ICJK
Chinese Journal of Information Fusion
http://dx.doi.org/10.62762/CJIF.2025.500710

REVIEW ARTICLE

Check for updates

A Systematic Review on Real-time Detection of Small Obstacles Based on Multidimensional Information Fusion

Yong Zhang[1,2], **Zhenke Qian**[1] and **Junyan Cai**[3,*]

[1] College of Electronics and Information Engineering, Shenzhen University, Shenzhen 518060, China
[2] Guangdong Provincial Key Laboratory of Intelligent Information Processing, Shenzhen 518060, China
[3] Dongguan New-Generation Artificial Intelligence Industry Technology Research Institute, Dongguan 523000, China

Abstract

Real-time detection of small obstacles is a critical challenge for autonomous systems such as self-driving vehicles, unmanned aerial vehicles (UAVs), and mobile robots. These small obstacles (e.g., road debris, fallen branches, cables) pose significant safety risks due to their low visibility and irregular appearances. This paper presents a comprehensive systematic review of 117 technical articles published between 2016 and 2025, focusing on the techniques and deployment strategies for real-time small obstacle detection using fused multidimensional information. We summarize and analyze developments in small obstacle definitions, sensing hardware, detection algorithms, fusion methods, and real-time optimization techniques. Our findings reveal a growing trend toward integrating multiscale learning, multimodal fusion, and lightweight models for deployment in resource-constrained environments. However, challenges such as performance consistency across platforms, lack of standard definitions, and insufficient deployment practices persist. This review identifies future research opportunities and provides recommendations to guide further advances in robust, scalable, and deployable small obstacle detection systems.

Keywords: small obstacle detection, real-time detection, multidimensional information, multimodal fusion, lightweight models.

1 Introduction

Small obstacle detection is a major challenge for autonomous vehicles, robots and unmanned aerial vehicles (UAVs). Despite advances in sensing systems, obstacles smaller than 20 centimeters in size (e.g., road debris, fallen tree branches, or construction debris) remain a significant threat. The NHTSA report, "Safety Challenges for Automated Driving Systems," states that between 2022 and 2023, approximately 84 percent to 89 percent of automated driving system (ADS) out-of-control events are caused by active human driver interventions. These interventions are often

Academic Editor:
Deqiang Han

Submitted: 05 June 2025
Accepted: 23 October 2025
Published: 10 November 2025

Vol. 2, **No.** 4, 2025.
10.62762/CJIF.2025.500710

***Corresponding author:**
Junyan Cai
junyancai2023@163.com

Citation
Zhang, Y., Qian, Z., & Cai, J. (2025). A Systematic Review on Real-time Detection of Small Obstacles Based on Multidimensional Information Fusion. *Chinese Journal of Information Fusion*, 2(4), 313–339.

driven by safety concerns, especially if the system is unable to reliably detect potential obstacles [1]. In addition, a study revealed that two-thirds of manned aircraft pilots failed to spot nearby drones, especially stationary ones, indicating the difficulty in detecting small, static obstacles [2]. These accidents highlight the urgent need for a robust real-time detection system that can address the unique challenges posed by small obstacles: low visibility (e.g., low light or occlusion), dynamic behavior (e.g., moving animals or wind-blown debris), and environmental disturbances (e.g., rain, fog, or uneven terrain).

Traditional single-modal approaches (e.g., camera or LIDAR) have inherent limitations in real-time small-scale obstacle detection [3]. Intramodal enhancement techniques such as multiscale feature extraction (e.g., hierarchical convolution) and attentional mechanisms (e.g., spatial channel weighting) can improve local detail capture and contextual awareness [4]. Multimodal fusion (e.g., vision-LIDAR-radar) compensates for sensor-specific blind spots, while multi-view systems (e.g., stereo cameras or UAV swarms) enhance spatial perception through cross-view analysis [5, 6]. Despite these advances, integrating multiscale learning, multimodal synergy, and multiview coordination under real-time constraints remains challenging due to computational complexity and hardware resource limitations [7].

Based on various research results, a systematic review is necessary to conclude the current phase of research. Existing reviews on real-time detection of small or tiny obstacles report the current status, but the summarization of some key points is still insufficient: 1) There is no accurate definition, classification, and summary of small or tiny obstacles. 2) There is no complete update of small or tiny obstacle target detection algorithms and improvement methods. 3) There is no summary of the current state of the art of accelerating the small or tiny obstacle detection model, and deploying the system in real-time. 4) There is no multimodal and multidimensional feature related to the summarization.

This paper presents a systematic review of real-time detection of small obstacles based on multidimensional information in response to current limitations, with the following objectives: 1) to review the evolution of small obstacle detection systems, with a focus on algorithmic advances and deployment challenges; 2) to assess trends in multidimensional approaches, including intra-modal multi-scale learning, cross-modal fusion, and multi-view coordination; 3) to identify key gaps in the existing literature, such as fragmented sensor co-analysis and inadequate real-time optimization frameworks; and 4) summarize effective methods for real-time detection of small obstacles. Therefore, the main contributions of this paper are 1) proposing the definition of small or tiny obstacles and refining the classification of obstacles; 2) summarizing the latest algorithms for real-time detection of small obstacles and the ideas for improvement; 3) comprehensively sorting out the system implementation of real-time detection of small or tiny obstacles; 4) analysing and summarizing the multi-dimensional information processing methods of small obstacles; and 5) summarizing the methods for improving real-time detection of small or tiny obstacles and the methods for deploying them.

The structure of this paper is organized as follows: Section 2 reviews existing survey literature on small or tiny obstacle detection and outlines the core research questions of this study. Section 3 details the methodology of this systematic review, including the literature search strategy and screening process. Section 4 addresses six research questions through a structured analysis, covering application scenarios, definitions and classifications of small or tiny obstacles, sensing and hardware systems, detection algorithms and enhancement strategies, multidimensional information processing, and real-time deployment and model optimization. Section 5 summarizes the main findings and challenges in the current research. Section 6 discusses the limitations of this review and future research directions. Finally, Section 7 concludes the paper.

2 Literature Review

2.1 Definition of small or tiny obstacles and related research

At present, for the way of defining small or tiny obstacles, the relevant definitions of small targets can be referred to, which can be roughly divided into two categories: one category is the absolute size definition, for example, in the MS-COCO dataset, the target smaller than 32×32 pixels is categorized as a small object; whereas in the USC-GRAD-STDB dataset, the target with a size smaller than 16×16 pixels is regarded as a small object [8]. Another category is the relative size definition, according to the standard of the International Society of Photo-Electrical Engineers (SPIE), a small target is defined as a target area smaller than 9×9 pixels in a 256×256 pixel image; i.e., if the size

Table 1. Overview of the related reviews.

Author	Year	Review years	Definition	Hardware and Sensors	Datasets	Deep Learning	Obstacle Description	Multimodal	application Domains	Deployment Analysis
Tong et al. [11]	2022	2019-2021	✓		✓	✓				
Liu et al. [12]	2021	2016-2020			✓	✓				
Muzammul et al. [13]	2021	2014-2021	✓		✓	✓				
Lu et al. [14]	2023	2010-2022		✓		✓	✓	✓		
Yadikar et al. [15]	2023	2018-2022				✓				✓
Badrloo et al. [16]	2022	2002-2022		✓		✓	✓			
Iqra et al. [17]	2024	2004-2024	✓		✓	✓			✓	✓
Ristić-Durrant et al. [18]	2021	2003-2020		✓	✓	✓		✓	✓	✓
Wei et al. [19]	2022	2001-2021		✓	✓	✓		✓		✓
Mirzaei et al. [20]	2023	1991-2023		✓	✓	✓		✓		
Ours	**2025**	**2016-2025**	✓	✓	✓	✓	✓	✓	✓	✓

of an object is less than 0.12% of the original image area, it is considered as a small target. In addition, Krishnal et al. [9] stated that when an object occupies less than 1% of the image area, it can also be considered as a small target.

However, the above definitions are still insufficient when dealing with the widespread presence of small obstacles in real scenes. For example, power lines, large protective nets, or tree branches with large pixel sizes but very low reflectance at the perceptual level are often not accurately recognized in an image; conversely, certain targets that appear at long distances but do not actually affect path planning (e.g., a vehicle a few hundred meters away) may be misidentified as obstacles. This type of misidentification not only affects the labeling of obstacle categories in the construction of the dataset, but may also mislead the obstacle avoidance and path planning system, which in turn affects the overall detection performance.

Therefore, a more reasonable definition should focus on targets that pose a potential stability threat to unmanned systems (e.g., self-driving vehicles, drones, or robots) within their effective sensing range. For example, one study used "detecting targets less than 5cm in height with a low error rate within 20 meters" as a criterion for determining small obstacles [10].

2.2 Related reviews

After performing a database search using the terms "small obstacle detection", "small target detection" and "small or tiny obstacle recognition", we selected 10 review papers on small obstacle detection systems in the past 5 years. Table 1 summarizes the research topics, research focus, and analysis of the selected reviews.

In terms of the application scenarios and domain challenges of small obstacles, existing research mainly focuses on the generic problem of small target detection. Liu et al. [12] systematically summarize the main challenges and strategies of small target detection, and analyze them in depth with the research in multiple fields, such as generic target detection, face detection, remote sensing image analysis, and instance segmentation, etc. Muzammul et al. [13] focus on the cross-domain adaptation problem, and verify the effectiveness of multiple domain adaptation methods on typical datasets such as Cityscapes and KITTI. Irqa et al. [17] comprehensively analyze the existing datasets covered in the previous review articles, and explore the diversified application scenarios of small-target detection in different industries, and Wei et al. [19] clearly define or tiny the task requirements of 2D and 3D target detection for the autonomous driving environment. They clearly define or tiny the task requirements of 2D and 3D target detection in autonomous driving environments. Although small object detection has been comprehensively summarized and researched in the above mentioned fields, there is a lack of systematic review and research on the detection of small obstacles specifically used in obstacle avoidance scenarios. This gap also highlights the urgent need and research value of this direction in practical applications.

In terms of definition and classification of tiny obstacles, Tong et al. [11] explored the quantification criteria of small/tiny targets in different scenarios based on 30 datasets, such as targets smaller than 32×32 pixels defined in COCO. Lu et al. [14] focused on MAV non-cooperative obstacle avoidance technology, and categorized obstacle representations into the categories of features, distance fields, point clouds, maps and events. Badrloo et al. [16] proposed a spatial definition of narrow obstacles (width <15cm) and emphasized that the algorithm should be able to recognize multiple types of obstacles, including moving and omnidirectional obstacles, but the current dynamic characterization is still insufficient. Existing studies mostly use relative or absolute scales to define or tiny small or tiny obstacles, but there are limitations. For example, distant targets need not be detected even though they are small, while large targets with low reflectivity such as protective nets and glass should instead be included in the detection range. In addition, there is no clear delineation of obstacle categories and a unified definitional framework is lacking.

In terms of sensors and hardware technologies, Lu et al. [14] systematically compared MAV obstacle avoidance sensors, pointing out that event cameras have faster response (10μs) than traditional RGB (30fps); LiDAR has high accuracy but is susceptible to weather, and ultrasound is low-cost but has limited range (<5m). Ristić-Durrant et al. [18] point out that obstacle detection sensors in railways are similar to those in cars, covering both active and passive sensors, while Wei et al. [19] analyze 2D/3D detection tasks in autonomous driving, emphasizing the complementary nature of radar's immunity to interference and vision's high resolution. Although studies have summarized commonly used sensor types, the hardware of detection systems also includes specific deployment environments and platforms. There is still a lack of systematic analysis of the hardware differences between autonomous driving and UAV system.

In terms of algorithm improvement and technology optimization, Tong et al. [11] systematically reviewed seven aspects of small target detection methods, including super-resolution, context information, multi-scale learning, anchor frame mechanism, training strategy, data enhancement, and loss function, etc. Liu et al. [12] evaluated the performances of YOLOv3, Faster R-CNN, and SSD on several small target datasets, and pointed out that the overall detection accuracy was low, with Faster R-CNN being the best, but did not mention its lack of real-time performance. Liu et al. [12] evaluated the performance of YOLOv3, Faster R-CNN and SSD on several small target datasets, and pointed out that the overall detection accuracy is low, among which Faster R-CNN is the best, but did not mention its real-time insufficiency. Muzammul et al. [13] reviewed the latest methods of deep domain adaptation and small target detection, and proposed the strategies of multi-scale fusion, data augmentation, and GAN assisting, but did not deal with the integration and optimization of the multidimensional information processing methods, such as multimodal.

In terms of multidimensional information processing strategies, there is a lack of systematic summarization of multimodal fusion, multiscale feature fusion, multisensor fusion and multiview fusion. Except for Wei et al. [19] who pointed out the complementary nature of radar and vision in autonomous driving (e.g., radar is strong against interference, and camera provides high-resolution images), the related research generally analyzes superficially and lacks in-depth discussion and practical comparison.

In terms of real-time optimization and deployment challenges, Tong et al. [11] only briefly mention lightweight models such as YOLOv3-tiny and do not systematically analyze model compression techniques. Yadikar et al. [15] provide an overview of knowledge distillation methods in target detection from the perspective of isomorphic and heteromorphic distillation, which covers the roles of three types of knowledge: features, responses, and relationships, and points out that background features are beneficial for detection, and that the main focus is on combining features and responses for distillation. distillation by combining features and responses. However, they did not explore the specific application of lightweight models (e.g., MobileNet, SqueezeDet) and strategies such as pruning and quantization in small target detection, and other reviews have similar gaps.

2.3 Research question

Based on the review in Section 2.2, to address the research shortcomings of the above review papers, the following research questions are proposed, and the discussion encompasses small or tiny obstacle application scenarios and challenges, small or tiny obstacle definitions and classifications, detection hardware systems, algorithms and improvement methods, multi-signal processing strategies, model acceleration and deployment, as shown in Table 2.

3 Search Strategy

A search across four online databases identified 476 records. After removing duplicates, 455 studies remained for title and abstract screening. Following full-text review, 117 articles met the inclusion criteria, as shown in Figure 1.

3.1 Search strings and source searched

Between 2016 and 2025, electronic database searches were conducted on ScienceDirect, Scopus, and IEEE Xplore. These databases were selected because they offer extensive, multidisciplinary coverage of high-quality, peer-reviewed literature in engineering,

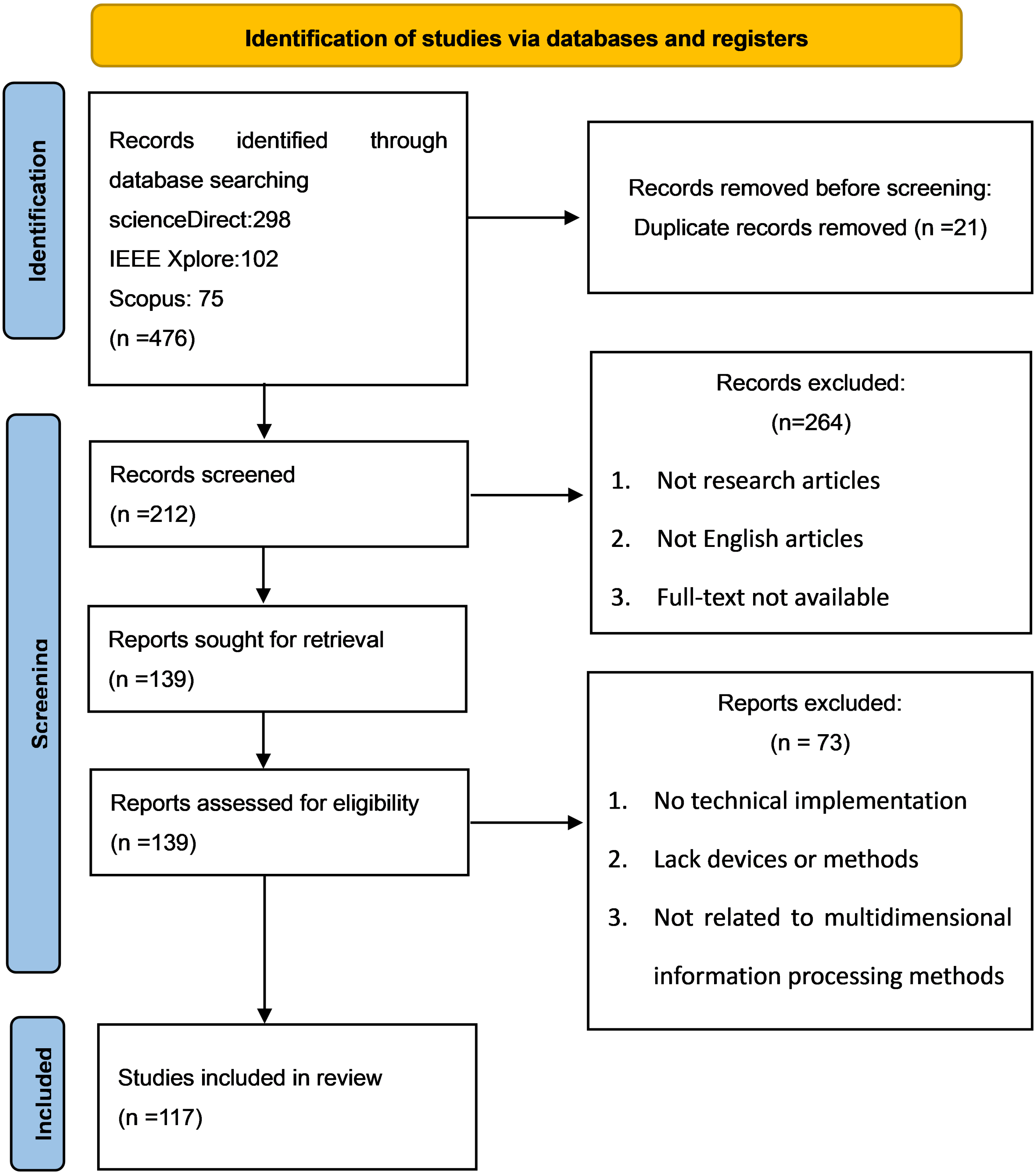

Figure 1. PRISMA flowchart.

computer science, and transportation research. Additionally, we manually searched the reference lists of the linked articles and databases using the terms listed in Table 3. It summarizes the number of retrieved papers from different databases using predefined search strings. ScienceDirect contributed the largest portion (298 papers), while IEEE Xplore and Scopus contributed 102 and 75 papers, respectively.

3.2 Inclusion and exclusion criteria

To systematically evaluate the selected articles, inclusion and exclusion criteria were established, as shown in Table 4. Only studies meeting all inclusion criteria and none of the following exclusion criteria were retained. An article was excluded if it met at least one exclusion condition.

Table 2. Research questions and motivations.

Research question	Motivation
RQ1: What are the core real-time detection tasks and application demands for small obstacle detection in autonomous driving and UAV scenarios?	To identify the core tasks and challenges of real-time tiny obstacle detection in autonomous driving and UAV scenarios.
RQ2: How are small or tiny obstacles defined and categorized?	To clarify the definition and categorization of small or tiny obstacles to facilitate the construction of rigorous datasets.
RQ3: What sensors and hardware are typically required for real-time detection of small or tiny obstacles?	To explore hardware systems for small or tiny obstacle detection.
RQ4: What algorithms and improvement ideas are currently used to detect small or tiny obstacles?	To analyse and summarize the latest algorithms to advance small or tiny obstacle detection.
RQ5: How to extract and fuse multidimensional information about small or tiny obstacles?	To summarize the multidimensional information fusion methods for small or tiny obstacles.
RQ6: How to accelerate real-time detection of small or tiny obstacles?	To summarize the deployment methodology for small or tiny obstacle detection.

Table 3. Search strings and results.

Database	Results	Keyword search and other applied filters
ScienceDirect	298	("Small objects detection" OR "Small obstacles avoidance") AND ("Multi-" OR "Fusion" OR "Multimodal" OR "Multi-scale") AND ("Autonomous Driving" OR "UAV") AND ("Real-time" OR "lightweight") AND affective Year: 2016–2025. Article type: Research articles
Scopus	75	("Small objects detection" OR "Small obstacles avoidance") AND ("Multi-" OR "Fusion" OR "Multimodal" OR "Multi-scale") AND ("Autonomous Driving" OR "UAV") AND ("Real-time" OR "lightweight") AND affective Year: 2016–2025. Article type: Research articles
IEEE Xplore	102	("Small objects detection" OR "Small obstacles avoidance") AND ("Multi-" OR "Fusion" OR "Multimodal" OR "Multi-scale") AND ("Autonomous Driving" OR "UAV") AND ("Real-time" OR "lightweight") AND affective Year: 2016–2025. Article type: Research articles

Table 4. Inclusion and exclusion criteria.

Inclusion criteria	Exclusion criteria
IC1: The publication year is within 10 years.	**EC1:** It's a duplicate paper.
IC2: Articles are written in English.	**EC2:** It is a review article, book chapter, patents, or conference paper.
IC3: Articles are related to small obstacle detection.	**EC3:** The full-text article is not available.
IC4: Articles use data or feature fusion approaches.	**EC4:** This article only focuses on normal size obstacle detection.
IC5: Articles are related to the real-time application scenarios	**EC5:** This article only focuses on normal size obstacle detection.

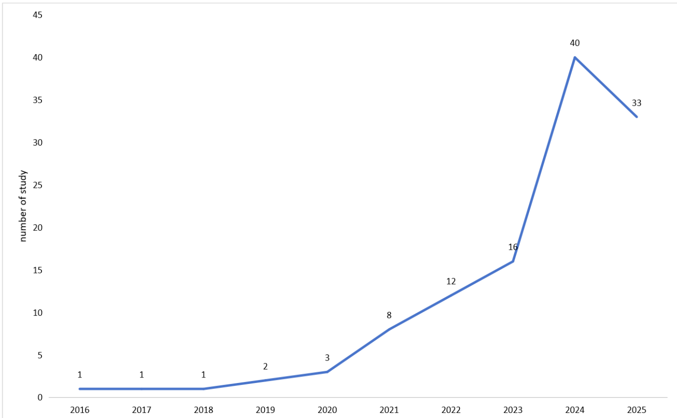

Figure 2. Research trends in small obstacle real-time detection.

4 Results

This section presents the results of the systematic review with reference to the research questions. These results are summarized by the following article, which focuses on the last 10 years. As shown in Figure 2, it can be seen that small target detection is a research hotspot in the past two years.

To facilitate a quick understanding of the review, Figure 3 presents a classification framework for real-time small obstacle detection based on existing references. The framework integrates the main aspects covered in this study—application demands, obstacle definitions and categories, sensor modalities, edge computing platforms, detection algorithms, and

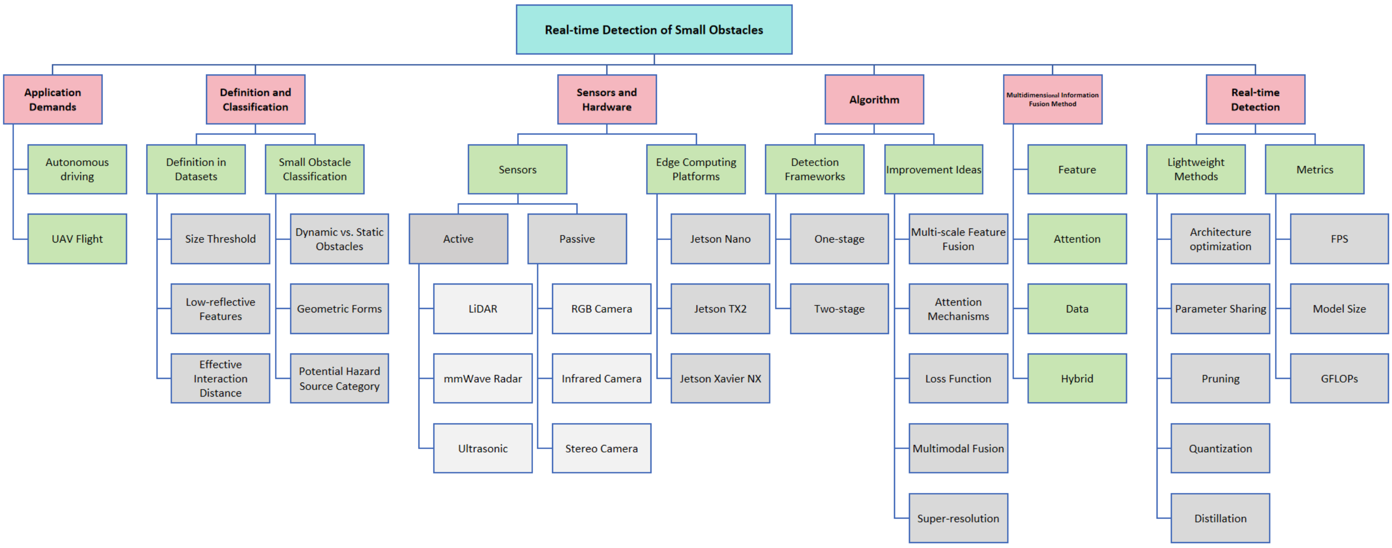

Figure 3. Taxonomy of real-time detection of small obstacles.

real-time considerations—while establishing links to the statistical results presented in later sections.

4.1 RQ1: What are the core real-time detection tasks and application demands for small obstacle detection in autonomous driving and UAV scenarios?

While most existing reviews provide a broad overview of small object detection across diverse domains—such as low-altitude surveillance, urban monitoring, fire inspection, and search and rescue—they often lack a focused summary of real-time detection tasks and specific application requirements related to tiny obstacle detection in autonomous driving and UAV environments. This gap highlights the need for a more targeted analysis of the operational challenges, performance constraints, and real-time processing needs unique to these critical scenarios.

Inside the selected reviews there are 80 articles focusing on autonomous driving application scenarios and 37 articles focusing on UAV scenarios, as shown in Figure 4.

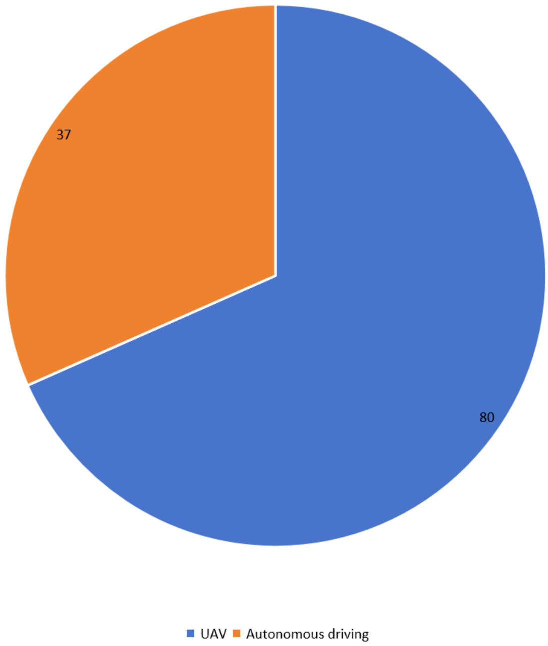

Figure 4. Integer of two application scenarios.

In the field of autonomous driving, real-time tiny obstacle detection mainly serves the two core tasks of driving safety and path planning. Studies have shown that multi-sensor fusion methods based on stereo vision and millimeter-wave radar are capable of detecting tiny targets such as ground debris and tire chips up to 5-10 cm within a range of 20 m, and meet the demand for low latency at high speeds with a frame rate of $\geq$20 Hz [21, 22]. In addition, modern deep learning networks (e.g., NVRadarNet) utilize sparse radar signals for real-time segmentation of dynamic obstacles and drivable areas under BEV viewpoints, which not only improves the detection accuracy of small targets, but also achieves beyond-real-time inference speed on embedded GPUs [21]. In low-speed or congested scenarios (parking lots, slow urban roads), lightweight RGB-D fused semantic segmentation networks (e.g., RFNet) can accurately identify low obstacles such as curbs, ground locks, and speed bumps, and assist the automatic parking and valet parking systems to complete fine positioning and obstacle avoidance [23].

In UAV scenarios, the core objective of small obstacle detection is to ensure vehicle navigation and obstacle avoidance, and reduce the risk of collision. Ultrasonic and small LiDAR sensors are commonly used to reduce on-board load and power consumption, while monocular/binocular vision and depth estimation algorithms are used to achieve 10-20 Hz detection and tracking of small targets in the range of 5-30 m [24, 25]. For example, the real-time obstacle avoidance method based on image processing can recognize small obstacles such as thin lines and cables in indoor and outdoor complex environments in a timely manner, and adjust the flight trajectory through the "Sense-and-Avoid" system, which significantly reduces the number of "blow-ups". The "Sense-and-Avoid" system adjusts the flight trajectory, which significantly reduces the occurrence rate of "blow-up" accidents [25]. Meanwhile, the lightweight target detection algorithm (e.g., the improved model based on YOLOv3) can achieve high recall and low false alarms for small static and dynamic obstacles in front of the flight path under the premise of ≥30 fps, which effectively improves the autonomous navigation capability of UAVs in GPS-restricted or dynamic environments [26]

4.2 RQ2: How are small or tiny obstacles defined and categorized?

The current review does not have a clear definition of small obstacles, this paper fills this gap and gives the following relevant definitions with reference to small target detection and gives common classifications.

4.2.1 Definition of small or tiny obstacles

In our review of 117 papers, we counted the datasets used in each study (see Figure 5): 22 self-built datasets, 52 VisDrone datasets, 17 UAVDT datasets, 5 DOTA datasets, 9 MS-COCO datasets, 9 Cityscapes datasets, 14 KITTI datasets, and 3 Lost and Found datasets.

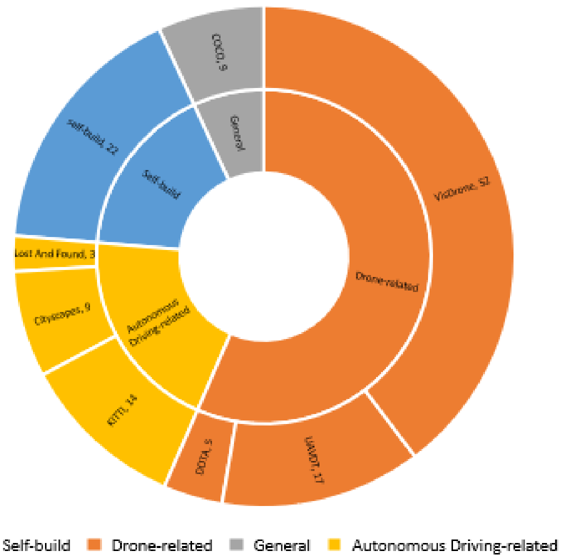

Figure 5. Datasets related to small or tiny obstacles.

These datasets have their own definitions of "small objects" (summarized in Table 5), and common criteria include: (i) pixel size less than 32 × 32 px (e.g., MS-COCO, DOTA, etc.); and (ii) the target area accounts for less than 1% of the total area of the image (e.g., VisDrone, UAVDT) [11]. In view of the high similarity between small or tiny obstacles and small targets in terms of size, texture, reflectivity, etc., this paper refers to the above definitions while noting that the current definition of small targets is still insufficient to cope with the wide range of small or tiny obstacles that are widely present in real scenes. For example, power lines, large protective nets, or tree branches that are large in pixel size but have very low reflectance at the perceptual level are often not accurately recognized in an image; conversely, certain targets that appear at a long distance but do not actually affect path planning (e.g., a vehicle a few hundred meters away) may be misidentified as obstacles. This kind of misclassification not only affects the labeling of obstacle categories in the construction of the dataset, but also may mislead the obstacle avoidance and path planning system, which in turn affects the overall detection performance. Therefore, a further definition of "small obstacles" is proposed:

- Size threshold: at 1920 × 1080 resolution, the surround box area is less than 2000 pixels (about 0.1%) and less than 50 pixels on either side of the length/width [27, 28];
- Low-reflective features: even if the pixel area exceeds the thresholds, objects like electric wires, large fences, and branches may still be considered small obstacles. Due to their material properties or angles, these objects often have low contrast or weak signals in grayscale or thermal images, making them difficult to detect and distinguish from the background using conventional models;
- Effective interaction distance: only targets that may be within the autopilot (5-20 m) or UAV obstacle avoidance buffer zone (2-10 m) are considered as obstacles to avoid false alarms for long-distance, no-real-impact targets [10].

4.2.2 Classification of Small or Tiny Obstacles

Based on the above definitions, we classify the small or tiny obstacles in different states into the following categories, as illustrated in Figure 6. Figure 6 provides visual examples of representative obstacles, helping

Table 5. An overview of some datasets about small or tiny objects.

Dataset	Definition of Small/Tiny Objects
MS-COCO [27]	The target occupies less than 32^2 pixels area in the image (area <1024 pixels).
VisDrone [28]	Target height in image < 20 pixels.
UAVDT [29]	Target height in image < 25 pixels.
DOTA [30]	Target actual pixel area < 10^2 (area < 100 pixels).
KITTI [31]	No specific definition of "small objects"; the assessment is divided by difficulty (easy/moderate/hard), with the hard category usually containing smaller and partially obscured objects.
Cityscapes [32]	The dataset does not categorize objects by size; it focuses on semantic segmentation without explicit small/medium/large object classification.
Lost And Found [10]	Small obstacles down to the height of 5cm can successfully be detected at 20m distance at low false positive rates.

to clarify the classification scheme described in the following subsections.

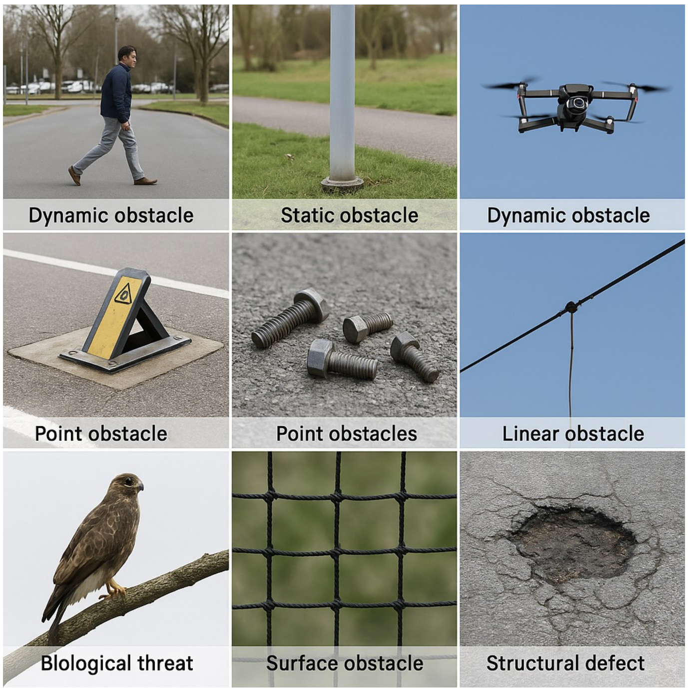

Figure 6. Representative examples of small or tiny obstacles. The nine categories include: (1) Dynamic obstacle — pedestrian; (2) Static obstacle — lamp post; (3) Dynamic obstacle — small drone; (4) Static obstacle — ground lock (parking wedge); (5) Point obstacle — bolts; (6) Linear obstacle — power line; (7) Biological threat — raptor (bird of prey); (8) Surface obstacle — net; (9) Structural defect — pothole.

Dynamic vs. Static Obstacles

- *Dynamic obstacles*. Refers to objects that have motion in the environment, such as pedestrians, animals, other vehicles, or drones. Due to its constantly changing position and speed, the detection system needs to track and predict its movement trajectory in real time to ensure safe avoidance.

- *Static obstacles*. Refers to objects that are stationary or have a constant position for a short period of time, such as curbs, ground locks, utility poles, etc. Despite their fixed position, these obstacles may still pose a challenge to the detection system due to their small size and low reflectivity.

Geometric Forms

- *Point obstacles*. Smaller in size, usually objects with a diameter or minimum side length of less than 10 centimeters, such as bolts, debris, etc. This type of obstacle is small in size, but may cause damage to the equipment, need to be detected and avoided.

- *Linear obstacles*. Slender objects whose length is much greater than their width, such as wires and cables. Due to its small cross-sectional area, and may be suspended in the air, detection is more difficult.

- *Surface obstacles*. Objects with a certain area but thin thickness, such as protective nets, plastic films, etc. Such obstacles may affect the normal operation of the sensor and require special attention.

Potential Hazard Source Category

- *Biological threats*. For animal obstacles, they can be further classified as non-aggressive (e.g., small birds, domestic animals) and potentially aggressive (e.g., wild boars, dog packs). The former may briefly affect the sensor's field of view, leading to perception barriers; while the latter may not only obscure the field of view, but also actively attack equipment or personnel, posing a direct threat and requiring special vigilance. Structural defects in road surfaces.

- *Ground defects*. such as potholes and cracks are static surface obstacles. These defects may be further enlarged by vehicle loads and environmental factors, and even lead to roadway collapse, seriously affecting traffic safety. Therefore, depending on the depth and width of the defects, their risk level to vehicle or drone traffic needs to be assessed in a timely manner and appropriate avoidance or repair measures taken.

4.3 RQ3: What sensors and hardware are typically required for real-time detection of small or tiny obstacles?

Very few reviews exist that comprehensively summarize hardware systems used for small obstacles, and this paper summarizes edge computing platforms in addition to input sensors.

4.3.1 Input Sensors

Real-time detection of small or tiny obstacles presents distinct technical challenges due to the low profile, weak reflectivity, and potential for occlusion of such objects. Sensor choice is critical to ensure accurate perception under varying environmental conditions and computational constraints. Based on their working principles, the sensors used in the reviewed studies can be broadly categorized into active and passive types, as shown in Figure 7. Then, Figure 8 summarizes their adoption frequency across the collected studies.

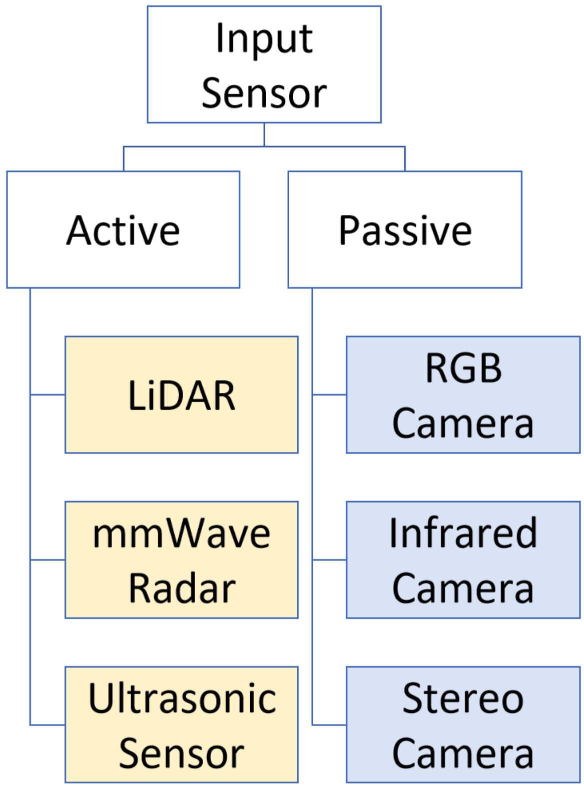

Figure 7. Input sensors for small or tiny obstacles detection.

Active Sensors

Active sensors emit energy signals (e.g., light, sound, electromagnetic waves) and analyze their reflections to reconstruct spatial and depth information. These

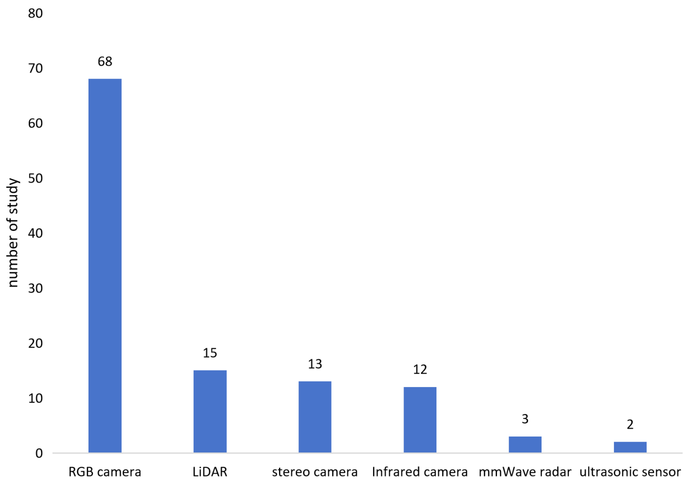

Figure 8. Input sensors for small or tiny obstacles detection.

sensors are particularly advantageous in detecting objects under variable lighting or environmental conditions.

- *LiDAR sensors* provide high-resolution 3D point clouds, making them particularly effective in identifying small obstacles with subtle vertical structures such as road debris, curbs, and dropped cargo. Their performance is robust across varying lighting conditions, but they are limited by high cost, relatively slow frame rates, and difficulty detecting low-reflectivity materials [19].
- *Millimeter-Wave Radar* sensors are valued for their robustness in adverse weather (fog, rain), but their relatively low spatial resolution limits their standalone application for detecting fine-grained or small static objects. In most reviewed studies, radar plays a supporting role in multimodal systems [19].
- *Ultrasonic sensors* offer simple and inexpensive obstacle detection in low-speed scenarios. However, they are highly limited in range and angular resolution, and thus rarely serve as primary sensors in dynamic or high-speed environments.

Passive Sensors

Passive sensors rely on ambient environmental signals (e.g., light, thermal radiation) and are widely used due to their low power consumption and compatibility with vision-based AI algorithms.

- *RGB cameras* is the most commonly used (68 studies, 60.2%) sensor type. RGB cameras provide high-resolution visual data that is well-suited to deep learning methods for object recognition. However, monocular setups lack depth perception

Table 6. List of commonly used GPU-based computing platform.

Specification	Jetson Nano	Jetson TX2	Jetson Xavier NX
AI Performance	0.5 TFLOPS	1.26–1.33 TFLOPS	Up to 21 TOPS
GPU	128-core Maxwell architecture	256-core Pascal architecture	384-core Volta architecture with 48 Tensor Cores
CPU	Quad-core ARM Cortex-A57	Dual-core Denver 2 + Quad-core ARM Cortex-A57	6-core NVIDIA Carmel ARMv8.2 64-bit
Memory	4GB LPDDR4, 25.6 GB/s	8GB LPDDR4, 59.7 GB/s	8GB LPDDR4x, 59.7 GB/s
Storage	microSD slot	16GB eMMC 5.1	16GB eMMC 5.1
Video Encode/Decode	4Kp30 encode/decode	4Kp60 encode/decode	8Kp30 decode, 4Kp60 encode
Camera Support	1x MIPI CSI-2 interface	Up to 6 cameras, 12x MIPI CSI-2 lanes	Up to 6 cameras, 14x MIPI CSI-2 lanes
Interfaces	USB 3.0, HDMI, GPIO, I2C, SPI	USB 3.0, HDMI, PCIe Gen2, GPIO, I2C, SPI	USB 3.1, HDMI 2.0, PCIe Gen3, GPIO, I2C, SPI
Power Consumption	5–10W	7.5–15W	10–15W (configurable)
Dimensions	69.6mm x 45mm	87mm x 50mm	69.6mm x 45mm
Suitable Applications	Education, lightweight robotics, low-power vision applications	Industrial robots, drones, embedded AI devices	High-performance edge AI, complex multi-sensor fusion systems

and can be unreliable in poor lighting or highly reflective environments. These weaknesses are often mitigated via multimodal sensor fusion or data augmentation.

- *Stereo cameras* offer both visual texture and depth estimation, making them useful for 3D localization of small obstacles. They are more affordable than LiDAR and suitable for embedded systems, although their performance suffers in textureless or low-light regions. Song et al. [34] utilized stereo vision to achieve end-to-end real-time obstacle detection by jointly learning semantic segmentation and parallax maps through the RODSNet network.

- *Infrared/Thermal cameras* are particularly useful for detecting living obstacles (e.g., animals, pedestrians) in low-light or nighttime conditions. However, their utility in detecting non-thermal small objects such as debris is limited, and they are often used in conjunction with other sensor modalities [33].

4.3.2 *Edge Computing Platforms*

In a real-time detection system for small obstacles, in addition to sensing hardware, edge computing platforms determine whether detection systems can perform real-time inference within the latency and energy constraints of embedded deployment. In the literature counted, there are several studies that explicitly report on the use of NVIDIA's Jetson family of edge computing modules. Seven of these articles deployed a small obstacle detection system to the Jetson Nano edge computing device, 2 to the Jetson TX2, and 9 to the Jetson Xavier NX, as shown in the Figure 9. The performance of each edge computing device is shown in Table 6.

- *Jetson Nano* is suitable for resource-constrained application scenarios such as educational programs, lightweight robotics, low-power vision tasks, and simple detection tasks. Its low cost and small size make it ideal for entry-level AI projects, but may be limited in computational power when working with complex deep learning models. Rawlley et al. [35] utilized the nano edge computing platform for traffic video stream analysis and obstacle detection tasks, improving accuracy by up to 20%, reducing training time to less than 60%, and increasing real-time inference speed by up to 30 FPS compared to the baseline model under different autonomous driving environments.

- *Jetson TX2* delivers greater computational performance than the Nano, as shown in Table 6, and is suitable for industrial applications

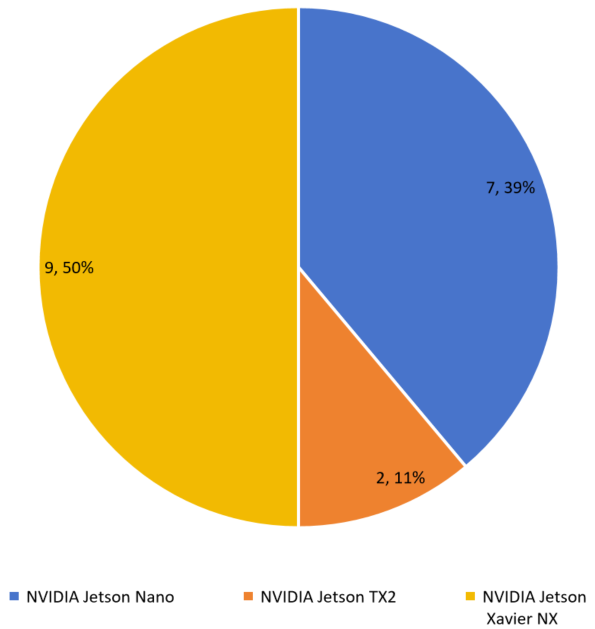

Figure 9. Edge computing platforms for small or tiny obstacles detection.

requiring moderate computational resources, such as drones, industrial robots, and embedded AI devices. Its higher memory bandwidth and diverse interface support allow it to excel in multi-sensor data processing. Wu et al. [36] utilized the Jetson TX2 edge computing platform for the insulator defect detection task in UAV-assisted power inspection, with a mAP of 81.3% in detection accuracy, and a frame rate FPS of 49.7 in real-time, realizing a low-latency and high-precision detection effect.

- Jetson Xavier NX delivers powerful AI computing capabilities while maintaining a small footprint for complex applications requiring high-performance edge computing, such as autonomous driver assistance systems, intelligent surveillance and multi-sensor fusion robotic systems. Its AI performance of up to 21 TOPS and rich interface support make it the first choice for high-end embedded AI projects. Other detailed parameters can be found in Table 6. Min et al. [37] utilized the Jetson Xavier NX platform for real-time UAV target detection. The LWUAVDet-S model has a mAP of 26.5% and an mAP@0.5 of 46.7% on the VisDrone dataset, and the detection speed reaches 44.8 FPS on this platform.

4.4 RQ4: What algorithms and improvement ideas are currently used to detect small or tiny obstacles?

Real-time detection algorithms for small obstacles have not been updated as summarized in several reviews, this paper summarizes the real-time detection algorithms in the last 10 years and introduces them into two categories: one-stage and two-stage, and also summarizes the latest and effective improvement strategies.

4.4.1 One-Stage Detection Framework

Among the 117 technical papers we researched, 100 of them adopted one-stage methods, accounting for about 85%, as shown in Figure 10 and Table 7. one-stage detectors have become the mainstream choice for real-time detection of small obstacles due to their end-to-end structure, which eliminates the need for a separate candidate frame generation phase, and possesses the advantages of compact modeling and fast inference. This class of methods is especially suitable for application scenarios with very high response time requirements, such as autonomous driving and robot navigation. The current one-stage models mainly include the YOLO series and the DETR series.

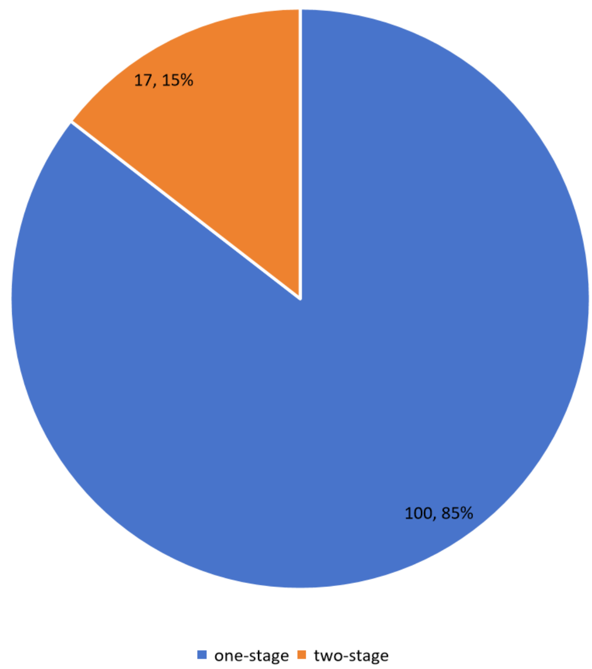

Figure 10. Detection framework for small or tiny obstacles detection.

YOLO (You Only Look Once) series, as the most representative one-stage method shown in Figure 11, has been widely used for small obstacle detection since YOLOv3 [42, 44, 45, 50, 52]. YOLOv4 [48, 53,

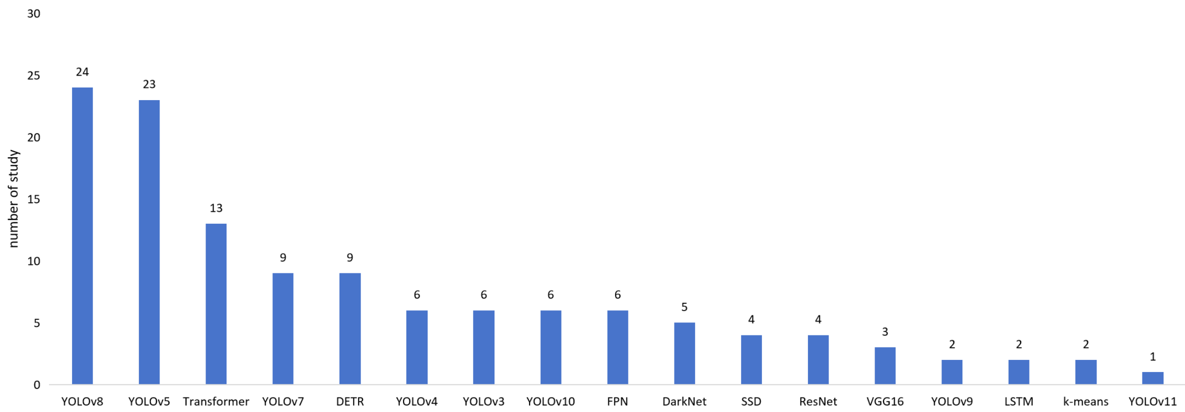

Figure 11. Algorithm models for small or tiny obstacles detection.

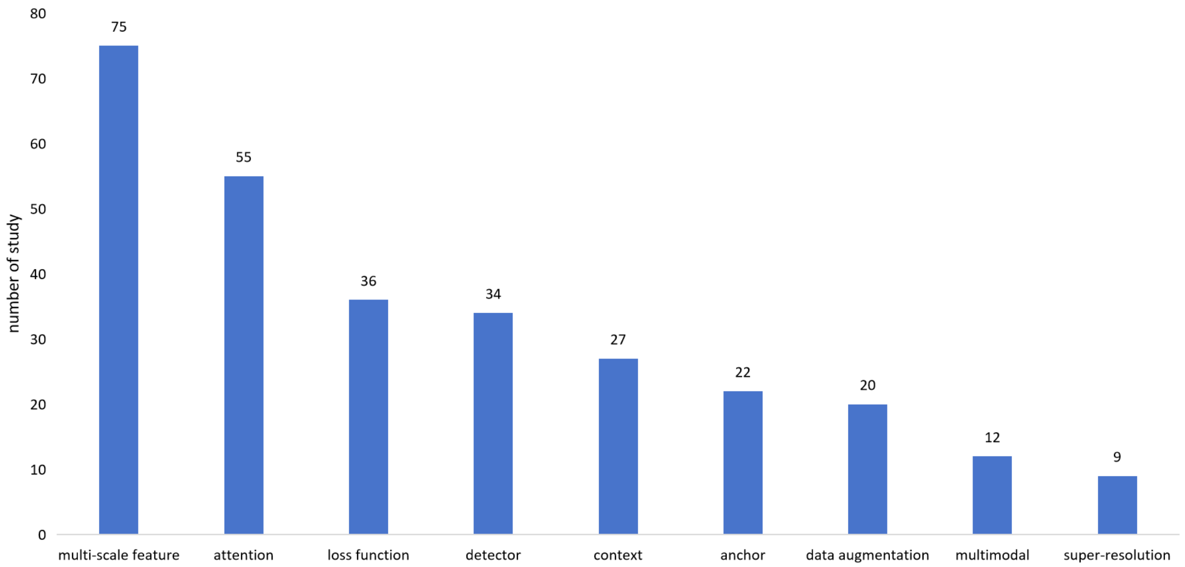

Figure 12. Improvement ideas for small or tiny obstacles detection.

Table 7. Research paper statistics for small obstacle detection algorithm framework.

Algorithm Framework	Studies
One-stage	[39], [40], [42], [44], [45], [47], [48], [50], [52]-[57], [59]-[77], [79]-[89], [91]-[117], [119]-[134], [136]-[143], [145], [148], [150]
Two-stage	[38], [41], [43], [46], [49], [51], [58], [78], [90], [118], [135], [144], [149]

54, 61] improves the accuracy while maintaining the real-time performance through the introduction of CSPDarkNet [42, 54], PANet, and Mish activation function; YOLOv5 [59, 62, 63, 66, 85] further optimizes the network structure and introduces techniques such as AutoAnchor mechanism and Mosaic data enhancement. YOLOv7 innovates in lightweight design and detection head structure to improve the adaptability of small target detection [68, 69, 72, 79]. And YOLOv8, as the latest generation, has been deeply optimized in terms of backbone, Neck structure and NMS strategy to make it more robust to tiny obstacles in complex scenes [82–84, 92, 93]. In small obstacle detection tasks, researchers usually combine multi-scale detection heads, attentional modules (e.g., CBAM, SE), and context-aware modules to enhance the YOLO series model's ability to perceive tiny targets. In addition, improving the loss function (e.g., CIoU, Focal Loss) has also become a key means to enhance the detection accuracy of YOLO small targets.

DETR (*Detection Transformer*) *series*. The DETR model introduces the Transformer structure into the target detection task for the first time, which no longer relies on the NMS with anchor frame mechanism, and realizes the direct prediction of the target through sequence modeling. Although the original DETR has certain advantages in accuracy, its slow convergence speed and insensitivity to small targets become the main problems limiting its practical application [89, 98, 143, 145].

4.4.2 Two-Stage Detection Framework

Compared with one-stage methods, two-stage methods focus more on detection accuracy and usually include a Region Proposal Network (RPN) with a classification regression network. Among the 117 papers we investigated, only 17 papers adopted the two-stage structure, accounting for about 15%. Although it is relatively weak in terms of real-time performance, it still shows advantages in certain high-precision detection tasks, especially in small-sample scenarios and applications requiring fine-grained feature recognition.

Faster R-CNN, as the earliest model to introduce an end-to-end RPN mechanism, establishes benchmark performance in a wide range of detection tasks. Its regression and classification of candidate regions through shared convolutional features enable significant improvement in detection accuracy. In small obstacle detection tasks, researchers usually combine the feature pyramid (FPN) to enhance feature extraction of high-resolution small targets, while adjusting the anchor frame size and positive and negative sample matching strategies of the RPN to make it more adaptable to the size distribution of tiny objects. In addition, its recognition ability for small targets is also further optimized by introducing post-processing methods such as IoU-balanced Sampling strategy or Soft-NMS [49, 90]. To solve the problem of regression error accumulation in traditional two-stage detectors, Cascade R-CNN adopts a multi-stage step-by-step optimization strategy, which raises the IoU threshold at each stage to optimize the quality of candidate frames layer by layer. This mechanism has a natural advantage in small-target detection, as it can refine the low-quality candidate region multiple times to improve the detection accuracy.

However, the two-stage approach is often not applicable to mobile or real-time systems with very high latency requirements due to its complex structure and high computational overhead. This is the main reason for its relatively low percentage in our research sample.

4.4.3 Improvement Ideas for Small Obstacles Detection

In order to overcome the challenges of small-size obstacles in the target detection task, such as small target size, blurred texture, strong background interference, and easy to be occluded, researchers have proposed various improvement strategies around mainstream detection models. These improvements mainly focus on feature extraction, information aggregation, loss modeling, structural design, and multi-source information utilization to improve the detection accuracy and robustness of the models in real-world scenarios. According to our statistics of 117 papers, the main improvement methods and frequency of use are shown in Figure 12.

Multi-scale feature fusion. Small obstacles often occupy only a very small pixel area in the image and are easily lost in the high-level semantic features, so fusion of features at different scales becomes one of the key strategies to improve detection performance. The most common approach is to introduce a feature pyramid structure (FPN), which combines shallow detail information with deep semantic information by means of bottom-up feature aggregation to enhance the perception of small targets [51, 94]. BiFPN further introduces a weighted fusion mechanism with bi-directional connectivity, which improves the efficiency of the feature flow and retention of key features [56, 59]. Nghiem et al. [137] improved the Feature Pyramid Network (FPN) to enhance small target detection by adding a P2 header and adjusting the number of ELAN blocks in the PAFPN architecture to extract small target features using a high-resolution feature map.

Attention Mechanisms. To enhance the network's ability to attend to small obstacles, attention mechanisms have been widely introduced into detection frameworks. Channel attention mechanisms such as Squeeze-and-Excitation (SE) [97] and Efficient Channel Attention (ECA) [112] enhance key information expression by enhancing the channel feature responses associated with small targets; spatial attention mechanisms such as CBAM (Convolutional Block Attention Module) [59, 65, 112, 121] and PAM (Position Attention Module) [52, 113] pay more attention to the spatial saliency localization of small targets, thus improving the localization accuracy. In addition, there have been studies designing joint attention modules that fuse channel and spatial attention for finer local area modeling. These mechanisms are often embedded into backbone or Neck modules, which work in concert with multiscale fusion strategies to significantly improve the detectability and classification discriminative power of small targets. Yao et al. [112] integrated four attention mechanisms, namely, Convolutional Block Attention Module (CBAM), Efficient Channel Attention (ECA), Stochastic Attention (SA), and Global Attention Mechanism (GAM), into a YOLOv8-based road indicator detection model, which enabled the model to achieve significant improvement in the S2TLD, BSTLD datasets in terms of the mAP@0.5 and mAP@0.5:0.95 metrics are significantly improved. Yu et al. [113] proposed a multi-path attention mechanism (MPAM) to enhance small target detection, which consists of multi-dimensional pooling and self-attentive feature extraction module, which enables the model to reach 86.1% mAP@0.5 value on TT100K dataset, which is a 9.6% improvement compared to YOLOv8n.

Loss functions in standard target detection (e.g., intersection and concurrency ratio IoU, categorical cross-entropy) suffer from positive and negative sample imbalance and large regression bias when dealing with small targets. Therefore, some studies have proposed loss function improvement strategies applicable to small obstacle detection [38, 39, 44, 45, 53]. Hu et al. [86] replaced the CIOU loss function of YOLOv5s with Inner-SIOU loss function to improve small target detection, and the mAP@0.5 of the improved model was improved by 0.6% and 5.2% on CCTSDB2021 and TT100K datasets, respectively, and the detection speed was accelerated. Niu et al. [138] used Task-Aligned Assigner to improve the loss function by combining BCE Loss, DFL Loss and CIOU Loss. Among them, BCE Loss measures the classification difference, DFL Loss improves the regression flexibility, and CIOU Loss considers multiple geometric factors to stabilize the training, so that the mAP@0.5 of VSTDet-l on VisDrone2019 and AI-TODv2 datasets reaches 50.3% and 63.3%, respectively, which effectively improves the detection of small targets.

Contextual semantic enhancement. Due to the lack of information about small obstacles themselves, leveraging the contextual information around them becomes an effective means of recognition. Common practices include the introduction of Dilated Convolution to expand the sensory field or the integration of context-aware modules (e.g., ASPP, Context Encoding Module) to model the semantic associations between the target and the environment. Some studies have also employed Graph Convolutional Networks (GCNs) to construct structural relationships between pixel nodes in the feature graph, thereby enhancing the fusion of small targets with their semantic regions [91, 92, 95, 109, 113]. Ahmad et al. [123] proposed DSFF-Net, which enhances the context-awareness of the network by introducing the DynamicConv (DC) module and generates an attention map using the Smooth Attention (SA) module to highlight target features and suppress background interference. The method achieves 76.6% and 87.8% mAP on DOTA v1.0 and HRSC2016 datasets respectively, which significantly outperforms most of the compared models, and performs well in the small target detection task.

Multimodal information fusion. Single visual modality has limited ability to recognize small obstacles in harsh environments such as low light, rain and fog, so multimodal information fusion has become a hot research topic in recent years. Common fusion approaches include the alignment fusion of LiDAR point cloud and RGB images, cross-attention modeling of millimeter-wave radar and visual features, and joint decoding of feature maps for infrared and visible images [49, 57, 82, 99]. By introducing structural complementary information or temperature-sensitive features, the model maintains good perceptual stability for small targets in complex scenarios such as nighttime, rainy day, and occlusion. Yang et al. [111] enhanced small target detection by fusing image and LiDAR data. Image semantics compensate for point cloud sparsity, while a cross-modal attention mechanism focuses on small targets and adaptively balances feature contributions. On the KITTI dataset, their method improved mAP from 52.3% to 61.7%, demonstrating a significant performance boost.

Image super-resolution. Since small obstacles have sparse pixels in the original image, it is often difficult for detectors to extract effective features. To solve this problem, some studies have used image super-resolution (Super-Resolution) techniques to preprocess the input image so that the small obstacles region can be enhanced in spatial resolution. Common methods include reducing low-resolution images to high-quality feature maps based on super-resolution reconstruction networks such as ESRGAN and EDSR, and performing target detection on this basis. This strategy is especially suitable for small obstacles detection tasks under working conditions such as long-distance shooting and supervision of low-altitude cities [120, 136, 144, 145]. Yang et al. [145] employed a super-resolution approach based on the EDSR network to enhance image resolution using residual learning, Pixel Shuffle, and multi-scale training. They also applied data augmentation techniques such as noise injection and geometric transformations to enrich training samples. This strategy improved small target detection performance, achieving a 3.8% gain in both mAP@0.5 and mAP@0.5:0.95, while also enhancing the model's adaptability to complex environments.

Overall, multi-scale feature fusion and attention mechanisms are undoubtedly the two most widely used and effective classes of methods in small obstacle detection at present, and they have been validated in most mainstream detection architectures. Other improvement strategies show unique advantages in specific scenarios, constituting a diverse and complementary research pattern.

4.5 RQ5: How to extract and fuse multidimensional information about small or tiny obstacles?

In the small obstacle field detection algorithms, there are many research papers using a variety of multi-dimensional information extraction and processing methods, but there is no review to systematically summarize, this paper on this gap to develop the part of the review.

Small obstacles usually have the characteristics of small size, weak texture, and easy to be occluded, leading to the fact that single modal information (e.g., RGB images) is often insufficient to support the detection task with high robustness and high accuracy. For this reason, more and more researchers have begun to try to fuse multi-dimensional information and utilize data from different sources to comprehensively portray small obstacles, in order to enhance the perceptual and generalization capabilities of detection models. As shown in Figure 13 is the statistical frequency of using multidimensional information processing strategies.

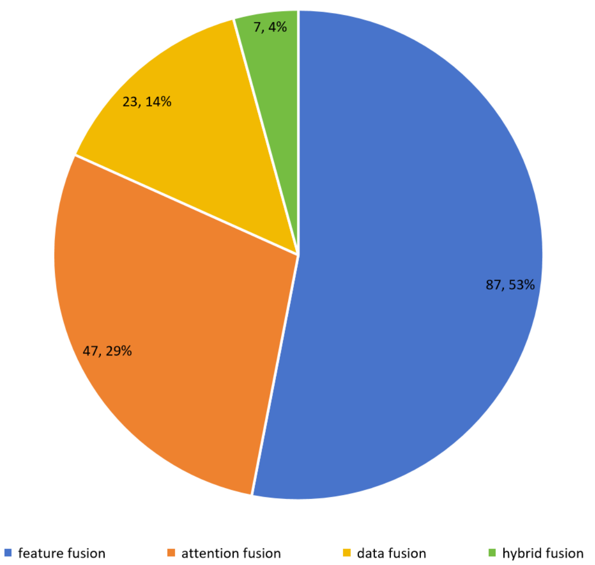

Figure 13. Fusion methods for small or tiny obstacles detection.

Multiscale feature fusion is currently the most commonly used method for extracting multidimensional information, with a total of 87 papers dealing with related strategies, accounting for nearly 53% of the total. The core idea of this method is to simultaneously utilize deep semantic information and shallow spatial details to construct cross-level multiscale representations, thus enhancing the model's responsiveness to targets of different sizes (especially small targets). Typical structures such as Feature Pyramid Network (FPN), PANet, BiFPN, etc., have been widely integrated into mainstream detection models such as YOLOv5/YOLOv8, RetinaNet and EfficientDet [125, 128, 136].

Attention mechanisms provide fine-grained perceptual capabilities for multidimensional information extraction, and have also been frequently employed in small obstacle detection tasks, with a total of 47 papers addressing them. Fusion attention strategies generally model feature importance in both the channel and spatial domains, highlighting the response of weak target regions. For example, lightweight channel attention mechanisms such as SE (Squeeze-and-Excitation) and ECA (Efficient Channel Attention) can be used to enhance the saliency of small targets, while CBAM (Convolutional Block Attention Module) and PAM (Position Attention Module) take into account both spatial position perception and feature enhancement. In practice, this class of methods proves to be particularly effective for small obstacle detection in extreme scenes such as visual blur, low contrast, and heavy occlusion [97, 112, 113].

Multimodal fusion methods use information from radar (LiDAR, millimeter-wave radar), infrared thermography, and depth maps to enhance the ability to model weakly textured or occluded targets. A total of 23 of the papers we researched use multimodal information fusion techniques. Among them, one class of methods uses early fusion, where data from different sensors are spliced or encoded at the input level (e.g., projecting a point cloud onto the image plane and fusing the channels); another class of methods uses intermediate fusion, where semantic information extracted from the image and the point cloud, respectively, is fused at the intermediate feature layer through an attention mechanism or a two-branch network; and some other work uses a late fusion strategy, where, at the level of detection results Joint decision making of multi-source information is performed to improve the stability and reliability of detection. The multimodal fusion method performs particularly well in complex weather, low-light, reflection interference and other scenes, and has gradually become a key component of highly robust detection systems [49, 57, 82, 99].

4.6 RQ6: How to accelerate real-time detection of small or tiny obstacles?

Current reviews rarely analyse and summarize the deployment of small obstacle detection algorithms, and the summarized lightweight methods are not comprehensive enough. In this paper, we first comprehensively summarize the lightweight methods of small obstacle detection algorithms in the past 10 years, and then further analyse and summarize the real-time performance of the current models.

In practical applications, the detection system of small obstacles not only requires high accuracy, but also must have good real-time performance, which is especially critical in time-sensitive scenarios such as autonomous driving, robot navigation and unmanned devices. In order to speed up model inference and reduce the consumption of computational resources, researchers have proposed and widely applied a variety of model lightweight techniques. According to the statistics of the literature in this review, as shown in Figure 14, a total of 69 articles use model structure optimization methods, 15 articles use parameter sharing techniques, 13 articles apply pruning strategies, 8 articles use model quantization, and only 3 articles use knowledge distillation. Each of these methods has its own advantages and disadvantages and is applicable to different practical demand scenarios.

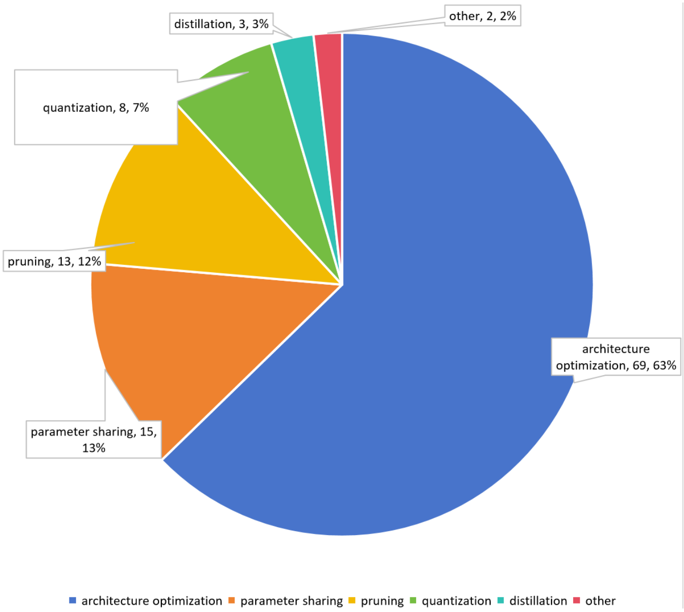

Figure 14. Lightweight methods for small or tiny obstacles detection.

Architecture optimization is one of the most mainstream real-time enhancement tools, and its core idea is to reduce computational complexity and model size by designing lightweight network structures. For example, network architectures such as MobileNet [111, 116], ShuffleNet, EfficientNet [121], and GhostNet [110, 119, 137] significantly reduce model size and inference time while maintaining high accuracy. In addition, techniques such as deep separable convolution, channel reorganization, and dynamic convolution have been introduced to further reduce computational redundancy. Structure optimization is widely adopted mainly because it reduces the computational burden from the source, is suitable for designing lightweight models from scratch or replacing the original complex model structure, and has good migrability and platform adaptability.

Parameter sharing achieves the goal of compressing model size and reducing storage overhead by reusing some parameters (e.g., convolution kernel, attention module, etc.) in the network [58, 126, 134]. This approach is more common in Transformer structure and multi-scale feature fusion networks, which can reduce memory consumption while maintaining a certain expressive power. Since this method imposes certain restrictions on model design and has limited optimization space, it is used relatively infrequently.

Pruning is a method to reduce model complexity by removing redundant parameters or channels under the premise of maintaining the original structure or performance of the model, which is mainly categorized into the following two types: structured pruning cuts channels, convolution kernels, layers, or feature maps as a unit, and the pruned model still has a clear structural advantage, which makes it easy to be deployed on hardware. For example, channels that contribute less to a convolutional layer can be evaluated and pruned to make the model lighter. This approach is widely used in real-world deployments and is especially suitable for embedded platforms. Unstructured pruning, on the other hand, sparsifies for individual weights or connections, and reduces the model computation by threshold filtering to zero out connections with smaller weights. Although this approach can achieve higher compression rates, it is less friendly to hardware due to its irregular sparse structure, and often needs to be combined with sparse matrix optimization tools to get the desired acceleration effect [128, 136, 139].

Model quantization is used to reduce storage and

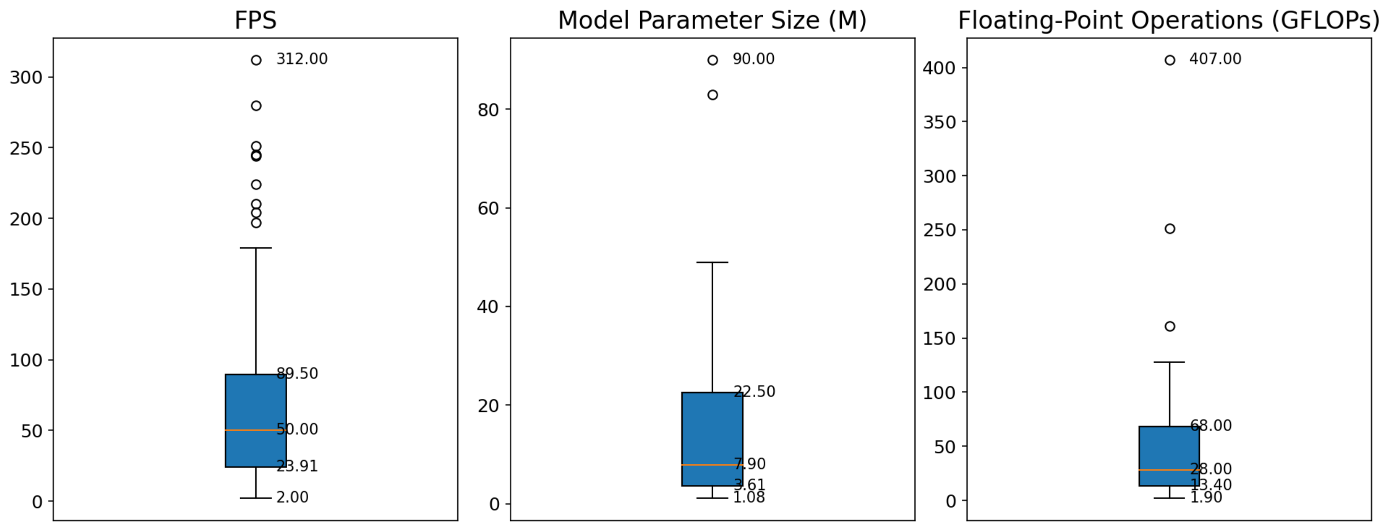

Figure 15. Box diagram of real-time metrics for small or tiny obstacles detection.

computation requirements by compressing model parameters from a high-bit (e.g., 32-bit float) to a low-bit (e.g., 8-bit int) representation. Common techniques include symmetric or asymmetric quantization of weights and activation values, or even exploring extreme compression methods with even lower bit counts (e.g., 4-bit, binarization). However, in small obstacle detection tasks, quantization may lead to degradation of detection performance, especially the ability to capture edge details with weak features [64, 120, 136]. As a result, the method is applied relatively rarely and is often combined with strategies such as quantization-aware training (QAT) to minimize the accuracy loss.

Knowledge distillation improves the performance of a small model (student model) by allowing it to imitate the output behavior of a large model (teacher model) [121]. This method is more common in image classification tasks, but in small obstacle detection, due to the sparse distribution of targets and the complex design of the distillation mechanism, there are fewer studies using this method, mainly focusing on its use in combination with structural optimization.

In order to further understand the practical deployment performance of current small obstacle detection methods, this review provides statistics on the model inference metrics disclosed in the collected literature, which include frame rate (FPS), model size (in MB) and floating point operations (GFLOPs). The statistical results are shown in the box plots in Figure 15, and the statistical characteristics are shown in Table 8.

Frame rate (FPS) is one of the key indicators to measure

Table 8. Statistical characterization of real-time metrics for small or tiny obstacles detection.

Metrics	FPS	Model Size(M)	GFLOPs
Count	93	63	48
Mean	74.4	16.16	52.58
Standard Deviation	69.22	19.29	70.77
Minimum	2	1.08	1.9
25th Percentile	23.91	3.61	13.40
Median	50	7.9	28
75th Percentile	89.5	22.5	68
Maximum	312	90	407

the real-time performance of a model. Overall, the average frame rate of the 93 papers is 74.4 FPS, which indicates that most models have basically met the real-time processing requirements on specific hardware platforms (more than 30 FPS is usually considered as real-time). However, the standard deviation reaches 69.22, the maximum value is as high as 312 FPS [75], and the minimum value is only 2 FPS, indicating that some high-precision models have the problem of slow inference, and it is still difficult to be applied to the scenarios that are highly sensitive to delay. Further observing the quartiles: 25% of the models have a frame rate lower than 23.91 FPS, which is suitable for quasi-real-time tasks; the median is 50 FPS; and 75% of the models have a frame rate higher than 89.5 FPS, which has the potential for efficient deployment. This distribution reflects that although some models are obviously lightweight, their reasoning ability has not been fully optimized, especially when introducing complex attention mechanisms or multi-scale processing in target detection, which is prone to cause frame rate degradation.

Model size. Among the 63 papers that disclose the model size, the average model size is 16.16 MB, the minimum is only 1.08 MB, and the maximum is 90 MB. the median is 7.9 MB, indicating that most models have good compression performance and can be deployed in edge devices or embedded platforms with limited memory resources. In general, smaller models tend to achieve higher inference speeds (measured in FPS) due to reduced computational complexity. However, model size alone does not fully determine inference performance. Lightweight models may still suffer from low frame rates if they lack proper operator-level optimization or hardware acceleration. Conversely, larger models are not necessarily slower—when well-optimized and executed on compatible hardware, they can also deliver real-time performance. Therefore, inference speed is ultimately influenced by a combination of model size, architecture design, optimization strategies, and hardware support, rather than being dictated by size alone.

GFLOPs (floating point operations per second) are used to measure the amount of theoretical computation in the inference phase of a model. 48 Relevant literatures show that the mean value of GFLOPs is 52.58, the median is 28, but the standard deviation is as high as 70.77, and the maximum value is 407 GFLOPs. This suggests that a large number of high-complexity structures in the current tiny obstacle detection models (e.g., the Transformer structures, feature pyramids, attention mechanisms) are used to enhance the recognition of tiny targets. However, in practical deployment, high GFLOPs models usually bring more pressure on latency and power consumption, so more and more researches are turning to designing model structures with low GFLOPs but efficient feature extraction capabilities, such as introducing dynamic feature fusion, channel selection, hierarchical cropping, and other mechanisms.

5 Discussion

This review systematically analyzes 117 technical research papers on real-time small obstacle detection based on multidimensional information over the past decade, covering a wide range of dimensions from definitions, sensors and hardware platforms, detection algorithms, improved methods, multimodal processing strategies to model acceleration and deployment optimization. The research results show that although the field has made significant progress in method innovation and system integration, it still faces many challenges in terms of robustness, efficiency, and generalization ability to cope with complex real-world environments.

First, in terms of the definition and categorization of small obstacles, existing studies mostly refer to standards in the field of small target detection, but often fail to consider the degree of influence of obstacles on path planning in practical applications. This study proposes a definitional framework that combines size thresholds, reflective properties and interaction distances, which can more effectively identify small obstacles that truly threaten the operation of unmanned systems, especially low-reflective thin lines, mesh structures or translucent objects.

Second, in terms of sensing hardware, RGB cameras are still the most commonly used sensors, but they perform poorly in extreme environments such as low light or rainy foggy conditions. Therefore, researchers have started to introduce a variety of sensors such as millimeter-wave radar, infrared thermography, binocular vision with LIDAR to build a multimodal fusion system in order to improve the system's ability to detect weak texture or occluded targets. However, how to realize the efficient fusion of multimodal data without adding too much computational burden is still a technical difficulty.

Third, in terms of detection algorithms, one-stage models (e.g., YOLO series) have become the mainstream choice due to their end-to-end structure and fast inference speed, especially performing well in resource-constrained scenarios. However, they are still inferior to some two-stage models in terms of detection accuracy. The former generally improves the perception of tiny obstacles by sacrificing a small amount of real-time to introduce the attention mechanism and multi-scale structure. In the future, we can consider further exploiting the lightweight potential of two-stage structures while maintaining real-time performance.

Fourth, for multidimensional information processing, multi-scale feature fusion and attention mechanisms are the most common and effective strategies. Especially, the attention mechanism (e.g., CBAM, SE, ECA, etc.) plays a key role in improving the model's responsiveness to small obstacle regions. Meanwhile, multimodal fusion (e.g., vision-radar, RGB-depth, infrared-visible, etc.) is gradually becoming a mainstream trend, which demonstrates stronger stability in complex environments. However, there is a lack of a unified fusion framework and a systematic comparison of the advantages and disadvantages of fusion strategies (early/mid/late) in different tasks.

Fifth, regarding real-time deployment and accelerated optimization, researchers have proposed rich lightweight methods, including structure optimization (e.g., MobileNet, GhostNet), parameter sharing, pruning, quantization and knowledge distillation, etc., and some of them have been successfully deployed on embedded platforms such as Jetson Nano, TX2 and Xavier NX. Although most models have achieved inference speeds above 30 FPS, they still suffer from elevated latency when introducing complex structures such as Transformer or BiFPN. In addition, there is still a lack of a deployment standard and performance evaluation system that is common across different hardware platforms.

In summary, although the current real-time small obstacle detection system based on multi-dimensional information has made significant progress, further in-depth research is still needed in defining standards, multi-modal fusion, model optimization and cross-platform deployment. The future development direction should be more focused on the unity of practicality and generalization ability, and promote the technology to really land in the real-world scenarios such as autonomous driving, low-altitude aircraft and mobile robots.

6 Limitations and future work

Despite the comprehensive review of real-time small obstacle detection systems based on multidimensional information over the past decade, several limitations still exist in this study: First, there may be blind spots in the literature collection. Although this review covered three major academic databases—ScienceDirect, Scopus, and IEEE Xplore—with carefully designed keyword search strategies, some unpublished or gray literature (e.g., technical reports, industrial white papers, patents) may not have been included. This could result in the omission of certain cutting-edge engineering practices.

Second, the evaluation of performance metrics lacks unified benchmarks. When analyzing the real-time performance of detection models (e.g., FPS, model size, GFLOPs), inconsistencies in hardware platforms, datasets, and evaluation protocols across studies limit direct comparisons. Although this paper attempts to mitigate such issues through statistical methods (e.g., medians and quartiles), it cannot entirely eliminate the performance variations caused by heterogeneous experimental setups. Third, there is a lack of quantitative analysis of fusion strategies. While this review categorizes multimodal fusion methods (e.g., early, mid, late fusion), most studies do not publicly share code or provide detailed ablation studies. Therefore, it is difficult to perform fine-grained quantitative comparisons of fusion strategies in typical real-world scenarios.

Fourth, practical deployment cases are not comprehensively summarized. Although this paper discusses deployments on platforms such as Jetson Nano, TX2, and Xavier NX, few industrial-grade cases are publicly available. As such, it remains difficult to present a complete picture of the transition from lab-scale prototypes to real-world engineering deployments.

Based on these limitations, future work could be advanced in the following directions. 1) Establish standardized evaluation benchmarks and open-source testing platforms; 2) Explore deeper cross-modal perception and fusion frameworks; 3) Promote platform-aware model adaptation; 4) Expand research on real-world deployment in practical scenarios.

Overall, the research on real-time small obstacle detection systems is evolving from model design to holistic system-level integration and deployment. The intersection of "multimodal fusion + lightweight optimization + scenario generalization" will likely define the next key direction of progress.

7 Conclusion

This paper presents a comprehensive systematic review of 117 studies on real-time small obstacle detection using multidimensional information over the past decade. It synthesizes the field's key developments across definition frameworks, sensor systems, detection algorithms, enhancement strategies, multimodal fusion techniques, and deployment optimization. By addressing six core research questions, this review not only outlines the strengths and trends of current solutions but also identifies critical limitations that hinder practical adoption.

Findings reveal that while single-modality systems suffer from robustness issues in complex environments, strategies such as multiscale feature fusion, attention mechanisms, and multimodal perception significantly improve detection performance for small and weakly textured targets. Meanwhile, lightweight model architectures and deployment on edge hardware platforms (e.g., Jetson series) have demonstrated encouraging real-time capabilities. However, balancing accuracy with speed and ensuring platform adaptability remain key bottlenecks for large-scale applications.

In conclusion, the field is steadily transitioning from model-centered research to system-level integration that includes perception, computation, and deployment. Future research should emphasize: (1) establishing standard definitions and evaluation metrics; (2) exploring deeper mechanisms of multimodal fusion; (3) designing efficient yet accurate architectures; and (4) advancing field deployment in real-world scenarios. As intelligent autonomous systems become increasingly pervasive, building robust, efficient, and deployable small obstacle perception systems will be a foundational capability across domains such as intelligent transportation, urban management, and industrial automation.

Data Availability Statement

Not applicable.

Funding

This work was supported by the Project of Shenzhen Science and Technology Plan under Grant KJZD20230923114405012 and Guangdong Provincial Key Laboratory under Grant 2023B1212060076.

Conflicts of Interest

Yong Zhang is an employee of Guangdong Provincial Key Laboratory of Intelligent Information Processing, Shenzhen 518060, China, Junyan Cai is an employee of Dongguan New-Generation Artificial Intelligence Industry Technology Research Institute, Dongguan 523000, China. The authors declare no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Khan, M., LeMaster, D., & Najm, W. G. (2024). Understanding Safety Challenges of Vehicles Equipped with Automated Driving Systems (ADS): Analysis of ADS Disengagements. *US Department of Transportation, Highly Automated Systems Safety Center of Excellence*, 2024-08.

[2] Werfelman, L. (2019, October 30). *Study: Two-thirds of manned aircraft pilots failed to spot nearby drones*. Flight Safety Foundation. Retrieved from https://flightsafety.org/study-two-thirds-of-manned-aircraft-pilots-failed-to-spot-nearby-drones/

[3] Song, J., Zhao, L., & Skinner, K. A. (2024). Lirafusion: Deep adaptive lidar-radar fusion for 3d object detection. In *2024 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 18250-18257). IEEE. [CrossRef]

[4] Laga, H., Jospin, L. V., Boussaid, F., & Bennamoun, M. (2020). A survey on deep learning techniques for stereo-based depth estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 44*(4), 1738-1764. [CrossRef]

[5] Drews, F., Feng, D., Faion, F., & Glaser, C. (2022). Deepfusion: A robust and modular 3d object detector for lidars, cameras and radars. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 560-567). IEEE. [CrossRef]

[6] Zhao, C., Song, Y., Yang, X., & Chen, Z. (2023). Target detection based on multi-scale feature fusion and cross-channel interactive attention mechanism. In *Journal of Physics: Conference Series* (Vol. 2562, No. 1, p. 012046). IOP Publishing. [CrossRef]

[7] Deng, Y., Hu, X., Li, B., & Liu, Y. (2023). Multi-scale self-attention-based feature enhancement for detection of targets with small image sizes. *Pattern Recognition Letters, 166*, 46-52. [CrossRef]

[8] Bosquet, B., Mucientes, M., & Brea, V. M. (2020). STDnet: Exploiting high resolution feature maps for small object detection. *Engineering Applications of Artificial Intelligence, 91*, 103615. [CrossRef]

[9] Krishna, H., & Jawahar, C. V. (2017, November). Improving small object detection. In *2017 4th IAPR Asian conference on pattern recognition (ACPR)* (pp. 340-345). IEEE. [CrossRef]

[10] Pinggera, P., Ramos, S., Gehrig, S., Franke, U., Rother, C., & Mester, R. (2016, October). Lost and found: detecting small road hazards for self-driving vehicles. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 1099-1106). IEEE. [CrossRef]

[11] Tong, K., & Wu, Y. (2022). Deep learning-based detection from the perspective of small or tiny objects: A survey. *Image and Vision Computing, 123*, 104471. [CrossRef]

[12] Liu, Y., Sun, P., Wergeles, N., & Shang, Y. (2021). A survey and performance evaluation of deep learning methods for small object detection. *Expert Systems with Applications, 172*, 114602. [CrossRef]

[13] Muzammul, M., & Li, X. (2021). A survey on deep domain adaptation and tiny object detection challenges, techniques and datasets. *arXiv preprint*

arXiv:2107.07927.

[14] Lu, L., Fasano, G., Carrio, A., Lei, M., Bavle, H., & Campoy, P. (2023). A comprehensive survey on non-cooperative collision avoidance for micro aerial vehicles: Sensing and obstacle detection. *Journal of Field Robotics*, *40*(6), 1697-1720. [CrossRef]

[15] Yadikar, N., & Ubul, K. (2023). A review of knowledge distillation in object detection. *IEEE Access*. [CrossRef]

[16] Badrloo, S., Varshosaz, M., Pirasteh, S., & Li, J. (2022). Image-based obstacle detection methods for the safe navigation of unmanned vehicles: A review. *Remote Sensing*, *14*(15), 3824. [CrossRef]

[17] Iqra, Giri, K. J., & Javed, M. (2024). Small object detection in diverse application landscapes: a survey. *Multimedia Tools and Applications*, *83*(41), 88645-88680. [CrossRef]

[18] Ristić-Durrant, D., Franke, M., & Michels, K. (2021). A review of vision-based on-board obstacle detection and distance estimation in railways. *Sensors*, *21*(10), 3452. [CrossRef]

[19] Wei, Z., Zhang, F., Chang, S., Liu, Y., Wu, Q., & Li, Z. (2022). Mmwave radar and vision fusion for object detection in autonomous driving: A review. *Sensors*, *22*(7), 2542. [CrossRef]

[20] Mirzaei, B., Nezamabadi-Pour, H., Raoof, A., & Lee, M. (2023). Small object detection and tracking: a comprehensive review. *Sensors*, *23*(15), 6887. [CrossRef]

[21] Popov, A., Gebhardt, P., Chen, K., & Feng, S. (2023). Nvradarnet: Real-time radar obstacle and free space detection for autonomous driving. In *2023 IEEE International Conference on Robotics and Automation* (*ICRA*) (pp. 6958-6964). IEEE. [CrossRef]

[22] Haris, M., & Hou, J. (2020). Obstacle Detection and Safely Navigate the Autonomous Vehicle from Unexpected Obstacles on the Driving Lane. *Sensors*, *20*(17), 4719. [CrossRef]

[23] Sun, L., Yang, K., Hu, X., Hu, W., & Wang, K. (2020). Real-time fusion network for RGB-D semantic segmentation incorporating unexpected obstacle detection for road-driving images. *IEEE robotics and automation letters*, *5*(4), 5558-5565. [CrossRef]

[24] Chandran, N. K., Sultan, M. T. H., Łukaszewicz, A., Shahar, F. S., Holovatyy, A., & Giernacki, W. (2023). Review on Type of Sensors and Detection Method of Anti-Collision System of Unmanned Aerial Vehicle. *Sensors*, *23*(15), 6810. [CrossRef]

[25] Merei, A., Mcheick, H., Ghaddar, A., & Al-Fuqaha, A. (2025). A Survey on Obstacle Detection and Avoidance Methods for UAVs. *Drones*, *9*(3), 203. [CrossRef]

[26] Du, Z., Feng, X., Li, F., & Zhang, Y. (2024). A Lightweight UAV Visual Obstacle Avoidance Algorithm Based on Improved YOLOv8. *Computers, Materials & Continua*, *81*(2). [CrossRef]

[27] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014, September). Microsoft coco: Common objects in context. In *European conference on computer vision* (pp. 740-755). Cham: Springer International Publishing. [CrossRef]

[28] Du, D., Zhu, P., Wen, L., Bian, X., Lin, H., Hu, Q., ... & Liu, Z. (2019, October). VisDrone-DET2019: The Vision Meets Drone Object Detection in Image Challenge Results. In *2019 IEEE/CVF International Conference on Computer Vision Workshop* (*ICCVW*) (pp. 213-226). IEEE. [CrossRef]

[29] Du, D., Qi, Y., Yu, H., Yang, Y., Duan, K., Li, G., ... & Lyu, S. (2018). The unmanned aerial vehicle benchmark: Object detection and tracking. In *Proceedings of the European Conference on Computer Vision* (*ECCV*) (pp. 370-386). [CrossRef]

[30] Xia, G. S., Bai, X., Ding, J., Zhu, Z., Belongie, S., Luo, J., ... & Zhang, L. (2018). DOTA: A large-scale dataset for object detection in aerial images. In *2018 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*) (pp. 1-10). [CrossRef]

[31] Geiger, A., Lenz, P., & Urtasun, R. (2012, June). Are we ready for autonomous driving? the kitti vision benchmark suite. In *2012 IEEE conference on computer vision and pattern recognition* (pp. 3354-3361). IEEE. [CrossRef]

[32] Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., ... & Schiele, B. (2016, June). The Cityscapes Dataset for Semantic Urban Scene Understanding. In *2016 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*) (pp. 3213-3223). IEEE. [CrossRef]

[33] Bhadoriya, A. S., Vegamoor, V., & Rathinam, S. (2022). Vehicle detection and tracking using thermal cameras in adverse visibility conditions. *Sensors*, *22*(12), 4567. [CrossRef]

[34] Song, T. J., Jeong, J., & Kim, J. H. (2022). End-to-end real-time obstacle detection network for safe self-driving via multi-task learning. *IEEE Transactions on Intelligent Transportation Systems*, *23*(9), 16318-16329. [CrossRef]

[35] Rawlley, O., Gupta, S., Mahajan, K., & Kumar, S. (2024). Green-EMulTO: a next generation edge-assisted multi-level traffic orchestrator for green computing in consumer autonomous vehicles. *IEEE Transactions on Consumer Electronics*. [CrossRef]

[36] Wu, J., Jing, R., Bai, Y., & Liu, Y. (2024). Small insulator defects detection based on multiscale feature interaction transformer for UAV-assisted power IoVT. *IEEE Internet of Things Journal*, *11*(13), 23410-23427. [CrossRef]

[37] Min, X., Zhou, W., Hu, R., Wu, Y., Pang, Y., & Yi, J. (2024). Lwuavdet: A lightweight uav object detection network on edge devices. *IEEE Internet of Things Journal*, *11*(13), 24013-24023. [CrossRef]

[38] Wang, J., Zhu, M., Sun, D., Wang, B., Gao, W., & Wei, H. (2019). MCF3D: Multi-Stage Complementary Fusion for Multi-Sensor 3D Object Detection. *IEEE Access*, *7*, 90801-90814. [CrossRef]

[39] Yang, J., Xie, X., & Yang, W. (2019). Effective Contexts for UAV Vehicle Detection. *IEEE Access*, *7*, 85042-85054. [CrossRef]

[40] Ma, W., Wu, Y., Cen, F., & Wang, G. (2020). MDFN: Multi-scale deep feature learning network for object detection. *Pattern Recognition*, *100*, 107149. [CrossRef]

[41] Sun, L., Yang, K., Hu, X., Hu, W., & Wang, K. (2020). Real-Time Fusion Network for RGB-D Semantic Segmentation Incorporating Unexpected Obstacle Detection for Road-Driving Images. *IEEE Robotics and Automation Letters*, *5*, 5558-5565. [CrossRef]

[42] Zhang, H., Qin, L., Li, J., Guo, Y., Zhou, Y., Zhang, J., & Xu, Z. (2020). Real-Time Detection Method for Small Traffic Signs Based on Yolov3. *IEEE Access*, *8*, 64145-64156. [CrossRef]

[43] Anand, B., Senapati, M., Barsaiyan, V., & Rajalakshmi, P. (2021). LiDAR-INS/GNSS-Based Real-Time Ground Removal, Segmentation, and Georeferencing Framework for Smart Transportation. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1-11. [CrossRef]

[44] Li, Q., Garg, S., Nie, J., Li, X., Liu, R. W., Cao, Z., & Hossain, M. S. (2020). A highly efficient vehicle taillight detection approach based on deep learning. *IEEE transactions on intelligent transportation systems*, *22*(7), 4716-4726. [CrossRef]

[45] Li, Z., Liu, X., Zhao, Y., Liu, B., Huang, Z., & Hong, R. (2021). A lightweight multi-scale aggregated model for detecting aerial images captured by UAVs. *Journal of Visual Communication and Image Representation*, *77*, 103058. [CrossRef]

[46] Li, Z., Zhou, A., Pu, J., & Yu, J. (2021). Multi-Modal Neural Feature Fusion for Automatic Driving Through Perception-Aware Path Planning. *IEEE Access*, *9*, 142782-142794. [CrossRef]

[47] Sun, C., Wen, M., Zhang, K., Meng, P., & Cui, R. (2021). Traffic sign detection algorithm based on feature expression enhancement. *Multimedia Tools and Applications*, *80*(25), 33593-33614. [CrossRef]

[48] Tao, Y., Zongyang, Z., Jun, Z., Xinghua, C., & Fuqiang, Z. (2021). Low-altitude small-sized object detection using lightweight feature-enhanced convolutional neural network. *Journal of Systems Engineering and Electronics*, *32*, 841-853. [CrossRef]

[49] Wu, Y., Jiang, X., Fang, Z., Gao, Y., & Fujita, H. (2021). Multi-modal 3D object detection by 2D-guided precision anchor proposal and multi-layer fusion. *Applied Soft Computing*, *108*, 107405. [CrossRef]

[50] Zeng, T., Wang, J., Cui, B., Wang, X., Wang, D., & Zhang, Y. (2021). The equipment detection and localization of large-scale construction jobsite by far-field construction surveillance video based on improving YOLOv3 and grey wolf optimizer improving extreme learning machine. *Construction and Building Materials*, *291*, 123268. [CrossRef]

[51] Chen, T. H., & Chang, T. S. (2022). RangeSeg: Range-Aware Real Time Segmentation of 3D LiDAR Point Clouds. *IEEE Transactions on Intelligent Vehicles*, *7*, 93-101. [CrossRef]

[52] Chen, X., Li, Y., & Nakatoh, Y. (2022). Pyramid attention object detection network with multi-scale feature fusion. *Computers and Electrical Engineering*, *104*, 108436. [CrossRef]

[53] Junos, M. H., Khairuddin, A. S. M., & Dahari, M. (2022). Automated object detection on aerial images for limited capacity embedded device using a lightweight CNN model. *Alexandria Engineering Journal*, *61*(8), 6023-6041. [CrossRef]

[54] Liu, X., Leng, C., Niu, X., Pei, Z., Cheng, I., & Basu, A. (2022). Find Small Objects in UAV Images by Feature Mining and Attention. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1-5. [CrossRef]

[55] Wang, K., Wang, Y., Zhang, S., Tian, Y., & Li, D. (2022). SLMS-SSD: Improving the balance of semantic and spatial information in object detection. *Expert Systems with Applications*, *206*, 117682. [CrossRef]

[56] Ye, T., Zhang, J., Li, Y., Zhang, X., Zhao, Z., & Li, Z. (2022). CT-Net: An Efficient Network for Low-Altitude Object Detection Based on Convolution and Transformer. *IEEE Transactions on Instrumentation and Measurement*, *71*, 1-12. [CrossRef]

[57] Yuan, S., Zhang, Q., Zhu, L., Wang, S., Zang, Y., & Zhao, X. (2022). Multi-level object detection by multi-sensor perception of traffic scenes. *Neurocomputing*, *514*, 486-499. [CrossRef]

[58] Yuan, Y., Cheng, H., & Sester, M. (2022). Keypoints-based deep feature fusion for cooperative vehicle detection of autonomous driving. *IEEE Robotics and Automation Letters*, *7*(2), 3054-3061. [CrossRef]

[59] Zhang, Q., Tang, J., Zheng, H., & Lin, C. (2022). Efficient object detection method based on aerial optical sensors for remote sensing. *Displays*, *75*, 102328. [CrossRef]

[60] Zhang, Y., Zhang, Y., Fu, R., Shi, Z., Zhang, J., Liu, D., & Du, J. (2022). Learning Nonlocal Quadrature Contrast for Detection and Recognition of Infrared Rotary-Wing UAV Targets in Complex Background. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-19. [CrossRef]

[61] Zhou, Y., Wen, S., Wang, D., Meng, J., Mu, J., & Irampaye, R. (2022). MobileYOLO: Real-time object detection algorithm in autonomous driving scenarios. *Sensors*, *22*(9), 3349. [CrossRef]

[62] Cao, Y., Li, C., Peng, Y., & Ru, H. (2023). MCS-YOLO: A Multiscale Object Detection Method for Autonomous Driving Road Environment

Recognition. *IEEE Access, 11*, 22342-22354. [CrossRef]

[63] Cui, M., Gong, G., Chen, G., Wang, H., Jin, M., Mao, W., & Lu, H. (2023). LC-YOLO: A lightweight model with efficient utilization of limited detail features for small object detection. *Applied Sciences, 13*(5), 3174. [CrossRef]

[64] Dong, Y., Ma, Y., Li, Y., & Li, Z. (2023). High-precision real-time UAV target recognition based on improved YOLOv4. *Computer Communications, 206*, 124-132. [CrossRef]

[65] Wang, J., Zhang, F., Zhang, Y., Liu, Y., & Cheng, T. (2023). Lightweight object detection algorithm for uav aerial imagery. *Sensors, 23*(13), 5786. [CrossRef]

[66] Kim, S., Hong, S. H., Kim, H., Lee, M., & Hwang, S. (2023). Small object detection (SOD) system for comprehensive construction site safety monitoring. *Automation in Construction, 156*, 105103. [CrossRef]

[67] Li, Z., Pang, C., Dong, C., & Zeng, X. (2023). R-YOLOv5: A Lightweight Rotational Object Detection Algorithm for Real-Time Detection of Vehicles in Dense Scenes. *IEEE Access, 11*, 61546-61559. [CrossRef]

[68] Liang, B., Su, J., Feng, K., Liu, Y., & Hou, W. (2023). Cross-Layer Triple-Branch Parallel Fusion Network for Small Object Detection in UAV Images. *IEEE Access, 11*, 39738-39750. [CrossRef]

[69] Liu, Y., Li, C., Huang, J., & Gao, M. (2023). Minesds: A unified framework for small object detection and drivable area segmentation for open-pit mining scenario. *Sensors, 23*(13), 5977. [CrossRef]

[70] Ma, C., Fu, Y., Wang, D., Guo, R., Zhao, X., & Fang, J. (2023). YOLO-UAV: Object Detection Method of Unmanned Aerial Vehicle Imagery Based on Efficient Multi-Scale Feature Fusion. *IEEE Access, 11*, 126857-126878. [CrossRef]

[71] Mahaur, B., Mishra, K. K., & Kumar, A. (2023). An improved lightweight small object detection framework applied to real-time autonomous driving. *Expert Systems with Applications, 234*, 121036. [CrossRef]

[72] She, F., Hong, Z., Zeng, Z., & Yu, W. (2023). Improved Traffic Sign Detection Model Based on YOLOv7-Tiny. *IEEE Access, 11*, 126555-126567. [CrossRef]

[73] Song, T., Zhang, X., Yang, D., Ye, Y., Liu, C., Zhou, J., & Song, Y. (2023). Lightweight detection network based on receptive-field feature enhancement convolution and three dimensions attention for images captured by UAVs. *Image and Vision Computing, 140*, 104855. [CrossRef]

[74] Wang, X., He, N., Hong, C., Wang, Q., & Chen, M. (2023). Improved YOLOX-X based UAV aerial photography object detection algorithm. *Image and Vision Computing, 135*, 104697. [CrossRef]

[75] Ye, T., Qin, W., Zhao, Z., Gao, X., Deng, X., & Ouyang, Y. (2023). Real-Time Object Detection Network in UAV-Vision Based on CNN and Transformer. *IEEE Transactions on Instrumentation and Measurement, 72*, 1-13. [CrossRef]

[76] Zhang, Y., Wu, C., Guo, W., Zhang, T., & Li, W. (2023). CFANet: Efficient Detection of UAV Image Based on Cross-Layer Feature Aggregation. *IEEE Transactions on Geoscience and Remote Sensing, 61*, 1-11. [CrossRef]

[77] Zheng, Y., Zhan, Y., Huang, X., & Ji, G. (2023). YOLOv5s FMG: An Improved Small Target Detection Algorithm Based on YOLOv5 in Low Visibility. *IEEE Access, 11*, 75782-75793. [CrossRef]

[78] Bai, J., Hu, H., Liu, X., Zhuang, S., & Wang, Z. (2024). UAV image object detection based on self-attention guidance and global feature fusion. *Image and Vision Computing, 151*, 105262. [CrossRef]

[79] Dong, C., Jiang, X., Hu, Y., Du, Y., & Pan, L. (2024). EL-Net: An efficient and lightweight optimized network for object detection in remote sensing images. *Expert Systems with Applications, 255*, 124661. [CrossRef]

[80] Du, Y., & Jiang, X. (2024). A Real-Time Small Target Vehicle Detection Algorithm with an Improved YOLOv5m Network Model. *Computers, Materials and Continua, 78*(1), 303-327. [CrossRef]

[81] Fang, H., Ding, L., Wang, X., Chang, Y., Yan, L., Liu, L., & Fang, J. (2024). SCINet: Spatial and Contrast Interactive Super-Resolution Assisted Infrared UAV Target Detection. *IEEE Transactions on Geoscience and Remote Sensing, 62*, 1-22. [CrossRef]

[82] Fu, Q., Zheng, Q., & Yu, F. (2024). LMANet: A Lighter and More Accurate Multiobject Detection Network for UAV Remote Sensing Imagery. *IEEE Geoscience and Remote Sensing Letters, 21*, 1-5. [CrossRef]

[83] Gu, C., Miao, X., & Zuo, C. (2024). TFDNet: A triple focus diffusion network for object detection in urban congestion with accurate multi-scale feature fusion and real-time capability. *Journal of King Saud University-Computer and Information Sciences, 36*(9), 102223. [CrossRef]

[84] He, A., Li, X., Wu, X., Su, C., Chen, J., Xu, S., & Guo, X. (2024). ALSS-YOLO: An Adaptive Lightweight Channel Split and Shuffling Network for TIR Wildlife Detection in UAV Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17*, 17308-17326. [CrossRef]

[85] Hoanh, N., & Pham, T. V. (2024). A Multi-Task Framework for Car Detection From High-Resolution UAV Imagery Focusing on Road Regions. *IEEE Transactions on Intelligent Transportation Systems, 25*, 17160-17173. [CrossRef]

[86] Hu, M., Pi, S., Zhou, J., Wan, X., & Liu, Q. (2024). High-Performance YOLOv5s: Traffic Sign Detection Algorithm for Small Target. *IEEE Access, 12*, 191527-191536. [CrossRef]

[87] Hui, Y., Wang, J., & Li, B. (2024). STF-YOLO: A

small target detection algorithm for UAV remote sensing images based on improved SwinTransformer and class weighted classification decoupling head. *Measurement*, *224*, 113936. [CrossRef]

[88] Khan, M. U., Dil, M., Alam, M. Z., Orakazi, F. A., Almasoud, A. M., Kaleem, Z., & Yuen, C. (2024). SafeSpace MFNet: Precise and Efficient MultiFeature Drone Detection Network. *IEEE Transactions on Vehicular Technology*, *73*, 3106-3118. [CrossRef]

[89] Kong, Y., Shang, X., & Jia, S. (2024). Drone-DETR: Efficient small object detection for remote sensing image using enhanced RT-DETR model. *Sensors*, *24*(17), 5496. [CrossRef]

[90] Li, J., Cheang, C. F., Liu, S., Tang, S., Li, T., & Cheng, Q. (2024). Dynamic-TLD: A Traffic Light Detector Based on Dynamic Strategies. *IEEE Sensors Journal*, *24*, 6677-6686. [CrossRef]

[91] Li, S., Liu, C., Tang, K., Meng, F., Zhu, Z., Zhou, L., & Chen, F. (2024). Improved YOLOv5s Algorithm for Small Target Detection in UAV Aerial Photography. *IEEE Access*, *12*, 9784-9791. [CrossRef]

[92] Liu, B., & Jiang, W. (2024). LA-YOLO: Bidirectional Adaptive Feature Fusion Approach for Small Object Detection of Insulator Self-Explosion Defects. *IEEE Transactions on Power Delivery*, *39*, 3387-3397. [CrossRef]

[93] Ning, T., Wu, W., & Zhang, J. (2024). Small object detection based on YOLOv8 in UAV perspective. *Pattern Analysis and Applications*, *27*(3), 103. [CrossRef]

[94] Ou, K., Dong, C., Liu, X., Zhai, Y., Li, Y., Huang, W., Qiu, W., Wang, Y., & Wang, C. (2024). Drone-TOOD: A Lightweight Task-Aligned Object Detection Algorithm for Vehicle Detection in UAV Images. *IEEE Access*, *12*, 41999-42016. [CrossRef]

[95] Qian, R., & Ding, Y. (2024). An efficient UAV image object detection algorithm based on global attention and multi-scale feature fusion. *Electronics*, *13*(20), 3989. [CrossRef]

[96] Qiu, X., Chen, Y., Sun, C., Li, J., & Niu, M. (2024). DMFF-YOLO: YOLOv8 Based on Dynamic Multiscale Feature Fusion for Object Detection on UAV Aerial Photography. *IEEE Access*, *12*, 125160-125169. [CrossRef]

[97] Song, G., Du, H., Zhang, X., Bao, F., & Zhang, Y. (2024). Small object detection in unmanned aerial vehicle images using multi-scale hybrid attention. *Engineering Applications of Artificial Intelligence*, *128*, 107455. [CrossRef]

[98] Sun, F., He, N., Li, R., Wang, X., & Xu, S. (2024). GD-PAN: a multiscale fusion architecture applied to object detection in UAV aerial images. *Multimedia Systems*, *30*(3), 143. [CrossRef]

[99] Sun, S., Mo, B., Xu, J., Li, D., Zhao, J., & Han, S. (2024). Multi-YOLOv8: An infrared moving small object detection model based on YOLOv8 for air vehicle. *Neurocomputing*, *588*, 127685. [CrossRef]

[100] Sun, T., Chen, H., Liu, H., Deng, L., Liu, L., & Li, S. (2024). DS-YOLOv7: Dense Small Object Detection Algorithm for UAV. *IEEE Access*, *12*, 75865-75872. [CrossRef]

[101] Tang, X., Ruan, C., Li, X., Li, B., & Fu, C. (2024). MSC-YOLO: Improved YOLOv7 Based on Multi-Scale Spatial Context for Small Object Detection in UAV-View. *Computers, Materials and Continua*, *79*(1), 983-1003. [CrossRef]

[102] Tian, D., Wang, C., Zhou, D., Yan, X., & Zeng, L. (2024). NFE-YOLO: A Lightweight and Efficient Detection Network for Low, Slow, and Small Drones. *IEEE Access*, *12*, 175458-175471. [CrossRef]

[103] Tian, D., Yan, X., Zhou, D., Wang, C., & Zhang, W. (2024). IV-YOLO: A lightweight Dual-Branch object detection network. *Sensors*, *24*(19), 6181. [CrossRef]

[104] Wang, X., Wang, X., Zhou, Z., & Song, Y. (2024). A Deep-Learning Method Based on the Multistage Fusion of Radar and Camera in UAV Obstacle Avoidance. *IEEE Transactions on Aerospace and Electronic Systems*, *60*(5), 6734-6751. [CrossRef]

[105] Wang, Y., Zhang, J., & Zhou, J. (2024). Urban traffic tiny object detection via attention and multi-scale feature driven in UAV-vision. *Scientific Reports*, *14*(1), 20614. [CrossRef]

[106] Wu, W., Liu, A., Hu, J., Mo, Y., Xiang, S., Duan, P., & Liang, Q. (2024). EUAVDet: An efficient and lightweight object detector for UAV aerial images with an edge-based computing platform. *Drones*, *8*(6), 261. [CrossRef]

[107] Xu, H., Ling, Z., Yuan, X., & Wang, Y. (2024). A video object detector with Spatio-Temporal Attention Module for micro UAV detection. *Neurocomputing*, *597*, 127973. [CrossRef]

[108] Xu, L., Zhao, Y., Zhai, Y., Huang, L., & Ruan, C. (2024). Small object detection in UAV images based on Yolov8n. *International Journal of Computational Intelligence Systems*, *17*(1), 223. [CrossRef]

[109] Xue, C., Xia, Y., Wu, M., Chen, Z., Cheng, F., & Yun, L. (2024). EL-YOLO: An efficient and lightweight low-altitude aerial objects detector for onboard applications. *Expert Systems with Applications*, *256*, 124848. [CrossRef]

[110] Yang, Y., Feng, F., Liu, G., & Di, J. (2024). MEL-YOLO: A Novel YOLO Network With Multi-Scale, Effective, and Lightweight Methods for Small Object Detection in Aerial Images. *IEEE Access*, *12*, 194280-194295. [CrossRef]

[111] Yang, Y., Yin, H., Chong, A. X., Wan, J., & Liu, Q. Y. (2023). SACINet: Semantic-aware cross-modal interaction network for real-time 3D object detection. *IEEE Transactions on Intelligent Vehicles*, *9*(2), 3917-3927. [CrossRef]

[112] Yao, Z., Hu, Z., Tian, P., & Sun, J. (2024). Adaptation

of Object Detection Algorithms for Road Indicator Lights in Complex Scenes. *Applied Sciences*, *14*(21), 10012. [CrossRef]

[113] Yu, B., Li, Z., Cao, Y., Wu, C., Qi, J., & Wu, L. (2024). YOLO-MPAM: Efficient real-time neural networks based on multi-channel feature fusion. *Expert Systems with Applications*, *252*, 124282. [CrossRef]

[114] Yu, Q., Huang, T., Zhang, W., Xu, J., & He, Y. (2024). Improving real-time small objects detection by fusion features of spatial coordinates. *Signal, Image and Video Processing*, *18*, 789-798. [CrossRef]

[115] Yuan, Z., Gong, J., Guo, B., Wang, C., Liao, N., Song, J., & Wu, Q. (2024). Small object detection in uav remote sensing images based on intra-group multi-scale fusion attention and adaptive weighted feature fusion mechanism. *Remote Sensing*, *16*(22), 4265. [CrossRef]

[116] Zhang, H., Gong, Y., Yao, F., & Zhang, Q. (2024). Research on Real-Time Detection Algorithm for Pedestrian and Vehicle in Foggy Weather Based on Lightweight XM-YOLOViT. *IEEE Access*, *12*, 7864-7883. [CrossRef]

[117] Zhang, H., Li, G., Wan, D., Wang, Z., Dong, J., Lin, S., ... & Liu, H. (2024). DS-YOLO: A dense small object detection algorithm based on inverted bottleneck and multi-scale fusion network. *Biomimetic Intelligence and Robotics*, *4*(4), 100190. [CrossRef]

[118] Zhang, W., Shi, H., Zhao, Y., Feng, Z., & Lovreglio, R. (2024). MMAF-Net: Multi-view multi-stage adaptive fusion for multi-sensor 3D object detection. *Expert Systems with Applications*, *242*, 122716. [CrossRef]

[119] Zhang, Y., Li, B., Shang, J., Huang, X., Zhai, P., & Geng, C. (2024). DSA-Net: An Attention-Guided Network for Real-Time Defect Detection of Transmission Line Dampers Applied to UAV Inspections. *IEEE Transactions on Instrumentation and Measurement*, *73*, 1-22. [CrossRef]

[120] Zheng, W., Xu, H., Li, P., Wang, R., & Shao, X. (2024). SAC-RSM: A High-Performance UAV-Side Road Surveillance Model Based on Super-Resolution Assisted Learning. *IEEE Internet of Things Journal*, *11*, 36066-36083. [CrossRef]

[121] Zhou, W., Wang, X., Fan, Y., Yang, Y., Wen, Y., Li, Y., ... & Wang, J. (2024). KDSMALL: A lightweight small object detection algorithm based on knowledge distillation. *Computer Communications*, *219*, 271-281. [CrossRef]

[122] Zhu, J., Ma, C., Rong, J., & Cao, Y. (2024). Bird and UAVs Recognition Detection and Tracking Based on Improved YOLOv9-DeepSORT. *IEEE Access*, *12*, 147942-147957. [CrossRef]

[123] Ahmad, I., Lu, W., Chen, S. B., Tang, J., & Luo, B. (2025). Lightweight oriented object detection with Dynamic Smooth Feature Fusion Network. *Neurocomputing*, *628*, 129725. [CrossRef]

[124] Bai, C., Zhang, K., Jin, H., Qian, P., Zhai, R., & Lu, K. (2025). SFFEF-YOLO: Small object detection network based on fine-grained feature extraction and fusion for unmanned aerial images. *Image and Vision Computing*, *156*, 105469. [CrossRef]

[125] Chen, X., & Lin, C. (2025). EVMNet: Eagle visual mechanism-inspired lightweight network for small object detection in UAV aerial images. *Digital Signal Processing*, *158*, 104957. [CrossRef]

[126] Chen, Y., & Liu, Z. (2025). DFTD-YOLO: Lightweight Multi-Target Detection From Unmanned Aerial Vehicle Viewpoints. *IEEE Access*, *13*, 24672-24680. [CrossRef]

[127] Chen, Y., Mu, Y., Ni, R., & Yang, B. (2025). 3D pedestrian detection based on hybrid multi-scale cascade fusion network. *Computers and Electrical Engineering*, *123*, 110139. [CrossRef]

[128] Fan, Q., Li, Y., Deveci, M., Zhong, K., & Kadry, S. (2025). LUD-YOLO: A novel lightweight object detection network for unmanned aerial vehicle. *Information Sciences*, *686*, 121366. [CrossRef]

[129] Feng, F., Yang, L., Zhou, Q., & Li, W. (2025). YOLO-Tiny: A lightweight small object detection algorithm for UAV aerial imagery. *IET Image Processing*, *19*(1), e13314. [CrossRef]

[130] Hu, J., Pang, T., Peng, B., Shi, Y., & Li, T. (2025). A small object detection model for drone images based on multi-attention fusion network. *Image and Vision Computing*, *155*, 105436. [CrossRef]

[131] Kim, S., Kim, C., & Kim, S. (2024). Improving Performance of Real-Time Object Detection in Edge Device Through Concurrent Multi-Frame Processing. *IEEE Access*, *13*, 1522-1533. [CrossRef]

[132] Liang, X., Li, M., Lin, Y. E., & Fang, X. (2025). GACFNet: A global attention cross-level feature fusion network for aerial image object detection. *Computers and Electrical Engineering*, *123*, 110042. [CrossRef]

[133] Liu, Y., Huang, Z., Song, Q., & Bai, K. (2025). PV-YOLO: A lightweight pedestrian and vehicle detection model based on improved YOLOv8. *Digital Signal Processing*, *156*, 104857. [CrossRef]

[134] Liu, Y., & Shi, Y. (2025). VRU-YOLO: A Small Object Detection Algorithm for Vulnerable Road Users in Complex Scenes. *IEEE Access*, *13*, 19996-20015. [CrossRef]

[135] Mushtaq, H., Deng, X., Alizadehsani, R., Iqbal, M. S., Khan, T., & Abbasi, A. A. (2025). SC3D: Semantic-guided and Class-adaptive cross-domain fusion for 3D object detection in autonomous vehicles. *Expert Systems with Applications*, *268*, 126359. [CrossRef]

[136] Muzammul, M., Assam, M., & Qahmash, A. (2024). Quantum-Inspired Multi-Scale Object Detection in UAV Imagery: Advancing Ultra-Small Object Accuracy and Efficiency for Real-Time Applications. *IEEE Access*, *13*, 2173-2186. [CrossRef]

[137] Nghiem, V. Q., Nguyen, H. H., & Hoang, M. S. (2025). LEAF-YOLO: Lightweight Edge-Real-Time Small Object Detection on Aerial Imagery. *Intelligent Systems with Applications, 25*, 200484. [CrossRef]

[138] Niu, Y., Lin, C., Jiang, X., & Qu, Z. (2025). VSTDet: A lightweight small object detection network inspired by the ventral visual pathway. *Applied Soft Computing, 171*, 112775. [CrossRef]

[139] Pan, W., & Yang, Z. (2025). A lightweight enhanced YOLOv8 algorithm for detecting small objects in UAV aerial photography. *The Visual Computer*, 1-17. [CrossRef]

[140] Qu, J., Li, Q., Pan, J., Sun, M., Lu, X., Zhou, Y., & Zhu, H. (2025). SS-YOLOv8: small-size object detection algorithm based on improved YOLOv8 for UAV imagery. *Multimedia Systems, 31*(1), 42. [CrossRef]

[141] Wang, G., Jiang, Q., Jin, X., Wozniak, M., Wang, P., & Yao, S. (2024). RTM-UAVDet: a real-time multimodal UAV detector. *IEEE Transactions on Aerospace and Electronic Systems, 61*(1), 473-489. [CrossRef]

[142] Wang, S., Li, G., He, B., Cheng, B., Ding, Y., & Li, W. (2025). LUMF-YOLO: a lightweight object detection network integrating UAV motion features. *Computing, 107*(1), 25. [CrossRef]

[143] Xue, R., Hua, S., & Xu, H. (2025). FECI-RTDETR a Lightweight Unmanned Aerial Vehicle Infrared Small Target Detector Algorithm Based on RT-DETR. *IEEE Access, 13*, 9578-9591. [CrossRef]

[144] Yan, Z., Chen, C., Wu, S., Wang, Z., Li, L., Sun, S., Yang, B., & Fu, J. (2025). RF-DET: Refocusing on the small-scale objects using aggregated context for accurate power transmitting components detection on UAV oblique imagery. *ISPRS Journal of Photogrammetry and Remote Sensing, 220*, 692-711. [CrossRef]

[145] Yang, H., Wang, J., Bo, Y., & Wang, J. (2025). ISTD-DETR: A deep learning algorithm based on DETR and Super-resolution for infrared small target detection. *Neurocomputing, 621*, 129289. [CrossRef]

[146] Yang, W., He, Q., & Li, Z. (2025). A lightweight multidimensional feature network for small object detection on UAVs. *Pattern Analysis and Applications, 28*(1), 29. [CrossRef]

[147] Yang, Y., Liu, Z., Chen, J., Gao, H., & Wang, T. (2025). Railway Foreign Object Intrusion Detection Using UAV Images and YOLO-UAT. *IEEE Access, 13*, 18498-18509. [CrossRef]

[148] Zhang, F., Leong, L. V., Yen, K. S., & Zhang, Y. (2025). An enhanced lightweight model for small-scale pedestrian detection based on YOLOv8s. *Digital Signal Processing, 156*, 104866. [CrossRef]

[149] Zhang, H., Xiao, P., Yao, F., Zhang, Q., & Gong, Y. (2025). Fusion of multi-scale attention for aerial images small-target detection model based on PARE-YOLO. *Scientific Reports, 15*. [CrossRef]

[150] Zhang, L., Yang, K., Han, Y., Li, J., Wei, W., Tan, H., Yu, P., Zhang, K., & Yang, X. (2025). TSD-DETR: A lightweight real-time detection transformer of traffic sign detection for long-range perception of autonomous driving. *Engineering Applications of Artificial Intelligence, 139*, 109536. [CrossRef]

Yong Zhang Ph.D., Professor at the College of Electronics and Information Engineering, Shenzhen University, and the Guangdong Provincial Key Laboratory of Intelligent Information Processing. His research interests include artificial intelligence and information fusion. (Email: yzhang@szu.edu.cn)

Zhenke Qian Master's degree candidate at the College of Electronics and Information Engineering, Shenzhen University, and the Guangdong Provincial Key Laboratory of Intelligent Information Processing. His research interests include artificial intelligence and information fusion. (Email: 1372231016@qq.com)

Junyan Cai Received the Master's degree from the College of Electronics and Information Engineering, Shenzhen University, and the Guangdong Provincial Key Laboratory of Intelligent Information Processing in 2025. He is currently with the Dongguan New-Generation Artificial Intelligence Industry Technology Research Institute. His research interests include artificial intelligence and information fusion. (Email: junyancai2023@163.com)