

# Pedestrian and Vehicle Detection Performance Analysis of Faster R-CNN under Extreme Weather Conditions

**Yuhan Wang**

Chang'an Dublin International College of Transportation, Chang'an University, Xi'an, Shanxi, 710018, China

yuhan.wang1@ucdconnect.ie

**Abstract.** Vehicle safety systems now have a significant problem in reliably and accurately detecting objects in adverse weather due to the ongoing advancements in autonomous driving technology. The main objective of this work is to assess how well the Faster R-CNN model performs when it comes to identifying cars and pedestrians in inclement weather, including rain, snow, and fog. In order to learn more about the accuracy and resilience of the model, the researchers trained and tested it using a variety of datasets, including a range of harsh weather conditions. The model's performance significantly degrades in poor visibility and small object identification, according to the testing data. This decline in effectiveness demonstrates how severely inclement weather affects detecting systems. The paper also addresses potential enhancements, such as parameter fine-tuning, model optimization, and the application of increased data augmentation methods to boost robustness. Expanding the dataset, enhancing image quality in adverse situations, and researching more sophisticated model architectures will be the main goals of future studies in order to increase detection accuracy and robustness in practical autonomous driving scenarios.

**Keywords:** Faster R-CNN, Extreme Weather Conditions, Pedestrian and Vehicle Detection Performance.

## 1. Introduction

In recent years, autonomous driving technology has developed rapidly, from the first simple assisted driving system to the current advanced fully automatic unmanned driving, autonomous driving has become an important development area of the automotive industry. Among them, autonomous vehicles are mainly composed of two systems: perception and decision-making [1]. The vision system is typically broken down into numerous subsystems that handle various tasks like traffic signal detection and identification, autonomous vehicle localization, road mapping, static obstacle mapping, and tracking of moving obstacles [2]. The decision-making system is often divided into numerous subsystems that handle various tasks like path planning, route planning, behavior selection, motion planning, obstacle avoidance, and control. However, there are several variations in the literature regarding this partitioning, and it is not entirely clear [2]. In all the above subsystems, the recognition of real-time road condition information such as pedestrians and vehicles on the road surface in the perception system is particularly important. With the development of deep learning models, especially convolutional neural networks, the accuracy of object detection has been further improved. In order to achieve high accuracy, He et al.

[3] suggested a lightweight convolutional network made up of pooling layers and simple convolution layers. Feature pyramid network (FPN) [4] and SSD were integrated by Lv et al. [5] to enhance remote sensing picture tiny item recognition performance. For vehicle detection, Wu et al. [6] created a unified architecture made up of segmentation, transfer learning, and active classification. In the meantime, object detection research is expanding quickly. By restricting the receptive field of multi-head attention operations to the fixed, Liu et al. [7] inserted the Transformer into the backbone and decreased model complexity, attaining the state-of-the-art performance of 60.6 AP on COCO test-dev [8].

From the beginning of using the basic R-CNN (Region-based Convolutional Neural Networks) for object detection, the basic principle is to select different candidate regions and apply convolutional neural networks on these selected candidate regions for object classification and detection. However, because R-CNN needs to extract independent features for each candidate region, the computational efficiency of R-CNN is reduced. In order to solve this problem, Fast R-CNN reduces double computation by sharing convolutional feature maps of the entire image, and improves the processing speed by introducing an ROI Pooling layer [9]. Later, Faster R-CNN realized end-to-end training and prediction of object detection by introducing the Regional Recommendation Network (RPN), which greatly improved the speed and accuracy of object detection [10]. With the continuous iteration of technology and the improvement of speed and accuracy, the popularity of object detection technology in the field of autonomous driving has deepened.

R-CNN innovatively introduces convolutional neural network into object detection, but due to its huge amount of computation, it still cannot meet the needs of real-time applications. Fast R-CNN reduces the computational cost and improves the detection speed by using the shared convolutional feature map, which lays the foundation for the real-time application of this method [9]. However, Fast R-CNN relies too much on the external region suggestion algorithm, resulting in a very limited improvement in detection speed. The subsequent Faster R-CNN solves the bottleneck problem of detection speed improvement by integrating candidate region generation and detection network by introducing RPN, and greatly improves the detection efficiency and accuracy [10]. These improvements in detection technologies have significantly improved the response speed and accuracy, making object detection technologies more widely used in practical scenarios such as autonomous driving and intelligent surveillance [11].

After discussing the progress of object detection technology, the weather environment at the time of target detection is also an important reference index. Studies have shown that the contrast and sharpness of the collected images vary under different weather conditions, and these changes are likely to skew the detection results of the inspection system [12]. After discussing the progress of object detection technology, the weather environment at the time of target detection is also an important reference index. Studies have shown that the contrast and sharpness of the collected images vary under different weather conditions, and these changes are likely to skew the detection results of the inspection system [12]. Especially under extreme weather conditions such as rain, snow, and haze, the difficulty of detection is greatly increased [13]. Extreme weather conditions such as heavy rain, heavy snow, dense fog, sandstorms, etc., may cause the contrast of the picture to decrease or the target boundary to be blurred, resulting in the degradation of image quality. Researchers have put forth a number of strategies to increase robustness, including data augmentation, adversarial training, and multimodal sensor fusion-based detection algorithms [14]. It can be seen that when under extreme weather conditions, the impact on target recognition is huge.

Therefore, the purpose of this paper is to discuss and explore the impact of extreme weather on Faster R-CNN-based object detection, and to discuss the causes and improvement methods based on the results and use a pavement information dataset under extreme weather conditions and train a perception system based on Faster R-CNN technology that can recognize road information such as pedestrians and vehicles under extreme weather conditions such as rain, snow, haze, etc. At the end of the article, the performance of the model is systematically evaluated and analyzed, and whether it can be effectively and quickly identified and suggestions for improvement are discussed. Contribute to the development of autonomous driving.

## 2. Methodology

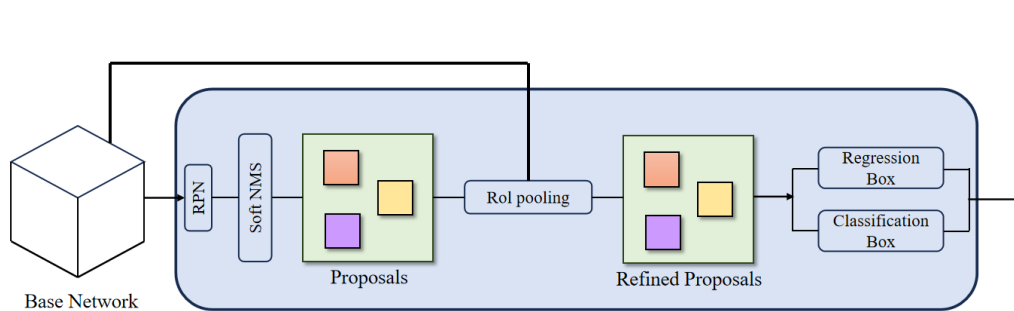
### 2.1. Preparing dataset

In this paper, this study used publicly available "Kaggle" images and self-collected crystalline images under extreme weather to train and evaluate the Faster R-CNN model. In order to ensure the diversity of the dataset, the dataset used in this training contains a total of 3,404 illustrations of various extreme weather conditions on different roads and in different regions. Moreover, the resolution of this dataset is uniform, and the image clarity is moderate, which is suitable for deep learning.

This research categorized the data after obtaining it according to its classifications. Six categories constitute our database: car, truck, person, bicycle, bus, and motorcycle. prior to deep learning. This paper split the dataset into three distinct sets, referred to as the training set and validation set, in accordance with a 9:1 ratio to guarantee the dependability of the final trained model. Additionally, this study makes use of three sets of data: the training set, which is used to go through machine learning, the validation set, which is used to modify the model's parameters, the test set, which yields the results.

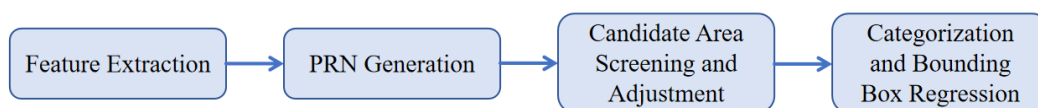
### 2.2. The working principle and flow of Faster R-CNN

By merging CNN and Region Proposal, Faster R-CNN is an object detection model that gradually advances from R-CNN to Fast R-CNN. This enhances both the speed and accuracy of object detection. The feature extraction network, RPN, and the classification and bounding box regression steps make up its three main components. First, using convolutional neural networks that have already been trained, like ResNet or VGG, Faster R-CNN extracts features from the input image. The image is transformed into a low-resolution feature map by the feature extraction network once the features have been extracted. After that, the RPN will go across this feature map to produce a number of candidate regions by producing an array of Ahchors with various aspect ratios and scales. After then, RPN will categorize these potential regions to ascertain if they are a part of the intended item. Lastly, Faster R-CNN will carry out edge regression and classification using the candidate areas produced by RPN with a fully connected network. At this stage, the model will categorize each potential region based on the various groups to which it belongs and then modify the bounding box's location to guarantee correctness [15].



**Figure 1.** Faster-RCNN model[16]

As shown in Figure 1, this research can disassemble and further simplify Fast R-CNN into the following few, four steps as shown in Figure 2.



**Figure 2.** Faster R-CNN workflow diagram

In the PRN Generation section, a loss function is used to optimize the accuracy of classification and regression. The loss function includes Cross-Entropy Loss and Smooth L1 Loss. the specific loss function is:

$$L(p_i, t_i) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (1)$$

Here  $L_{cls}$  is the classification loss,  $L_{reg}$  is the regression loss,  $p_i$  is the predicted category probability,  $t_i$  is the predicted bounding box parameter, and  $p_i^*$  and  $t_i^*$  are the true category and bounding box parameters.

For the training speed of the model, to ensure the training efficiency, first of all, the images used for training are standardized to ensure that each image is uniformly scaled to the same size during training. In this training this paper scaled the images to 600×600 pixels uniformly, as well as the pixel value normalization of the images adjusted to the [0,1] interval. The rest of the basic parameters used for the related training are shown in Table 1.

**Table 1.** Basic parameters

Parameters	Set values
Base network	VGG19
Batch size	2
Optimizer type	Adam
Learning rate	0.0001

### 3. Result

#### 3.1. Overall model performance

After 150 rounds of training for about 16 hours, the model output is shown in Table 2.

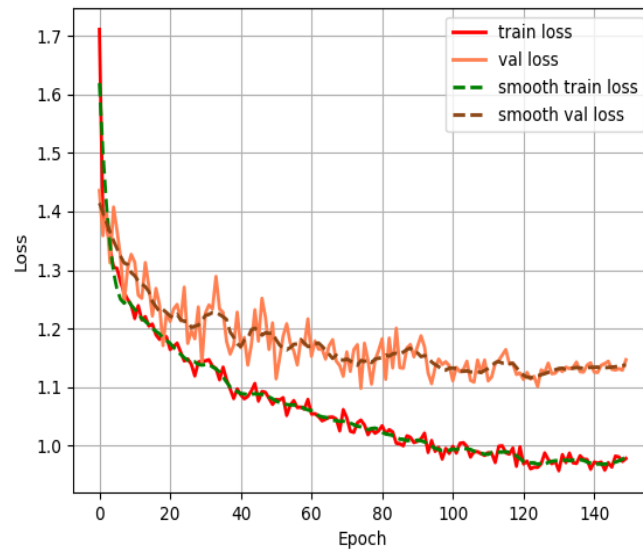
**Table 2.** Training result

Classes	Score Threshold	F1	Recall	Precision
Bicycle	0.5	0.29	33.96%	24.66%
Bus	0.5	0.30	22.50%	45.00%
Car	0.5	0.47	67.56%	36.49%
Motorcycle	0.5	0.30	38.89%	25.00%
Person	0.5	0.36	44.62%	30.80%
Truck	0.5	0.45	46.71%	43.83%
MAP = 33.40%				

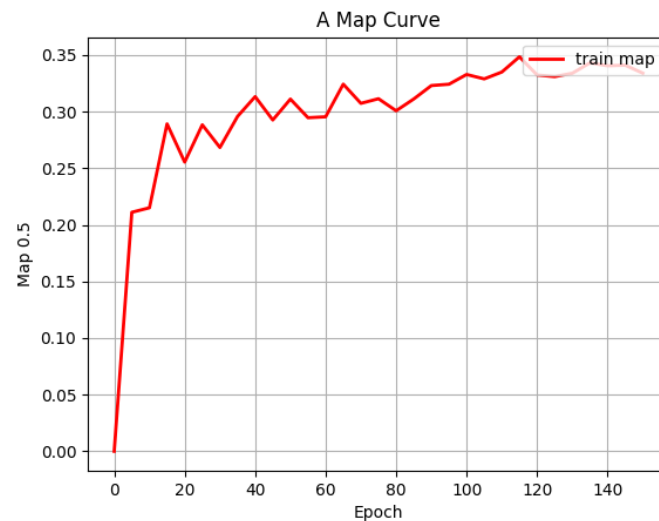
Table 2 illustrates that the model trained on this test set by Faster R-CNN is not perfect, with an average accuracy (MAP) of just 33.40%. Because of this, there are significant variations amongst various classes as well. For instance, the Car model has a greater recall and F1 of 67.56% and 0.47, respectively, although the Bus model has a higher accuracy of 45%. However, the model did not perform well for the Bicycle and Motorcycle, with accuracy of only 24.66% and 25.00%, respectively. These results preliminarily show that the target detection ability of Faster R-CNN is very limited under extreme weather conditions, and also indicate that the model also has great differences in dealing with targets of different sizes. Especially in the detection of small targets, the model effect is poor.

#### 3.2. Chart display

In order to better understand the performance of this model, This research plot Figure 3 Training loss vs. validation loss curves diagram and Figure 4 MAP Graph to show the trend during training.



**Figure 3.** Training loss vs. validation loss curves diagram



**Figure 4.** MAP graph

Figure 3 demonstrates the variation of the model's loss during the training process, where the red line represents the training loss and the orange indicates the validation loss. The model's fit to the data is getting better, as demonstrated by the figure, which shows that the training loss is generally on the down. The curves exhibit some oscillation around the middle and end of training before progressively leveling off.

Figure 4 illustrates how the model's MAP changed as it was being trained. The MAP rises quickly in the early stages of training, as seen in the figure, and then tends to settle progressively before stabilizing at approximately 33.40%, suggesting that the model has achieved the top limit of the training impact under these conditions.

#### 4. Discussion

Combining the above results, it is easy to see that the training results are not very satisfying. It can be seen that extreme weather has a significant impact on model learning, especially for the model trained based on Faster R-CNN.

This result may be due to the fact that the detected targets in the dataset may be blurred due to extreme weather on one hand, and on the other hand, the model parameters may be more empirical, resulting in some of the model parameters not being set reasonably. In addition, too much and too detailed target classification may also be the reason for the low accuracy of the model. The limited modeling capability of Faster R-CNN cannot be excluded.

The next steps in this study will concentrate on the following enhancements in order to handle this scenario and enhance the model's performance. Initially, in order to make up for the effects of extreme weather on the photos, the model would keep adding photographs to its dataset and enhancing their quality. Furthermore, by altering the optimization model or the learning rate, this study will keep optimizing the appropriate model parameters, which should further lower error rates and raise accuracy. Subsequently, the number of categories in the photos can be decreased and only basic preliminary categorization of the pedestrian and vehicle categories can be identified if the MAP improvement of the model is still restricted. Finally, additional machine learning methods might be chosen or coupled if the aforementioned enhancements are insufficient to guarantee the stability and accuracy of the recognition.

## 5. Conclusion

The main objective of this research is to detect the effect of extreme weather conditions on the Faster R-CNN based model using a deep learning approach. Compared to bicycle and motorcycle, the model has a higher recognition rate for the automobile category, indicating that the model has a better recognition rate for bulky objects. However, the final MAP is only 33.40%, which indicates that the model has obvious deficiencies and is difficult to apply in practice, and also reflects the significant interference effect of extreme weather on the Faster R-CNN-based target recognition model. By analyzing the results, This paper found that it is very likely that the extreme weather blurred the target boundary, leading to a decrease in the recognition rate. It is also possible that the unreasonable setting of model parameters is the reason for the poor performance of the results. Based on this, the next step can be to optimize the dataset, improve the quality and quantity of images, and adjust the model parameters to improve the robustness and detection accuracy of Faster R-CNN in complex environments. In conclusion, this paper shows the adverse effects of extreme weather on the training of the Faster R-CNN algorithm based on the target detection model trained for extreme weather and proposes a feasible solution accordingly. Next, the model can be further optimized to enhance its stability and generalization ability in practical applications.

## References

- [1] Paden B Cap M Yong S Z Yershov D Frazzoli E 2016 A Survey of Motion Planning and Control Techniques for Self-Driving Urban Vehicles IEEE transactions on intelligent vehicles vol 1 no 1 p 33-55
- [2] Alam M K Ahmed A Salih R Al Asmar A F S Khan M A Mustafa N Mursaleen M and Islam S 2023 Faster RCNN based robust vehicle detection algorithm for identifying and classifying vehicles Journal of Real-Time Image Processing vol 20 p 93
- [3] He Z 2020 Traffic Sign Recognition Based on Convolutional Neural Network Model Chinese CAC p 155-158
- [4] Lin T Piotr D 2017 Feature pyramid networks for object detection Proceedings of the IEEE conference on computer vision and pattern recognition p 2117-2125
- [5] Lv S Lei Z H U and Wang W 2020 Improving SSD for detecting a small target in Remote Sensing Image CAC p 567-571
- [6] Wu X Wei L Qian D 2020 Vehicle detection of multi-source remote sensing data using active fine-tuning network ISPRS p39-53
- [7] Liu Z Lin Y Guo B 2013 Swin transformer: Hierarchical vision transformer using shifted windows arXiv preprint arXiv 2103
- [8] Lin T Y Michael M Serge B James H Pietro P Deva R Piotr D 2014 Microsoft coco: Common objects in context In European conference on computer vision p 740-755

- [9] Girshick R 2015 Fast R-CNN ICCV p 1440-1448.
- [10] Ren S He K Girshick R Sun J 2016 Faster R-CNN Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence vol 39 no 6 p 1137-1149
- [11] Zhao Z Q Zheng P Xu S T Wu X 2019 Object Detection with Deep Learning: A Review. IEEE Transactions on Neural Networks and Learning Systems vol 30 no 11 p 3212-3232
- [12] Cordts M Omran M Ramos S Rehfeld T Enzweiler M Benenson R Schiele B 2016 The Cityscapes Dataset for Semantic Urban Scene Understanding. CVPR p 3213-3223
- [13] Ding Y Luo X Zhang J Li Y 2024 SDNIA-YOLO: A Robust Object Detection Model for Extreme Weather Conditions. arXiv preprint arXiv:2401.12345.
- [14] Sun P Zhang R Jiang Y Kong T Xu C Zhan W 2021 High-Performance Object Detection: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence.
- [15] Zhao X Wang L Zhang Y Han X Deveci M and Parmar M 2024 A review of convolutional neural networks in computer vision Artificial Intelligence Review vol 57 no 4 p 99
- [16] Alam M K Ahmed A Salih R Al Asmari AFS Khan M A Mustafa N Mursaleen M and Islam S 2023 Faster RCNN based robust vehicle detection algorithm for identifying and classifying vehicles. Journal of Real-Time Image Processing vol 20 no 5 p 93