Automatic quantification of traffic safety with multiple object tracking using deep learning

Moustafa Elhagaly
Supervisors: Klaas Dijkstra & Lucas Ramos

Winter 2022

Abstract

We train an object tracker using different experimental setups where we use single, combined and sequenced datasets. The best IDF1-score (82.9%) is achieved when training on the TID dataset. The lowest IDF1-score (0.0%)was achieved when training on MOT due to different in video characteristic. Combining SCD and TID results in 80.8%. Fine-tuning TID on MOT results in an IDF1-score of 82.3%. Finally, training SCD and fine-tuning on MOT and fine-tune that on TID results in an IDF1-score of 81.8%.

Introduction

- Approximately 1.3mln people die each year as a result of road traffic accidents[1]
- Current traffic safety evaluation require expert knowledge analysing all footage and materials making it a time consuming task, therefore, conclusion are often drawn based on the analysis of shorter video fragments
- Crossing a red light is an important traffic safety indicator
- Advances in object detection have been widely used in object tracking, commonly referred as tracking-by-detection
- We test the effect of domain adaptation by running experiments using single and combined dataset using public datasets (MOT17 and SCD) and TID dataset
- We test and measure the effect of catastrophic forgetting by applying sequence learning where a model is trained on one dataset and fine-tuned on another dataset
- We explore the effect of training orders to see how if influences the performance of the tracker when changing dataset order

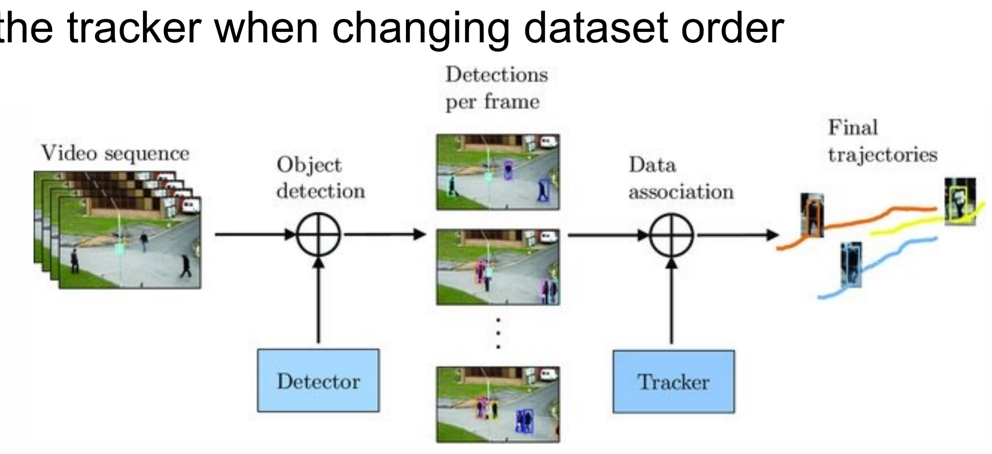

Figure 1. Tracking-by-detection paradigm

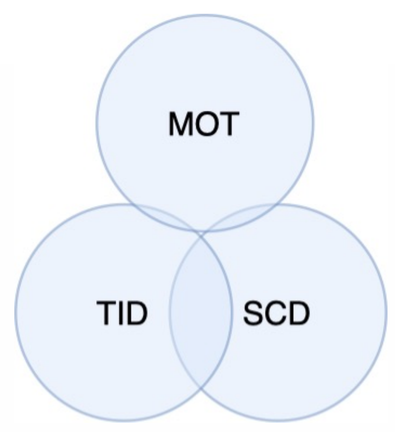

Figure 2. Domain overlap

Materials and Methods

Network:
- SiamMOT[2] that uses Faster-RCNN for object detection and Explicit Motion Modelling that explicitly learns a matching function between the same instance in sequential frames

Datasets:
- Multiple Object Tracking (MOT17)
- Specialized Cyclists Dataset (SCD)
- Traffic Intersection Dataset (TID)

Figure 3. Images from SCD (T), MOT (L) and TID (R)

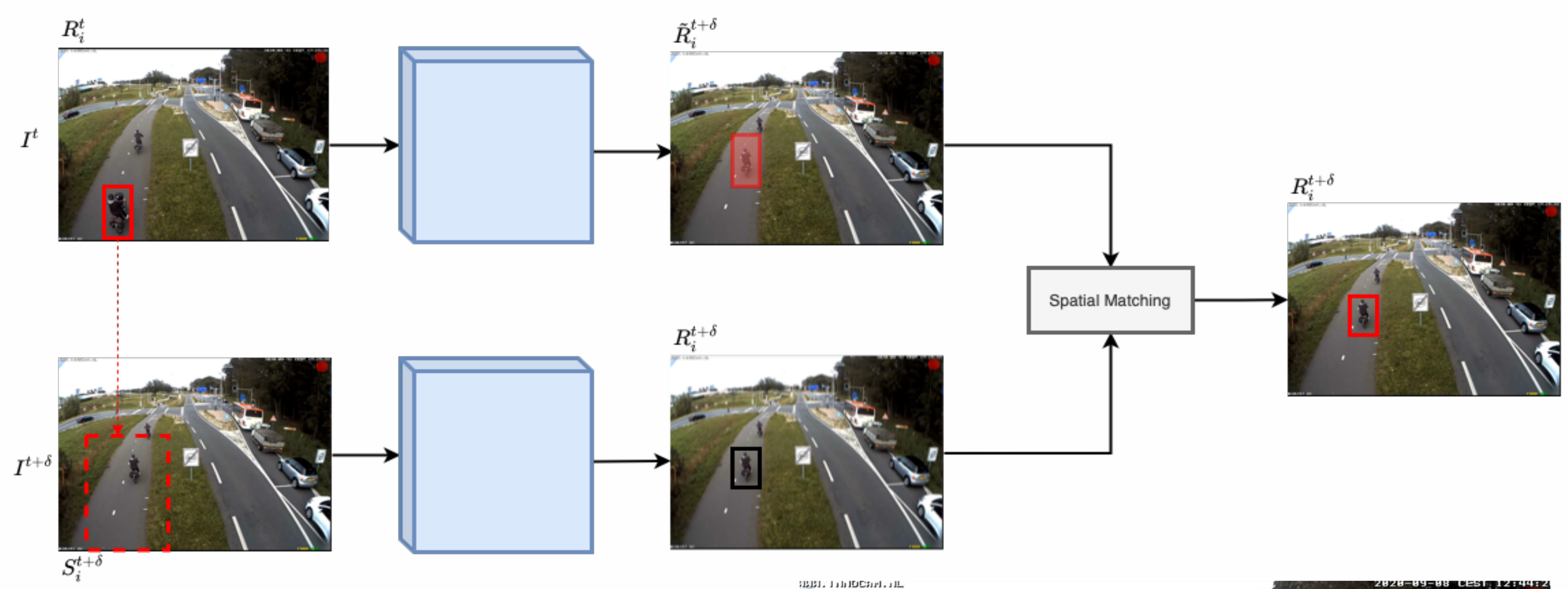

Figure 4. Architectural overview of SiamMOT

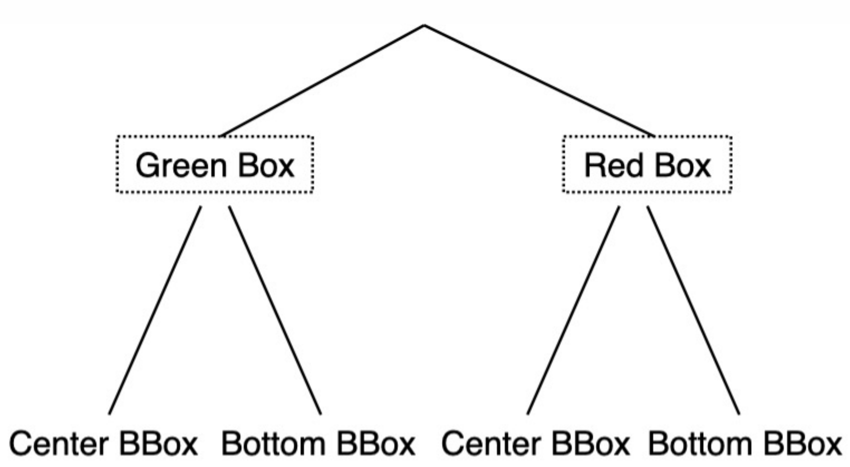

Figure 5. Combinations for determining the most reliable traffic light crossing point

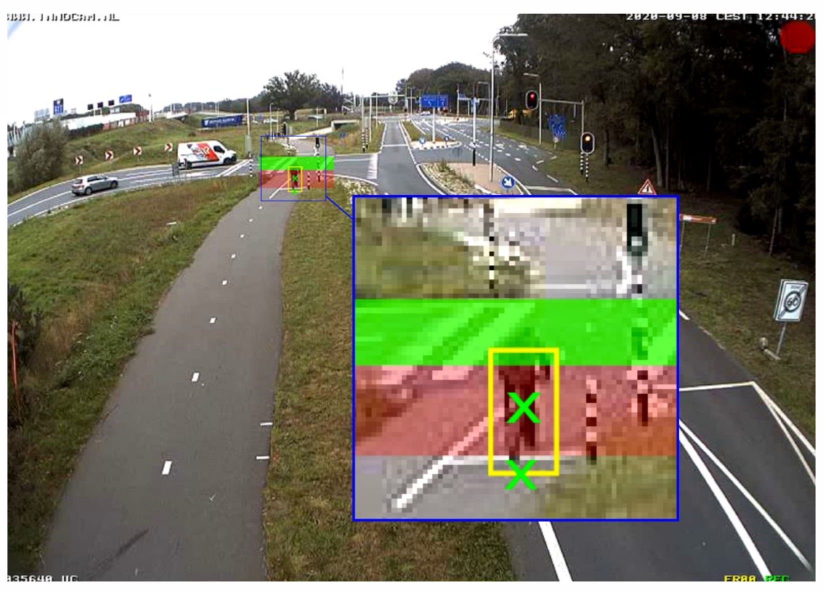

Figure 6. Identification of crossing a red-light area

Experiments and Results

- **Tracker and dataset permutations**
  - The best IDF1-score (82.9%) is achieved by training on the TID dataset;
  - The effect of MOT is visible when comparing the single and combined experiments;
  - The best result of the sequence experiment is when fine-tuning MOT → TID
  - The importance of dataset order when training is important as shown when training SCD → MOT → TID resulting in an IDF1-score of 81.8% and reversing the order of the MOT and TID results in an IDF1-score of 31.8%;

Experiment	Trainset	IDF1 ↑
Single	SCD	27.2%
Single	MOT	0.0%
Single	TID	**82.9%**
Combined	SCD + MOT	11.1%
Combined	MOT + TID	79.9%
Combined	SCD + TID	**80.8%**
Combined	SCD + MOT + TID	54.7%
Sequence	SCD → MOT	6.3%
Sequence	SCD → TID	68.7%
Sequence	MOT → TID	**82.3%**
Sequence	MOT → SCD	33.2%
Sequence	TID → SCD	39.3%
Sequence	TID → MOT	41.8%
Sequence	SCD → MOT → TID	**81.8%**
Sequence	SCD → TID → MOT	31.8%
Sequence	MOT → SCD → TID	78.9%
Sequence	MOT → TID → SCD	32.9%
Sequence	TID → SCD → MOT	46.8%
Sequence	TID → MOT → SCD	31.0%

Table 1. IDF1 score off the different permutations

- **Tracker evaluation with traffic light detection**
  - The best model (single SCD) is used and results in long tracks allowing for the tracker to track objects past the traffic light which in conjunction with traffic light state allows for crossing red light detection;

Figure 7. Results of crossing red light results in daylight and dusk

- **Tracker evaluation on different times of day**
  - The best model is able of tracking the bicycle instances in daylight, dusk/twilight;
  - The model is capable of tracking some of the bicycles during nighttime;

Figure 8. Results of tracker during nighttime

Conclusions

- The best results are achieved when training the tracker on the TID dataset resulting in an IDF1-score of 82.9%;
- Combining TID + MOT and TID + SCD results in a high IDF1 of 79.9% and 80.8% respectively;
- The effect of the SCD and MOT datasets can clearly be seen from the results:
  - MOT and SCD datasets has a few dissimilarities compared to TID
- The sequence experiment shows the importance of training order as seen when training SCD → MOT → TID (IDF1-score: 81.8%) and SCD → TID → MOT (IDF1-score: 31.8%)

Acknowledgements

This project is financially supported by Regieorgaan SIA (part of NWO) under the KIEM SI project STATS

NHL STENDEN computer vision & data science

Regieorgaan SIA

MEETEL

PROVINCIE UTRECHT

References

[1] Global status report on road safety 2018. Retrieved from http://luchemos.org.ar/images/Extranjero/Globalstatusreportonroadsafety2018.pdf

[2] Bing Shuai, Andrew G. Berneshawi, Xinyu Li, Davide Modolo, and Joseph Tighe. Siammot: Siamese multi-object tracking. In IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021, pages 12372–12382. Computer Vision Foundation / IEEE, 2021.