

A YOLOv8-based Analysis of Image Augmentation Techniques for Vehicle Detection in Adverse Weather Conditions

Alexander Van Hecke (852631385) Frederik Lefever (838836963)

October 30, 2023

Abstract

To do

1 Introduction

Adverse weather conditions such as rainfall and snow are widely considered to have an effect on traffic flow. Traffic breakdown occurs when demand exceeds capacity in some part of a transportation network and in [11] it is shown that the odds of traffic breakdown at bottleneck locations are significantly increased by rainfall. The ensuing economic damage is considerable and provides a strong incentive to mitigate this problem as much as possible.

Part of a mitigation strategy are traffic monitoring and dynamic flow control. The current technology behind traffic monitoring depends essentially on vehicle detection in images. Yet in turn, vehicle detection in images is itself not indifferent to weather conditions.

We propose to investigate the effect of image augmentation on the robustness of YOLOv8-based vehicle detection models. In particular, we propose to investigate the effect of artificially adding rain, fog and snow to real clear weather images of motorways. In comparison with the standard YOLOv8 model, we expect to see an improvement of the Mean Average Precision (mAP) for vehicle detection in models derived from YOLOv8 by training on the augmented images.

The immediate relevance of our findings can be explained in terms of the above mentioned applications to traffic monitoring. But it is not unlikely that some results could be of value to other weather-sensitive arenas, such as surveillance of skiers on mountain slopes.

2 Literature review

In the context of machine learning, image augmentation can broadly be defined as the automated creation of variation in actual image sets. This technique can be used to avoid a learning algorithm overfitting the data. Overfitting occurs when a learning algorithm learns a function in such a way that it perfectly models the training data. Large volumes of data with sufficient variation can alleviate the problem of overfitting. Yet sometimes sampling data from an application domain is nontrivial, eg. when learning from medical images.

A comprehensive survey of modern image augmentation techniques is presented in [9]. Several approaches such as geometric transformations, color space augmentations, kernel filters, mixing images, random erasing, feature space augmentation, adversarial training, generative adversarial networks, neural style transfer, and meta-learning, are treated with mild technicality. [13] offers a more recent survey of image augmentation techniques as used in deep learning. This survey introduces a novel taxonomy, where image augmentation algorithms are classified as either model-free, modelbased, or optimizing policy-based. The objectives of image augmentation are explained from an analysis of challenges encountered when deploying deep learning models for computer vision. A theoretical framework for understanding data augmentation is described in [2]. First a general model of augmentation as a Markov process is provided, and it is shown that kernels appear naturally with respect to this model. Next, a more directly analysis of the effect of augmentation on kernel classifiers is offered.

The effectiveness of image augmentation on the classification of images with deep learning is discussed in [8]. Simple techniques, such as cropping, rotating, and flipping images are compared. Additionally this paper reports on experiments with generative adversarial neural networks to learn augmentation strategies.

Research similar to the proposed research can be found in [5]. The central question in [5] is whether YOLOv8 can be improved through transfer learning to detect objects in adverse weather. This paper supports the hypothesis that training with actual images of adverse weather conditions significantly improves the detection performance compared to the standard YOLOv8 model. An enhanced YOLO-based model for vehicle detection in foggy weather conditions is the focus of [6]. The approach used to create the enhanced model is interesting in that the trainingset is obtained from 350 traffic images that are augmented with a fog-effect and subsequently dehazed with the multi-scale retinex with color restoration (MSRCR) algorithm. The final trainingset is then composed of the original images, fogged images and dehazed images. Whereas results in [6] are reported with autonomous vehicles in mind, [10] specifically targets vehicle detection and counting in highway management. A new segmentation method is proposed to divide a depicted highway road surface into a distal and a proximal area. Using this separation a YOLOv3 model is trained to detect the type and location of vehicles. To estimate vehicle count and trajectories the Oriented FAST and Rotated BRIEF (ORB) algorithm is added to the image processing pipeline.

3 Research questions

The image augmentation techniques described in the literature are numerous and varied, ranging from simple geometric transformations to the addition of noise and the changing of color, saturation and hue parameters. We propose to investigate the effect of thematic image augmentation consisting of the addition of rain, fog and snow to clear weather images. At the more general level, we aim to answer following question:

Given that the standard YOLOv8 model is used as a comparative baseline and trained with thematically augmented images, will the resulting model then be more accurate in predicting the presence of vehicles in actual images of traffic in adverse weather conditions?

Furthermore, to appreciate the value of image augmentation in contexts where data is not necessarily scarce, we pose an additional question:

Given that the standard YOLOv8 model is trained with actual images of traffic in adverse weather conditions, how then does the accuracy of the resulting model compare to the accuracy of a model derived from standard YOLOv8 with augmented images?

4 Research method

4.1 Measuring model accuracy

To estimate the robustness of vehicle detection models, we use the mean average precision (mAP) metric. The mAP is an aggregate metric based on the confusion matrix, the intersection over union (IoU), recall and precision. We consider three relevant classes of vehicles, nl. 'car', 'bus' and 'truck'.

4.1.1 Calculation of the mAP

The mAP is calculated by deviding the sum of the average precision (AP) per class by the number of classes N .

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i$$

In summary, the AP of a model is obtained by following steps:

1. Use model to generate prediction scores
2. Map prediction scores onto class labels
3. Construct the confusion matrix
4. Calculate precision and recall for a set of IoU thresholds
5. Calculate area under precision-recall curve
6. Calculate average precision

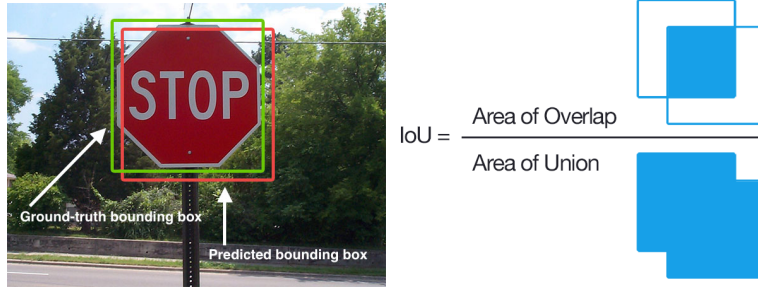


Figure 1: Images from Wikipedia (https://en.wikipedia.org/wiki/Jaccard_index)

The IoU metric is used to evaluate the performance of object detection by quantifying the fit between the ground truth bounding box and the predicted bounding box.

By using a lower bound for the value of the IoU metric, one can discriminate between positive and negative predictions. The IoU metric can then be used to calculate recall and precision. In the context of traffic management, it seems reasonable to slightly favour recall over precision and choose an IoU threshold to reflect this preference. However, choosing a good IoU threshold is itself an optimization process. For comparing models, optimizing the IoU threshold is unnecessary. Instead, we calculate AP as an average of AP’s calculated for a set of IoU thresholds per class.

4.2 Data collection

Two open-source datasets are used for this research. The DAWN dataset [1] is a collection of 1000 images from real-traffic environments in adverse weather conditions. The images are divided into four sets of weather conditions: fog, snow, rain and sandstorms. They are labelled according to the categories relevant for current research (“car”, “bus”, “truck”) and annotated with object bounding boxes. We use the DAWN dataset for both validation and training.

The second dataset is the UA-DETRAC [12] dataset. This dataset provides 140K traffic images taken at Beijing and Tianjin (China). Images of the UA-DETRAC dataset are divided into four weather categories: “cloudy”, “night”, “sunny” and “rainy”. They are labelled “car”, “bus”, “van” and “others”; and object bounding boxes are provided. The dataset is made available as a bi-partition (DETRAC-Train-Images and DETRAC-Test-Images). From the union of the train and test datasets, we only use images of the category “sunny” for data augmentation and training. No other images are used from the UA-DETRAC dataset.

4.3 Model selection

For vehicle detection we chose an open source convolutional neural network called YOLOv8 [3]. We use this model in transfer learning and as a reference for derived models. YOLOv8 is trained on the Microsoft COCO dataset [7] and capable of detecting object categories “car”, “bus” and “truck”.

4.4 Experiments

Since the chosen augmentation software imgaug [4] (version 0.4.0) lacks the capability to add sandstorm effects to images, we start by creating a subset of the DAWN dataset by excluding all images of the weather category “sandstorm”. The resulting subset is then randomly partitioned with a 80/20 ratio. The larger part is denoted by DAWN-train and the smaller part by DAWN-test.

Next we establish a baseline measurement, by presenting the DAWN-test set to the standard YOLOv8 model and determine its mAP over the “car”, “bus” and “truck” classes. Then, we compose the union of the DETRAC-Train-Images and DETRAC-Test-Images and from it extract images of the weather category “sunny”. The extracted images are used to produce three distinct trainingsets by augmenting them with fog, rain and snow using imgaug [4]. The obtained trainingsets are denoted resp. DETRAC-fog, DETRAC-rain and DETRAC-snow. We then train the standard YOLOv8 model with the union of the DETRAC-fog, DETRAC-rain and DETRAC-snow datasets. Finally, we determine the mAP of the new model over the “car”, “bus” and “truck” classes, by presenting it the DAWN-test set. Comparing the mAP of the baseline with the mAP of the new model should provide some ground for answering the first research question.

To answer the second research question, we train the standard YOLOv8 model with the DAWN-train set and measure the mAP of the derived model. Like before, the mAP is calculated over the “car”, “bus” and “truck” classes for the DAWN-test set.

5 Proposed time line

References

- [1] Mourad A. KENK and Mahmoud Hassaballah. *DAWN: Vehicle Detection in Adverse Weather Nature*. 2020. DOI: 10.21227/bw1x-yh39. URL: <https://dx.doi.org/10.21227/bw1x-yh39>.
- [2] Tri Dao et al. “A Kernel Theory of Modern Data Augmentation”. In: *Proceedings of the 36th International Conference on Machine Learning*. PMLR, May 2019, pp. 1528–1537. (Visited on 10/26/2023).
- [3] Glenn Jocher, Ayush Chaurasia, and Jing Qiu. *Ultralytics YOLOv8*. Version 8.0.0. 2023. URL: <https://github.com/ultralytics/ultralytics>.

- [4] Alexander B. Jung. *imgaug*. Version 0.4.0. 2020. URL: <https://github.com/aleju/imgaug>.
- [5] Debasis Kumar and Naveed Muhammad. “Object Detection in Adverse Weather for Autonomous Driving through Data Merging and YOLOv8”. In: *Sensors (Basel, Switzerland)* 23.20 (2023), pp. 8471–. ISSN: 1424-8220. DOI: 10.3390/s23208471.
- [6] Wei Li. “Vehicle Detection in Foggy Weather Based on an Enhanced YOLO Method”. In: *Journal of Physics: Conference Series* 2284.1 (June 2022), p. 012015. ISSN: 1742-6588, 1742-6596. DOI: 10.1088/1742-6596/2284/1/012015. (Visited on 09/23/2023).
- [7] Tsung-Yi Lin et al. *Microsoft COCO: Common Objects in Context*. Feb. 2015. DOI: 10.48550/arXiv.1405.0312. arXiv: 1405.0312 [cs]. (Visited on 10/29/2023).
- [8] Luis Perez and Jason Wang. *The Effectiveness of Data Augmentation in Image Classification Using Deep Learning*. Dec. 2017. DOI: 10.48550/arXiv.1712.04621. arXiv: 1712.04621 [cs]. (Visited on 10/23/2023).
- [9] Connor Shorten and Taghi M. Khoshgoftaar. “A Survey on Image Data Augmentation for Deep Learning”. In: *Journal of Big Data* 6.1 (Dec. 2019), p. 60. ISSN: 2196-1115. DOI: 10.1186/s40537-019-0197-0. (Visited on 10/27/2023).
- [10] Huansheng Song et al. “Vision-Based Vehicle Detection and Counting System Using Deep Learning in Highway Scenes”. In: *European Transport Research Review* 11.1 (Dec. 2019), p. 51. ISSN: 1866-8887. DOI: 10.1186/s12544-019-0390-4. (Visited on 10/28/2023).
- [11] Wouter J. H. Stralen, Simeon C. Calvert, and Eric J. E. Molin. “The Influence of Adverse Weather Conditions on Probability of Congestion on Dutch Motorways”. In: *European Journal of Transport and Infrastructure Research* 15.4 (2015). ISSN: 1567-7141. DOI: 10.18757/ejtir.2015.15.4.3093.
- [12] Longyin Wen et al. “UA-DETRAC: A New Benchmark and Protocol for Multi-Object Detection and Tracking”. In: *Computer Vision and Image Understanding* (2020).
- [13] Mingle Xu et al. “A Comprehensive Survey of Image Augmentation Techniques for Deep Learning”. In: *Pattern Recognition* 137 (May 2023), p. 109347. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2023.109347. (Visited on 10/23/2023).