Improving the Robustness of Object Detection Under Hazardous Conditions

Rui Li
University of Michigan
Ann Arbor, MI
ruili@umich.edu

Yuqing Luo
University of Michigan
Ann Arbor, MI
yuqingll@umich.edu

Inbum Park
University of Michigan
Ann Arbor, MI
ibpark@umich.edu

Ziyang Xuan
University of Michigan
Ann Arbor, MI
xuanzy@umich.edu

***Abstract*—Advanced Driver Assistance Systems (ADAS) demand high accuracy and robustness in object detection, particularly under adverse weather conditions such as rain, snow, fog, and sandstorms. Transformer-based detectors like DETR have demonstrated strong performance under clear weather, but their effectiveness deteriorates under low-visibility scenarios. In this project, we evaluate DETR's baseline performance on the DAWN dataset and propose improvements through fine-tuning to better extract low-level features and handle noisy inputs. Additionally, we integrate deweathering techniques to enhance model generalization. Our results show that fine-tuning on domain-specific datasets improves detection robustness, but challenges remain in achieving consistent performance across diverse hazardous conditions. Future work includes integrating de-weathering modules to further enhance detection reliability.**

I. Introduction

Advanced Driver Assistance Systems (ADAS) rely heavily on accurate and robust object detection to support critical functionalities such as emergency braking, collision avoidance, and lane departure warning. While modern object detectors, particularly transformer-based models like DEtection TRansformers (DETR) [1], have achieved strong performance under clear weather conditions, their robustness significantly degrades when faced with adverse environments such as rain, snow, fog, and sandstorms. These hazardous conditions introduce challenges like reduced visibility, motion blur, and low-contrast scenes, making object detection a much harder task even for human drivers.

Improving detection robustness under these conditions is crucial for ensuring the safety and reliability of autonomous systems. Collecting large-scale real-world datasets across different weather scenarios is both costly and logistically challenging, which motivates the exploration of domain-specific datasets such as DAWN [2], an image collection focusing on vehicle detection under adverse weather. Furthermore, simple fine-tuning of existing models often falls short, as models may overfit to specific conditions without achieving true generalization across diverse environments.

In this project, we aim to improve the robustness of transformer-based object detectors under hazardous weather conditions. We first evaluate DETR's baseline performance on the DAWN dataset and identify key weaknesses. We then propose a combination of fine-tuning the CNN backbone and transformer decoder modules, along with de-weathering techniques. Through these efforts, we seek to enhance detection accuracy and model generalization under adverse weather conditions. Our findings provide insights into the challenges and opportunities for advancing perception systems in safety-critical applications.

II. Related Work

Object detection for ADAS has been primarily driven by deep learning-based methods, with YOLO, Faster R-CNN, and Single-Shot Detector (SSD) among the most widely adopted. YOLO offers real-time performance with relatively high accuracy, making it suitable for embedded automotive systems. However, it tends to struggle under low-visibility conditions due to its sensitivity to image noise and contrast loss. Faster R-CNN provides higher detection precision, especially for small objects, but is slower and more computationally demanding, which limits its real-time application in vehicles. SSD balances speed and accuracy better than Faster R-CNN but similarly degrades in adverse weather scenarios. Recent transformer-based approaches, like DETR, have shown promise in handling complex scenes but often require large datasets and extensive computation, making them challenging for deployment on resource-limited ADAS platforms.

DETR have made a breakthrough by removing the need for traditional hand-crafted components such as anchor box design and non-maximum suppression (NMS). DETR reframes object detection as a direct set prediction problem, using a transformer encoder-decoder architecture to model global relationships across the entire image. The self-attention mechanisms in the transformer allow DETR to reason about all object instances simultaneously, enabling better handling of complex scenes with overlapping objects. DETR simplifies the training pipeline by requiring only bipartite matching between predictions and ground-truth objects, resulting in a simple and end-to-end optimization process compared to two-stage detectors like Faster R-CNN.

This fully attention-based approach benefits ADAS applications with higher robustness to occlusions, better contextual understanding, and the potential for integrating multi-modal sensor inputs. Furthermore, DETR's ability to model long-range dependencies can be particularly useful for identifying partially visible or small objects, which are critical for safe autonomous driving. However, DETR has limitations under adverse weather. Its detection performance drops when visibility and contrast are reduced, such as in rain, fog, or snow, due to weaker feature extraction. It is also computationally heavy and slow to train, making real-time deployment in vehicles difficult, which remains a challenge for ADAS.

The DAWN dataset was introduced to address the lack of adverse-weather-specific benchmarks. It provides annotated images under various weather conditions, including snow, fog, rain, and sandstorms, simulating real-world driving environments. However, the dataset size remains relatively small compared to general object detection datasets, posing challenges for generalization and fine-tuning.

Related efforts on de-weathering such as TransWeather [3] and Restormer [4] explore the use of transformer architectures for image restoration and enhancement, potentially increase image visibility by reducing the adverse weather effects. Restormer proposes an efficient transformer model for high-resolution image restoration and has achieved state-of-the-art results on several key image restoration tasks, including image deraining, single-image motion deblurring, defocus deblurring, and image denoising. It makes key improvements into the traditional transformer architectures such that it can capture long-range pixel interactions in large images. TransWeather uses its novel encoder to enhance attention inside the image patches to effectively remove weather degradations and introduces its transformer decoder with learnable weather type embeddings to adjust to the weather degradation at hand. It saves efforts from redundant weather-specific encoders in CNN-based models. Both Restormer and Transweather suggests the potential of resolving low visibility due to adverse weathers and integrating de-weathering modules with object detectors to further improve detection robustness.

While previous work has focused either on improving detection in clear conditions or on restoring images degraded by weather, our project aims to directly enhance transformer-based object detection models to operate more reliably under hazardous conditions without relying solely on image restoration preprocessing.

III. Methodology

Our methodology consists of three major stages: Baseline Evaluation, two Fine-tuning DETRs, and Weather-aware Data Augmentation. Our code can be found in our github repository.

A. Weather-aware Data Augmentation

The DAWN dataset contains images under 7 classes of adverse weathers (e.g. fog, snow storm, dust tornado, sand storm, mist, haze, and rain storm). The original method has already applied basic data augmentations, such as horizontal flipping, to expand the training set. In our experiments, we use the dataset as provided, without applying further augmentations due to study focus and time limitation. The lack of more sophisticated weather-specific transformations may limit the model's ability to generalize to unseen weather conditions.

B. Baseline Evaluation

We use the pre-trained DETR model with a ResNet-50 backbone (`facebook/detr-resnet-50`), provided by the Hugging Face `transformers` library. The model is initialized with weights pre-trained on the COCO dataset and deployed on a GPU when available. Input images are processed using the corresponding `DetrImageProcessor`, maintaining aspect ratio and resizing according to the default DETR configuration.

For each image, the model performs inference. We set the detection confidence threshold to 0.5. The resulting bounding boxes and class predictions are post-processed and formatted to match the COCO evaluation format.

Evaluation is conducted separately for each weather condition (e.g. fog, snow storm, dust tornado, sand storm, mist, haze, and rain storm) using the DAWN dataset annotations and the `pycocotools` library. Performance metrics, including mean Average Precision (mAP), are computed following the standard COCO evaluation protocol with IoU type set to `bbox`.

C. Fine-tuning DETR

To enhance DETR's performance under adverse weather conditions, we explore two fine-tuning strategies.

The first strategy involves directly fine-tuning the original pre-trained DETR model (`facebook/detr-resnet-50`) without freezing any layers. We use the DAWN dataset, with weather-specific data augmentations applied during training. The model is fine-tuned using the original DETR loss functions, which consist of a bipartite matching loss (Hungarian loss) and a bounding box regression loss. The goal is to allow the model to adapt its feature extraction and object detection capabilities to adverse weather scenarios without altering its underlying architecture.

The second strategy is motivated by the observation that the DAWN dataset contains only 8 object categories. To better align the model's output space with the dataset, we modify DETR's classification head to predict only 8 categories instead of the original 91 COCO categories. By restricting the prediction space, we aim to reduce false positive detections on irrelevant classes and improve the model's precision under adverse weather conditions. During training, to mitigate the variance introduced by the new classification head, we initially freeze all backbone and transformer layers for the first 100 iterations, allowing only the newly initialized classification layers to be updated. After this stabilization phase, we unfreeze the entire network and continue fine-tuning all parameters jointly for an additional 200 iterations.

D. Deweathering

We also attempt the approach of deweathering to recover image features and enhance the object detection accuracy under adverse weather conditions. We use the pre-trained Restormer model for deraining task with corresponding weights and parameters and apply it on the dawn dataset to achieve enhanced DAWN dataset images. Each image is individually processed and inferenced by the Restormer model to enhance visibility. The de-weathered output images are later streamed into the DETR model without additional retraining.

IV. Experiment and Results

We compare the baseline results against the finetuned version of our model denoted as "Finetuned", and the de-weathering approach denoted as "De-weathered". While there are eight weather types in the DAWN dataset, we omit 'dusttornado' which only has two samples, and also a duplicate of one of the weather types. The results for each weather type is summarized as a table or figure, and there are a total of six weather types. The categories are each 'rain storm', 'sand storm', 'haze', 'mist', 'foggy', and 'snow storm' with 28, 21, 7, 7, 11, and 14 samples respectively. We follow the same metric evaluation as in the baseline, which comprises of six Average Precision (AP) metrics and six Average Recall (AR) metrics. Note that the difference of rows 4-6 and rows 10-12 are both based on the area (small/medium/large), and rows 7-9 differ in the maximum number of detections (1/10/100). IoU=0.5:0.95 means that the result is sampled equally spaced between IoU=0.5 and IoU=0.95 and averaged.

Metric	Baseline	Finetuned	De-weathered
AP @ IoU=0.5:0.95	0.026	0.036	0.026
AP @ IoU=0.5	0.044	0.056	0.044
AP @ IoU=0.75	0.029	0.038	0.029
AP @ IoU=0.5:0.95	0.006	0.015	0.006
AP @ IoU=0.5:0.95	0.028	0.038	0.029
AP @ IoU=0.5:0.95	0.61	0.078	0.061
AR @ IoU=0.5:0.95	0.008	0.009	0.009
AR @ IoU=0.5:0.95	0.024	0.032	0.025
AR @ IoU=0.5:0.95	0.029	0.041	0.030
AR @ IoU=0.5:0.95	0.008	0.023	0.008
AR @ IoU=0.5:0.95	0.032	0.044	0.033
AR @ IoU=0.5:0.95	0.071	0.085	0.071

TABLE I
Rain Storm

Metric	Baseline	Finetuned	De-weathered
AP @ IoU=0.5:0.95	0.018	0.058	0.061
AP @ IoU=0.5	0.041	0.089	0.113
AP @ IoU=0.75	0.014	0.068	0.034
AP @ IoU=0.5:0.95	0.016	0.038	0.009
AP @ IoU=0.5:0.95	0.021	0.068	0.048
AP @ IoU=0.5:0.95	0.028	0.036	0.136
AR @ IoU=0.5:0.95	0.009	0.019	0.049
AR @ IoU=0.5:0.95	0.021	0.060	0.063
AR @ IoU=0.5:0.95	0.021	0.062	0.064
AR @ IoU=0.5:0.95	0.015	0.038	0.014
AR @ IoU=0.5:0.95	0.027	0.074	0.050
AR @ IoU=0.5:0.95	0.028	0.035	0.135

TABLE II
Sand Storm

We also qualitatively evaluate the baseline and the finetuned model on the DAWN dataset. We visualize the ground truth bounding boxes as the green boxes, and the red boxes signify the predictions for each of the models in Fig 1.

Metric	Baseline	Finetuned	De-weathered
AP @ IoU=0.5:0.95	0.041	0.010	0.042
AP @ IoU=0.5	0.071	0.020	0.074
AP @ IoU=0.75	0.063	0.009	0.062
AP @ IoU=0.5:0.95	0.103	0.007	0.103
AP @ IoU=0.5:0.95	0.010	0.011	0.010
AP @ IoU=0.5:0.95	0.009	0.011	0.009
AR @ IoU=0.5:0.95	0.035	0.003	0.035
AR @ IoU=0.5:0.95	0.041	0.008	0.041
AR @ IoU=0.5:0.95	0.043	0.011	0.043
AR @ IoU=0.5:0.95	0.106	0.010	0.106
AR @ IoU=0.5:0.95	0.010	0.012	0.010
AR @ IoU=0.5:0.95	0.008	0.011	0.008

TABLE III
Mist

Metric	Baseline	Finetuned	De-weathered
AP @ IoU=0.5:0.95	0.012	0.026	0.012
AP @ IoU=0.5	0.021	0.049	0.021
AP @ IoU=0.75	0.013	0.021	0.013
AP @ IoU=0.5:0.95	0.003	0.006	0.003
AP @ IoU=0.5:0.95	0.021	0.040	0.021
AP @ IoU=0.5:0.95	0.007	0.007	0.007
AR @ IoU=0.5:0.95	0.005	0.012	0.005
AR @ IoU=0.5:0.95	0.012	0.026	0.012
AR @ IoU=0.5:0.95	0.012	0.027	0.012
AR @ IoU=0.5:0.95	0.004	0.007	0.004
AR @ IoU=0.5:0.95	0.022	0.041	0.021
AR @ IoU=0.5:0.95	0.006	0.006	0.005

TABLE IV
Haze

V. Discussion

We find a general trend of improvement in the finetuned model's experimental results. We do not list the results of the second approach of finetuning (altering the classification layers first), as they have similar results to simply finetuning the entire model. Especially, one can clearly see that the images in Fig 1 correctly find vehicles in harsh weather by noticing cues like headlights or the outlines of the vehicle. Also, the finetuned model has less ghost predictions compared to the baseline model's predictions. However, we also note that this is not always the case and there are failure examples where the baseline model performs better than the finetuned model. Especially, the pretrained DETR model outperforms our finetuned model in certain weather conditions, such as 'mist' or 'rain storm'. Some of these causes come from the class imbalance in the dataset, or the fact that there are such little information in the image for the model to infer any existing vehicles. We also note that the finetuned model has improved itself in much more harsher conditions such as 'sand storm' or 'haze' in a short amount of time, and thus there is potential in the model to consistently outperform in various hazardous weather conditions.

VI. Limitations

One major limitation of this project arises from the DAWN dataset itself. First, the dataset contains numerous annotation

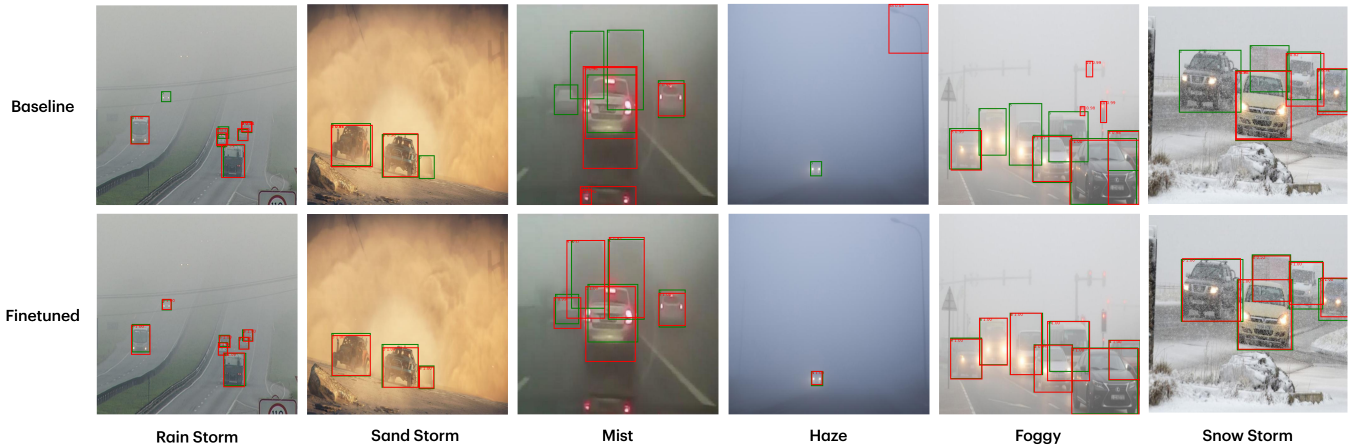

Fig. 1. Qualitative Examples

Metric	Baseline	Finetuned	De-weathered
AP @ IoU=0.5:0.95	0.004	0.017	0.004
AP @ IoU=0.5	0.006	0.030	0.006
AP @ IoU=0.75	0.005	0.008	0.005
AP @ IoU=0.5:0.95	0.000	0.001	0.000
AP @ IoU=0.5:0.95	0.003	0.022	0.003
AP @ IoU=0.5:0.95	0.012	0.022	0.012
AR @ IoU=0.5:0.95	0.001	0.011	0.001
AR @ IoU=0.5:0.95	0.003	0.016	0.003
AR @ IoU=0.5:0.95	0.003	0.016	0.003
AR @ IoU=0.5:0.95	0.000	0.001	0.000
AR @ IoU=0.5:0.95	0.003	0.021	0.003
AR @ IoU=0.5:0.95	0.011	0.022	0.011

TABLE V
FOGGY

Metric	Baseline	Finetuned	De-weathered
AP @ IoU=0.5:0.95	0.027	0.031	0.026
AP @ IoU=0.5	0.048	0.052	0.047
AP @ IoU=0.75	0.029	0.027	0.027
AP @ IoU=0.5:0.95	0.020	0.024	0.020
AP @ IoU=0.5:0.95	0.023	0.027	0.021
AP @ IoU=0.5:0.95	0.060	0.076	0.059
AR @ IoU=0.5:0.95	0.014	0.014	0.014
AR @ IoU=0.5:0.95	0.029	0.037	0.028
AR @ IoU=0.5:0.95	0.029	0.042	0.029
AR @ IoU=0.5:0.95	0.020	0.041	0.019
AR @ IoU=0.5:0.95	0.026	0.030	0.026
AR @ IoU=0.5:0.95	0.064	0.077	0.062

TABLE VI
SNOW STORM

errors, such as missing labels for certain vehicles, which can negatively impact both training and evaluation. Second, the dataset size is relatively small, limiting the diversity of scenarios the model can learn from. Furthermore, the distribution of weather conditions within the dataset is highly imbalanced, with some weather types represented by only two test cases such as 'dusttornado'. In addition to the imbalance in weather conditions, the category distribution is also heavily skewed. The majority of annotations correspond to category ID 3 (cars), while other vehicle categories, such as trains, are extremely underrepresented, with only a single instance found in the entire dataset. Finally, the data augmentation techniques applied by the model were relatively simple as we only opt for horizontal flipping.

VII. FUTURE WORK

Based on the limitations observed in this study, several directions for future work are worth exploring.

First, more advanced data augmentation techniques could be employed to address the lack of diverse adverse weather data. In particular, it may be feasible to leverage generative models to simulate various adverse weather effects (e.g., rain, fog, snow) on images originally captured in normal conditions. This approach could substantially expand the dataset without requiring additional manual annotations.

Second, given a sufficiently large amount of training data, future models could incorporate weather-awareness directly into the detection pipeline. For example, an auxiliary branch could be added after the CNN backbone to predict the weather condition of the input image. The predicted weather feature could then be fused into the Transformer module, allowing the model to adapt its object detection strategy based on the environmental context.

VIII. STATEMENT OF CONTRIBUTION

Rui contributed to the baseline visualization for proposal, website deployment, and report sections [FILL]. Yuqing contributed to the baseline evaluation and visualization, deweathering experiments, and report sections [FILL]. Inbum contributed to finetuning DETR directly, making the poster, and report sections IV and V. Ziyang contributed to finetuning DETR with modified classification head and report sections [FILL].

REFERENCES

[1] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-end object detection with transformers," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2020. [Online]. Available: https://arxiv.org/abs/2005.12872

[2] M. Kenk and M. Hassaballah, "Dawn: Vehicle detection in adverse weather nature dataset," *arXiv preprint arXiv:2008.07016*, 2020. [Online]. Available: https://arxiv.org/abs/2008.07016

[3] J. M. J. Valanarasu, R. Yasarla, and V. M. Patel, "Transweather: Transformer-based restoration of images degraded by adverse weather conditions," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. [Online]. Available: https://arxiv.org/abs/2111.14813

[4] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *CVPR*, 2022.