

# Advances in Vehicle Object Detection Algorithms for Adverse Weather Conditions

Wen Su<sup>1,a</sup>, Guifang Guo<sup>2,b,\*</sup>

<sup>1</sup>School of Cybersecurity, Xizang Minzu University, Xianyang, China, 712082

<sup>2</sup>School of Engineering, Xizang Minzu University, Xianyang, China, 712082

<sup>a</sup>1622553823@qq.com, <sup>b</sup>18822055926@163.com

**Abstract:** With the explosive growth of global vehicle ownership, traffic accidents caused by adverse weather conditions have become a critical public safety issue worldwide. Vehicle object detection, as the core perception technology of intelligent driver assistance systems (IDAS) and autonomous driving systems, directly determines the safety and reliability of road traffic. This paper systematically reviews the evolution of object detection algorithms from traditional handcrafted feature-based methods to deep learning-driven architectures, including two-stage, single-stage, and Transformer-based frameworks. It focuses on the latest research advances (2025-2026) in vehicle detection under typical low-visibility environments (night, fog, rain, snow) and analyzes the key technical challenges restricting practical deployment. Finally, this review summarizes the limitations of existing methods and prospects future research directions, aiming to provide a comprehensive reference for developing high-precision, lightweight, and all-weather robust vehicle detection algorithms.

**Keywords:** Vehicle object detection; Deep learning; Low-visibility environments; Adverse weather; Autonomous driving

## 1. Introduction

As global vehicle ownership surges, China—the world's largest automobile market—has exceeded 460 million vehicles and 500 million licensed drivers by 2025 (Figure 1). However, this growth has brought frequent traffic accidents: over 60,000 annual fatalities in China, with more than 30% occurring in adverse weather. Most weather-related accidents stem from drivers' delayed perception of hazards due to reduced visibility. Therefore, developing all-weather real-time vehicle detection systems is critical for traffic safety and autonomous driving deployment<sup>[1]</sup>.

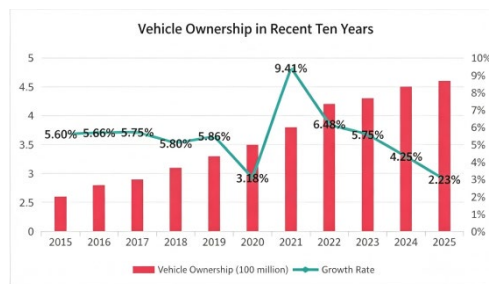


Figure 1. Car Ownership Chart of 2015-2025

Object detection technology has undergone a paradigm shift from traditional manual feature engineering to deep learning over the past two decades. With the breakthrough of AlexNet in 2012 <sup>[2]</sup>, convolutional neural networks (CNN) have become the mainstream backbone for object detection tasks. Currently, deep learning-based object detection algorithms are mainly divided into three categories: two-stage algorithms represented by Faster R-CNN, single-stage algorithms represented by YOLO and SSD, and Transformer-based algorithms represented by DETR and RT-DETR. Although these algorithms have achieved excellent performance in standard sunny traffic scenes, their detection accuracy drops by 30%-50% when facing low-visibility environments such as night, fog, rain, and snow<sup>[3]</sup>. In these environments, image quality is severely degraded through insufficient illumination, atmospheric scattering, or weather interference, leading to blurred target edges, missing detail features, low contrast between targets and backgrounds, and the introduction of a large amount of structured

noise. Typical adverse weather scenarios in traffic are shown in Figure 2.



Figure 2. Bad Weather in Traffic Scene

Furthermore, most existing studies focus on optimizing algorithms for a single specific adverse weather condition, lacking robustness to dynamically changing weather in real-world traffic. For example, algorithms optimized for night scenes often perform poorly in foggy or rainy conditions. In addition, many studies only pursue detection accuracy while ignoring model size, inference speed, and memory usage, which are critical factors for deploying autonomous driving systems on resource-constrained edge devices. These limitations have become the main bottlenecks restricting the practical application of vehicle detection technology. With the intensification of global climate change, severe weather events occur more frequently, making the research on all-weather robust vehicle detection algorithms more urgent and important.

## 2. Evolution of Object Detection Algorithms

### 2.1. Traditional Object Detection Methods

Before the deep learning era, traditional object detection methods dominated by combining handcrafted features with machine learning classifiers. Their core pipeline included three steps: candidate region generation (sliding window/selective search), feature extraction (Haar-like, HOG, SIFT), and classification with bounding box regression (SVM). Representative works include the 2001 VJ detector (first real-time face detection)<sup>[4]</sup>, 2005 HOG descriptor (breakthrough in pedestrian detection)<sup>[5]</sup>, and 2008 DPM (improved robustness via part modeling)<sup>[6]</sup>. However, these methods suffer from low-level feature expression, high computational redundancy, and lack of end-to-end optimization, failing to meet the high-accuracy and real-time requirements of autonomous driving, especially in degraded adverse weather conditions. The overall process of traditional target detection methods is shown in Figure 3.

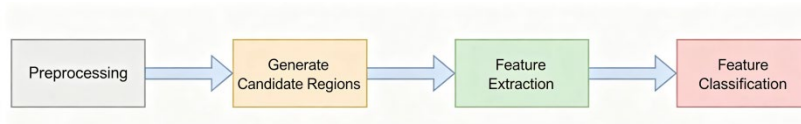


Figure 3. Overall Process of Traditional Target Detection Method

### 2.2. Deep Learning-Based Object Detection Methods

The success of AlexNet in the ImageNet classification competition in 2012 revolutionized the field of object detection. Deep learning-based methods use CNNs to automatically learn hierarchical features from raw data, achieving end-to-end training and inference, which have significantly improved both detection accuracy and speed compared with traditional methods. Currently, deep learning-based object detection algorithms are mainly divided into three categories: two-stage detection algorithms, single-stage detection algorithms, and Transformer-based detection algorithms, all of which have been extensively adapted for adverse weather detection.

#### 2.2.1. Two-Stage Detection Algorithms

Two-stage detection algorithms, also known as region-based algorithms, divide the detection process into two stages: first generating a set of candidate regions that may contain targets, and then classifying and regressing these regions to obtain the final detection results. The origin of two-stage algorithms can be traced back to R-CNN proposed by Girshick et al. in 2014<sup>[7]</sup>, which first applied

CNN to region-based object detection and significantly improved detection accuracy compared with traditional methods. However, R-CNN requires separate CNN processing for each candidate region, resulting in huge computational overhead and slow inference speed. To solve this problem, Girshick proposed Fast R-CNN in 2015<sup>[8]</sup>, which inputs the entire image into the CNN to extract feature maps once, then maps candidate regions to the feature maps and uses Region of Interest (RoI) pooling to unify features of different sizes into fixed-length vectors, enabling end-to-end training and greatly reducing computational complexity.

The real breakthrough of two-stage algorithms came with Faster R-CNN proposed by Ren et al. in 2015<sup>[9]</sup>, which completely abandoned the traditional selective search method and introduced the Region Proposal Network (RPN) to generate candidate regions. RPN is a fully convolutional network that shares convolutional features with the subsequent detection network, achieving truly end-to-end object detection and becoming the classic baseline of two-stage algorithms. Since then, researchers have proposed a series of improved algorithms based on Faster R-CNN, such as Feature Pyramid Network (FPN) for multi-scale feature fusion<sup>[10]</sup>, Mask R-CNN for instance segmentation<sup>[11]</sup>, and Cascade R-CNN for high-quality detection<sup>[12]</sup>. For adverse weather detection, two-stage algorithms have been adapted by adding weather-specific feature enhancement modules and robust loss functions. However, their two-stage pipeline leads to relatively slow inference speed, making them less suitable for real-time autonomous driving applications compared with single-stage algorithms. The evolution of R-CNN series two-stage detection algorithms is shown in Figure 4.

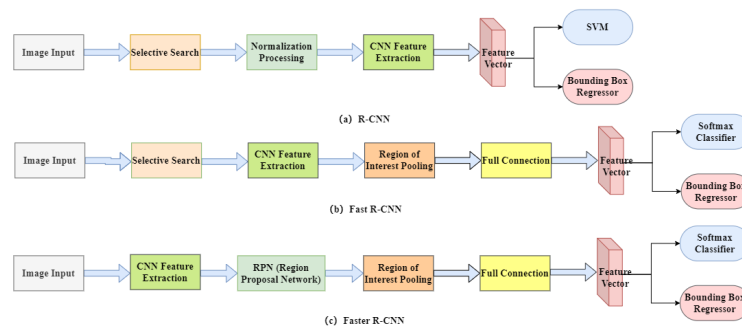


Figure 4. R-CNN Series Two-stage Target Detection Algorithm Flow Evolution Diagram

### 2.2.2. Single-Stage Detection Algorithms

Single-stage detection algorithms skip the candidate region generation stage and directly predict target categories and bounding boxes in a single network. With simpler structures and faster inference than two-stage methods, they dominate real-time autonomous driving applications, especially latency-critical adverse weather scenarios. The 2016 SSD<sup>[13]</sup> achieves speed-accuracy balance via multi-scale feature prediction, while 2017 RetinaNet<sup>[14]</sup> introduced Focal Loss to resolve class imbalance, boosting accuracy on low-quality images.

The YOLO series is the most influential framework for adverse weather vehicle detection. Since its 2016 debut<sup>[15]</sup>, it has evolved continuously: YOLOv3<sup>[16]</sup> adopted Darknet-53 and multi-scale prediction; YOLOv7<sup>[17]</sup> delivered state-of-the-art performance via advanced training techniques; YOLOv10<sup>[18]</sup> eliminated NMS to cut inference latency by 50%.

The latest 2025-2026 advances in the YOLO series have specifically targeted adverse weather detection challenges. In 2025, Tian et al. proposed YOLOv12, which adopts an attention-centric architecture that integrates spatial and channel attention mechanisms into the backbone and neck networks, improving small target detection accuracy by 12% in foggy conditions while maintaining 85 FPS inference speed on RTX 4090<sup>[19]</sup>. In early 2026, researchers from Tsinghua University proposed YOLOv13, which introduces hypergraph modeling (HyperACE) to capture high-order spatial relationships between vehicle targets, achieving 92.5% mAP on the KITTI vehicle detection dataset and showing strong robustness to partial occlusions caused by weather interference. For adverse weather detection, AWD-YOLO proposed in 2026 uses a dual-backbone architecture that fuses features from both raw and weather-enhanced images, improving detection accuracy by 15.7% on the BDD100K adverse weather subset without significant speed loss. Another 2026 study proposed Weather-YOLO, which introduces a dynamic weather adaptation module that can automatically adjust feature extraction strategies according to input weather conditions, achieving consistent performance across night, fog, rain, and snow scenarios. The flow charts of typical single-stage detection algorithms are shown in Figure 5.

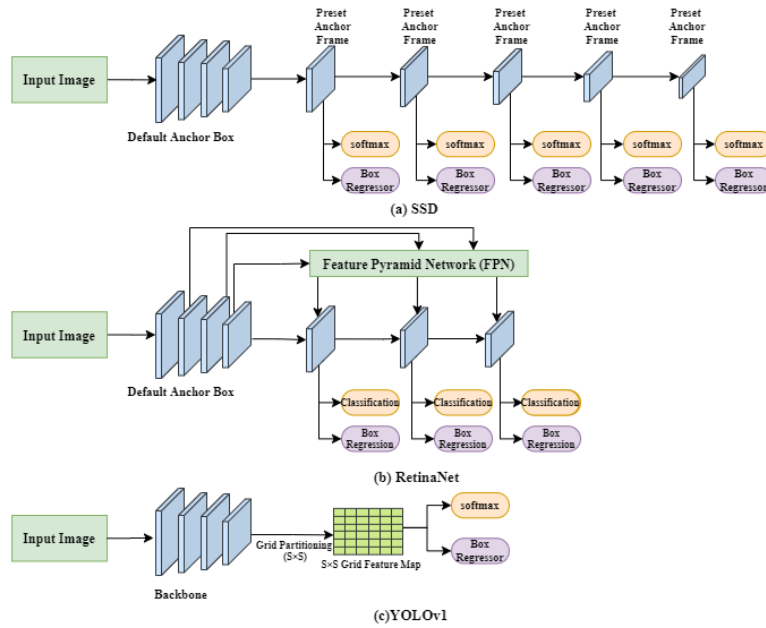


Figure 5. Flow Chart of Single-stage Target Detection Algorithm

### 2.2.3. Transformer-Based Detection Algorithms

Building on Transformer's success in NLP, researchers introduced it to computer vision for its global self-attention mechanism, which captures long-range dependencies and outperforms CNNs in feature expression for adverse weather-degraded images. The 2020 DETR<sup>[20]</sup> achieved anchor-free, NMS-free end-to-end detection with Faster R-CNN-level accuracy but suffered from slow convergence and poor small-target performance. Deformable DETR<sup>[21]</sup> addressed this by focusing only on key positions, speeding up convergence 10-fold. The 2023 RT-DETR<sup>[22]</sup> marked a real-time breakthrough, outperforming YOLO for the first time. Latest 2025-2026 advances include DynamicConv RT-DETR (54.3% COCO AP, 74 FPS on T4), edge-optimized MD-MoE-RT-DETR (50% GFLOPs reduction), and weather-specific RD-DETR (18.2%/22.3% accuracy gains on foggy/snowy ACDC datasets). The architecture of Transformer-based detection algorithms is shown in Figure 6.

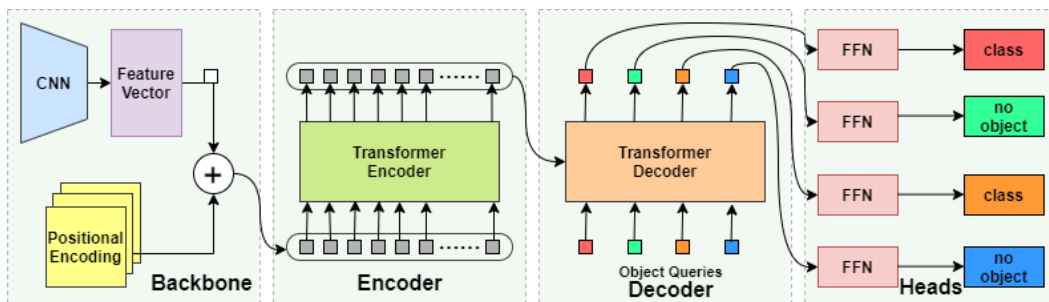


Figure 6. Flow Chart of Detection Transformer Target Detection Algorithm

## 3. Vehicle Detection in Adverse Weather Environments

Adverse weather environments such as night, fog, rain, and snow severely degrade image quality through various physical mechanisms, leading to significant performance drops in standard object detection algorithms. In recent years, researchers have carried out extensive research on vehicle detection in these environments, and the latest 2025-2026 studies have increasingly focused on end-to-end collaborative optimization and multi-modal fusion to improve detection robustness.

Night, the most common adverse weather, suffers from low brightness, uneven illumination, and severe glare, causing blurred edges, missing details, and overexposure that hinder standard detection. Traditional methods rely on separate low-light enhancement preprocessing: Hou et al.'s dynamic RGB adjustment network improved YOLOv3 night detection by 21%<sup>[23]</sup>, and 2023's Retinexformer achieved state-of-the-art enhancement via illumination-guided Transformer blocks<sup>[24]</sup>. The 2025-2026 research

has shifted to end-to-end frameworks and multi-modal fusion: VLM-augmented detection reduced partial occlusion misses by 25%; thermal imaging (deployed in 20+ vehicle models by 2026) extends detection range to 300m in total darkness; MIT's visible-thermal-radar fusion framework achieved 94% mAP on FLIR dataset, 32% higher than single-modality detection.

Fog reduces image contrast and blurs distant targets via light scattering, with visibility dropping below 50m in dense fog. Traditional atmospheric scattering-based dehazing methods suffer from high complexity and artifacts. Early deep learning solutions include DehazeNet (first end-to-end CNN dehazing)<sup>[25]</sup> and 2023's MB-TaylorFormer (detail-preserving dehazing)<sup>[26]</sup>. Recent 2025–2026 advances focus on unified dehazing-detection pipelines to avoid error accumulation: ECE-VDTDA improved foggy detection by 17.45% on DAWN dataset; Shanghai Jiao Tong University's shape-semantic YOLOv8 achieved 70.46% precision in dense fog; a lightweight edge model runs at 62 FPS on Jetson Xavier NX with 65% mAP in moderate fog.

Rain introduces raindrops, streaks, and road reflections, causing target confusion and false detections. Traditional deraining preprocessing includes Li et al.'s context aggregation network (effective for heavy rain)<sup>[27]</sup> and 2021's unsupervised DerainCycleGAN<sup>[28]</sup>. The 2025–2026 advances adopt contrastive learning and diffusion data augmentation: the SCLA framework improved rainy detection mAP50 by 8.9% on Nordic dataset; diffusion-generated training data boosted accuracy by 12.3%; USTC's multi-scale integrated model achieved 78% mAP in heavy rain, 15% better than previous state-of-the-art.

Snow is the most challenging environment, with dynamic snowflakes causing random occlusions and structured noise, plus low contrast from snow-covered backgrounds. Early methods used separate desnowing: DesnowNet's dual-module design removed both translucent and opaque snow<sup>[29]</sup>, and 2024's dual-branch network integrated restoration and detection<sup>[30]</sup>. The 2025–2026 breakthroughs come from Transformer architectures and multi-modal fusion: RD-DETR suppressed snow noise via reaction-diffusion dynamics, improving accuracy by 22.3% on ACDC dataset; DSERT-RoLL's 5-sensor fusion (RGB, event, thermal, 4D radar, dual LiDAR) achieved 91.2% mAP in heavy snow, establishing multi-modal fusion as the mainstream direction for all-weather perception.

#### 4. Conclusions

This paper systematically reviews the evolution of vehicle object detection algorithms and their latest advances in adverse weather conditions, covering traditional handcrafted feature methods, deep learning-based CNN and Transformer architectures, and targeted solutions for night, fog, rain, and snow scenarios. Despite significant progress, three key challenges still hinder practical deployment: most methods are optimized for single weather conditions and lack robustness to dynamic weather changes; many prioritize detection accuracy over model size, inference speed, and memory efficiency, making them incompatible with resource-constrained automotive edge devices; and the shortage of large-scale, high-quality multi-weather datasets (especially those capturing continuous weather transitions and extreme events) limits the development of general-purpose models.

Future research should focus on five critical directions: developing general-purpose all-weather detection models via unified feature learning and domain adaptation; designing lightweight edge-friendly architectures using pruning, quantization, and knowledge distillation; promoting end-to-end collaborative optimization of image enhancement and object detection; advancing multi-modal sensor fusion for extreme low-visibility perception; and constructing comprehensive multi-weather datasets with accurate annotations. These efforts will drive breakthroughs in all-weather perception technology, providing robust support for safe and efficient intelligent transportation systems.

#### Acknowledgments

I would like to acknowledge the financial support from the Scientific Research Innovation and Practice Project of Xizang Minzu University (Grant No. Y2026135), which supported this research.

#### References

[1] Zhu Q. *Tort Liability for Traffic Accidents of Driverless Vehicles[J]. Price Supervision and Anti-monopoly in China, 2024(05): 80-82.*

- [2] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. *Advances in Neural Information Processing Systems*, 2012, 25.
- [3] Lin M, Zhou G, Yang Y W, et al. Overview of target detection methods under special weather conditions[J]. *Computer Engineering and Applications*, 2022, 58(13): 36-47.
- [4] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[C]//2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'01). IEEE, 2001, 1: 511-518.
- [5] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). IEEE, 2005, 1: 886-893.
- [6] Felzenszwalb P, McAllester D, Ramanan D. A discriminatively trained, multiscale, deformable part model[C]//2008 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2008: 1-8.
- [7] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 580-587.
- [8] Girshick R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. 2015: 1440-1448.
- [9] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *Advances in Neural Information Processing Systems*, 2015, 28.
- [10] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 2117-2125.
- [11] He K, Gkioxari G, Dollár P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. 2017: 2961-2969.
- [12] Cai Z, Vasconcelos N. Cascade R-CNN: Delving into high quality object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 6154-6162.
- [13] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]//European Conference on Computer Vision (ECCV). Springer, 2016: 21-37.
- [14] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. 2017: 2980-2988.
- [15] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 779-788.
- [16] Redmon J, Farhadi A. YOLOv3: An incremental improvement[J]. *arXiv preprint arXiv:1804.02767*, 2018.
- [17] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 7464-7475.
- [18] Wang A, Chen H, Liu L, et al. YOLOv10: Real-time end-to-end object detection[J]. *Advances in Neural Information Processing Systems*, 2024, 37: 107984-108011.
- [19] Tian Y, Ye Q, Doermann D. YOLOv12: Attention-centric real-time object detectors[J]. *arXiv preprint arXiv:2502.12524*, 2025.
- [20] Carion N, Massa F, Synnaeve G, et al. End-to-end object detection with transformers[C]//European Conference on Computer Vision. Springer, 2020: 213-229.
- [21] Zhu X, Su W, Lu L, et al. Deformable DETR: Deformable transformers for end-to-end object detection[J]. *arXiv preprint arXiv:2010.04159*, 2020.
- [22] Zhao Y, Lv W, Xu S, et al. DETRs beat YOLOs on real-time object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 16965-16974.
- [23] Hou R, Ma B, Chang H, et al. Interaction-and-aggregation network for person re-identification[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 9317-9326.
- [24] Cai Y, Bian H, Lin J, et al. Retinexformer: One-stage Retinex-based transformer for low-light image enhancement[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 12504-12513.
- [25] Cai B, Xu X, Jia K, et al. DehazeNet: An end-to-end system for single image haze removal[J]. *IEEE Transactions on Image Processing*, 2016, 25(11): 5187-5198.
- [26] Qiu Y, Zhang K, Wang C, et al. MB-TaylorFormer: Multi-branch efficient transformer expanded by Taylor formula for image dehazing[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 12802-12813.
- [27] Li X, Wu J, Lin Z, et al. Recurrent squeeze-and-excitation context aggregation net for single image

*deraining[C]//European Conference on Computer Vision (ECCV). 2018: 254-269.*

[28] Wei Y, Zhang Z, Wang Y, et al. *DerainCycleGAN: Rain attentive CycleGAN for single image deraining and rainmaking[J]. IEEE Transactions on Image Processing, 2021, 30: 4788-4801.*

[29] Liu Y F, Jaw D W, Huang S C, et al. *DesnowNet: Context-aware deep network for snow removal[J]. IEEE Transactions on Image Processing, 2018, 27(6): 3064-3073.*

[30] Wang J, Xu M, Xue H, et al. *Joint image restoration for object detection in snowy weather[J]. IET Computer Vision, 2024, 18(6): 759-771.*