# Reproducing: Robust Object Detection in Challenging Weather Conditions

Diego Fernandez
University of Illinois Urbana-Champaign
https://drive.google.com/file/d/1IXf8a16K4cDZ_QMo35TXBQ9ATZcPSGaV/view?usp=sharing
diegoaf2@illinois.edu

## Abstract

*As a graduate student of Autonomy and Robotics this project synergizes well with my current course work, which is primarily focused on autonomous vehicles and mobile robots. Object detection is a core component of motion planning, whether to avoid obstacles, pick up objects or updating the plan based on signs. I have not worked with any of the datasets or YOLO models in the past and using them in this project will benefit my future coursework and professional career. Adverse weather conditions like snow, fog, and rain significantly challenge the reliability of object detection models, increasing the risk of accidents and operational failures when implemented carelessly. This project investigates strategies to enhance deep learning-based object detection under adverse weather conditions. I focus on reproducing baseline results and evaluating the efficacy of style transfer augmentation for simulating adverse weather scenarios. Three approaches are explored: (1) training on real-world clear weather datasets (2) training on real-world all-weather datasets, (3) training on real-world all-weather datasets augmenting training data with synthetic foggy weather patterns generated through style transfer. Using the BDD100K dataset for training and unseen adverse weather images from the DAWN dataset for evaluation, I analyze the performance of these strategies. Preliminary findings support the findings* of *Gupta et al. [1], that models trained on real-world all-weather. datasets achieve superior robustness compared to those trained on clear weather datasets. Due to computational limits this project was unable to complete training with the augmented dataset, however the process is explored and can be completed in another iteration.*

## 1. Introduction

Autonomous vehicle systems depend on the accurate identification of pedestrians, vehicles, traffic signals, and obstacles to navigate complex urban and highway environments. While deep learning has revolutionized object detection under ideal conditions, adverse weather such as snow, fog and rain introduces noise, occlusions, and distortions that negatively impact model performance in real time scenarios.

To address these challenges, Gupta et al. [1] have proposed three main strategies in their paper, 'Robust Object Detection in Challenging Weather Conditions'. The first involves training models on real-world datasets including diverse weather conditions. This is presented as effective; however, adverse weather datasets suffer from insufficient coverage or imbalanced annotations across weather scenarios. The second strategy uses synthetic data augmentation, employing techniques such as physics-based rendering, generative adversarial networks (GANs), and style transfer to simulate adverse weather effects on clear images. These methods are presented as inconsistent in preserving content integrity and realism, which is a large trade-off in model training. The third approach involves image denoising as a preprocessing step, attempting to mitigate weather-induced noise and improve detection accuracy.

This project seeks to evaluate these strategies, focusing on style transfer augmentation as a tool for simulating adverse weather. Specifically, we aim to reproduce the results from Gupta et al. [1] and extend the analysis by attempting additional tuning of the style transfer augmentations. Increasing convolution layers, correcting based on color channels and using two style source images instead of one are proposed as techniques to improve style transfer. Using the BDD100K [2] dataset for training and the DAWN [3] dataset for evaluation, I seek to address the following key questions: How effective are basic image augmentations for training robust detection models? Can synthetic weather augmentations, particularly those generated by style transfer, match the robustness achieved by real-world all-weather datasets?

Through this exploration, I aim to advance robust object detection strategies applicable to real-world scenarios, contributing to the field of autonomous systems and robotics.

## 2. Approach

The proposed approach utilizes the YOLOv5 [4] object detection model to evaluate performance on the DAWN dataset, particularly under challenging weather conditions like fog, rain, and snow. The system is fine-tuned to detect two major object categories: vehicles (combining car, truck, bus, motorcycle, and bicycle) and persons (pedestrian and rider).

## 2.1. Datasets

To use the BDD100k dataset with YOLOv5 some manipulation of the dataset was necessary. In the original dataset vehicle and person related classes are treated as separate classes, to replicate the results from Gupta *et al.* [1]. Cars, buses, trucks, and trains were mapped to class 0 while pedestrians and riders were mapped to class 1. Label files were updated to reflect this mapping and bounding boxes were normalized and centered to match the expected YOLOv5 standards as per Equations 1-4.

$$x = \frac{box[0] + box[2]}{2 \times w} \tag{1}$$

$$y = \frac{box[1] + box[3]}{2 \times w} \tag{2}$$

$$w = \frac{box[2] - box[0]}{w} \tag{3}$$

$$h = \frac{box[3] - box[1]}{w} \tag{4}$$

Images and labels were structured into separate folders nested within respective training and validation folders. Two base imagesets were generated from the. BDD100k dataset depending on the image annotated weather conditions.

1. imageset1: clear, overcast, partly cloudy, unknown
2. imageset2: imageset1 + foggy, rainy, snowy

The BDD100k dataset is made up of 100k annotated images split into 70k images for training, 10k images for validation and 20k images for testing. All 70k images were used for training and 10k images for validation. For testing and evaluating, the DAWN dataset was used instead.

The DAWN dataset is a collection of 1000 images of adverse weather conditions for traffic environments. The dataset includes snow, rain, fog and sand as adverse weather. Sand images were not used as the BDD100k dataset does not have sand weather conditions. The DAWN dataset was also mapped to the vehicle and person related classes, 0 and 1. The dataset was given the following structure to individually evaluate weather conditions:

```
/DAWN/Test/{Weather}/
  images/
    image1.jpg
    image2.jpg
    ...
  labels/
    image1.txt
    image2.txt
    ...
```

## 2.2. Style transfer augmentation

Additionally, imagesets that include the style transfer augmentations are proposed:

1. imageset3: imageset2 + style transfer augmented images
2. imageset4-6: imageset2 + expanded style transfer augmented images

Generating the style transfer augmented images is computationally intensive so only a small sample were generated in this project, specifically for foggy weather. The style transfer augmented images are generated with the finetuned VGG19 model and weights [3] from the original paper. The expanded style transfer augmented images are attempts to improve the style transfer through several processes. These include:

1. Increasing the convolution layers from 5 to 6
2. Mean squared error across the color channels between the generated and original image
3. Using two source style images to transfer style instead of one

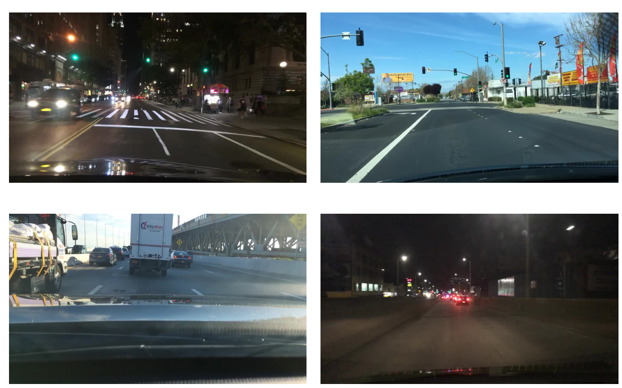
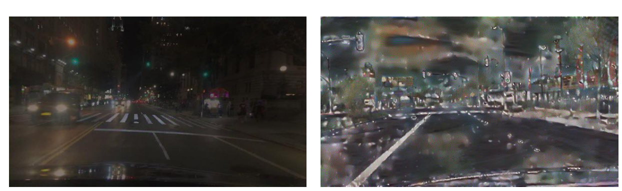


Fig 1. Original images



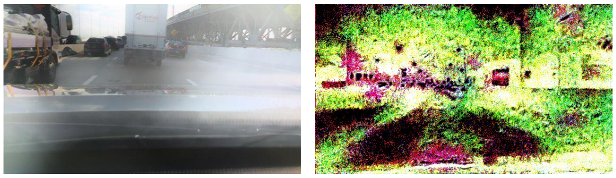


Fig 2. Style transfer augmented images using original paper finetuned VGG19

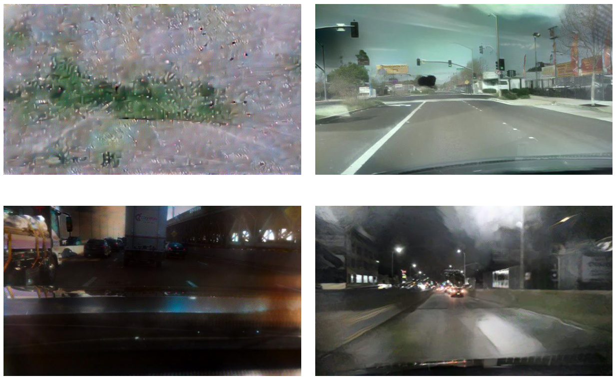


Fig 3. Style transfer augmented images using 6 convolution layers

Fig 4. Style transfer augmented images using mean squared error across color channels

Fig 5. Style transfer augmented images using two source style images + reduced contrast + desaturation

Figures 1-4 present the original images and results of the different style transfer augmentations proposed to augment the training imagesets.

Figure 5 presents the results of using two source style images when performing style transfer augmentation along with reducing the contrast and desaturating the style images. This technique results in the best augmentation at a visual level as none of the images significantly loose detail or have suffer from severe coloring/noise issues. The expectation is for this technique to work best when used to augment training datasets for foggy conditions.

This technique was implemented with the following high-level breakdown:

1. Function to load content and two style images:
   i) Resize content and style images to the same size.
   ii) Reduce contrast and desaturate style images
   iii) Convert images to tensors.
2. Modified model preparation function:
   i) Accept two style images.
   ii) Compute style losses for both images at specified layers.
   iii) Add style losses for the two style images to separate lists.
3. Update the style transfer function:
   i) Accept weights for both style images (style_weight1 and style_weight2).
   ii) Compute loss contributions for the two style images.
   iii) Combine two style losses and content loss into a single loss function.
   iv) Optimize input image to minimize the combined loss.

### 2.3. Training

Two base image augmentation sets were utilized to train the object detection models.

1. Imageaug1: Translations, scaling and. vertical flipping
2. Imageaug2: Imageaug1 + HSV jittering, greyscaling, CLACHE, median blur and mixup.

Imageaug1 augmentations were applied with 0.5 probability. Imageaug2 specific augmentations were applied with 0.01 probability as in the original paper.

The YOLOv5 training pipeline used an image size of 640x640, batch size 60, 50 epochs, SGD optimizer with a learning rate of 0.01 to 0.001 with cosine learning rate decay, momentum 0.937, and weight decay 0.0005. The training was performed on Google Colab Pro with an A100 GPU. Each training run completed in approximately 4 hours with a total of 4 models trained.

| | | | P | R | mAP50 |
|---|---|---|---|---|---|
| imageset1 | Imageaug1 | Fog | 0.819 | 0.52 | 0.597 |
| | | Rain | 0.662 | 0.552 | 0.599 |
| | | Snow | 0.786 | 0.526 | 0.609 |
| | Imageaug2 | Fog | 0.815 | 0.539 | 0.614 |
| | | Rain | 0.78 | 0.557 | 0.608 |
| | | Snow | 0.783 | 0.484 | 0.566 |
| imageset2 | Imageaug1 | Fog | 0.823 | 0.547 | 0.62 |
| | | Rain | 0.781 | 0.625 | 0.686 |
| | | Snow | 0.814 | 0.57 | 0.671 |
| | Imageaug2 | Fog | 0.782 | 0.572 | 0.633 |
| | | Rain | 0.84 | 0.598 | 0.682 |
| | | Snow | 0.819 | 0.603 | 0.671 |

Table 1. Object detection results of 4 YOLOv5 models evaluated on the DAWN dataset

## 3. Results

The results of training 4 YOLOv5 models are presented in Table 1. The precision P, recall R, and mean average precision mAP50 for the different combinations of imagesets and image augmentations are given for three adverse weather conditions (Fog, Rain, and Snow). Comparing these results with the original paper there is a significant difference in recall metrics across all models and weather conditions. In these experimental results the range of recall values goes from 0.484 to 0.625 while the range is 0.690 to 0.788 in the paper.

The best mAP obtained is 0.686 for Rain weather conditions using imageset2 with Imageaug1 (trained with all weather conditions). Figure 5 presents some example predictions from this model. Overall imageset2 resulted in consistently better precision and recall metrics across all adverse weather conditions.

Fig 5. Prediction examples for rainy weather with imsageset2

## 4. Discussion

This project aimed to reproduce the results presented in the paper 'Robust Object Detection in Challenging Weather Conditions' by Gupta et al. [1]. Due to computational limitations, this was only partially achieved, with the baseline metrics of four YOLOv5 models trained with real-world clear-weather condition images and real-world all-weather condition images. While the metrics presented in this work do not exactly match the ones presented in the inspiration paper, the overall metrics suggest the same findings, real-world all-weather datasets generate the best model for object detection in adverse weather conditions. The offset in metrics between this work and the paper likely stems from the fact that the paper results utilized a manually. curated subset of the BDD100K dataset for training. This was not possible to reproduce in this work as there was a time constraint.

From the proposed changes to the style transfer approach for foggy weather, the most promising is utilizing two source style images along with reduced contrast and desaturation preprocessing of the style images.

In future work, the dataset should be augmented with this technique and additional YOLOv5 models should be trained to assess the effectiveness of the technique.

## 5. Conclusion

In conclusion, while the full original scope of the work was not completed, due to computational and time constraints, the style transfer technique augmentation resulted in promising images for training models in future work. This work further denotes the potential in synthetic image augmentations when real-world datasets are insufficient for reliable and safe object detection in adverse weather conditions.

## 6. Statement of contributions

As the sole member of this project, Diego Fernandez was responsible for the following tasks: Dataset preparation, reproducing baseline results with the described training pipeline, applying image augmentation techniques to the BDD100K dataset, enhancing image style transfer augmentation technique.

## References

[1] Gupta, Himanshu and Kotlyar, Oleksandr and Andreasson, Henrik and Lilienthal, Achim J. Robust Object Detection in Challenging Weather Conditions. *2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 7523–7532, 2024.

[2] Fisher Yu, Haofeng Chen, Xin Wang, Wenqi Xian, Yingying Chen, Fangchen Liu, Vashisht Madhavan, and Trevor Darrell. Bdd100k: A diverse driving dataset for heterogeneous multitask learning. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 2636–2645, 2020

[3] Mourad KENK. Dawn: Vehicle detection in adverse weather nature dataset, 2020. 1

[4] Glenn Jocher. YOLOv5 by Ultralytics, May 2020