Real-Time Traffic Sign Recognition on the GTSDB dataset

Abstract

This paper deals with traffic sign recognition using Yolov3 and Darknet on Tensorflow. At first, an introduction into the field of neural networks is given which will explain, what is desirable, and which are the factors for performance evaluation.

Following the introduction, related work will be presented, showing the achievements of comparable networks and giving an overview about how yolo performs in comparison to those other networks. Also, the predecessors of Yolov3 are briefly presented and a short overview over the changes is given. Every competitor is presented in a short manner to give a better understanding of which key aspects the individual networks present.

After that, the utilized main data set ‘GTSDB’ is presented. It contains 900 images of which 600 are to be used for training and 300 for validation purposes. It contains ground truth annotations in the form of class id and bounding boxes. In total, there are 43 classes to this set. We also mention ‘GTSRB’ which is used for better classification in Experiment II:

The methodology chapter focusses more on the features of yolov3 compared to yolov2 and how feature maps and upscaling is used to increase performance in detecting small, medium and big objects.

Afterwards, the experimental results are discussed. They contain different aspects such as the experiment setup being Yolov3 with Darknet 53 using Python, Tensorflow, Keras and CUDA. Hardware specifications are mentioned for comparison purposes. It is also explained, why the ‘GTSDB’ was used. Then, it is shown how data must be prepared to fit the existing network and its interfaces.

Leading onwards to the evaluation of the network’s performance, the paper shows that the network fulfills its purpose. The results are not in the range of top-tier-networks, but show that the concept works. The mAP of the network is discussed as well as its runtime per image.

Thereafter, the experimental results are discussed and what might have led to the network’s poor performance when being compared to top-tier-networks. Also, some statistics are shown on how the loss of the network within a section of training develops. Additionally, class distribution in general is shown and how classes are detected by the trained network.

The conclusion shows that our networks worked, even though there is potential for a lot of improvement. We suggest to use another, bigger dataset to gain better training results or to implement a custom classifier for better object detection.