Bike helmet detection using resizing

Luco Buitink

Supervisors: Maya Aghaei Gavari, Willem Dijkstra

Summer 2023

Introduction

- Many cyclists in the Netherlands are exposed to the dangers of traffic, with cars and trucks posing a significant risk to their safety.
- The low rate of helmet usage increases the risk of traffic incidents, especially among elderly cyclists who have slower responsiveness.
- Training on an open image dataset provides a diverse range of helmet images, improving generalization and avoiding data collection biases.
- Optimizing the resize parameter values during training enhances the model's ability to detect helmets accurately and efficiently.

Materials and Methods

- Two datasets are used in this research: Open Image v7 and a custom dataset.
- The Open Image v7 dataset is used to train helmet detection models.
- A custom-made dataset consisting of real-world data is used to test the model's performance in real-life scenarios.
- To train the model, a YOLOv5 object detector is employed [1].
- Various experiments with various resize parameter values have been done.

Abstract

We provide an object detection model that demonstrates the usability of deep learning in helmet detection. Two different datasets are utilized. The Open Image v7 dataset and a custom-made dataset. The custom-made dataset comprises three different scenarios that receive different outcomes. Each of the scenarios comprises three different cameras. The different cameras per scenario are all focused on the same place. Different resize parameter values are chosen based on the mean of the data and the maximum values. Results show that resizing to the mean values is the better option. Besides that, the results show that there is a significant difference between the cameras per scenario.

Figure 1. Open Image v7 dataset image

Figure 2. Custom dataset image

- The Open Image v7 dataset is used to choose the best YOLOv5 model for the experiments.
- The resize parameter values of the model need to be divisible by 2 at least 7 times. The closest value to the mean of the heights is 768 and for the widths it is 1024.

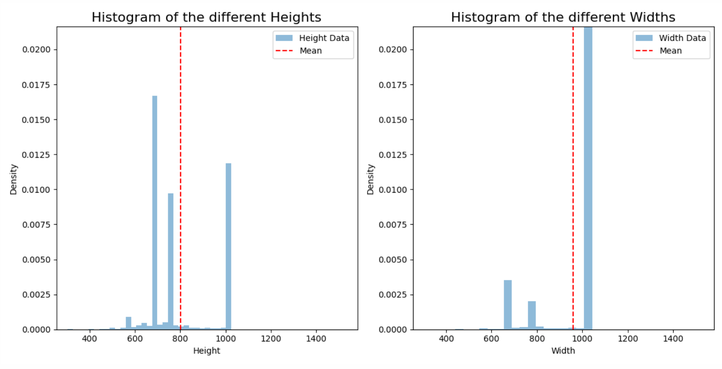

Figure 3. The distribution of the resize parameter values.

	Images	Helmets
Bikeroad	1579	363
Cam 1	817	184
Cam 2	310	117
Cam 3	452	62
Roundabout	2268	1216
Cam 1	1066	765
Cam 2	470	179
Cam 3	732	272
Campus	7285	3936
Cam 1	2431	1251
Cam 2	2428	1550
Cam 3	2426	1135
Total	11132	5515

Figure 4. The amount of images per location per camera.

Experiments and Results

Experiment: Different resizing parameter values.

- Each model is tested on the three different cameras per location.
- Baseline F1-score on Open Image Dataset for 768 x 1024 = 81,5% while the F1-score for 1024 x 1024 = 76,6%

	768p	1024p
Bikeroad	**F1score**	**F1score**
cam1	32,00%	25,00%
cam2	5,40%	0
cam3	61,50%	28,60%
Roundabout		
cam1	2,90%	0
cam2	19,50%	5,30%
cam3	12,50%	0
Campus		
cam1	45,90%	37,90%
cam2	12,50%	15,40%
cam3	46,70%	8,30%

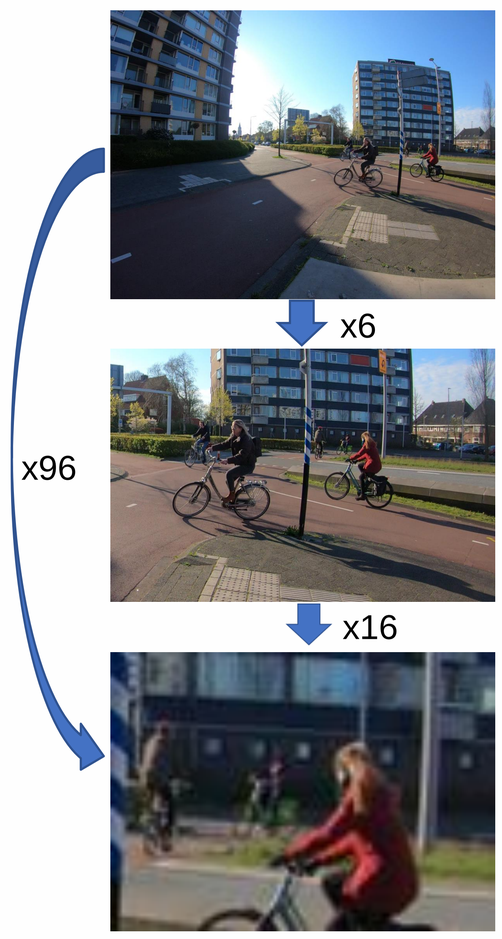

Figure 5. Example of difficult helmets

- The experiment shows that resizing the parameters to 768x1024 respectively, performs better than resizing to 1024x1024. The amount of interpolation for the smaller images outweighs the compressing for the bigger images.
- The best performing camera is cam3 of the bikeroad scenario while the worst performing camera, within the same resize parameter values, is cam1 from the roundabout scenario. The following figures demonstrate a comparison between the two cameras.

Figure 6. A result of camera 3 of the bikeroad scenario

Figure 7. A result of camera 1 of the roundabout scenario

Conclusions

- Resizing to the mean dimensions is better than resizing to the maximum dimensions.
- Helmets within a larger distance of the camera are significantly more difficult to detect.
- Straight camera view is the best.

Acknowledgements

- *This project is a collaboration between the Computer Vision & Data Science professorship and Provinsje Fryslân within the ROF (Regionaal Orgaan verkeersveiligheid Fryslân) programme.*

NHL STENDEN computer vision & data science

provinsje fryslân provincie fryslân

ROF REGIONAAL ORGAAN VERKEERSVEILIGHEID FRYSLÂN

References

[1] A Forest Fire Detection System Based on Ensemble Learning - Scientific Figure on ResearchGate. https://www.researchgate.net/figure/The-network-architecture-of-Yolov5-It-consists-of-three-parts-1-Backbone-CSPDarknet_fig1_349299852