang, Wentong Liao, Xinbo Li, and Bodo Rosenhahn. 2018. Deep learning for vehicle detection in aerial images. In *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 3079–3083.

[47] Rongjie Yu, Yin Zheng, and Xiaobo Qu. 2021. Dynamic driving environment complexity quantification method and its verification. *Transportation Research Part C: Emerging Technologies* 127 (2021), 103051.

[48] Quanlong Zheng, Xiaotian Qiao, Ying Cao, and Rynson WH Lau. 2019. Distraction-aware shadow detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 5167–5176.

[49] Haibo Zhou, Wenchao Xu, Jiacheng Chen, and Wei Wang. 2020. Evolutionary V2X Technologies Toward the Internet of Vehicles: Challenges and Opportunities. *Proc. IEEE* 108, 2 (2020), 308–323. https://doi.org/10.1109/JPROC.2019.2961937

[50] Barbara Zitova and Jan Flusser. 2003. Image registration methods: a survey. *Image and vision computing* 21, 11 (2003), 977–1000.