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

Alexander Van Hecke (852631385) Frederik Lefever (838836963)

November 8, 2023

Abstract

Vehicle detection is an important aspect of semi-automated traffic monitoring and surveillance. We propose to investigate whether the technique of image augmentation can be used to increase the accuracy of vehicle detection in adverse weather conditions. The study will evaluate whether the use of image augmentation to train a YOLOv8 model increases accuracy compared to a baseline model. We also propose to compare the effect of training with augmentated images to that of training with actual images of traffic in adverse weather conditions. To this extent we will train a YOLOv8 model with actual images and compare with the same baseline model and the accuracy obtained from the first model. Our main hypothesis is that image augmentation can indeed be used to improve the accuracy of vehicle detection in adverse weather conditions. Our secondary hypothesis is that the accuracy of a model derived from augmented images falls within statistically insignificant margins when compared to a model obtained from actual images.

1 Introduction

Adverse weather conditions such as rainfall, snow and fog 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. This can lead to considerable economic damage and provides a strong incentive to mitigate this problem as much as possible.

Traffic monitoring and dynamic flow controle can be part of a mitigation strategy. The current technology behind traffic monitoring largely uses vehicle detection in images captured from CCTV cameras positioned next to highways. However, vehicle detection in images can be influenced by adverse 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 highways. Compared to 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. We chose YOLOv8 as our base model because of its modernity and popularity. Also, YOLOv8 is well-documented, both as a ready-to-use product and in scientific literature. Finally, YOLOv8 is an open-source software, licensed under AGPL-3.0. All artifacts produced by the proposed study will hence be made available through a public GitHub repository.

The immediate relevance of our findings can be explained in terms of the above mentioned applications to traffic monitoring. But it is possible that some results could be of value in other weather-sensitive applications, 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 datasets. 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, but is unable to generalize beyond 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. Image augmentation can be used to artificially create additional images, or to add some sort of noise to images in order to make the resulting model more robust.

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 explained. [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, model-based, or optimizing policy-based. The objectives of image augmentation are explained by analyzing the 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 given, and it is shown that kernels appear naturally with respect to this model. Next, a more direct 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 YOLOv8-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 in the addition of rain, fog and snow to clear weather images. At the more general level, we aim to answer following questions:

RQ 1 *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 as measured by the mean average precision (mAp) at 0.50-0.95 IoU, in predicting the location 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 two additional questions:

RQ 2 *Given that the standard YOLOv8 model is used as a comparative baseline and trained with actual images of traffic in adverse weather conditions, will the resulting model then be more accurate as measured by the mean average precision (mAP) at 0.50-0.95 IoU, in predicting the location of vehicles in actual images of traffic in adverse weather conditions?*

RQ 3 *Given a model obtained by training the standard YOLOv8 model with thematically augmented images of traffic in adverse weather conditions, and a model obtained by training the standard YOLOv8 model with actual images of traffic in adverse weather conditions, will both models then predict the location of vehicles in actual images of traffic in adverse weather condition with similar accuracy as measured by their respective mean average precision (mAP) at 0.50-0.95 IoU?*

4 Research method

4.1 Measuring model accuracy

To estimate the robustness of vehicle detection models, we use the mean average precision (mAP) at 0.50-0.95 IoU, henceforth simply referred to as “mAP”. The mAP is an aggregate metric based on the confusion matrix, the intersection over union (IoU), recall and precision. In this study, we only consider two relevant classes of vehicles, namely “car” and “bus”.

4.1.1 Calculation of the mAP at 0.50-0.95 IoU

The mAP is calculated by dividing the sum of the average precision (AP) per class by the number of classes N. In our study $N = 2$ (“car” and “bus”).

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

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 (0.50-0.95)
5. Calculate area under precision-recall curve
6. Calculate average precision

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.

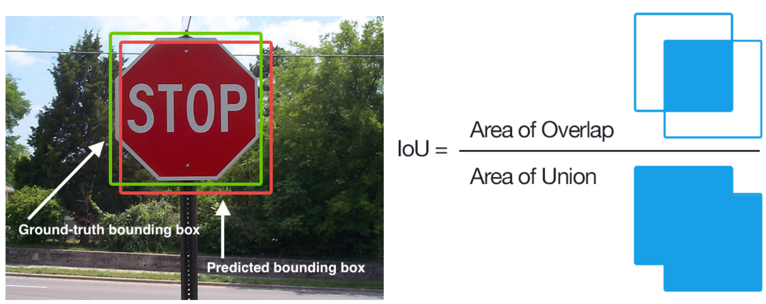

Figure 1: Images from Wikipedia (`https://en.wikipedia.org/wiki/Jaccard_index`)

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" and "bus") 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 annotated with object bounding boxes. The dataset is split into a training dataset (DETRAC-Train-Images) and an evaluation dataset (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. See Table 1 for an overview.

dataset	labels	filter	number of images
DAWN	car, bus	fog, snow, rain	
EU-DETRAC	car, bus	sunny	

Table 1: Datasets used

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" and "bus".

4.4 Experiments

The chosen augmentation software imgaug [4] (version 0.4.0) cannot add sandstorm effects to images, so 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" and "bus" 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" and "bus" 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 dataset and measure the mAP of the derived model. Like before, the mAP is calculated over the “car” and “bus” classes for the DAWN-test set.

A schematic summary of the proposed research process is given in figure 2 and tables 2 and 3.

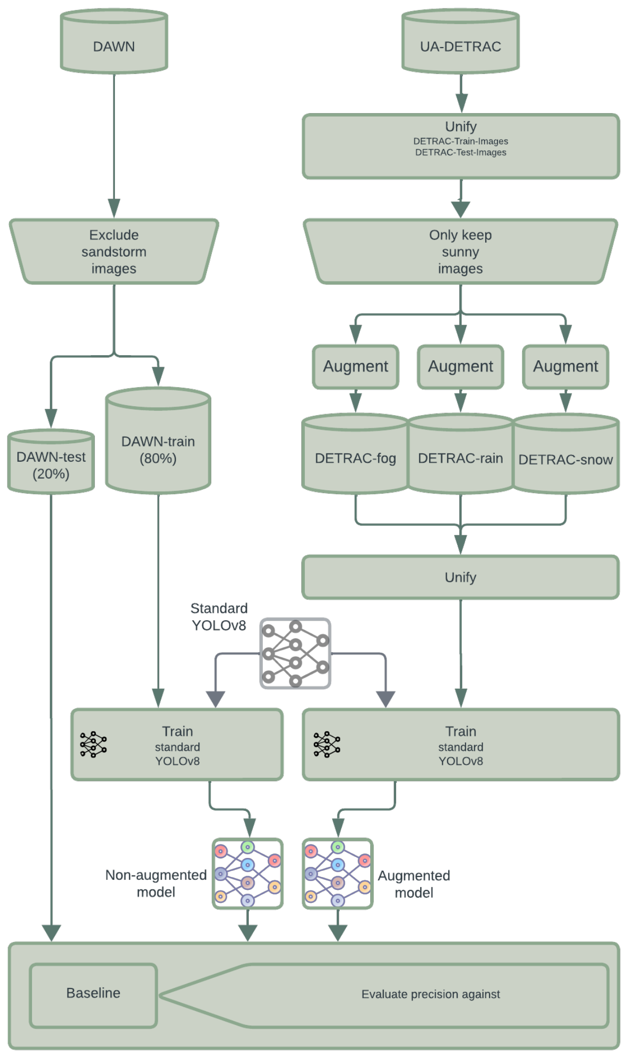

Figure 2: Schematic representation of proposed research

measurement	**training**	**data set**
baseline	none	DAWN-test
augmented images	DETRAC-train	DAWN-test

Table 2: setup RQ 1

measurement	**training**	**data set**
baseline	none	DAWN-test
actual images	DAWN-train	DAWN-test

Table 3: setup RQ 2

5 Data analysis

Our approach to the proposed study will be that of a single case mechanism experiment. More specifically, we investigate the effect of differences in an independent variable X being a model derived by transfer learning from YOLOv8 on a dependent variable mAP. Since we only consider three models (standard YOLOv8, a model derived from augmented images and a model derived from actual images), the study does not produce a volume of data large enough to use statistical techniques.

Our scientific hypothesis related to research question 1 is that, for the detection of cars and busses in images of traffic in adverse weather conditions, the mAP at 0.50-0.95 IoU of a model obtained by transfer learning from the standard YOLOv8 with thematically augmented images, is higher than the mAP at at 0.50-0.95 IoU of the standard YOLOv8.

Related to research question 2, our scientific hypothesis is similar to the hypothesis related to research question 1. Ie, for the detection of cars and busses in images of traffic in adverse weather conditions, the mAP at 0.50-0.95 IoU of a model obtained by transfer learning from the standard YOLOv8 with *actual* images, is higher than the mAP at 0.50-0.95 IoU of the standard YOLOv8.

The statistical significans of the observed differences in accuracy must be left undecided in this study. Yet we will attempt to give an explanation of the causal relationship between image augmentation and improved accuracy of prediction.

To conclusively answer research question 3 in which we compare the accuracy of a model learned from augmented images with that of a model learned from actual images, our study provides insufficient evidence. Nevertheless, we will shortly present our intuitions based on our understanding of neural networks.

6 Proposed time line

The following tasks will have to be addressed (see Figure 3) :

- Dataset preparation
- Images augmentation
- Training models

- Evaluation of models
- Data analysis
- Write the report

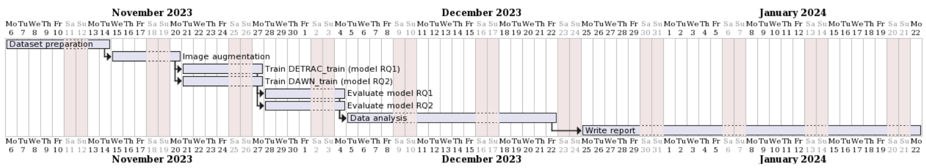

Figure 3: Proposed timeline

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).