

Supplementary Information to  
Computational Desire Line Analysis of Cyclists  
on the Dybbølsbro Intersection in Copenhagen

Simon Martin Breum<sup>a</sup>, Bojan Kostic<sup>a</sup>, and Michael Szell<sup>\*a,b,c</sup>

<sup>a</sup>NETwoRks, Data, and Society (NERDS), IT University of Copenhagen, 2300  
Copenhagen, Denmark

<sup>b</sup>ISI Foundation, 10126 Turin, Italy

<sup>c</sup>Complexity Science Hub Vienna, 1080 Vienna, Austria

## S1 Object detection and object tracking

Although the analysis of the object detection and tracking algorithms is not part of our study, as we already start with the given trajectories, it is useful to know how those trajectories were generated. The objects (bicycles) for the trajectories were detected with YOLOv5, trained on a 10 hour high-definition 4K video ( $3840 \times 2160$ ) from the same day (2022-06-09 07:00-17:00), where the studied trajectories come from the first hour of the video. The trajectories were then tracked with SORT, see below.

We used transfer learning and a pre-trained object detection model to train the detection model to fit our use case. The pre-trained model was trained on the Microsoft COCO (Common Objects in COntext) data set (Lin et al., 2014), which contains about 2.5 million labeled instances in 328,124 images, namely 165,482 train, 81,208 validation, and 81,434 test images, with 91 classes such as dog, cat, and several traffic-related classes, such as person, bicycle, car, motorcycle, bus, truck. To detect all traffic participants with particular local characteristics, we trained the model to detect the following nine classes: pedestrian, bicycle, cargo bike, motorcycle, car, van, light truck, heavy truck, bus.

The YOLOv5 (<https://github.com/ultralytics/yolov5>) (Redmon & Farhadi, 2018) object detection algorithm was used. It is one of the most used algorithm due to its speed and accuracy. YOLO (“You Only Look Once”) works by dividing the image into a grid system, where each grid cell detects objects inside it. YOLOv5 is provided in several versions, with different model architectures and image processing sizes. We used version YOLOv5l6, which is a large YOLOv5 model (76.8 M parameters) with image sizes of  $1280 \times 1280$  pixels. We chose the larger model as it yields better accuracy, which is a key factor influencing tracking performance.

The SORT (Simple Online and Realtime Tracking) (Wojke et al., 2017) multiple object tracking algorithm was then used to track the objects that were previously detected by YOLOv5. The output of the tracking algorithm are multiple object trajectories. The task of correctly associating detected objects throughout video frames is a key challenge in tracking. SORT handles that by combining Kalman filters (KF) for motion prediction of the bounding boxes, and the Hungarian algorithm for data association.

---

\*Corresponding author. Email: [misz@itu.dk](mailto:misz@itu.dk)

## S2 Path-clustering with dynamic time warping

Dynamic time warping (DTW) is a distance measure that can be used on two temporal sequences which may vary in speed (Berndt & Clifford, 1994; Müller, 2007). The method allows warping time to account for variances in speed. Similar paths will thus have a low distance and very different paths will have a very large distance. For any two given time-series of observations, the DTW-distance is the summed distance between pairs of data points such that:

- Each observation can be paired only with observations from the other time-series and must be paired with at least one data point (but can be paired with multiple).
- The start points of the time-series are paired (they may be paired with other points too).
- The end points of the time-series are paired (they may be paired with other points too).
- Pairings must be continuous — For any pairing such that observation  $o$  in the time series  $S_1$  is paired with observation  $p$  in time series  $S_2$ , any observation in  $S_1$  that is before  $o$  cannot be paired with an observation in  $S_2$  that is after  $p$ , and vice versa.

This paper uses DTW distance to calculate similarity between paths with similar source and destination. The distance is used to hierarchically cluster paths. For details on the implementation see our code repository.

## References

- Berndt, D. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. *KDD workshop*, 10(16), 359–370.
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. *European conference on computer vision*, 740–755. [https://doi.org/https://doi.org/10.1007/978-3-319-10602-1\\_48](https://doi.org/https://doi.org/10.1007/978-3-319-10602-1_48)
- Müller, M. (2007). Dynamic time warping. *Information Retrieval for Music and Motion*, 2, 69–84. [https://doi.org/10.1007/978-3-540-74048-3\\_4](https://doi.org/10.1007/978-3-540-74048-3_4)
- Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. <https://doi.org/https://doi.org/10.48550/arXiv.1804.02767>
- Wojke, N., Bewley, A., & Paulus, D. (2017). Simple online and realtime tracking with a deep association metric. *2017 IEEE International Conference on Image Processing (ICIP)*, 3645–3649. <https://doi.org/10.1109/ICIP.2017.8296962>