

Supplemental material Contents:

- A** Additional content of the Benchmark
 - A-A** Mini-test
 - A-B** The adaptability of the benchmark
 - A-C** Corresponding config files of the selected detectors
 - A-D** Explanation about the selected objects
 - A-E** The necessity of the benchmark
 - A-E1** Utilities of the benchmark
 - A-E2** Potential applications of the benchmark
 - A-F** Limitations and potential impacts
- B** Additional content of the Experiments
 - B-A** Generated data for ablation studies
 - B-B** A detailed illustration of the performance gap
 - B-B1**
 - B-C** Supplemented experiments analysis and discussion
 - B-C1** Detection perspective
 - B-C2** Attack perspective
 - B-D** Additional overall experimental experiments
 - B-E** Additional ablation experiments
 - B-E1** Ablation study on physical dynamics
 - B-E2** Ablation study on training dataset
 - B-E3** Ablation study on 2D and 3D perturbations
- C** User feedback

APPENDIX A

ADDITIONAL CONTENT OF THE BENCHMARK

A. Mini-test

We kindly invite the reviewers and readers to participate in a mini-test to discriminate the real-world images and the simulated images as shown in Fig. 11, the answer is revealed in its caption.

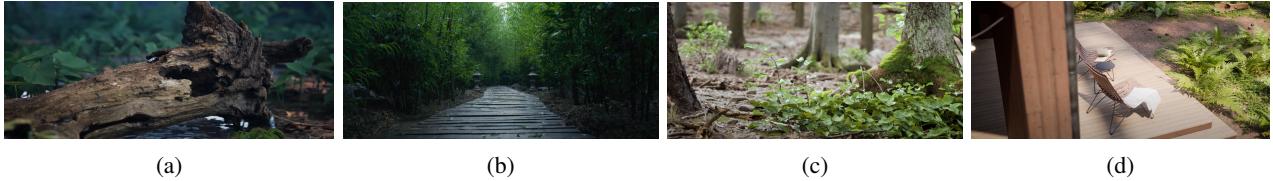


Fig. 11: Which are simulated images? Surprisingly, they were all generated by Unreal Engine, a popular game engine. The visual quality of the simulated images is so high that it is hard to find any deficiencies. This mini-test demonstrates the potential of the simulated environment in the research field.

TABLE IV: The selected detectors and their corresponding config files. Full list of the optional maps, where Town8 and Town9 are unseen for competition. Please refer to CARLA [75] documentary for more details.

Town	Description
Town1	A small, simple town with a river and several bridges.
Town2	A small simple town with a mixture of residential and commercial buildings.
Town3	A larger, urban map with a roundabout and large junctions.
Town4	A small town embedded in the mountains with a special "figure of 8" infinite highway.
Town5	Squared-grid town with cross junctions and a bridge. It has multiple lanes per direction. Useful to perform lane changes.
Town6	Long many lane highways with many highway entrances and exits. It also has a Michigan left.
Town7	A rural environment with narrow roads, corn, barns and hardly any traffic lights.
Town8	Secret "unseen" town used for the Leaderboard challenge.
Town9	Secret "unseen" town used for the Leaderboard challenge.
Town10	A downtown urban environment with skyscrapers, residential buildings and an ocean promenade.
Town11	A Large Map that is undecorated. Serves as a proof of concept for the Large Maps feature.
Town12	A Large Map with numerous different regions, including high-rise, residential and rural environments.

B. The adaptability of the benchmark

We provide a detailed illustration of the scene diversity of the benchmark in Table IV and Fig. 12, where the optional maps are listed with their descriptions. In addition, we display the extendable vehicles, pedestrians, and traffic signs in Fig. 13, Fig. 14, and Fig. 15, respectively, which can be easily extended to evaluate other objects in the benchmark. The users are also allowed to export any customized scenes and objects to the benchmark as needed, which can be easily integrated into the benchmark.

C. Corresponding config files of the selected detectors

The corresponding config files of the selected detectors are listed in Table V. Specifically, 1-25 and 26-40 are CNN-based One-stage and Two-stage object detectors, respectively. 41-48 are Transformer-based object detectors. The corresponding config files of the detectors are available in our codebase or MMDetection [126] toolbox.

D. Explanation about the selected objects

According to a survey [82] published in TPAMI 2024, most physical attacks against object detection are optimized for specific target categories, such as vehicles, persons, and a few for traffic signs. In line with this, we have chosen vehicles and pedestrians as the representative target categories, to evaluate the robustness of object detectors against physical attacks. .

In order to ensure the validity of our benchmark for different types of objects, we have demonstrated that our benchmark can be easily extended to other target categories, as shown by the experiments conducted on traffic sign in Table XX. The benchmark is designed to evaluate the robustness of

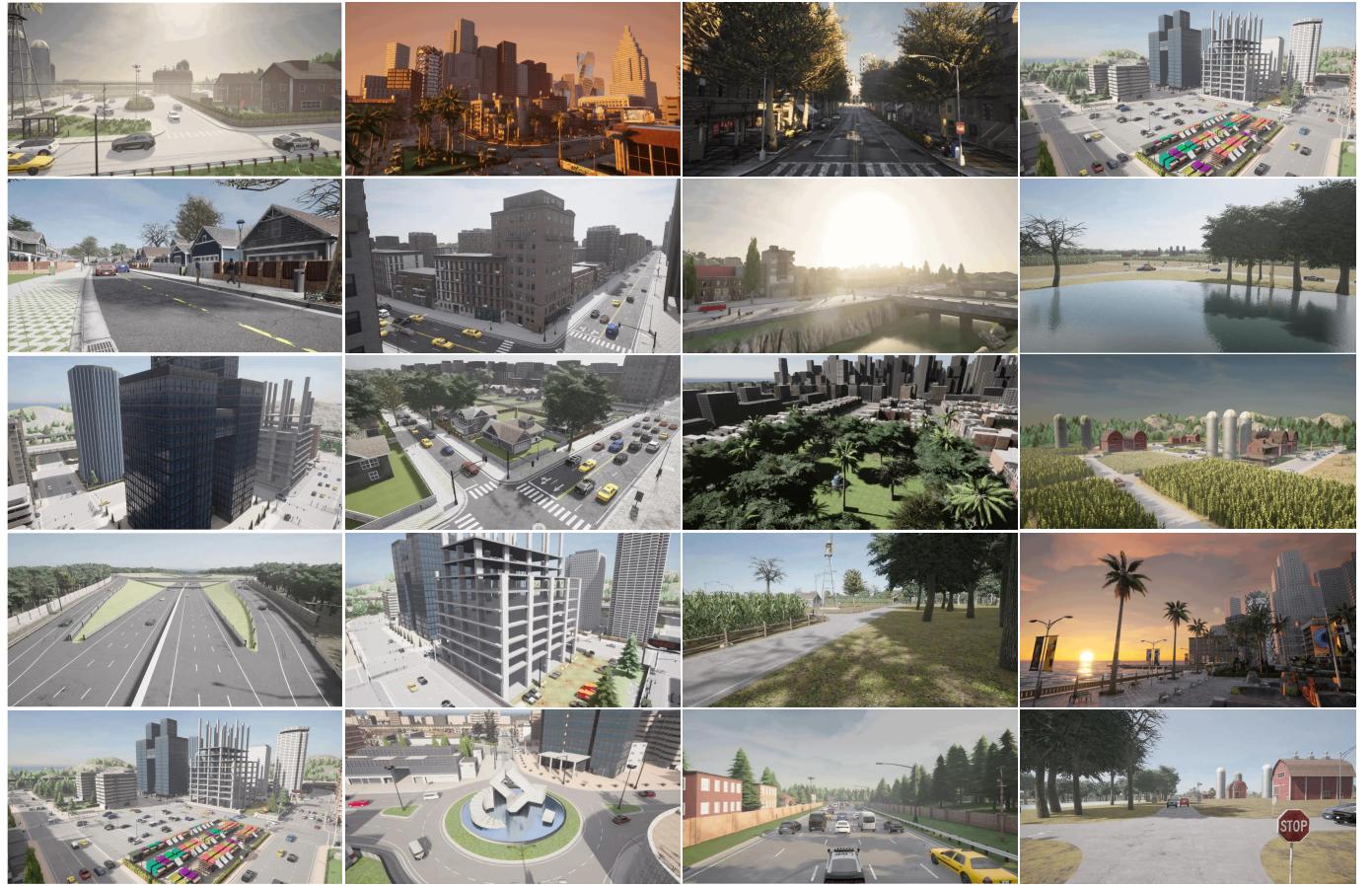


Fig. 12: Illustration of the extensible scenes of the benchmark.

object detectors against physical attacks in various aligned scenarios for ensuring fairness. It can be extended to other target categories with minimal modifications.

We have thoroughly reviewed over forty physical attack methods, and we found that most of these methods conducted experiments under unaligned conditions and without fair comparisons. This lack of clarity hinders the accurate assessment of the progress of physical adversarial attacks and the development of physical adversarial robustness. Therefore, we are motivated to establish a comprehensive and rigorous benchmark for physical attacks to address these limitations and provide a solid foundation for future research.

E. The utility of the benchmark

In this section, we summarize our motivation and provide the potential applications of the benchmark.

1) Utilities of the benchmark:

- **Standardization and Fair Evaluation:** The primary utility of PADetBench lies in its ability to standardize the evaluation of physical attacks against object detection models. By ensuring that all evaluations are conducted under the same physical dynamics, PADetBench eliminates inconsistencies found in real-world experiments, making it a fair and rigorous benchmark.
- **Comprehensive Coverage:** PADetBench includes 23 physical attack methods and evaluates 48 state-of-the-

art object detectors, providing a comprehensive coverage that enables researchers to compare and contrast various models and attack strategies.

2) Potential applications of the benchmark:

- **Research and Development:** Researchers developing robust object detection models or physical attack strategies need a benchmark to evaluate and compare their approaches.
- **Security Assessments:** Security teams need to assess the robustness of deployed object detection systems in critical infrastructure.
- **Regulatory Compliance:** Regulatory bodies require evidence of robustness and security for autonomous systems.
- **Product Testing:** Companies developing autonomous vehicles or security systems need to test their products under various physical attack scenarios.
- **Educational Purposes:** Educators and students need resources to understand the vulnerabilities of object detection models.

F. Limitations and potential impacts

Limitations

For now, PADetBench primarily focuses on evaluating the robustness of object detection models against physical attacks. In the future, we plan to extend the benchmark to include other vision tasks, such as instance segmentation, 3D object



Fig. 13: Illustration of the extensible vehicles of the benchmark.

detection, and depth estimation. This expansion will provide a more comprehensive evaluation framework that covers a broader range of computer vision applications.

Potential Impacts

1) Positive Impacts: The in-depth understanding gained through PADetBench will contribute significantly to the development of more robust object detection models. By identifying vulnerabilities and limitations, researchers and practitioners can design improved algorithms that are better equipped to handle physical adversarial attacks. This enhanced robustness is crucial for real-world applications where reliability and accuracy are paramount.

2) Negative Impacts: While the benchmark provides valuable insights, there is a risk that it could be misused to conduct physical attacks in real-life scenarios. Such misuse

could threaten the security of critical applications involving intelligent visual perception systems. Therefore, it is essential to promote responsible use of the benchmark and to emphasize the importance of ethical considerations in research and development.

APPENDIX B

ADDITIONAL CONTENT OF THE EXPERIMENTS

A. Generated data for ablation studies

We provide the generated data samples for the ablation studies in Fig. 16.

B. A detailed illustration of the performance gap

1) Performance gap between the benchmark and the original papers: In this section, we provide an explanation for

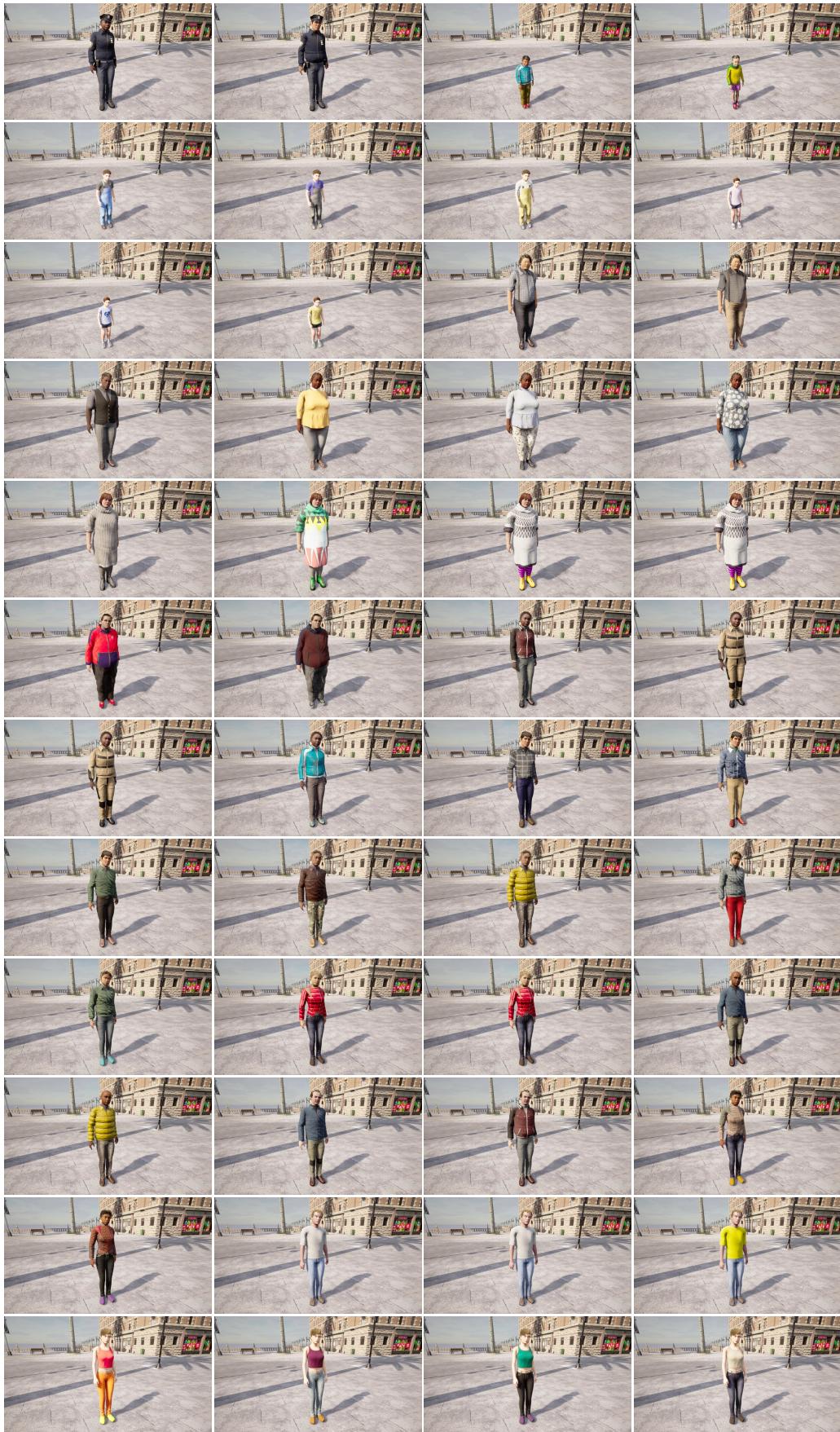


Fig. 14: Illustration of the extensible walkers of the benchmark.



Fig. 15: Illustration of the extensible traffic signs of the benchmark.

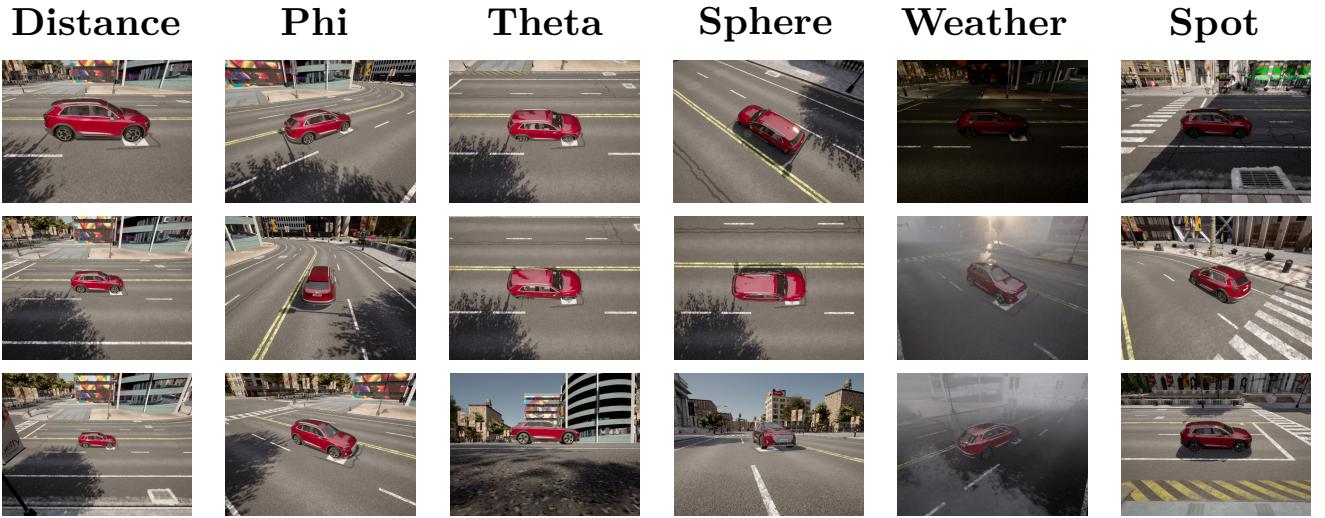


Fig. 16: The randomly selected samples for the ablation studies of six different dynamics.

the performance gap between the reported attack performance in the original papers and the results in our benchmark. Our benchmark encompasses a wide range of physical dynamics, whereas previous validation settings are often limited to a few specific scenarios. The comprehensive physical dynamics in our benchmark reveal the shortcomings of existing object detectors and physical attacks and this is the main motivation of our work. Therefore, our benchmark might not be captured by previous validation settings, leading to the discrepancy between our results and the reviewer's individual experiences.

In comparison, we removed various physical dynamics including weather (rain, snow, fog), lighting (nighttime), and distance (far positions), and reproduced the results of several attack methods on YOLOv7 as reported in ACTIVE [9], which are listed in Table XXIII. It is worth noting that these reported results are also included in our benchmark with particular evaluation settings.

Contrary to the simplified settings of these reproduced experiments, more comprehensive physical dynamics incorporated into our benchmark significantly highlight the inef-

fectiveness of existing physical attacks. These aspects may not have been adequately captured by previous validation settings. As illustrated in Table XXIII, when we exclude various dynamics, the effectiveness of physical attacks notably increases, thereby reducing the performance of object detectors. Therefore, our benchmark strive to encompass and align these physical dynamics for comprehensive and equitable comparisons.

2) *Performance gap between attacks against vehicle and person detection:* For the gap between attacks against vehicle and person detection, one reason is that these attacks are optimized to fool object detectors in particular target detection during training process. Consequently, we follow the attack purpose of the original works in this benchmark to attack specified target category accordingly for fairness, which partially accounts for the phenomenon that pedestrian detection performance is less affected regarding various attacks in comparison with car detection.

Another potential reason is that the stronger physical perturbations are optimized with consideration of 3D space and

TABLE V: The selected detectors and their corresponding config files. 1-25 and 26-40 are CNN-based One-stage and Two-stage object detectors, respectively. 41-48 are Transformer-based object detectors. The corresponding config files of the detectors are available in our codebase or MMDetection [126] toolbox.

Number	Config Files	Detectors
1	atss_r50_fpn_1x_coco	ATSS [92]
2	autoassign_r50-caffe_fpn_1x_coco	AutoAssign [93]
3	centernet-update_r50-caffe_fpn_ms-1x_coco	CenterNet [95]
4	centripetalnet_hourglass104_16xb6-crop511-210e-mstest_coco	CentripetalNet [104]
5	cornernet_hourglass104_10xb5-crop511-210e-mstest_coco	CornerNet [96]
6	ddod_r50_fpn_1x_coco	DDOD [98]
7	atss_r50_fpn_dyhead_1x_coco	DyHead [26]
8	retinanet_effb3_fpn_8xb4-crop896-1x_coco	EfficientNet [99]
9	fcos_x101-64x4d_fpn_gn-head_ms-640-800-2x_coco	FCOS [100]
10	fovea_r50_fpn_4xb4-1x_coco	FoveaBox [101]
11	freeanchor_r50_fpn_1x_coco	FreeAnchor [102]
12	fsaf_r50_fpn_1x_coco	FSAF [105]
13	gfl_r50_fpn_1x_coco	GFL [94]
14	ld_r50-gflv1-r101_fpn_1x_coco	LD [103]
15	retinanet_r50_nasfpn_crop640-50e_coco	NAS-FPN [130]
16	paa_r50_fpn_1x_coco	PAA [97]
17	retinanet_r50_fpn_1x_coco	RetinaNet [110]
18	rtmdet_s_8xb32-300e_coco	RTMDet [106]
19	tood_r50_fpn_1x_coco	TOOD [107]
20	vfnet_r50_fpn_1x_coco	VarifocalNet [108]
21	yolov5_l-p6-v62_syncbn_fast_8xb16-300e_coco	YOLOv5 [34]
22	yolov6_l_syncbn_fast_8xb32-300e_coco	YOLOv6 [35]
23	yolov7_l_syncbn_fast_8x16b-300e_coco	YOLOv7 [36]
24	yolov8_l_syncbn_fast_8xb16-500e_coco	YOLOv8 [37]
25	yolox_l_fast_8xb8-300e_coco	YOLOX [109]
26	faster-rcnn_r50_fpn_1x_coco	Faster R-CNN [23]
27	cascade-rcnn_r50_fpn_1x_coco	Cascade R-CNN [111]
28	cascade-rpn_faster-rcnn_r50-caffe_fpn_1x_coco	Cascade RPN [112]
29	dh-faster-rcnn_r50_fpn_1x_coco	Double Heads [26]
30	faster-rcnn_r50_fpg_crop640-50e_coco	FPG [113]
31	grid-rcnn_r50_fpn_gn-head_2x_coco	Grid R-CNN [24]
32	ga-faster-rcnn_x101-32x4d_fpn_1x_coco	Guided Anchoring [119]
33	faster-rcnn_hrnetv2p-w18-1x_coco	HRNet [115]
34	libra-retinanet_r50_fpn_1x_coco	Libra R-CNN [25]
35	faster-rcnn_r50_pafpn_1x_coco	PAFPN [114]
36	reppoints-moment_r50_fpn_1x_coco	RepPoints [120]
37	faster-rcnn_res2net-101_fpn_2x_coco	Res2Net [117]
38	faster-rcnn_s50_fpn_syncbn-backbone+head_ms-range-1x_coco	ResNeSt [116]
39	sabl-faster-rcnn_r50_fpn_1x_coco	SABL [118]
40	sparse-rcnn_r50_fpn_1x_coco	Sparse R-CNN [28]
41	detr_r50_8xb2-150e_coco	DETR [40]
42	conditional-detr_r50_8xb2-50e_coco	Conditional DETR [124]
43	ddq-detr-4scale_r50_8xb2-12e_coco	DDQ [122]
44	dab-detr_r50_8xb2-50e_coco	DAB-DETR [41]
45	deformable-detr_r50_16xb2-50e_coco	Deformable DETR [123]
46	dino-4scale_r50_8xb2-12e_coco	DINO [43]
47	retinanet_pvt-t_fpn_1x_coco	PVT [121]
48	retinanet_pvtv2-b0_fpn_1x_coco	PVTv2 [121]

accommodate more complex physical dynamics, while physical attacks aiming to fool person detectors are commonly performed with optimized 2D patches, which work well in particular physical dynamics, as evidenced by the ablation experiments in B-B1, which empirically demonstrate the pressing need and necessity of a comprehensive and rigorous benchmark for physical attacks.

C. Supplemented experiments analysis and discussion

1) Detection perspective: Vehicle Detection Perspective:

Physical attacks on vehicle detection systems pose a substantial challenge due to the specialized nature of the perturba-

tions crafted to deceive these models. These attacks can lead to a drastic decline in average recall rates, reaching as low as 50%. This high level of vulnerability is largely attributed to the complex dynamics in the 3D environment where vehicles operate. Physical attacks on vehicle detectors exploit this three-dimensional context, introducing perturbations that consider real-world factors such as lighting, perspective, occlusion, and motion, making them more effective in disrupting the model's performance.

On the other hand, pedestrians, operating in a somewhat simpler 2D plane, seem to be less affected by similar adver-

sarial attacks, with a decrease in average recall rates of less than 20%. Adversarial examples targeting pedestrian detection typically involve 2D patches, which might be more straightforward to apply in specific scenarios but may not account for the full range of real-world complexities. As a result, there is an urgent demand to establish a comprehensive and stringent benchmark to systematically evaluate the resilience of these models against physical attacks, facilitating research and development towards more secure systems.

Pedestrian Detection Perspective:

In contrast to vehicle detection, pedestrian detectors exhibit a certain level of inherent robustness, potentially due to the simpler constraints imposed on the recognition process. Nevertheless, as seen across various detectors, the extent of this robustness varies widely. Models like EfficientNet, YOLO series, RTMDet (one-stage detectors), and DDQ (transformer-based detectors) demonstrate commendable resistance to physical attacks. The superior performance of DDQ could be linked to the attention mechanisms inherent to transformer architectures, which are capable of capturing global spatial dependencies, thus mitigating the impact of adversarial perturbations.

However, it is evident from the benchmark results that not all state-of-the-art (SOTA) detectors offer comparable adversarial robustness. Many detectors exhibit varying degrees of vulnerability, indicating that peak accuracy in standard detection tasks does not automatically guarantee resilience against adversarial threats. Consequently, this benchmarking framework not only identifies areas of weakness for refinement but also contributes to a better understanding of the interplay between detection performance and adversarial robustness in real-world deployments.

In conclusion, understanding and mitigating the effects of physical attacks in both vehicle and pedestrian detection domains can greatly benefit deep learning and computer vision research. By developing more robust models resistant to such attacks, we can enhance the safety and reliability of autonomous systems that rely on accurate object detection, ultimately fostering advancements in the fields of automotive technology, smart city infrastructure, and robotics. Furthermore, this benchmark would encourage researchers to explore defensive techniques and novel architectures that better withstand both digital and physical adversarial threats, pushing the boundaries of deep learning and computer vision capabilities.

2) Attack perspective: From the attacker's viewpoint, the effectiveness of physical attacks on deep learning-based vehicle detection systems is highly variant. Certain methodologies, such as ACTIVE, achieve astonishingly high success rates in defeating the detectors, with ASR values surpassing 70%. However, the majority of current attacks struggle to maintain comparable performance, often failing to reach even 20% ASR. This discrepancy can be partly attributed to the rapid advancements in detection algorithms, with the latest state-of-the-art models like EfficientNet, YOLO series, and RTMDet demonstrating increased resilience against known attacks. This disparity in the evolutionary pace between attackers and defenders underscores the importance of continuous research and innovation in adversarial attacks to keep pace with the evolving landscape of detection techniques.

Moreover, this evolving dynamic underscores a critical need for a more dynamic and collaborative ecosystem in deep learning and computer vision research. By closing the gap between attack methods and detector capabilities, the field will likely see increased robustness and security measures, ultimately benefiting automotive safety and other real-world applications relying on these systems.

On the other hand, when evaluating person detection, the outcome of physical attacks exhibits a different pattern, with ASR values typically remaining below 20%, and often even below 0%, which indicates that the attack method is less effective than random guessing and the eye-catching perturbation may arouse more attention than the object itself. Additionally, the variable transferability of these attack methodologies across different detectors leads to a wide disparity in ASR values. In certain instances, this manifests as negative ASR figures, indicative of a backfiring effect where the detectors become more adept at identifying targets in the presence of attempted attacks.

The significantly lower effectiveness of these attacks on pedestrian detection models highlights the comparative advantages of their 2D nature against primarily 2D adversarial perturbations. Nevertheless, the AdvTexture method, despite being a 2D approach, manages to incorporate 3D considerations, achieving higher ASRs compared to other attacks. This underscores the pivotal role of incorporating 3D awareness into attack strategies to exploit the vulnerabilities of pedestrian detectors more effectively.

These contrasting observations highlight the need for more sophisticated attack methods in the domain of pedestrian detection. By advancing the understanding of how 2D techniques can be adapted or combined with 3D concepts, attackers can create more potent adversarial samples, driving defender-side innovation to fortify models further. Such advancements will ultimately contribute to the progression of the field by promoting the design of more secure and reliable computer vision systems, particularly relevant in surveillance, autonomous navigation, and smart city infrastructures.

In summary, the diverse outcomes of physical attacks on both vehicle and person detection emphasize the importance of ongoing research and competition between attack and defense approaches. As the attacks become more intricate and align with the complex nature of real-world scenarios, deep learning and computer vision models will adapt, increasing their resilience and overall functionality. This continuous push-and-pull between adversaries and protectors fosters the evolution of robust, secure, and accurate object-detection technologies essential for numerous applications, including automotive safety, surveillance, and urban automation.

D. Additional overall experimental results

Due to space constraints, we provide additional overall experimental results in this part, as shown in Table VI, VII, VIII, IX, X, XI, XII, and XIII. In addition, the visualized evaluation results are shown in Fig. 17, 18, 19, 20, 21, 22, and 23.

TABLE VI: Overall experimental results of vehicle detection in the metric of mAP50(%).

	Clean	Random	ACTIVE	DTA	FCA	APPA	POOPatch	3D2Fool	CAMOU	RPAU
ATSS	0.83	0.477	0.231	0.545	0.606	0.502	0.434	0.678	0.235	0.532
AutoAssign	0.786	0.574	0.37	0.559	0.589	0.609	0.415	0.722	0.487	0.648
CenterNet	0.839	0.558	0.297	0.552	0.57	0.58	0.426	0.742	0.412	0.521
CentripetalNet	0.78	0.685	0.558	0.725	0.687	0.725	0.527	0.801	0.501	0.648
CornerNet	0.748	0.586	0.438	0.652	0.593	0.653	0.458	0.779	0.429	0.582
DDOD	0.838	0.708	0.433	0.695	0.686	0.745	0.548	0.785	0.694	0.646
DyHead	0.876	0.611	0.385	0.566	0.725	0.614	0.574	0.671	0.402	0.73
EfficientNet	0.881	0.711	0.506	0.687	0.721	0.764	0.638	0.763	0.71	0.665
FCOS	0.933	0.804	0.676	0.838	0.795	0.855	0.658	0.894	0.76	0.824
FoveaBox	0.814	0.597	0.294	0.514	0.645	0.548	0.467	0.649	0.469	0.618
FreeAnchor	0.81	0.51	0.381	0.611	0.563	0.643	0.431	0.638	0.336	0.582
FSAF	0.788	0.51	0.233	0.566	0.559	0.537	0.432	0.661	0.364	0.529
GFL	0.852	0.509	0.202	0.456	0.626	0.485	0.485	0.63	0.201	0.602
LD	0.825	0.554	0.305	0.563	0.658	0.548	0.463	0.664	0.29	0.591
NAS-FPN	0.87	0.623	0.473	0.662	0.673	0.695	0.5	0.764	0.382	0.655
PAA	0.808	0.582	0.474	0.621	0.605	0.619	0.501	0.685	0.567	0.64
RetinaNet	0.85	0.511	0.349	0.565	0.653	0.568	0.479	0.684	0.43	0.584
RTMDet	0.875	0.717	0.625	0.771	0.733	0.753	0.736	0.821	0.676	0.72
TOOD	0.781	0.495	0.353	0.522	0.584	0.557	0.462	0.615	0.37	0.572
VarifocalNet	0.874	0.472	0.205	0.424	0.628	0.529	0.419	0.573	0.208	0.538
YOLOv5	0.886	0.76	0.744	0.762	0.807	0.821	0.788	0.857	0.812	0.75
YOLOv6	0.907	0.824	0.747	0.834	0.833	0.887	0.776	0.896	0.851	0.784
YOLOv7	0.906	0.774	0.762	0.829	0.822	0.866	0.786	0.903	0.774	0.803
YOLOv8	0.929	0.791	0.761	0.812	0.838	0.873	0.793	0.917	0.74	0.803
YOLOX	0.908	0.766	0.683	0.783	0.817	0.851	0.794	0.849	0.702	0.782
Faster R-CNN	0.772	0.375	0.141	0.338	0.509	0.46	0.369	0.612	0.268	0.438
Cascade R-CNN	0.802	0.483	0.297	0.488	0.574	0.607	0.407	0.673	0.334	0.532
Cascade RPN	0.805	0.53	0.291	0.452	0.588	0.55	0.441	0.662	0.452	0.548
Double Heads	0.797	0.521	0.295	0.539	0.537	0.621	0.404	0.713	0.445	0.516
FPG	0.846	0.678	0.486	0.714	0.671	0.749	0.503	0.806	0.47	0.65
Grid R-CNN	0.795	0.472	0.244	0.513	0.529	0.65	0.399	0.699	0.392	0.494
Guided Anchoring	0.904	0.723	0.555	0.747	0.748	0.781	0.524	0.828	0.738	0.717
HRNet	0.76	0.547	0.29	0.52	0.512	0.571	0.336	0.738	0.462	0.511
Libra R-CNN	0.78	0.49	0.294	0.527	0.563	0.492	0.486	0.664	0.334	0.54
PAFPN	0.8	0.497	0.206	0.463	0.562	0.54	0.413	0.682	0.383	0.457
RepPoints	0.847	0.576	0.222	0.525	0.547	0.557	0.427	0.712	0.565	0.523
Res2Net	0.874	0.64	0.494	0.698	0.72	0.74	0.482	0.797	0.601	0.716
ResNeSt	0.837	0.502	0.352	0.535	0.587	0.499	0.538	0.493	0.407	0.555
SABL	0.796	0.46	0.262	0.501	0.563	0.535	0.423	0.648	0.359	0.484
Sparse R-CNN	0.774	0.469	0.257	0.398	0.604	0.418	0.44	0.518	0.316	0.532
DETR	0.636	0.114	0.048	0.033	0.351	0.17	0.333	0.198	0.025	0.339
Conditional DETR	0.793	0.554	0.408	0.575	0.644	0.617	0.525	0.7	0.457	0.671
DDQ	0.809	0.55	0.457	0.531	0.649	0.676	0.631	0.626	0.453	0.629
DAB-DETR	0.838	0.391	0.194	0.308	0.616	0.473	0.488	0.576	0.163	0.526
Deformable DETR	0.827	0.642	0.371	0.528	0.641	0.67	0.46	0.662	0.525	0.626
DINO	0.78	0.351	0.236	0.322	0.543	0.56	0.423	0.526	0.217	0.49
PVT	0.828	0.719	0.355	0.648	0.711	0.802	0.547	0.853	0.592	0.52
PVTv2	0.845	0.666	0.494	0.763	0.621	0.844	0.425	0.803	0.704	0.476

E. Additional ablation experiments

1) *Ablation study on physical dynamics:* Due to space constraints, we provide additional ablation experimental results in this part, as shown in Table XIV, XV, XVI, XVII, XVIII, and XIX. In addition, the visualized evaluation results are shown in Fig. 24, 25, and 26.

2) *Ablation study on training dataset:* To further investigate the impact of the training dataset on the physical attacks, we collected ten physical attacks for fooling person detection, and the results are shown in Table XXI, where the Median ASR represents the median attack success rate across the 48 detec-

tors. It can be observed that physical attacks trained on the INRIA and COCO datasets achieve comparable performance in general.

3) *Ablation study on 2D and 3D perturbations:* Physical attacks that evaluate 2D adversarial patches from a frontal perspective have a significant limitation, as they do not account for the effects of multiple viewing angles in a 3D environment. Our study aims to bridge this gap by developing a comprehensive benchmark for assessing physical attacks from various angles and incorporating a broader range of physical dynamics. During our investigation, we noted a substantial

TABLE VII: Overall experimental results of vehicle detection in the metric of mAP50:95(%).

	Clean	Random	ACTIVE	DTA	FCA	APPA	POOPatch	3D2Fool	CAMOU	RPAU
ATSS	0.238	0.156	0.077	0.182	0.183	0.164	0.133	0.212	0.087	0.166
AutoAssign	0.238	0.183	0.126	0.189	0.182	0.199	0.139	0.236	0.164	0.212
CenterNet	0.238	0.167	0.093	0.178	0.165	0.182	0.14	0.23	0.126	0.156
CentripetalNet	0.228	0.215	0.164	0.225	0.206	0.22	0.161	0.245	0.155	0.194
CornerNet	0.218	0.184	0.132	0.198	0.175	0.199	0.142	0.237	0.129	0.173
DDOD	0.242	0.21	0.127	0.221	0.202	0.234	0.174	0.242	0.211	0.193
DyHead	0.26	0.196	0.121	0.18	0.221	0.191	0.18	0.21	0.126	0.227
EfficientNet	0.252	0.225	0.168	0.217	0.219	0.239	0.206	0.243	0.232	0.21
FCOS	0.277	0.251	0.211	0.266	0.246	0.272	0.214	0.285	0.236	0.256
FoveaBox	0.235	0.184	0.089	0.165	0.191	0.17	0.145	0.196	0.149	0.193
FreeAnchor	0.241	0.159	0.111	0.19	0.179	0.205	0.138	0.205	0.098	0.184
FSAF	0.231	0.171	0.077	0.187	0.181	0.181	0.141	0.213	0.129	0.179
GFL	0.244	0.169	0.064	0.152	0.192	0.163	0.157	0.201	0.069	0.192
LD	0.235	0.176	0.095	0.181	0.205	0.176	0.146	0.208	0.094	0.187
NAS-FPN	0.256	0.199	0.146	0.209	0.213	0.215	0.16	0.251	0.124	0.208
PAA	0.237	0.178	0.139	0.189	0.178	0.195	0.157	0.208	0.173	0.194
RetinaNet	0.249	0.169	0.108	0.194	0.209	0.189	0.166	0.227	0.155	0.191
RTMDet	0.254	0.227	0.185	0.235	0.224	0.232	0.239	0.241	0.209	0.221
TOOD	0.231	0.155	0.109	0.159	0.173	0.176	0.14	0.183	0.118	0.169
VarifocalNet	0.248	0.144	0.063	0.127	0.188	0.164	0.132	0.175	0.062	0.162
YOLOv5	0.259	0.227	0.223	0.23	0.237	0.249	0.244	0.253	0.246	0.221
YOLOv6	0.256	0.245	0.218	0.251	0.242	0.262	0.238	0.263	0.25	0.229
YOLOv7	0.265	0.241	0.225	0.252	0.246	0.262	0.241	0.272	0.227	0.243
YOLOv8	0.276	0.246	0.239	0.254	0.252	0.269	0.256	0.283	0.236	0.246
YOLOX	0.263	0.233	0.212	0.236	0.237	0.253	0.248	0.251	0.217	0.233
Faster R-CNN	0.212	0.117	0.042	0.11	0.159	0.145	0.111	0.193	0.087	0.137
Cascade R-CNN	0.232	0.152	0.088	0.15	0.182	0.19	0.135	0.214	0.115	0.169
Cascade RPN	0.229	0.157	0.083	0.132	0.175	0.157	0.137	0.201	0.122	0.166
Double Heads	0.238	0.168	0.09	0.179	0.171	0.205	0.143	0.231	0.15	0.164
FPG	0.247	0.222	0.149	0.224	0.21	0.233	0.161	0.265	0.152	0.216
Grid R-CNN	0.231	0.154	0.078	0.173	0.17	0.208	0.135	0.225	0.127	0.164
Guided Anchoring	0.269	0.239	0.176	0.244	0.235	0.243	0.174	0.265	0.244	0.229
HRNet	0.219	0.171	0.091	0.164	0.159	0.173	0.114	0.236	0.148	0.165
Libra R-CNN	0.234	0.161	0.094	0.18	0.179	0.169	0.159	0.222	0.115	0.175
PAFPN	0.219	0.147	0.057	0.145	0.17	0.168	0.122	0.205	0.129	0.139
RepPoints	0.251	0.18	0.068	0.167	0.172	0.185	0.141	0.234	0.189	0.166
Res2Net	0.25	0.195	0.154	0.217	0.212	0.22	0.162	0.244	0.192	0.22
ResNeSt	0.23	0.156	0.101	0.158	0.174	0.154	0.169	0.142	0.126	0.162
SABL	0.233	0.146	0.08	0.159	0.173	0.177	0.139	0.199	0.123	0.155
Sparse R-CNN	0.238	0.159	0.095	0.137	0.195	0.145	0.154	0.172	0.128	0.168
DETR	0.186	0.047	0.017	0.01	0.113	0.059	0.111	0.065	0.009	0.105
Conditional DETR	0.236	0.183	0.129	0.183	0.199	0.205	0.172	0.211	0.154	0.212
DDQ	0.232	0.164	0.133	0.152	0.185	0.2	0.189	0.182	0.136	0.187
DAB-DETR	0.239	0.115	0.057	0.09	0.178	0.145	0.143	0.157	0.05	0.143
Deformable DETR	0.229	0.187	0.106	0.147	0.177	0.197	0.136	0.188	0.158	0.171
DINO	0.232	0.111	0.075	0.104	0.172	0.18	0.139	0.161	0.078	0.156
PVT	0.229	0.229	0.106	0.213	0.224	0.254	0.177	0.26	0.199	0.163
PVTv2	0.24	0.214	0.154	0.252	0.195	0.275	0.147	0.244	0.237	0.149

drop in performance (detection rate: $\frac{n_{detected}}{n_{total}}$) when adversarial patches were only applied to the frontal view of objects. To ensure a fair comparison and enhance the efficacy of the attacks, we expanded the application of these patches to cover the entirety of the object's surface. Additional experiments were conducted to assess the impact of adversarial patches on frontal views using several object detection algorithms. The results are summarized in Table XXII. The 'Entire Surface' column highlights cases where the adversarial patch was applied across the entire surface of an object. The values in parentheses indicate the relative decrease in performance

compared to full-surface patching.

APPENDIX C USER FEEDBACK

To ensure ease of use, we have addressed potential barriers by user feedback, such as CARLA deployment and customizing adversarial objects, by providing a comprehensive Docker installation guide for CARLA and a tutorial on customizing adversarial objects in our documentation. These resources enable users to install CARLA and customize objects in just a few minutes. We also conducted usability testing with five

TABLE VIII: Overall experimental results of **vehicle** detection in the metric of **mAR50(%)**.

	Clean	Random	ACTIVE	DTA	FCA	APPA	POOPatch	3D2Fool	CAMOU	RPAU
ATSS	0.973	0.808	0.518	0.84	0.883	0.811	0.732	0.924	0.549	0.826
AutoAssign	0.946	0.846	0.703	0.876	0.849	0.888	0.763	0.938	0.856	0.901
CenterNet	0.976	0.926	0.674	0.893	0.901	0.897	0.806	0.97	0.873	0.895
CentripetalNet	0.943	0.867	0.753	0.908	0.901	0.918	0.749	0.958	0.81	0.853
CornerNet	0.948	0.8	0.692	0.905	0.847	0.902	0.697	0.959	0.722	0.829
DDOD	0.95	0.936	0.756	0.923	0.9	0.934	0.863	0.948	0.949	0.909
DyHead	0.977	0.836	0.606	0.792	0.903	0.845	0.838	0.843	0.685	0.927
EfficientNet	0.977	0.974	0.912	0.975	0.974	0.97	0.951	0.973	0.966	0.948
FCOS	0.981	0.974	0.93	0.968	0.951	0.975	0.931	0.972	0.939	0.959
FoveaBox	0.955	0.873	0.623	0.806	0.896	0.832	0.748	0.881	0.768	0.878
FreeAnchor	0.973	0.841	0.855	0.928	0.864	0.939	0.781	0.89	0.68	0.887
FSAF	0.95	0.836	0.567	0.886	0.823	0.84	0.632	0.918	0.659	0.8
GFL	0.978	0.837	0.575	0.81	0.908	0.839	0.836	0.913	0.549	0.903
LD	0.98	0.893	0.579	0.862	0.927	0.871	0.743	0.917	0.676	0.91
NAS-FPN	0.975	0.916	0.768	0.932	0.944	0.946	0.817	0.937	0.766	0.925
PAA	0.966	0.923	0.861	0.952	0.905	0.937	0.895	0.938	0.951	0.92
RetinaNet	0.98	0.859	0.764	0.915	0.934	0.89	0.775	0.921	0.729	0.9
RTMDet	0.982	0.954	0.943	0.971	0.958	0.971	0.975	0.98	0.99	0.937
TOOD	0.908	0.763	0.666	0.83	0.834	0.826	0.717	0.847	0.622	0.853
VarifocalNet	0.977	0.802	0.486	0.77	0.9	0.832	0.671	0.866	0.466	0.829
YOLOv5	0.975	0.979	0.967	0.974	0.977	0.974	0.974	0.972	0.985	0.966
YOLOv6	0.985	0.982	0.968	0.974	0.983	0.979	0.974	0.984	0.978	0.973
YOLOv7	0.962	0.924	0.931	0.948	0.939	0.958	0.932	0.96	0.932	0.931
YOLOv8	0.975	0.919	0.903	0.93	0.942	0.96	0.916	0.965	0.885	0.917
YOLOX	0.955	0.877	0.853	0.902	0.91	0.945	0.926	0.918	0.868	0.891
Faster R-CNN	0.846	0.479	0.268	0.493	0.595	0.593	0.417	0.752	0.337	0.532
Cascade R-CNN	0.854	0.539	0.382	0.591	0.65	0.689	0.448	0.78	0.368	0.602
Cascade RPN	0.973	0.884	0.741	0.933	0.898	0.957	0.885	0.967	0.868	0.891
Double Heads	0.849	0.597	0.416	0.654	0.621	0.725	0.459	0.819	0.488	0.594
FPG	0.912	0.763	0.624	0.848	0.761	0.866	0.587	0.907	0.61	0.748
Grid R-CNN	0.873	0.585	0.39	0.672	0.653	0.794	0.488	0.836	0.495	0.614
Guided Anchoring	0.975	0.962	0.914	0.966	0.962	0.966	0.839	0.961	0.929	0.956
HRNet	0.844	0.655	0.429	0.687	0.627	0.732	0.423	0.875	0.532	0.609
Libra R-CNN	0.959	0.828	0.599	0.824	0.892	0.775	0.805	0.92	0.563	0.858
PAFPN	0.856	0.61	0.337	0.591	0.661	0.659	0.49	0.805	0.434	0.569
RepPoints	0.978	0.903	0.595	0.874	0.885	0.884	0.833	0.945	0.883	0.848
Res2Net	0.911	0.692	0.541	0.757	0.763	0.793	0.511	0.852	0.646	0.761
ResNeSt	0.929	0.646	0.497	0.727	0.701	0.691	0.685	0.802	0.515	0.716
SABL	0.856	0.545	0.387	0.646	0.656	0.636	0.488	0.774	0.407	0.571
Sparse R-CNN	0.959	0.877	0.513	0.759	0.913	0.746	0.733	0.882	0.605	0.871
DETR	0.746	0.328	0.232	0.256	0.468	0.383	0.457	0.369	0.234	0.468
Conditional DETR	0.962	0.831	0.73	0.931	0.865	0.934	0.881	0.964	0.839	0.917
DDQ	0.983	0.976	0.972	0.979	0.975	0.975	0.977	0.974	0.983	0.969
DAB-DETR	0.98	0.909	0.928	0.97	0.948	0.968	0.946	0.924	0.91	0.934
Deformable DETR	0.954	0.907	0.704	0.902	0.879	0.924	0.766	0.905	0.91	0.88
DINO	0.975	0.895	0.883	0.953	0.923	0.958	0.922	0.953	0.912	0.924
PVT	0.948	0.953	0.827	0.936	0.957	0.956	0.901	0.963	0.954	0.886
PVTv2	0.973	0.942	0.884	0.952	0.934	0.973	0.835	0.967	0.941	0.867

researchers from a well-known University and got feedback from them in the form of a survey questionnaire as shown in Table XXV. The users consistently found the benchmark easy to use and provided positive feedback on its usability.

TABLE IX: Overall experimental results of vehicle detection in the metric of mAR50:95(%).

	Clean	Random	ACTIVE	DTA	FCA	APPA	POCPatch	3D2Fool	CAMOU	RPAU
ATSS	0.374	0.318	0.206	0.341	0.342	0.321	0.296	0.371	0.219	0.325
AutoAssign	0.385	0.341	0.293	0.366	0.335	0.367	0.32	0.39	0.345	0.374
CenterNet	0.396	0.365	0.275	0.373	0.353	0.373	0.342	0.408	0.344	0.361
CentripetalNet	0.378	0.362	0.313	0.386	0.365	0.384	0.31	0.404	0.321	0.35
CornerNet	0.387	0.337	0.286	0.382	0.343	0.38	0.292	0.401	0.282	0.336
DDOD	0.366	0.363	0.302	0.371	0.347	0.379	0.357	0.383	0.366	0.358
DyHead	0.378	0.338	0.254	0.32	0.35	0.338	0.342	0.336	0.255	0.364
EfficientNet	0.387	0.397	0.377	0.395	0.393	0.394	0.399	0.402	0.406	0.38
FCOS	0.401	0.4	0.382	0.399	0.389	0.405	0.389	0.418	0.37	0.397
FoveaBox	0.371	0.337	0.246	0.327	0.344	0.333	0.304	0.353	0.303	0.344
FreeAnchor	0.384	0.332	0.333	0.367	0.345	0.38	0.324	0.365	0.244	0.357
FSAF	0.37	0.337	0.229	0.362	0.33	0.342	0.271	0.377	0.257	0.325
GFL	0.375	0.337	0.238	0.336	0.354	0.341	0.348	0.373	0.226	0.357
LD	0.372	0.352	0.236	0.351	0.367	0.348	0.314	0.367	0.268	0.36
NAS-FPN	0.38	0.367	0.307	0.373	0.378	0.375	0.34	0.384	0.296	0.373
PAA	0.399	0.371	0.337	0.385	0.357	0.39	0.376	0.388	0.383	0.369
RetinaNet	0.392	0.343	0.298	0.371	0.378	0.367	0.34	0.39	0.301	0.364
RTMDet	0.357	0.322	0.293	0.345	0.319	0.351	0.34	0.362	0.304	0.311
TOOD	0.345	0.293	0.261	0.332	0.312	0.329	0.283	0.327	0.236	0.323
VarifocalNet	0.376	0.31	0.188	0.304	0.352	0.329	0.277	0.338	0.173	0.321
YOLOv5	0.364	0.364	0.366	0.371	0.358	0.378	0.377	0.373	0.354	0.355
YOLOv6	0.357	0.361	0.343	0.363	0.352	0.37	0.359	0.37	0.351	0.343
YOLOv7	0.366	0.356	0.339	0.36	0.355	0.371	0.358	0.376	0.338	0.352
YOLOv8	0.376	0.358	0.354	0.369	0.357	0.376	0.368	0.385	0.336	0.354
YOLOX	0.362	0.33	0.325	0.345	0.338	0.368	0.367	0.358	0.309	0.332
Faster R-CNN	0.303	0.187	0.099	0.194	0.228	0.236	0.163	0.302	0.134	0.203
Cascade R-CNN	0.316	0.213	0.147	0.232	0.255	0.272	0.185	0.312	0.145	0.239
Cascade RPN	0.375	0.333	0.276	0.351	0.341	0.358	0.353	0.376	0.307	0.341
Double Heads	0.32	0.239	0.165	0.27	0.249	0.299	0.2	0.333	0.193	0.238
FPG	0.347	0.31	0.249	0.35	0.302	0.352	0.239	0.379	0.249	0.306
Grid R-CNN	0.326	0.23	0.155	0.273	0.259	0.312	0.205	0.33	0.188	0.244
Guided Anchoring	0.38	0.396	0.373	0.39	0.385	0.384	0.35	0.399	0.363	0.379
HRNet	0.303	0.239	0.159	0.252	0.228	0.264	0.167	0.333	0.188	0.226
Libra R-CNN	0.383	0.326	0.233	0.334	0.357	0.315	0.333	0.385	0.223	0.337
PAFPN	0.304	0.226	0.127	0.228	0.249	0.254	0.185	0.304	0.167	0.214
RepPoints	0.387	0.356	0.235	0.355	0.349	0.372	0.358	0.393	0.364	0.336
Res2Net	0.341	0.261	0.208	0.302	0.284	0.305	0.21	0.338	0.237	0.292
ResNeSt	0.341	0.257	0.189	0.274	0.261	0.27	0.268	0.302	0.201	0.271
SABL	0.318	0.213	0.151	0.26	0.253	0.257	0.204	0.303	0.16	0.226
Sparse R-CNN	0.395	0.367	0.219	0.324	0.38	0.32	0.312	0.381	0.249	0.361
DETR	0.334	0.144	0.105	0.106	0.203	0.171	0.204	0.158	0.085	0.203
Conditional DETR	0.387	0.341	0.32	0.404	0.342	0.406	0.367	0.391	0.328	0.374
DDQ	0.393	0.389	0.388	0.394	0.39	0.389	0.403	0.391	0.374	0.389
DAB-DETR	0.404	0.366	0.384	0.406	0.385	0.413	0.404	0.376	0.36	0.383
Deformable DETR	0.378	0.352	0.277	0.351	0.333	0.372	0.3	0.351	0.339	0.34
DINO	0.388	0.353	0.353	0.39	0.364	0.388	0.372	0.384	0.346	0.37
PVT	0.349	0.371	0.319	0.365	0.37	0.379	0.365	0.382	0.387	0.34
PVTv2	0.372	0.372	0.359	0.386	0.365	0.399	0.354	0.39	0.384	0.343

TABLE X: Overall experimental results of person detection in the metric of mAP50(%).

	Clean	Random	AdvCam	UPC	NatPatch	MTD	LAP	InvisCloak	DAP	AdvTshirt	AdvTexture	AdvPatch	AdvPattern	AdvCat
ATSS	0.54	0.517	0.498	0.428	0.419	0.473	0.522	0.468	0.495	0.458	0.385	0.454	0.492	0.514
AutoAssign	0.491	0.466	0.454	0.314	0.36	0.423	0.456	0.43	0.403	0.41	0.346	0.427	0.453	0.484
CenterNet	0.524	0.476	0.477	0.408	0.39	0.469	0.483	0.436	0.437	0.45	0.372	0.43	0.474	0.524
CentripetalNet	0.526	0.53	0.524	0.48	0.349	0.524	0.508	0.51	0.471	0.473	0.405	0.472	0.515	0.526
CornerNet	0.517	0.51	0.505	0.403	0.295	0.488	0.494	0.449	0.444	0.42	0.345	0.414	0.48	0.506
DDOD	0.481	0.48	0.448	0.359	0.416	0.421	0.47	0.445	0.453	0.433	0.329	0.409	0.442	0.45
DyHead	0.474	0.483	0.485	0.406	0.402	0.464	0.501	0.454	0.473	0.433	0.4	0.433	0.467	0.474
EfficientNet	0.457	0.431	0.442	0.398	0.418	0.406	0.431	0.399	0.407	0.403	0.394	0.396	0.422	0.433
FCOS	0.45	0.438	0.429	0.364	0.383	0.41	0.433	0.407	0.404	0.407	0.356	0.409	0.425	0.448
FoveaBox	0.543	0.53	0.536	0.473	0.475	0.482	0.509	0.481	0.523	0.483	0.374	0.458	0.498	0.533
FreeAnchor	0.537	0.522	0.493	0.414	0.396	0.431	0.492	0.443	0.446	0.411	0.331	0.447	0.48	0.513
FSAF	0.554	0.551	0.529	0.444	0.439	0.485	0.538	0.496	0.515	0.457	0.379	0.479	0.512	0.527
GFL	0.57	0.541	0.532	0.431	0.453	0.495	0.509	0.478	0.495	0.48	0.398	0.459	0.52	0.547
LD	0.57	0.54	0.524	0.401	0.397	0.486	0.517	0.484	0.489	0.477	0.385	0.484	0.519	0.535
NAS-FPN	0.442	0.436	0.433	0.365	0.391	0.4	0.451	0.399	0.394	0.398	0.32	0.399	0.413	0.435
PAA	0.464	0.464	0.457	0.402	0.389	0.43	0.451	0.432	0.447	0.402	0.322	0.409	0.441	0.463
RetinaNet	0.522	0.543	0.497	0.438	0.425	0.459	0.505	0.483	0.478	0.454	0.384	0.475	0.489	0.528
RTMDet	0.533	0.482	0.515	0.52	0.466	0.46	0.495	0.437	0.472	0.47	0.459	0.449	0.466	0.5
TOOD	0.474	0.5	0.475	0.384	0.376	0.453	0.503	0.453	0.486	0.441	0.371	0.435	0.457	0.475
VarifocalNet	0.492	0.505	0.481	0.387	0.395	0.443	0.504	0.444	0.469	0.445	0.368	0.436	0.47	0.506
YOLOv5	0.481	0.46	0.472	0.448	0.403	0.453	0.484	0.454	0.485	0.418	0.35	0.42	0.452	0.459
YOLOv6	0.467	0.445	0.461	0.456	0.435	0.438	0.444	0.446	0.459	0.437	0.423	0.432	0.444	0.449
YOLOv7	0.463	0.438	0.48	0.446	0.343	0.401	0.438	0.403	0.457	0.402	0.372	0.366	0.429	0.437
YOLOv8	0.434	0.421	0.431	0.432	0.402	0.415	0.421	0.416	0.429	0.416	0.405	0.409	0.415	0.417
YOLOX	0.448	0.436	0.457	0.457	0.382	0.432	0.46	0.425	0.46	0.433	0.393	0.412	0.426	0.437
Faster R-CNN	0.541	0.547	0.497	0.416	0.425	0.456	0.532	0.468	0.456	0.448	0.341	0.432	0.512	0.534
Cascade R-CNN	0.559	0.551	0.539	0.431	0.445	0.488	0.551	0.463	0.508	0.479	0.355	0.454	0.523	0.55
Cascade RPN	0.538	0.537	0.528	0.389	0.407	0.482	0.508	0.472	0.483	0.461	0.335	0.445	0.508	0.532
Double Heads	0.552	0.526	0.531	0.419	0.408	0.46	0.533	0.442	0.489	0.454	0.359	0.44	0.496	0.527
FPG	0.462	0.473	0.451	0.413	0.395	0.415	0.466	0.408	0.424	0.405	0.317	0.409	0.444	0.45
Grid R-CNN	0.512	0.502	0.492	0.397	0.404	0.449	0.517	0.462	0.471	0.43	0.363	0.424	0.481	0.488
Guided Anchoring	0.497	0.537	0.504	0.427	0.396	0.47	0.525	0.479	0.454	0.449	0.375	0.452	0.494	0.502
HRNet	0.498	0.489	0.495	0.457	0.404	0.453	0.489	0.47	0.442	0.472	0.419	0.437	0.443	0.497
Libra R-CNN	0.535	0.517	0.479	0.452	0.404	0.446	0.468	0.433	0.447	0.431	0.374	0.421	0.442	0.494
PAFPN	0.539	0.534	0.529	0.429	0.438	0.468	0.522	0.477	0.47	0.458	0.349	0.447	0.516	0.559
RepPoints	0.572	0.559	0.53	0.434	0.475	0.478	0.535	0.47	0.504	0.455	0.38	0.451	0.525	0.56
Res2Net	0.449	0.437	0.435	0.403	0.301	0.402	0.458	0.406	0.407	0.386	0.324	0.387	0.428	0.451
ResNeSt	0.443	0.455	0.409	0.396	0.396	0.405	0.432	0.385	0.364	0.374	0.358	0.374	0.416	0.451
SABL	0.563	0.559	0.525	0.418	0.471	0.491	0.534	0.503	0.498	0.496	0.382	0.484	0.534	0.552
Sparse R-CNN	0.492	0.481	0.477	0.38	0.347	0.434	0.484	0.386	0.396	0.407	0.352	0.389	0.455	0.493
DETR	0.553	0.497	0.481	0.337	0.318	0.467	0.49	0.466	0.432	0.473	0.343	0.444	0.475	0.51
Conditional DETR	0.535	0.497	0.453	0.378	0.351	0.424	0.449	0.401	0.44	0.416	0.294	0.412	0.422	0.485
DDQ	0.449	0.454	0.452	0.377	0.388	0.424	0.451	0.426	0.408	0.422	0.34	0.421	0.439	0.451
DAB-DETR	0.441	0.429	0.428	0.344	0.36	0.371	0.407	0.357	0.373	0.374	0.276	0.336	0.392	0.44
Deformable DETR	0.475	0.462	0.475	0.345	0.389	0.421	0.46	0.442	0.418	0.424	0.286	0.416	0.471	0.45
DINO	0.419	0.421	0.416	0.337	0.283	0.385	0.415	0.402	0.375	0.391	0.316	0.378	0.394	0.423
PVT	0.474	0.465	0.418	0.378	0.368	0.383	0.405	0.359	0.369	0.393	0.381	0.391	0.41	0.426
PVTv2	0.51	0.431	0.41	0.403	0.4	0.395	0.414	0.389	0.41	0.384	0.347	0.382	0.392	0.436

TABLE XI: Overall experimental results of person detection in the metric of mAP50:95(%).

	Clean	Random	AdvCam	UPC	NatPatch	MTD	LAP	InvisCloak	DAP	AdvTshirt	AdvTexture	AdvPatch	AdvPattern	AdvCat
ATSS	0.157	0.148	0.147	0.121	0.113	0.127	0.155	0.138	0.143	0.128	0.098	0.132	0.138	0.151
AutoAssign	0.149	0.137	0.13	0.082	0.098	0.12	0.136	0.126	0.113	0.114	0.089	0.123	0.131	0.143
CenterNet	0.161	0.145	0.147	0.12	0.111	0.141	0.151	0.132	0.132	0.135	0.102	0.124	0.142	0.166
CentripetalNet	0.143	0.145	0.144	0.142	0.083	0.145	0.139	0.142	0.131	0.128	0.109	0.127	0.138	0.143
CornerNet	0.142	0.141	0.138	0.12	0.07	0.133	0.141	0.128	0.125	0.115	0.091	0.111	0.132	0.14
DDOD	0.142	0.141	0.129	0.102	0.119	0.115	0.139	0.133	0.131	0.129	0.085	0.114	0.126	0.128
DyHead	0.135	0.137	0.144	0.113	0.109	0.127	0.149	0.132	0.141	0.12	0.106	0.122	0.131	0.134
EfficientNet	0.123	0.115	0.119	0.11	0.113	0.106	0.116	0.105	0.11	0.106	0.104	0.104	0.109	0.115
FCOS	0.126	0.117	0.117	0.1	0.102	0.108	0.118	0.109	0.106	0.107	0.094	0.106	0.109	0.123
FoveaBox	0.173	0.163	0.166	0.135	0.133	0.141	0.162	0.141	0.155	0.136	0.098	0.131	0.15	0.161
FreeAnchor	0.165	0.155	0.147	0.122	0.106	0.121	0.149	0.134	0.132	0.108	0.083	0.127	0.14	0.152
FSAF	0.175	0.174	0.166	0.135	0.13	0.142	0.169	0.155	0.158	0.131	0.097	0.144	0.159	0.167
GFL	0.177	0.164	0.162	0.123	0.13	0.138	0.16	0.143	0.157	0.144	0.108	0.133	0.153	0.164
LD	0.179	0.164	0.162	0.114	0.109	0.143	0.162	0.147	0.148	0.142	0.101	0.145	0.157	0.162
NAS-FPN	0.127	0.119	0.122	0.107	0.11	0.106	0.133	0.112	0.111	0.106	0.076	0.11	0.111	0.122
PAA	0.132	0.132	0.13	0.112	0.103	0.116	0.135	0.128	0.129	0.108	0.081	0.12	0.122	0.129
RetinaNet	0.155	0.167	0.151	0.135	0.123	0.13	0.158	0.152	0.142	0.134	0.106	0.144	0.144	0.154
RTMDet	0.152	0.133	0.148	0.158	0.129	0.12	0.14	0.112	0.137	0.129	0.126	0.117	0.121	0.138
TOOD	0.131	0.145	0.139	0.108	0.099	0.121	0.154	0.13	0.143	0.122	0.093	0.122	0.123	0.132
VarifocalNet	0.141	0.152	0.145	0.11	0.111	0.127	0.157	0.133	0.144	0.125	0.095	0.132	0.137	0.154
YOLOv5	0.135	0.124	0.132	0.129	0.104	0.122	0.136	0.124	0.142	0.108	0.082	0.106	0.122	0.124
YOLOv6	0.13	0.121	0.129	0.139	0.119	0.118	0.126	0.123	0.128	0.118	0.115	0.116	0.119	0.123
YOLOv7	0.125	0.115	0.129	0.117	0.083	0.101	0.115	0.105	0.129	0.101	0.091	0.091	0.111	0.115
YOLOv8	0.12	0.113	0.12	0.128	0.111	0.111	0.115	0.111	0.119	0.109	0.106	0.106	0.111	0.112
YOLOX	0.13	0.122	0.129	0.138	0.106	0.117	0.129	0.12	0.137	0.121	0.103	0.112	0.117	0.123
Faster R-CNN	0.159	0.166	0.146	0.114	0.11	0.119	0.164	0.136	0.133	0.122	0.083	0.121	0.152	0.158
Cascade R-CNN	0.171	0.166	0.165	0.12	0.122	0.141	0.17	0.136	0.15	0.136	0.09	0.131	0.154	0.165
Cascade RPN	0.163	0.161	0.153	0.105	0.109	0.138	0.153	0.14	0.142	0.13	0.082	0.129	0.151	0.162
Double Heads	0.169	0.161	0.164	0.12	0.11	0.134	0.165	0.128	0.144	0.128	0.085	0.126	0.151	0.16
FPG	0.133	0.136	0.132	0.121	0.109	0.112	0.141	0.115	0.124	0.105	0.073	0.112	0.123	0.126
Grid R-CNN	0.151	0.145	0.145	0.109	0.11	0.123	0.15	0.13	0.139	0.116	0.091	0.114	0.137	0.144
Guided Anchoring	0.142	0.162	0.153	0.119	0.107	0.137	0.163	0.136	0.132	0.123	0.094	0.128	0.146	0.151
HRNet	0.147	0.139	0.146	0.132	0.105	0.117	0.141	0.133	0.122	0.132	0.109	0.117	0.122	0.144
Libra R-CNN	0.155	0.145	0.136	0.133	0.119	0.119	0.134	0.123	0.126	0.121	0.102	0.115	0.12	0.135
PAFPN	0.163	0.157	0.158	0.118	0.12	0.124	0.156	0.137	0.137	0.125	0.086	0.121	0.149	0.165
RepPoints	0.178	0.177	0.166	0.124	0.14	0.131	0.167	0.135	0.146	0.132	0.099	0.129	0.154	0.174
Res2Net	0.123	0.115	0.122	0.112	0.075	0.106	0.132	0.113	0.115	0.1	0.086	0.099	0.111	0.12
ResNeSt	0.125	0.132	0.118	0.11	0.11	0.109	0.125	0.108	0.102	0.101	0.091	0.1	0.118	0.127
SABL	0.177	0.169	0.166	0.118	0.133	0.145	0.166	0.151	0.147	0.143	0.096	0.141	0.161	0.172
Sparse R-CNN	0.135	0.134	0.138	0.104	0.092	0.114	0.141	0.106	0.111	0.109	0.095	0.103	0.122	0.131
DETR	0.171	0.151	0.151	0.101	0.088	0.142	0.152	0.146	0.126	0.138	0.095	0.128	0.142	0.151
Conditional DETR	0.157	0.138	0.128	0.1	0.09	0.114	0.129	0.112	0.123	0.108	0.068	0.105	0.111	0.135
DDQ	0.119	0.119	0.124	0.099	0.099	0.112	0.124	0.116	0.113	0.11	0.085	0.113	0.117	0.12
DAB-DETR	0.118	0.111	0.117	0.092	0.094	0.099	0.111	0.095	0.101	0.098	0.071	0.087	0.103	0.117
Deformable DETR	0.136	0.129	0.142	0.093	0.107	0.118	0.134	0.128	0.12	0.12	0.073	0.114	0.137	0.121
DINO	0.112	0.11	0.112	0.086	0.068	0.097	0.113	0.106	0.098	0.101	0.083	0.097	0.1	0.112
PVT	0.14	0.134	0.117	0.107	0.099	0.108	0.115	0.099	0.102	0.111	0.106	0.109	0.113	0.12
PVTv2	0.156	0.132	0.119	0.119	0.112	0.108	0.126	0.118	0.122	0.11	0.097	0.108	0.111	0.129

TABLE XII: Overall experimental results of person detection in the metric of mAR50(%).

	Clean	Random	AdrCam	UPC	NatPatch	MTD	LAP	InvisCloak	DAP	AdvTshirt	AdvTexture	AdvPatch	AdvPattern	AdvCat
ATSS	0.835	0.827	0.823	0.802	0.782	0.823	0.818	0.789	0.788	0.798	0.741	0.786	0.823	0.844
AutoAssign	0.854	0.837	0.841	0.794	0.827	0.832	0.826	0.814	0.817	0.827	0.793	0.811	0.832	0.851
CenterNet	0.848	0.835	0.84	0.849	0.809	0.842	0.829	0.823	0.822	0.828	0.794	0.82	0.848	0.858
CentripetalNet	0.854	0.858	0.838	0.809	0.777	0.85	0.83	0.86	0.82	0.814	0.737	0.841	0.852	0.854
CornerNet	0.871	0.87	0.853	0.814	0.778	0.853	0.841	0.848	0.826	0.805	0.74	0.817	0.857	0.858
DDOD	0.737	0.729	0.739	0.723	0.746	0.716	0.732	0.713	0.725	0.72	0.678	0.694	0.72	0.742
DyHead	0.725	0.731	0.747	0.702	0.7	0.723	0.745	0.711	0.711	0.725	0.724	0.716	0.727	0.744
EfficientNet	0.794	0.771	0.789	0.768	0.751	0.758	0.781	0.762	0.754	0.762	0.748	0.752	0.766	0.793
FCOS	0.897	0.905	0.89	0.864	0.86	0.891	0.894	0.893	0.858	0.897	0.82	0.898	0.899	0.908
FoveaBox	0.824	0.814	0.836	0.836	0.794	0.825	0.813	0.799	0.819	0.81	0.785	0.776	0.818	0.836
FreeAnchor	0.792	0.801	0.788	0.791	0.772	0.786	0.789	0.768	0.769	0.781	0.717	0.784	0.793	0.809
FSAF	0.812	0.814	0.818	0.803	0.79	0.807	0.822	0.808	0.808	0.803	0.785	0.803	0.805	0.812
GFL	0.827	0.803	0.805	0.796	0.782	0.806	0.784	0.781	0.763	0.785	0.743	0.777	0.804	0.823
LD	0.809	0.797	0.802	0.785	0.762	0.787	0.786	0.764	0.77	0.777	0.735	0.763	0.797	0.813
NAS-FPN	0.773	0.765	0.775	0.726	0.75	0.74	0.768	0.744	0.732	0.741	0.7	0.751	0.76	0.778
PAA	0.816	0.819	0.821	0.808	0.791	0.805	0.805	0.792	0.784	0.78	0.73	0.783	0.807	0.837
RetinaNet	0.838	0.823	0.843	0.817	0.795	0.823	0.822	0.802	0.796	0.818	0.777	0.791	0.823	0.848
RTMDet	0.976	0.978	0.974	0.973	0.972	0.975	0.974	0.966	0.971	0.97	0.966	0.976	0.977	0.976
TOOD	0.73	0.732	0.731	0.708	0.691	0.734	0.737	0.716	0.723	0.724	0.69	0.714	0.723	0.733
VarifocalNet	0.794	0.772	0.786	0.767	0.738	0.765	0.77	0.744	0.755	0.758	0.729	0.745	0.771	0.791
YOLOv5	0.814	0.832	0.836	0.773	0.782	0.825	0.831	0.841	0.814	0.829	0.817	0.836	0.831	0.843
YOLOv6	0.96	0.963	0.957	0.951	0.939	0.959	0.943	0.964	0.943	0.95	0.921	0.956	0.963	0.966
YOLOv7	0.729	0.73	0.744	0.723	0.7	0.716	0.731	0.712	0.721	0.731	0.721	0.723	0.722	0.743
YOLOv8	0.677	0.667	0.686	0.67	0.669	0.676	0.676	0.679	0.687	0.674	0.675	0.674	0.668	0.679
YOLOX	0.685	0.697	0.702	0.702	0.702	0.7	0.708	0.706	0.709	0.707	0.731	0.71	0.682	0.692
Faster R-CNN	0.69	0.682	0.691	0.625	0.659	0.679	0.674	0.653	0.645	0.679	0.615	0.639	0.67	0.685
Cascade R-CNN	0.699	0.689	0.703	0.644	0.652	0.679	0.695	0.66	0.673	0.677	0.623	0.651	0.68	0.7
Cascade RPN	0.785	0.791	0.789	0.759	0.758	0.78	0.776	0.748	0.751	0.774	0.715	0.758	0.784	0.797
Double Heads	0.696	0.688	0.7	0.636	0.639	0.677	0.678	0.664	0.654	0.679	0.623	0.643	0.676	0.696
FPG	0.656	0.659	0.665	0.598	0.598	0.643	0.659	0.636	0.622	0.642	0.588	0.635	0.659	0.661
Grid R-CNN	0.677	0.664	0.67	0.638	0.635	0.666	0.669	0.654	0.655	0.659	0.632	0.644	0.657	0.661
Guided Anchoring	0.74	0.732	0.748	0.729	0.709	0.738	0.745	0.733	0.711	0.736	0.72	0.718	0.739	0.739
HRNet	0.665	0.669	0.679	0.645	0.644	0.671	0.673	0.661	0.646	0.676	0.65	0.645	0.649	0.678
Libra R-CNN	0.829	0.813	0.831	0.822	0.766	0.81	0.811	0.781	0.776	0.797	0.754	0.778	0.819	0.848
PAFPN	0.685	0.682	0.697	0.644	0.647	0.678	0.679	0.664	0.658	0.682	0.614	0.651	0.679	0.698
RepPoints	0.826	0.813	0.822	0.816	0.813	0.804	0.807	0.799	0.808	0.801	0.783	0.779	0.813	0.828
Res2Net	0.636	0.63	0.638	0.599	0.543	0.627	0.623	0.614	0.602	0.623	0.578	0.618	0.63	0.634
ResNeSt	0.648	0.651	0.625	0.61	0.61	0.636	0.619	0.614	0.586	0.607	0.593	0.609	0.64	0.654
SABL	0.711	0.704	0.707	0.65	0.66	0.69	0.678	0.675	0.675	0.688	0.635	0.661	0.696	0.706
Sparse R-CNN	0.702	0.685	0.694	0.67	0.668	0.672	0.688	0.65	0.644	0.667	0.649	0.651	0.68	0.694
DETR	0.893	0.9	0.877	0.761	0.798	0.904	0.887	0.891	0.827	0.882	0.789	0.889	0.895	0.892
Conditional DETR	0.73	0.729	0.717	0.706	0.708	0.697	0.703	0.67	0.691	0.683	0.643	0.692	0.69	0.712
DDQ	0.774	0.774	0.768	0.768	0.758	0.746	0.757	0.728	0.744	0.748	0.725	0.745	0.746	0.773
DAB-DETR	0.743	0.739	0.745	0.737	0.76	0.729	0.742	0.705	0.73	0.717	0.703	0.716	0.722	0.743
Deformable DETR	0.703	0.699	0.707	0.68	0.693	0.671	0.692	0.666	0.675	0.673	0.591	0.67	0.679	0.693
DINO	0.739	0.748	0.744	0.735	0.747	0.732	0.747	0.723	0.72	0.727	0.718	0.734	0.73	0.736
PVT	0.703	0.705	0.7	0.692	0.673	0.693	0.696	0.679	0.669	0.692	0.681	0.691	0.688	0.699
PVTv2	0.738	0.733	0.714	0.736	0.731	0.712	0.733	0.715	0.705	0.725	0.7	0.716	0.709	0.731

TABLE XIII: Overall experimental results of person detection in the metric of mAR50:95(%).

	Clean	Random	AdvCam	UPC	NatPatch	MTD	LAP	InvisCloak	DAP	AdvTshirt	AdvTexture	AdvPatch	AdvPattern	AdvCat
ATSS	0.326	0.313	0.318	0.301	0.298	0.304	0.306	0.291	0.302	0.297	0.269	0.286	0.307	0.322
AutoAssign	0.317	0.309	0.309	0.288	0.306	0.299	0.305	0.297	0.3	0.301	0.282	0.297	0.306	0.314
CenterNet	0.351	0.344	0.342	0.341	0.315	0.342	0.332	0.326	0.325	0.332	0.309	0.329	0.346	0.352
CentripetalNet	0.332	0.337	0.326	0.32	0.293	0.333	0.319	0.338	0.311	0.307	0.277	0.321	0.331	0.332
CornerNet	0.331	0.333	0.325	0.306	0.281	0.322	0.314	0.314	0.306	0.297	0.27	0.298	0.321	0.326
DDOD	0.273	0.267	0.275	0.262	0.28	0.256	0.267	0.256	0.264	0.262	0.245	0.251	0.262	0.273
DyHead	0.283	0.284	0.294	0.267	0.273	0.275	0.293	0.273	0.279	0.279	0.27	0.274	0.279	0.291
EfficientNet	0.29	0.283	0.295	0.283	0.284	0.277	0.291	0.278	0.284	0.28	0.274	0.272	0.276	0.291
FCOS	0.339	0.341	0.341	0.323	0.331	0.331	0.339	0.337	0.322	0.337	0.296	0.335	0.333	0.344
FoveaBox	0.304	0.298	0.316	0.309	0.29	0.295	0.293	0.285	0.301	0.292	0.276	0.279	0.294	0.309
FreeAnchor	0.289	0.289	0.29	0.283	0.278	0.28	0.284	0.276	0.278	0.277	0.249	0.28	0.284	0.292
FSAF	0.307	0.305	0.31	0.298	0.304	0.299	0.307	0.299	0.307	0.294	0.279	0.297	0.301	0.305
GFL	0.311	0.298	0.304	0.29	0.29	0.292	0.29	0.283	0.285	0.284	0.267	0.281	0.295	0.308
LD	0.302	0.293	0.298	0.283	0.28	0.285	0.291	0.278	0.286	0.283	0.263	0.276	0.29	0.297
NAS-FPN	0.294	0.28	0.293	0.265	0.278	0.27	0.285	0.271	0.272	0.271	0.242	0.272	0.278	0.288
PAA	0.321	0.322	0.323	0.308	0.303	0.312	0.32	0.312	0.313	0.303	0.269	0.304	0.318	0.33
RetinaNet	0.317	0.32	0.327	0.314	0.305	0.312	0.316	0.309	0.307	0.312	0.295	0.304	0.312	0.327
RTMDet	0.246	0.242	0.248	0.235	0.233	0.236	0.244	0.229	0.237	0.238	0.233	0.233	0.234	0.238
TOOD	0.277	0.276	0.278	0.262	0.258	0.271	0.277	0.266	0.274	0.269	0.249	0.265	0.268	0.278
VarifocalNet	0.295	0.287	0.296	0.281	0.277	0.28	0.286	0.271	0.28	0.277	0.262	0.274	0.284	0.294
YOLOv5	0.241	0.239	0.243	0.235	0.222	0.237	0.242	0.238	0.241	0.229	0.179	0.223	0.238	0.235
YOLOv6	0.241	0.237	0.24	0.232	0.233	0.231	0.236	0.232	0.238	0.231	0.228	0.228	0.232	0.233
YOLOv7	0.242	0.236	0.242	0.229	0.207	0.221	0.237	0.218	0.235	0.226	0.215	0.207	0.232	0.233
YOLOv8	0.244	0.241	0.244	0.235	0.235	0.237	0.241	0.238	0.242	0.237	0.232	0.233	0.239	0.238
YOLOX	0.242	0.241	0.243	0.233	0.217	0.237	0.241	0.237	0.239	0.236	0.227	0.234	0.239	0.237
Faster R-CNN	0.256	0.255	0.261	0.228	0.242	0.246	0.253	0.24	0.244	0.25	0.212	0.233	0.247	0.255
Cascade R-CNN	0.262	0.255	0.267	0.238	0.246	0.248	0.261	0.244	0.256	0.249	0.224	0.24	0.249	0.258
Cascade RPN	0.279	0.278	0.28	0.26	0.266	0.27	0.274	0.261	0.266	0.268	0.241	0.262	0.273	0.281
Double Heads	0.26	0.255	0.262	0.233	0.232	0.243	0.253	0.24	0.245	0.244	0.209	0.233	0.249	0.258
FPG	0.248	0.243	0.251	0.217	0.211	0.233	0.247	0.235	0.235	0.231	0.196	0.229	0.242	0.241
Grid R-CNN	0.25	0.244	0.251	0.235	0.24	0.243	0.25	0.241	0.248	0.246	0.234	0.238	0.24	0.244
Guided Anchoring	0.271	0.266	0.271	0.254	0.252	0.262	0.272	0.259	0.257	0.265	0.244	0.253	0.264	0.268
HRNet	0.253	0.248	0.256	0.243	0.233	0.246	0.25	0.246	0.243	0.252	0.239	0.234	0.238	0.254
Libra R-CNN	0.321	0.314	0.319	0.319	0.293	0.302	0.307	0.294	0.3	0.303	0.285	0.291	0.309	0.326
PAFPN	0.257	0.253	0.263	0.232	0.243	0.245	0.255	0.243	0.251	0.249	0.212	0.235	0.25	0.259
RepPoints	0.314	0.31	0.319	0.304	0.306	0.295	0.301	0.291	0.305	0.298	0.28	0.287	0.306	0.316
Res2Net	0.237	0.231	0.24	0.22	0.193	0.229	0.234	0.227	0.226	0.228	0.216	0.222	0.23	0.232
ResNeSt	0.244	0.248	0.236	0.224	0.226	0.236	0.236	0.231	0.217	0.226	0.214	0.221	0.241	0.245
SABL	0.271	0.26	0.27	0.24	0.251	0.254	0.256	0.247	0.257	0.252	0.228	0.245	0.26	0.265
Sparse R-CNN	0.268	0.262	0.267	0.25	0.253	0.251	0.263	0.246	0.247	0.253	0.247	0.244	0.259	0.265
DETR	0.42	0.42	0.391	0.334	0.352	0.399	0.39	0.399	0.351	0.383	0.328	0.392	0.404	0.397
Conditional DETR	0.293	0.302	0.29	0.278	0.281	0.285	0.28	0.268	0.269	0.268	0.236	0.275	0.278	0.286
DDQ	0.322	0.32	0.32	0.303	0.304	0.301	0.313	0.29	0.29	0.305	0.284	0.299	0.3	0.322
DAB-DETR	0.301	0.301	0.303	0.3	0.31	0.295	0.296	0.277	0.294	0.285	0.27	0.28	0.292	0.303
Deformable DETR	0.281	0.271	0.278	0.258	0.282	0.256	0.27	0.252	0.261	0.26	0.221	0.257	0.261	0.267
DINO	0.303	0.309	0.303	0.29	0.294	0.294	0.302	0.287	0.276	0.293	0.276	0.293	0.295	0.299
PVT	0.273	0.275	0.27	0.264	0.253	0.268	0.271	0.26	0.254	0.271	0.266	0.267	0.266	0.275
PVTv2	0.284	0.283	0.277	0.278	0.277	0.271	0.281	0.276	0.266	0.279	0.264	0.273	0.273	0.28

TABLE XIV: **Ablation** experimental results (**weather**) of **vehicle** detection in the metric of **mAR50(%)**.

	Clean	Random	ACTIVE	DTA	FCA	APPA	POCPatch	3D2Fool	CAMOU	RPAU
ATSS	0.975	0.963	0.863	0.963	0.95	1.0	0.863	0.963	0.887	0.875
AutoAssign	0.975	1.0	0.975	1.0	1.0	1.0	0.975	0.988	1.0	0.988
CenterNet	0.95	1.0	0.963	1.0	0.975	1.0	0.887	1.0	1.0	0.925
CentripetalNet	0.975	1.0	0.938	1.0	0.95	1.0	0.975	1.0	0.988	0.988
CornerNet	1.0	1.0	0.95	1.0	1.0	1.0	1.0	1.0	1.0	0.95
DDOD	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.988
DyHead	0.988	1.0	1.0	1.0	1.0	1.0	1.0	0.988	0.975	1.0
EfficientNet	0.938	1.0	1.0	1.0	1.0	0.988	1.0	1.0	1.0	0.975
FCOS	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.988	1.0	0.975
FoveaBox	1.0	1.0	0.95	1.0	1.0	1.0	0.988	1.0	1.0	1.0
FreeAnchor	1.0	0.863	0.812	0.975	0.863	1.0	0.8	0.875	0.875	0.887
FSAF	0.988	0.887	0.887	0.988	0.925	1.0	0.925	0.975	0.9	0.912
GFL	0.988	0.875	0.875	1.0	0.912	1.0	0.887	0.875	0.875	0.875
LD	1.0	0.9	0.887	1.0	0.975	1.0	0.887	1.0	0.9	0.975
NAS-FPN	1.0	0.912	1.0	1.0	0.963	1.0	0.975	1.0	0.975	0.963
PAA	0.988	1.0	0.988	1.0	0.975	1.0	0.988	1.0	1.0	1.0
RetinaNet	0.988	0.938	0.925	0.963	0.925	1.0	0.863	0.988	0.863	0.925
RTMDet	1.0	1.0	1.0	1.0	1.0	0.988	1.0	1.0	1.0	1.0
TOOD	1.0	1.0	0.912	1.0	1.0	1.0	0.988	1.0	1.0	1.0
VarifocalNet	0.988	0.938	0.85	0.887	0.95	1.0	0.925	0.85	0.875	0.887
YOLOv5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
YOLOv6	1.0	1.0	1.0	1.0	1.0	0.963	1.0	1.0	1.0	1.0
YOLOv7	1.0	1.0	1.0	1.0	1.0	0.925	1.0	1.0	1.0	1.0
YOLOv8	1.0	1.0	1.0	1.0	1.0	0.975	1.0	1.0	1.0	1.0
YOLOX	1.0	1.0	1.0	1.0	1.0	0.988	1.0	1.0	1.0	1.0
Faster R-CNN	0.988	0.85	0.562	0.775	0.85	0.938	0.762	0.887	0.725	0.875
Cascade R-CNN	1.0	0.863	0.738	0.875	0.912	0.975	0.787	0.925	0.787	0.863
Cascade RPN	1.0	1.0	0.988	1.0	0.988	0.975	0.988	1.0	0.988	0.988
Double Heads	0.975	0.9	0.812	0.875	0.825	1.0	0.838	0.975	0.838	0.875
FPG	0.988	0.988	0.95	1.0	0.988	1.0	0.9	0.988	0.938	0.925
Grid R-CNN	1.0	0.863	0.738	0.963	0.85	0.988	0.838	1.0	0.875	0.875
Guided Anchoring	1.0	1.0	1.0	1.0	1.0	0.988	1.0	1.0	1.0	0.975
HRNet	0.938	0.875	0.938	0.95	0.875	0.963	0.775	0.9	0.95	0.925
Libra R-CNN	0.988	0.938	0.938	0.963	0.963	1.0	0.975	1.0	0.9	0.975
PAFPN	0.988	0.863	0.637	0.825	0.887	0.963	0.838	0.963	0.825	0.863
RepPoints	0.988	0.975	0.875	1.0	0.9	1.0	0.975	0.975	0.975	0.912
Res2Net	0.975	1.0	0.925	0.988	0.988	1.0	0.912	0.938	0.988	1.0
ResNeSt	0.975	0.863	0.825	0.988	1.0	1.0	1.0	1.0	0.863	0.963
SABL	1.0	0.875	0.775	0.938	0.887	0.988	0.787	0.95	0.85	0.863
Sparse R-CNN	0.975	0.912	0.85	1.0	0.988	0.988	0.95	0.95	0.838	0.9
DETR	0.95	0.537	0.388	0.237	0.8	0.713	0.812	0.5	0.487	0.8
Conditional DETR	0.988	0.887	0.975	0.975	0.975	1.0	1.0	1.0	0.875	0.912
DDQ	0.988	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
DAB-DETR	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.975	1.0
Deformable DETR	0.963	0.975	0.9	0.938	0.95	0.975	0.9	0.975	0.912	0.938
DINO	1.0	1.0	0.988	0.925	0.975	0.988	1.0	1.0	0.9	1.0
PVT	0.988	1.0	0.75	0.988	1.0	1.0	0.988	1.0	1.0	0.925
PVTv2	1.0	0.988	0.988	1.0	1.0	1.0	0.887	1.0	1.0	0.912

TABLE XV: **Ablation** experimental results (**spot**) of **vehicle** detection in the metric of **mAR50(%)**.

	Clean	Random	ACTIVE	DTA	FCA	APPA	POCPatch	3D2Fool	CAMOU	RPAU
ATSS	1.0	0.979	0.885	0.99	0.969	0.99	0.844	0.969	0.875	0.875
AutoAssign	0.958	0.99	0.99	0.99	0.99	0.99	0.969	0.99	0.99	0.99
CenterNet	0.99	0.979	0.99	0.99	0.979	0.99	0.885	0.979	0.99	0.938
CentripetalNet	1.0	0.99	0.979	0.99	0.99	0.99	0.979	1.0	0.99	0.979
CornerNet	0.99	0.99	0.958	0.99	0.99	0.99	0.979	1.0	0.969	0.99
DDOD	1.0	1.0	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
DyHead	1.0	1.0	0.979	1.0	0.99	1.0	1.0	0.979	0.969	1.0
EfficientNet	0.958	0.99	0.969	0.99	0.99	0.99	0.979	1.0	0.99	0.979
FCOS	1.0	0.99	0.99	0.99	0.99	0.99	0.979	0.99	0.99	0.958
FoveaBox	0.99	0.99	0.875	0.99	0.99	0.99	0.99	0.99	0.99	0.979
FreeAnchor	0.99	0.865	0.833	0.979	0.885	0.99	0.792	0.938	0.885	0.865
FSAF	1.0	0.906	0.812	0.99	0.938	0.99	0.865	0.938	0.875	0.885
GFL	0.99	0.927	0.969	0.99	0.938	0.99	0.885	0.927	0.917	0.896
LD	0.99	0.948	0.896	0.99	0.958	0.99	0.823	0.979	0.896	0.99
NAS-FPN	1.0	0.885	0.969	0.979	0.948	1.0	0.979	1.0	0.927	0.917
PAA	0.99	0.99	0.979	0.99	0.99	0.99	0.958	0.99	0.99	0.99
RetinaNet	0.99	0.865	0.865	0.917	0.948	0.979	0.854	0.917	0.802	0.875
RTMDet	1.0	1.0	0.99	1.0	0.99	0.979	0.99	0.99	1.0	0.99
TOOD	0.99	0.99	0.927	0.99	0.99	0.99	0.99	0.99	0.99	0.99
VarifocalNet	1.0	0.958	0.802	0.865	0.938	0.99	0.885	0.823	0.833	0.875
YOLOv5	1.0	0.99	0.99	1.0	0.99	1.0	1.0	0.99	0.99	0.969
YOLOv6	1.0	1.0	1.0	1.0	1.0	0.979	1.0	1.0	1.0	1.0
YOLOv7	1.0	1.0	1.0	0.99	0.99	0.917	1.0	1.0	0.99	0.99
YOLOv8	1.0	1.0	1.0	1.0	1.0	0.958	1.0	1.0	0.99	1.0
YOLOX	1.0	0.99	1.0	0.99	0.99	0.958	0.99	1.0	0.99	0.99
Faster R-CNN	0.99	0.865	0.74	0.969	0.885	0.99	0.792	0.948	0.833	0.885
Cascade R-CNN	0.99	0.865	0.854	0.969	0.885	0.99	0.823	0.917	0.844	0.854
Cascade RPN	0.99	0.958	0.938	0.99	0.938	0.969	0.906	0.99	0.99	0.958
Double Heads	0.99	0.896	0.927	0.969	0.823	0.99	0.865	0.99	0.844	0.865
FPG	1.0	0.948	0.896	0.979	0.958	0.99	0.833	0.99	0.969	0.906
Grid R-CNN	0.979	0.906	0.885	0.979	0.917	0.979	0.875	0.99	0.917	0.896
Guided Anchoring	1.0	0.979	0.99	0.99	0.99	0.917	0.979	0.99	0.99	0.979
HRNet	0.969	0.906	0.958	0.958	0.854	0.969	0.812	0.958	0.969	0.927
Libra R-CNN	1.0	0.958	0.917	0.99	0.948	0.99	0.948	0.99	0.854	0.948
PAFPN	0.99	0.875	0.74	0.938	0.938	0.969	0.927	0.979	0.844	0.844
RepPoints	0.99	0.979	0.823	0.99	0.885	0.99	0.948	0.979	0.927	0.896
Res2Net	0.979	0.979	0.979	0.99	0.99	0.99	0.927	0.99	0.979	0.979
ResNeSt	0.958	0.865	0.938	0.979	0.99	0.99	0.99	0.979	0.917	0.969
SABL	0.99	0.917	0.854	0.99	0.875	0.99	0.833	0.99	0.865	0.865
Sparse R-CNN	0.99	0.969	0.917	0.99	0.99	0.99	0.958	0.979	0.875	0.948
DETR	0.99	0.677	0.448	0.302	0.885	0.875	0.854	0.74	0.594	0.844
Conditional DETR	0.99	0.958	0.938	0.979	0.979	0.948	0.969	0.99	0.906	0.896
DDQ	1.0	0.99	1.0	1.0	0.99	1.0	1.0	0.99	1.0	1.0
DAB-DETR	1.0	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Deformable DETR	1.0	0.948	0.875	0.917	0.917	0.969	0.917	0.979	0.979	0.938
DINO	1.0	0.99	0.948	0.938	0.979	0.979	0.979	0.99	0.927	0.99
PVT	0.938	0.979	0.625	0.979	0.969	0.99	0.927	0.99	0.99	0.812
PVTv2	1.0	0.979	0.99	0.99	0.979	0.99	0.802	0.99	0.99	0.875

TABLE XVI: **Ablation** experimental results (**distance**) of **vehicle** detection in the metric of **mAR50(%)**.

	Clean	Random	ACTIVE	DTA	FCA	APPA	POCPatch	3D2Fool	CAMOU	RPAU
ATSS	0.979	0.979	0.958	0.99	0.99	0.99	0.948	0.979	0.979	0.958
AutoAssign	0.969	0.99	0.99	0.979	0.979	0.99	0.958	0.99	0.979	0.99
CenterNet	0.99	0.979	1.0	0.99	0.979	1.0	0.99	1.0	1.0	0.969
CentripetalNet	0.979	0.99	0.969	0.99	0.99	0.99	0.99	0.99	0.99	0.99
CornerNet	1.0	0.99	0.99	0.99	0.99	0.979	0.99	0.99	0.99	0.99
DDOD	0.99	0.979	0.958	0.99	0.979	0.99	0.969	0.99	0.99	0.979
DyHead	0.99	0.99	0.99	0.99	0.99	0.979	0.99	0.99	0.979	0.99
EfficientNet	0.969	0.99	0.99	0.979	0.979	0.917	0.99	0.99	0.99	0.979
FCOS	0.99	0.99	0.99	0.979	0.99	0.99	0.969	0.99	0.99	0.969
FoveaBox	0.99	0.99	0.979	0.99	0.99	0.99	0.99	0.99	0.99	0.979
FreeAnchor	0.979	0.917	0.865	0.979	0.927	0.99	0.802	0.969	0.802	0.917
FSAF	1.0	0.917	0.885	0.99	0.958	0.979	0.896	0.99	0.917	0.906
GFL	0.979	0.969	0.948	0.99	0.969	0.979	0.958	0.99	0.896	0.969
LD	0.979	0.969	0.969	0.99	0.979	0.99	0.958	0.99	0.969	0.979
NAS-FPN	0.99	0.969	0.979	0.979	0.979	0.99	0.99	0.99	0.948	0.958
PAA	0.99	1.0	0.979	0.99	0.99	0.979	0.969	0.99	0.99	0.99
RetinaNet	0.979	0.917	0.917	0.969	0.979	0.979	0.875	0.979	0.875	0.927
RTMDet	0.99	0.99	1.0	0.938	0.99	0.958	1.0	1.0	0.99	0.99
TOOD	0.99	0.99	0.948	0.979	0.979	0.99	0.99	0.99	0.99	0.979
VarifocalNet	0.979	0.938	0.854	0.958	0.969	0.979	0.906	0.979	0.833	0.896
YOLOv5	0.99	1.0	1.0	0.99	1.0	0.979	0.99	1.0	0.99	1.0
YOLOv6	0.99	0.99	0.99	0.969	0.99	0.99	0.99	0.99	0.979	0.99
YOLOv7	0.979	0.99	0.99	0.99	0.99	0.969	0.99	0.99	0.99	0.99
YOLOv8	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
YOLOX	0.979	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Faster R-CNN	0.979	0.885	0.76	0.99	0.958	0.979	0.802	0.958	0.875	0.896
Cascade R-CNN	0.99	0.906	0.823	0.979	0.948	0.99	0.792	0.958	0.875	0.896
Cascade RPN	0.99	0.979	0.958	0.979	0.979	0.99	0.927	0.99	0.979	0.958
Double Heads	0.99	0.948	0.885	0.99	0.896	0.99	0.875	0.969	0.927	0.885
FPG	0.979	0.99	0.979	0.979	0.979	0.979	0.906	0.99	0.958	0.958
Grid R-CNN	0.969	0.958	0.917	0.979	0.99	0.979	0.885	0.99	0.979	0.896
Guided Anchoring	0.99	0.99	0.99	0.99	0.99	0.958	0.99	0.99	0.99	0.979
HRNet	0.979	0.927	0.875	0.979	0.938	0.979	0.885	0.969	0.99	0.969
Libra R-CNN	0.99	0.969	0.969	0.979	0.979	0.979	0.969	0.99	0.958	0.938
PAFPN	0.979	0.875	0.781	0.969	0.938	0.979	0.823	0.958	0.885	0.885
RepPoints	0.99	0.958	0.896	0.99	0.948	0.99	0.979	1.0	1.0	0.938
Res2Net	0.979	0.99	0.979	0.99	0.979	0.979	0.979	0.99	0.99	0.979
ResNeSt	0.979	0.927	0.958	0.969	0.969	0.979	0.979	0.969	0.917	0.969
SABL	0.99	0.885	0.844	0.979	0.938	0.979	0.781	0.948	0.958	0.896
Sparse R-CNN	0.979	0.979	0.865	0.99	0.99	0.979	0.938	0.99	0.906	0.979
DETR	0.979	0.771	0.615	0.604	0.896	0.708	0.885	0.635	0.625	0.906
Conditional DETR	0.979	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
DDQ	0.979	0.99	0.99	0.99	0.979	0.979	0.99	0.99	0.99	0.99
DAB-DETR	0.99	0.99	0.99	0.99	0.99	0.99	1.0	0.979	0.99	0.99
Deformable DETR	0.99	0.99	0.969	0.979	0.969	0.969	0.948	1.0	0.979	0.979
DINO	0.99	0.948	0.979	0.99	0.969	0.99	0.99	0.979	0.979	0.99
PVT	0.958	0.979	0.792	0.958	0.969	0.979	0.875	0.979	0.969	0.958
PVTv2	0.979	0.99	0.917	1.0	1.0	0.99	0.875	0.99	0.99	0.948

TABLE XVII: **Ablation** experimental results (ϕ) of **vehicle** detection in the metric of **mAR50(%)**.

	Clean	Random	ACTIVE	DTA	FCA	APPA	POPatch	3D2Fool	CAMOU	RPAU
ATSS	1.0	0.98	0.86	0.98	0.99	1.0	0.92	1.0	0.9	0.91
AutoAssign	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.99
CenterNet	0.99	1.0	0.98	1.0	1.0	1.0	0.98	1.0	0.99	0.98
CentripetalNet	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.99	1.0
CornerNet	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
DDOD	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
DyHead	1.0	1.0	0.99	1.0	1.0	1.0	1.0	1.0	1.0	1.0
EfficientNet	1.0	1.0	0.98	1.0	1.0	1.0	1.0	1.0	1.0	1.0
FCOS	1.0	1.0	1.0	1.0	1.0	0.99	1.0	1.0	1.0	1.0
FoveaBox	1.0	1.0	0.95	1.0	1.0	1.0	1.0	1.0	1.0	1.0
FreeAnchor	1.0	0.87	0.93	0.96	0.89	1.0	0.89	1.0	0.98	0.9
FSAF	1.0	0.98	0.94	0.99	0.96	1.0	0.96	1.0	0.95	0.91
GFL	1.0	0.95	0.93	0.99	1.0	1.0	0.95	1.0	0.86	0.9
LD	0.99	1.0	1.0	1.0	1.0	1.0	0.98	1.0	0.93	1.0
NAS-FPN	1.0	1.0	0.98	0.99	0.98	1.0	0.98	1.0	0.99	0.97
PAA	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
RetinaNet	1.0	0.97	0.98	0.99	1.0	1.0	0.96	1.0	0.91	0.89
RTMDet	1.0	1.0	1.0	0.97	1.0	0.97	1.0	1.0	1.0	1.0
TOOD	1.0	1.0	0.96	1.0	1.0	1.0	1.0	1.0	1.0	1.0
VarifocalNet	0.99	0.94	0.97	0.98	0.99	1.0	0.94	1.0	0.88	0.91
YOLOv5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
YOLOv6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
YOLOv7	1.0	1.0	1.0	1.0	1.0	0.97	1.0	1.0	1.0	1.0
YOLOv8	1.0	1.0	1.0	1.0	1.0	0.99	1.0	1.0	1.0	1.0
YOLOX	1.0	1.0	1.0	1.0	1.0	0.98	1.0	1.0	1.0	1.0
Faster R-CNN	1.0	0.86	0.77	0.95	0.91	1.0	0.83	1.0	0.83	0.85
Cascade R-CNN	1.0	0.89	0.91	0.99	0.97	1.0	0.93	0.99	0.88	0.9
Cascade RPN	1.0	1.0	0.97	1.0	1.0	0.99	0.96	1.0	1.0	0.93
Double Heads	1.0	1.0	0.95	1.0	0.97	1.0	0.95	1.0	0.92	0.91
FPG	1.0	1.0	0.96	0.99	1.0	0.99	0.97	1.0	1.0	0.99
Grid R-CNN	0.98	0.97	0.89	1.0	0.95	0.99	0.95	1.0	0.94	0.9
Guided Anchoring	1.0	1.0	1.0	1.0	1.0	0.99	1.0	1.0	1.0	1.0
HRNet	1.0	0.98	1.0	0.99	0.94	0.99	0.84	1.0	0.98	0.95
Libra R-CNN	1.0	1.0	0.91	1.0	0.97	1.0	1.0	1.0	0.96	0.96
PAFPN	1.0	0.93	0.83	0.97	0.99	1.0	0.95	1.0	0.9	0.9
RepPoints	1.0	1.0	0.89	1.0	0.95	1.0	1.0	1.0	0.96	0.95
Res2Net	0.98	1.0	1.0	0.99	1.0	1.0	1.0	1.0	0.93	1.0
ResNeSt	0.96	0.85	0.94	0.99	1.0	0.96	0.99	0.99	0.9	0.99
SABL	1.0	1.0	0.89	0.99	0.96	1.0	0.92	1.0	0.89	0.89
Sparse R-CNN	1.0	1.0	0.88	0.99	1.0	0.98	1.0	1.0	0.92	0.95
DETR	1.0	0.84	0.72	0.49	0.93	0.93	0.96	0.87	0.81	0.94
Conditional DETR	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
DDQ	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
DAB-DETR	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Deformable DETR	0.99	1.0	0.99	1.0	0.99	1.0	0.99	1.0	0.98	1.0
DINO	1.0	1.0	0.99	1.0	1.0	1.0	1.0	1.0	0.98	1.0
PVT	1.0	0.98	0.7	0.98	0.96	1.0	0.92	1.0	0.96	0.85
PVTv2	1.0	1.0	1.0	1.0	1.0	1.0	0.82	1.0	1.0	0.97

TABLE XVIII: **Ablation** experimental results (θ) of **vehicle** detection in the metric of **mAR50(%)**.

	Clean	Random	ACTIVE	DTA	FCA	APPA	POPatch	3D2Fool	CAMOU	RPAU
ATSS	0.98	0.96	0.47	0.73	0.94	0.85	0.77	1.0	0.76	0.7
AutoAssign	1.0	0.86	0.74	0.92	0.91	0.97	0.71	0.93	0.87	0.98
CenterNet	1.0	1.0	0.62	0.98	0.98	0.99	0.71	0.91	0.92	1.0
CentripetalNet	1.0	1.0	0.78	1.0	0.96	1.0	0.74	1.0	0.88	0.98
CornerNet	0.99	0.79	0.49	0.75	0.93	0.78	0.51	0.95	0.66	0.6
DDOD	0.99	0.97	0.61	0.99	0.85	0.98	1.0	1.0	0.96	0.97
DyHead	1.0	0.81	0.52	0.53	0.98	0.81	0.61	0.55	0.62	0.93
EfficientNet	1.0	0.79	0.81	1.0	0.84	1.0	0.78	1.0	0.97	1.0
FCOS	1.0	0.94	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
FoveaBox	1.0	1.0	0.49	0.96	0.87	0.99	0.67	0.87	0.96	0.92
FreeAnchor	1.0	0.82	1.0	0.98	0.9	1.0	0.99	0.98	0.88	1.0
FSAF	1.0	1.0	0.64	1.0	0.82	1.0	0.99	1.0	1.0	1.0
GFL	1.0	0.97	0.45	0.95	0.93	0.96	0.94	0.98	0.73	1.0
LD	1.0	0.99	0.56	1.0	1.0	0.93	0.98	0.88	0.96	1.0
NAS-FPN	1.0	0.92	0.52	0.86	1.0	0.79	0.72	0.94	0.65	0.92
PAA	1.0	1.0	0.97	0.99	0.94	1.0	0.97	1.0	1.0	1.0
RetinaNet	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
RTMDet	1.0	1.0	0.99	0.98	1.0	1.0	1.0	1.0	1.0	0.88
TOOD	0.83	0.85	0.52	0.87	0.76	0.84	0.8	0.97	0.96	0.8
VarifocalNet	1.0	0.77	0.48	0.55	0.86	0.83	0.64	0.62	0.61	0.65
YOLOv5	1.0	1.0	0.74	1.0	1.0	1.0	0.92	1.0	1.0	1.0
YOLOv6	1.0	1.0	0.73	1.0	1.0	1.0	1.0	1.0	1.0	1.0
YOLOv7	0.98	0.76	0.78	0.72	0.91	1.0	0.78	1.0	0.75	0.79
YOLOv8	1.0	0.92	0.64	0.82	1.0	0.82	0.64	1.0	0.83	0.97
YOLOX	1.0	0.9	0.62	0.92	0.95	0.99	0.88	1.0	0.97	0.98
Faster R-CNN	0.85	0.49	0.46	0.52	0.66	0.73	0.53	0.63	0.55	0.49
Cascade R-CNN	0.75	0.64	0.47	0.54	0.67	0.67	0.62	0.65	0.55	0.6
Cascade RPN	1.0	0.92	0.51	0.79	0.88	1.0	0.96	1.0	0.85	0.82
Double Heads	0.76	0.72	0.46	0.57	0.68	0.71	0.63	0.87	0.6	0.62
FPG	1.0	0.89	0.5	0.93	0.97	0.9	0.72	0.91	0.96	0.53
Grid R-CNN	0.91	0.68	0.48	0.68	0.71	0.77	0.63	0.85	0.6	0.56
Guided Anchoring	1.0	1.0	0.87	0.99	1.0	1.0	0.8	1.0	1.0	0.9
HRNet	0.96	0.78	0.55	0.59	0.73	0.83	0.62	0.81	0.62	0.62
Libra R-CNN	1.0	1.0	0.67	1.0	1.0	1.0	1.0	1.0	1.0	1.0
PAFPN	0.74	0.63	0.4	0.51	0.63	0.71	0.62	0.66	0.57	0.44
RepPoints	1.0	1.0	0.63	1.0	0.94	0.77	1.0	1.0	1.0	0.9
Res2Net	0.9	0.64	0.57	0.55	0.73	0.65	0.66	0.77	0.66	0.72
ResNeSt	0.97	0.61	0.5	0.73	0.72	0.96	0.61	0.93	0.61	0.58
SABL	0.77	0.53	0.45	0.54	0.66	0.66	0.61	0.66	0.56	0.47
Sparse R-CNN	1.0	0.96	0.63	0.91	1.0	0.72	0.8	0.78	0.87	0.85
DETR	0.77	0.56	0.39	0.45	0.71	0.64	0.51	0.55	0.55	0.57
Conditional DETR	1.0	0.88	0.67	0.84	1.0	0.8	0.86	1.0	0.84	0.95
DDQ	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
DAB-DETR	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.99	1.0	0.98
Deformable DETR	1.0	1.0	0.94	0.99	1.0	1.0	1.0	1.0	1.0	1.0
DINO	1.0	1.0	0.93	1.0	1.0	1.0	1.0	1.0	1.0	1.0
PVT	1.0	1.0	0.95	1.0	1.0	1.0	1.0	1.0	1.0	0.96
PVTv2	1.0	0.93	0.64	0.94	0.87	1.0	0.76	1.0	0.76	0.6

TABLE XIX: **Ablation** experimental results (**sphere**) of **vehicle** detection in the metric of **mAR50(%)**.

	Clean	Random	ACTIVE	DTA	FCA	APPA	POPatch	3D2Fool	CAMOU	RPAU
ATSS	0.97	0.71	0.45	0.79	0.84	0.85	0.68	0.98	0.71	0.73
AutoAssign	0.98	0.9	0.74	0.92	0.94	1.0	0.83	0.98	0.91	0.99
CenterNet	0.98	0.91	0.58	0.91	0.92	0.89	0.93	0.9	0.88	0.94
CentripetalNet	0.98	0.78	0.66	0.79	0.9	0.86	0.7	0.85	0.71	0.86
CornerNet	0.96	0.75	0.61	0.87	0.93	0.86	0.7	0.94	0.76	0.82
DDOD	0.99	0.96	0.74	0.99	0.86	1.0	0.82	1.0	0.97	0.94
DyHead	1.0	0.79	0.53	0.79	0.92	0.8	0.79	0.73	0.77	0.99
EfficientNet	1.0	0.94	0.86	0.99	0.98	1.0	0.98	0.98	0.96	0.93
FCOS	1.0	0.99	0.99	1.0	1.0	1.0	1.0	1.0	1.0	1.0
FoveaBox	0.94	0.84	0.5	0.76	0.87	0.81	0.72	0.8	0.76	0.91
FreeAnchor	0.97	0.73	0.79	0.88	0.81	0.98	0.74	0.85	0.84	0.88
FSAF	0.93	0.75	0.49	0.8	0.72	0.81	0.67	0.89	0.74	0.84
GFL	1.0	0.8	0.48	0.9	0.95	0.95	0.89	0.92	0.81	0.9
LD	0.98	0.86	0.52	0.9	0.92	0.87	0.75	0.94	0.85	0.93
NAS-FPN	1.0	0.92	0.66	0.96	0.98	0.97	0.73	0.82	0.75	0.88
PAA	0.99	0.96	0.9	0.97	0.93	0.99	0.86	0.99	1.0	0.95
RetinaNet	1.0	0.81	0.67	0.85	0.89	0.86	0.81	0.84	0.83	0.8
RTMDet	1.0	1.0	0.93	0.99	1.0	0.98	0.99	1.0	1.0	0.98
TOOD	0.9	0.72	0.48	0.9	0.73	0.79	0.72	0.84	0.74	0.86
VarifocalNet	0.99	0.75	0.45	0.66	0.87	0.83	0.67	0.8	0.67	0.79
YOLOv5	1.0	1.0	0.98	1.0	1.0	1.0	1.0	1.0	0.97	0.99
YOLOv6	1.0	1.0	0.97	1.0	1.0	1.0	1.0	1.0	0.99	1.0
YOLOv7	1.0	0.94	0.97	0.94	0.98	1.0	0.96	1.0	0.98	0.99
YOLOv8	1.0	0.95	0.87	0.93	1.0	0.98	0.89	0.99	0.84	0.99
YOLOX	1.0	0.97	0.77	0.97	0.99	1.0	0.98	0.92	0.94	0.98
Faster R-CNN	0.85	0.45	0.35	0.47	0.54	0.59	0.49	0.61	0.44	0.49
Cascade R-CNN	0.82	0.51	0.4	0.52	0.58	0.64	0.53	0.62	0.48	0.51
Cascade RPN	1.0	0.91	0.62	0.93	0.96	0.99	0.96	1.0	0.86	0.96
Double Heads	0.83	0.61	0.44	0.58	0.58	0.72	0.54	0.8	0.57	0.57
FPG	1.0	0.78	0.5	0.93	0.82	0.99	0.7	0.94	0.9	0.71
Grid R-CNN	0.89	0.51	0.37	0.6	0.58	0.73	0.55	0.76	0.47	0.55
Guided Anchoring	1.0	0.97	0.98	0.99	0.99	1.0	0.89	1.0	0.97	0.97
HRNet	0.89	0.6	0.52	0.59	0.54	0.73	0.51	0.84	0.56	0.6
Libra R-CNN	0.94	0.76	0.49	0.8	0.9	0.74	0.85	0.81	0.72	0.93
PAFPN	0.84	0.55	0.37	0.52	0.59	0.65	0.58	0.71	0.5	0.54
RepPoints	0.97	0.89	0.52	0.84	0.83	0.84	0.8	0.95	0.94	0.84
Res2Net	0.88	0.59	0.52	0.66	0.77	0.83	0.54	0.79	0.55	0.73
ResNeSt	0.91	0.63	0.43	0.8	0.78	0.91	0.68	0.92	0.64	0.62
SABL	0.82	0.49	0.4	0.52	0.56	0.57	0.53	0.71	0.46	0.49
Sparse R-CNN	0.99	0.89	0.55	0.79	0.95	0.81	0.83	0.9	0.82	0.86
DETR	0.71	0.44	0.3	0.32	0.57	0.52	0.63	0.43	0.38	0.5
Conditional DETR	1.0	0.9	0.71	0.94	1.0	0.92	0.8	0.98	0.9	0.94
DDQ	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
DAB-DETR	1.0	0.94	1.0	0.99	0.99	1.0	0.98	0.96	1.0	0.99
Deformable DETR	1.0	0.9	0.74	0.86	0.87	0.93	0.95	0.96	0.86	0.91
DINO	1.0	0.92	0.95	0.99	0.96	1.0	1.0	1.0	0.92	1.0
PVT	0.98	0.95	0.78	0.87	0.96	0.97	0.91	0.94	0.87	0.8
PVTv2	1.0	0.99	0.77	0.97	1.0	1.0	0.78	0.99	0.98	0.8

TABLE XX: **Ablation** experimental results (**distance**) of **Traffic sign** detection in the metric of **mAR50(%)**.

	Clean	AdvCam	RP ₂	ShapeShifter
ATSS	0.929	0.93	0.892	0.919
AutoAssign	0.921	0.943	0.896	0.915
CenterNet	0.903	0.91	0.875	0.914
CentripetalNet	0.951	0.946	0.924	0.951
CornerNet	0.942	0.951	0.929	0.947
DDOD	0.916	0.926	0.9	0.915
DyHead	0.927	0.933	0.866	0.921
EfficientNet	0.921	0.913	0.887	0.919
FCOS	0.939	0.932	0.913	0.929
FoveaBox	0.915	0.917	0.871	0.913
FreeAnchor	0.921	0.919	0.877	0.912
FSAF	0.901	0.907	0.864	0.899
GFL	0.927	0.942	0.887	0.908
LD	0.933	0.931	0.88	0.92
NAS-FPN	0.93	0.942	0.883	0.925
PAA	0.928	0.921	0.886	0.91
RetinaNet	0.921	0.912	0.863	0.899
RTMDet	0.929	0.942	0.862	0.927
TOOD	0.922	0.93	0.895	0.921
VarifocalNet	0.929	0.932	0.902	0.921
YOLOv5	0.941	0.943	0.882	0.945
YOLOv6	0.94	0.954	0.901	0.936
YOLOv7	0.945	0.942	0.89	0.942
YOLOv8	0.942	0.943	0.868	0.944
YOLOX	0.923	0.922	0.866	0.906
Faster R-CNN	0.891	0.897	0.863	0.861
Cascade R-CNN	0.929	0.924	0.891	0.895
Cascade RPN	0.927	0.93	0.887	0.901
Double Heads	0.887	0.895	0.847	0.88
FPG	0.921	0.935	0.859	0.897
Grid R-CNN	0.913	0.911	0.865	0.91
Guided Anchoring	0.928	0.921	0.899	0.925
HRNet	0.923	0.915	0.91	0.904
Libra R-CNN	0.921	0.922	0.885	0.905
PAFPN	0.901	0.89	0.85	0.882
RepPoints	0.915	0.908	0.865	0.887
Res2Net	0.91	0.897	0.861	0.899
ResNeSt	0.929	0.901	0.872	0.889
SABL	0.92	0.913	0.88	0.898
Sparse R-CNN	0.931	0.925	0.9	0.927
DETR	0.908	0.904	0.878	0.933
Conditional DETR	0.93	0.931	0.907	0.924
DDQ	0.949	0.95	0.901	0.932
DAB-DETR	0.931	0.94	0.902	0.919
Deformable DETR	0.943	0.955	0.893	0.933
DINO	0.94	0.944	0.885	0.936
PVT	0.906	0.886	0.868	0.861
PVTv2	0.915	0.899	0.877	0.905

TABLE XXI: Ablation study on training dataset.

Physical attacks	Training datasets	Median ASR
AdvCam	ImageNet	0
AdvCaT	376 self-collected images	0
MTD	-	2
LAP	INRIA	2
AdvPattern	Market1501	2
AdvTshirt	40 self-collected videos	3
DAP	INRIA	5
NaTPatch	INRIA	5
InvisCloak	COCO	5
AdvTexture	INRIA	7

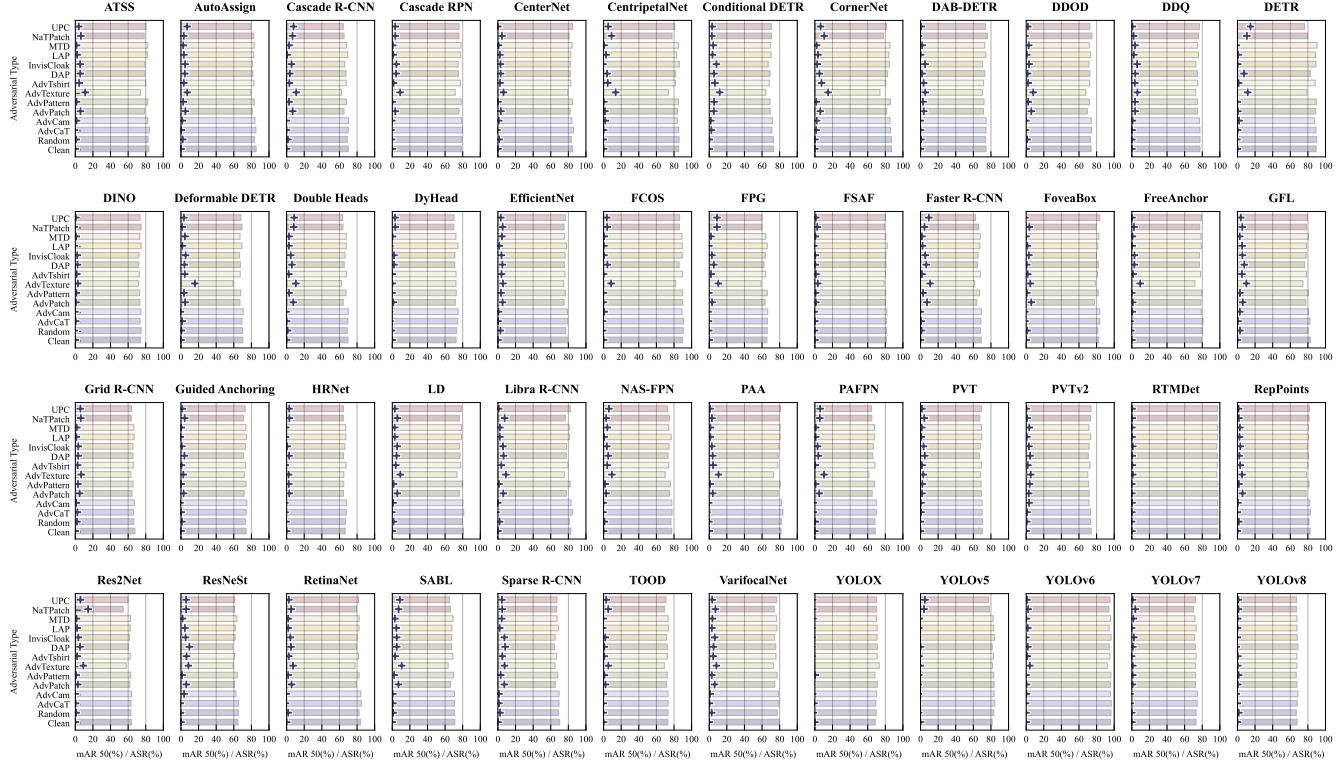


Fig. 17: Overall experimental results of person detection in the metric of mAR50(%), please zoom in for better view.

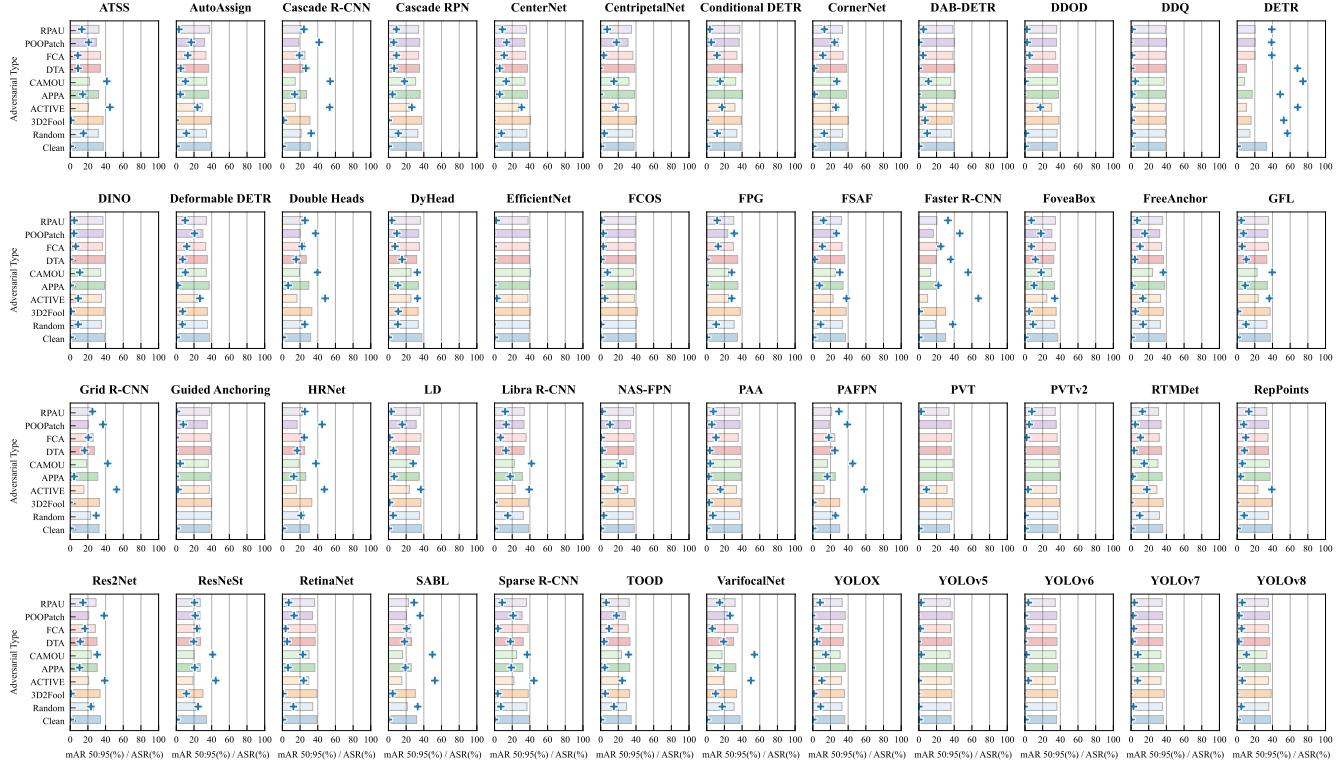


Fig. 18: Overall experimental results of vehicle detection in the metric of mAR50:95(%).

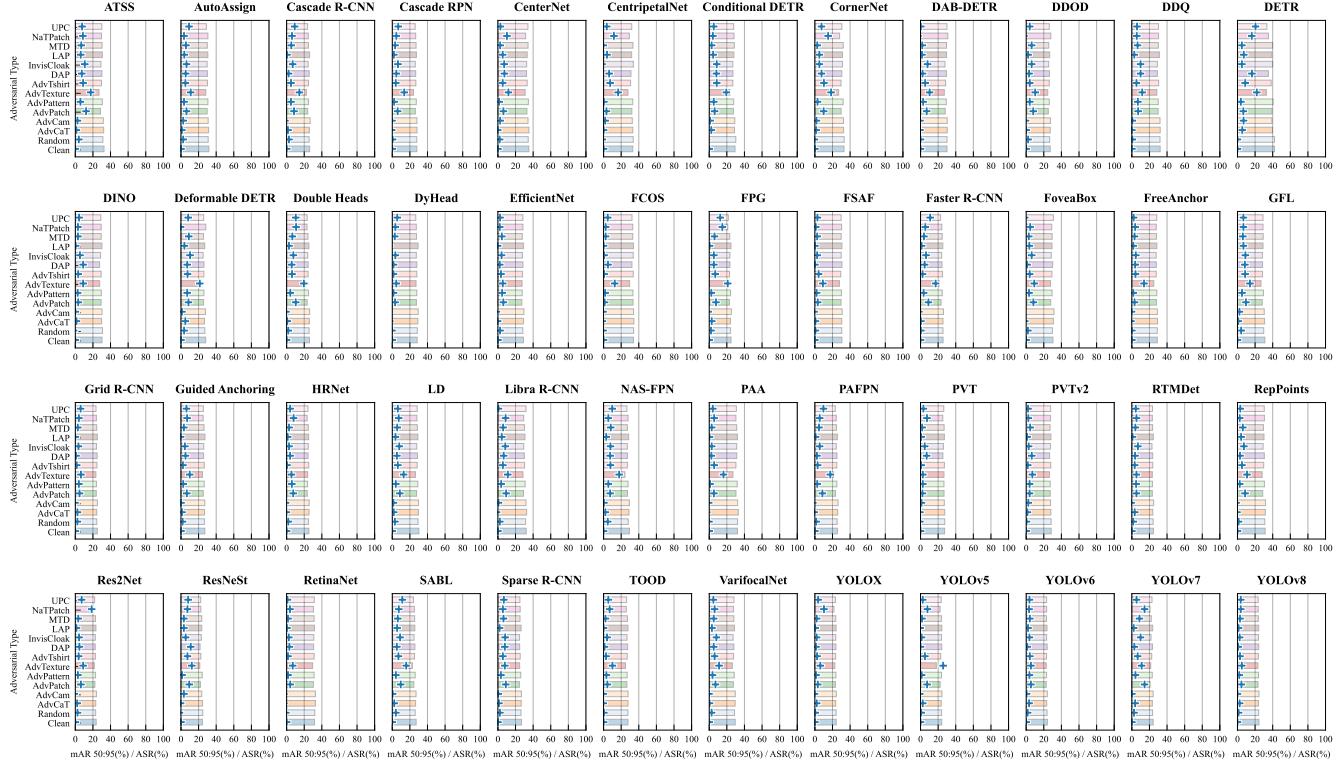


Fig. 19: Overall experimental results of person detection in the metric of mAR50:95(%).

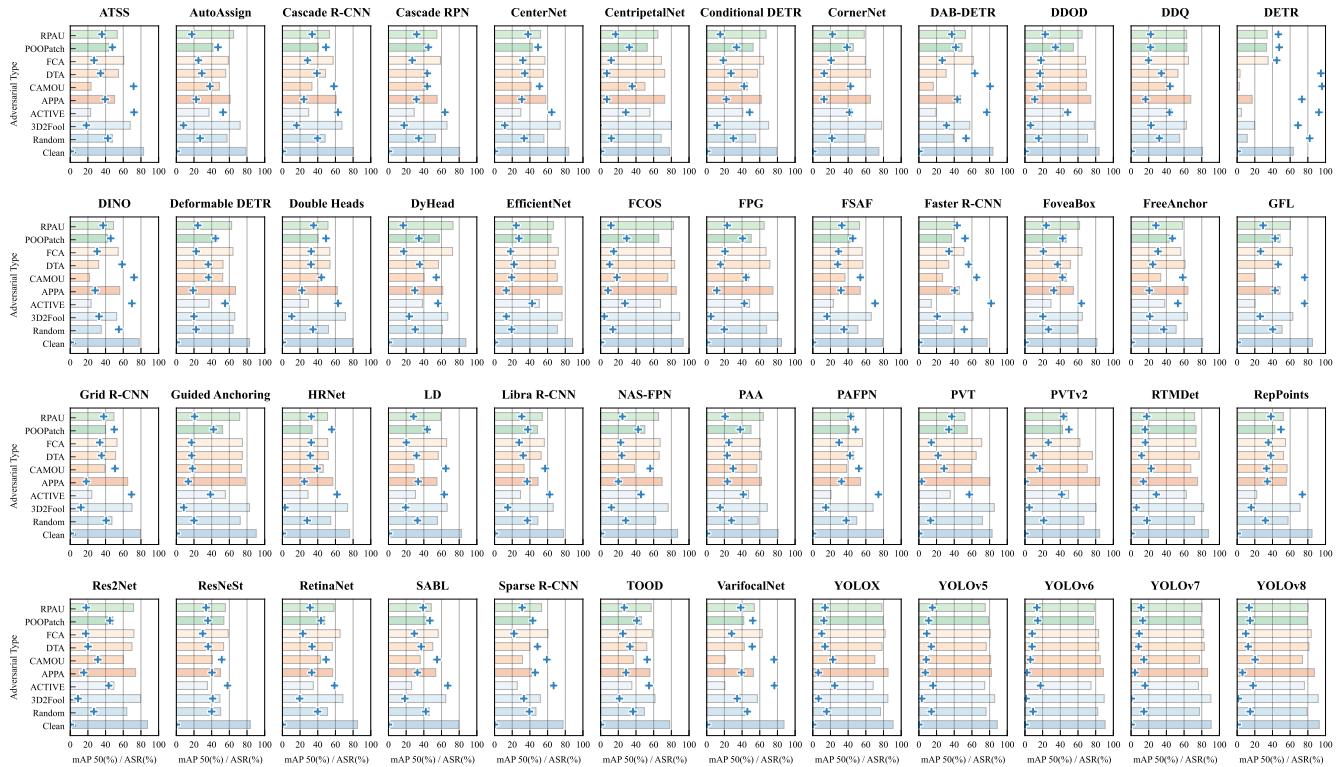


Fig. 20: Overall experimental results of vehicle detection in the metric of mAP50(%).

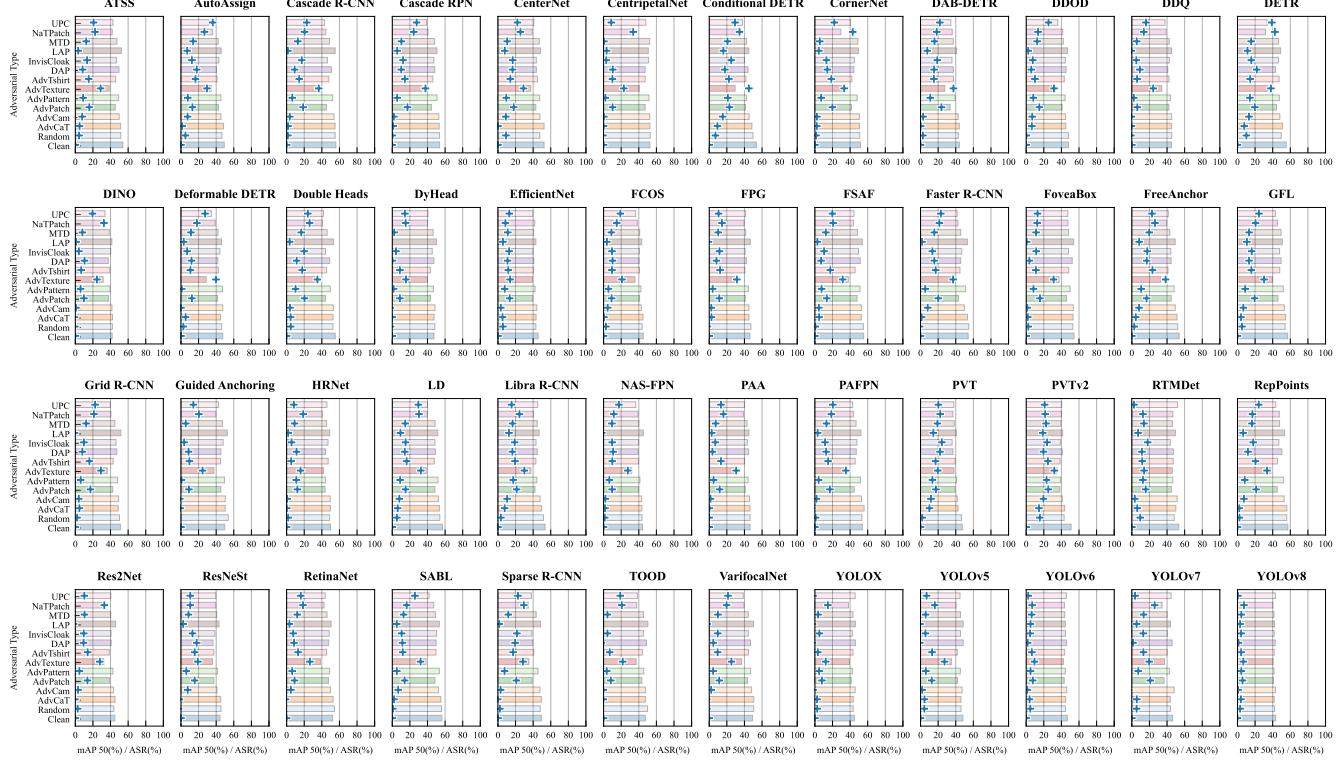


Fig. 21: Overall experimental results of person detection in the metric of mAP50(%).

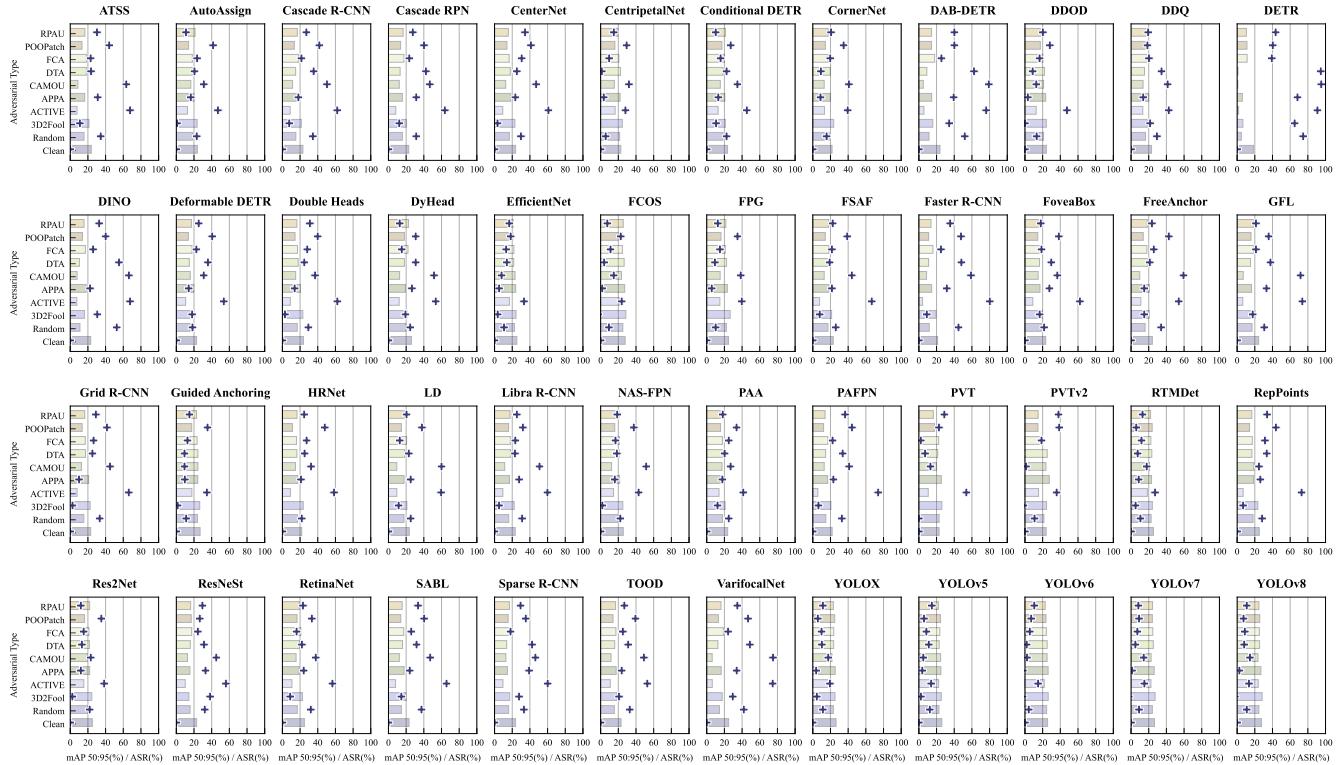


Fig. 22: Overall experimental results of vehicle detection in the metric of mAP50:95(%).

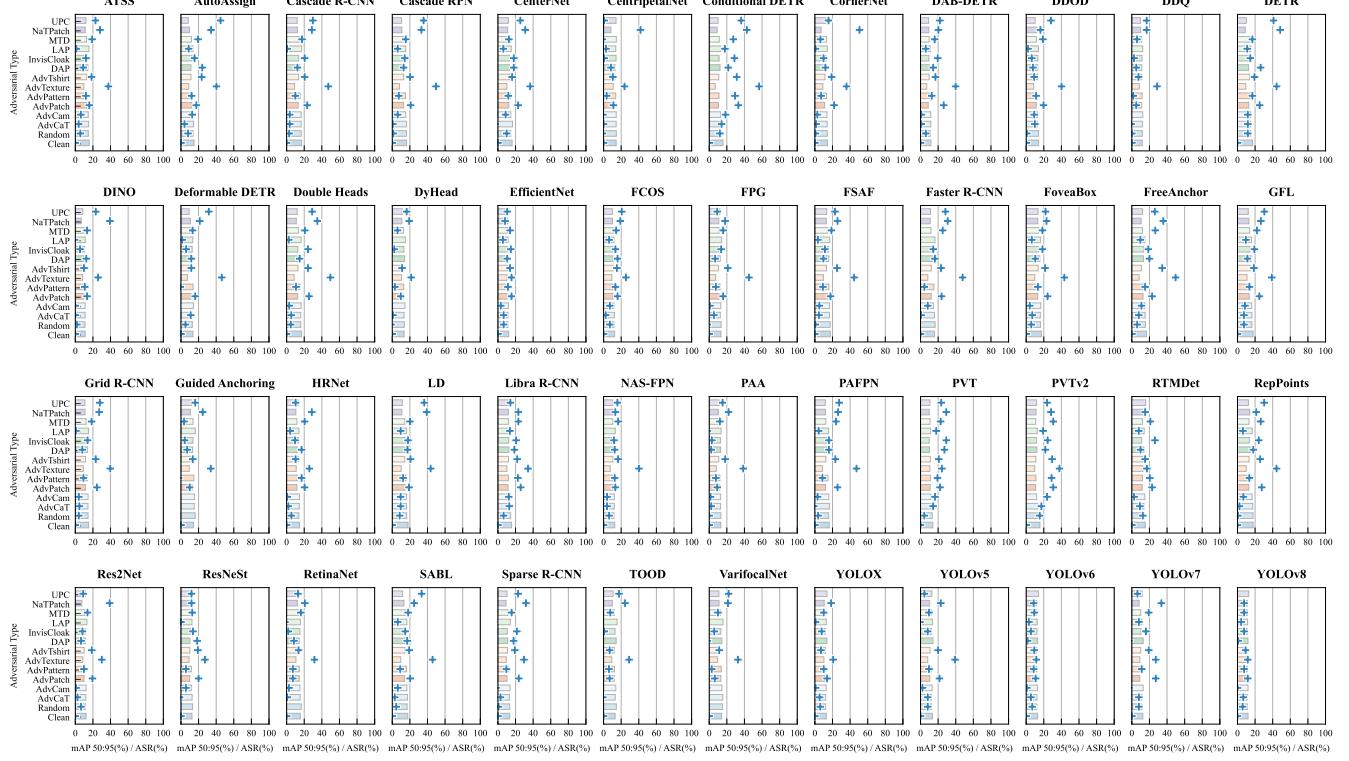


Fig. 23: Overall experimental results of person detection in the metric of mAP50:95(%).

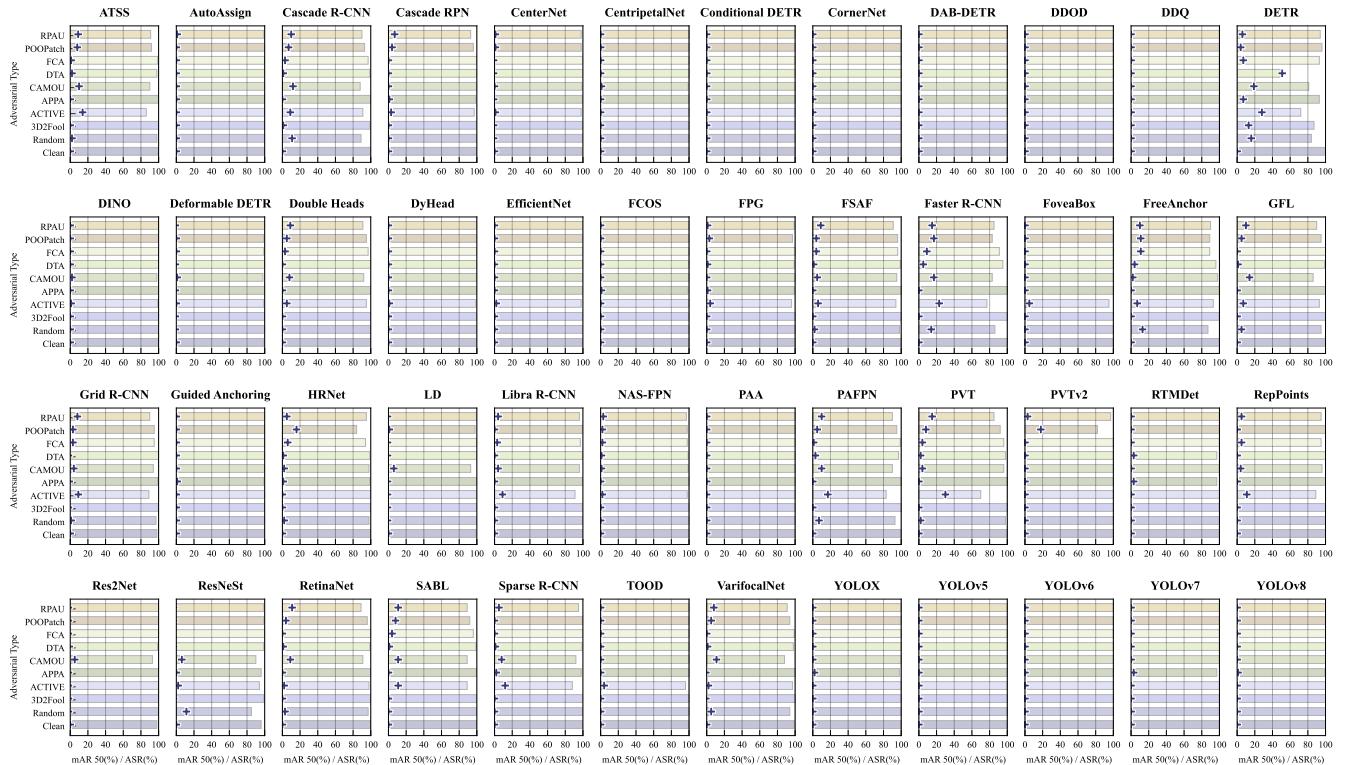


Fig. 24: The ablation experimental results of vehicle detection on Azimuth angle (ϕ) in the metric of mAR50(%).

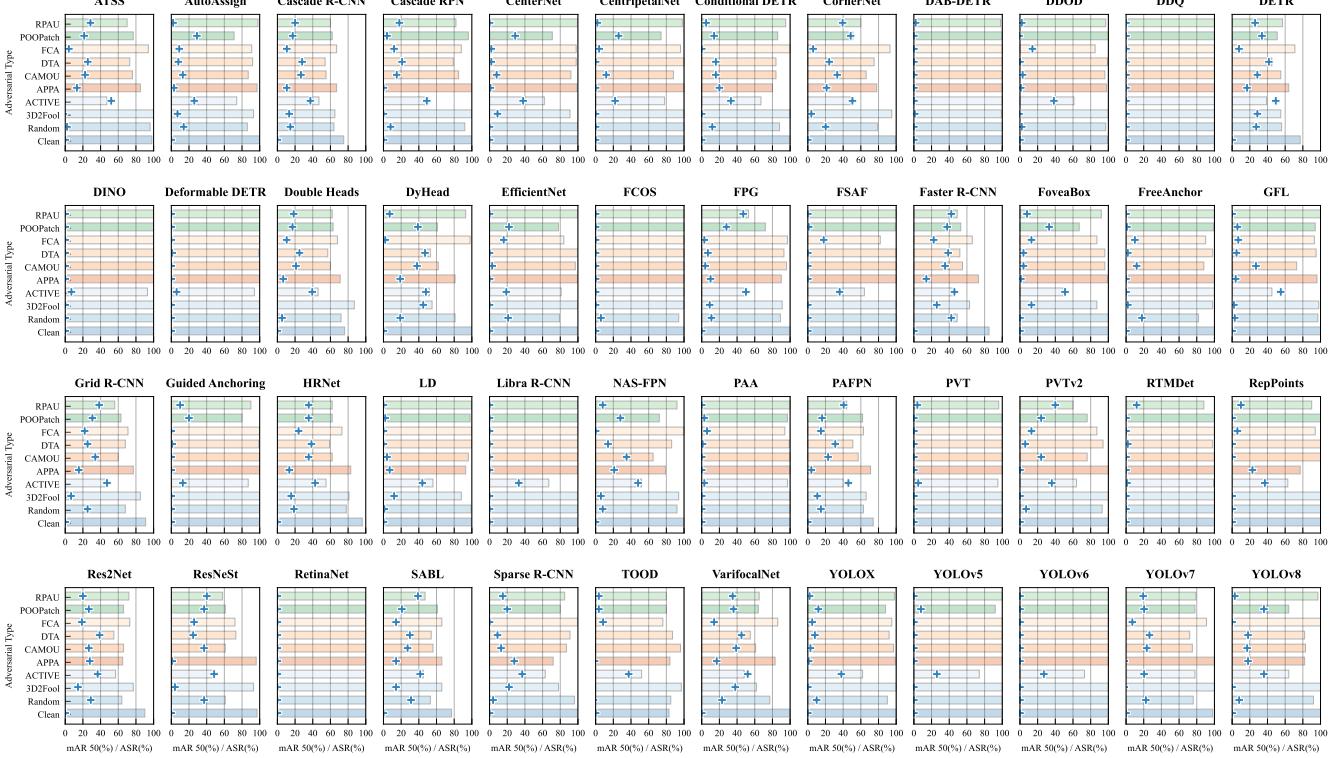


Fig. 25: The ablation experimental results of vehicle detection on Altitude angle (θ) in the metric of mAR50%.

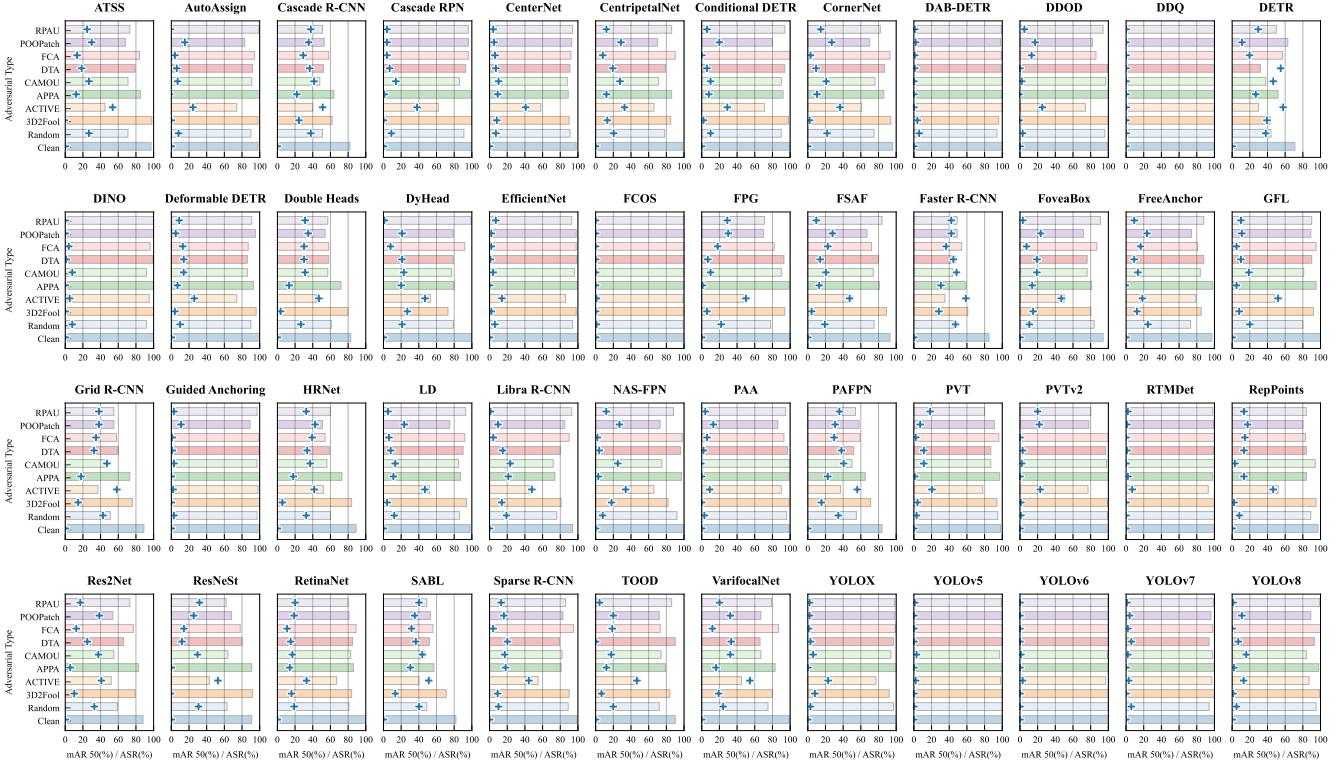


Fig. 26: The ablation experimental results of vehicle detection on Ball-space in the metric of mAR50%.

TABLE XXII: Ablation study on 2D and 3D perturbations.

Perturbations	Entire surface	CornerNet	VarifocalNet
Clean	-	87	80
Random	-	87	77
AdvTexture	✓	74	73
AdvTexture	✗	81(7)	77(4)
AdvPatch	✓	82	75
AdvPatch	✗	85(3)	79(4)
NatPatch	✓	78	74
NatPatch	✗	83(5)	77(3)

TABLE XXIII: Comparison of reported and reproduced results.

		Clean	Random	CAMOU	DTA	ACTIVE
YOLOv3	Reported	86	67	60	32	23
	Reproduced	86	66	62	33	23
YOLOv7	Reported	93	86	83	59	42
	Reproduced	93	85	83	60	41
PVT	Reported	89	78	69	56	52
	Reproduced	89	78	69	56	51

TABLE XXIV: User feedback survey.

Number	Questions
Q1	How easy was it to follow the Docker installation guide for CARLA? (Rating 1-5)
Q2	How helpful was the tutorial on customizing adversarial objects in the documentation? (Rating 1-5)
Q3	Were you able to successfully deploy CARLA using the provided resources? (Yes or No)
Q4	Were you able to successfully customize adversarial objects using the provided resources? (Yes or No)
Q5	Overall, how satisfied are you with the ease of CARLA deployment and customizing adversarial objects? (Rating 1-5)

TABLE XXV: User feedback survey.

Questions	User1	User2	User3	User4	User5
Q1	4	5	5	4	5
Q2	5	5	5	5	5
Q3	Yes	Yes	Yes	Yes	Yes
Q4	Yes	Yes	Yes	Yes	Yes
Q5	4	5	5	4.5	5