Denne artikel er udgivet i det elektroniske tidsskrift
Artikler fra Trafikdage på Aalborg Universitet
(Proceedings from the Annual Transport Conference at Aalborg University)
ISSN 1603-9696
https://journals.aau.dk/index.php/td

Feasibility of detecting cyclists' phone use with computer vision

Felix Wilhelm Siebert, felix@dtu.dk, Technical University of Denmark
Christoffer Riis, chrrii@dtu.dk, Technical University of Denmark
Jakob Kristensen, s173174@student.dtu.dk, Technical University of Denmark
Oguzhan Gül, s194273@student.dtu.dk, Technical University of Denmark
Kira Hyldekær Janstrup, kija@dtu.dk, Technical University of Denmark
Hanhe Lin, h.lin2@rgu.ac.uk, Robert Gordon University
Francisco Camara Pereira, camara@dtu.dk, Technical University of Denmark
Frederik Boe Hüttel, fbohy@dtu.dk, Technical University of Denmark

Abstract
Denmark has good cycling infrastructure, but cycling lanes are getting more crowded, increasing the need for safe behaviour in the cycling lane. Mobile phone use increases the risk for crashes significantly, but there is no automated way to detect it. In this study, the feasibility of computer-vision based algorithms for the detection of phone use was investigated. An observation camera was placed in the vicinity of cycle lanes of the greater Copenhagen area, collecting video data at three observation sites. A subset of this data was randomly selected and screened for cyclists. To test the feasibility of the automated detection of cyclists' phone use, cyclists were annotated, i.e., active cyclists, and their phone use were registered in a machine-readable format. A state-of-the-art object detection algorithm was then trained to automatically detect cyclists and their phone use in video data, after which the trained algorithm was applied to a test data set, containing video data which had been excluded in the training process. The results show a high accuracy in the general detection of active bicycles. Of all 126 bikes present in the test data 92.4.% were detected, i.e., only 8 out of the 126 bikes were not detected. In addition, in 14% of the detections, a bike is detected by the algorithm, although there was no bike present. Of the two instances of phone use in the test data set, both were detected (recall=1.0). Due to a liberal detection criterion, two additional instances of phone use were detected by the algorithm, although the cyclists did not use a phone (precision=0.5). This study shows that the automated detection of cyclists' phone use is generally possible, using state-of-the-art computer vision algorithms. Further research needs to be conducted to make the detection more robust and decrease the number of false positive detections.

Background
Denmark has one of the best cycling environments of Europe, as dedicated infrastructure is available that separates bicyclists from motorised vehicles. At the same time, bicycle lanes are getting more crowded with large cargo bikes, delivery service bikes, and new forms of mobility such as e-scooters and e-bikes. This general trend has been further intensified by the COVID-19 pandemic, which has been driving an increase in citizens' bike use in Denmark and in the whole EU (Buehler & Pucher, 2021). While the use of cycling

infrastructure by many road users is a positive development, it poses new safety-challenges for the cyclists in Denmark – crowded cycle infrastructure is less forgiving for maladjusted road user behaviour within the cycling lane (Useche, Alonso, Montoro, & Esteban, 2019).

Phone use during cycling has been linked to decreases in cycling performance and significant increases in cyclists crash risk (De Waard, Lewis-Evans, Jelijs, Tucha, & Brookhuis, 2014; Gioldasis, Christoforou, & Seidowsky, 2021; Stavrinos, Pope, Shen, & Schwebel, 2018). Further, it is especially frequent in young road users (Goldenbeld, Houtenbos, Ehlers, & De Waard, 2012), which are already the highest risk group on the road (Martínez-Ruiz, et al., 2014). With ever increasing ownership of smartphones, cyclists' phone use is unlikely to decrease in the near future. And while handheld use of phones for calls has decreased due to the ready availability of hands-free-calling equipment (e.g., Bluetooth earphones), handheld use for e.g., texting or navigating, has increased significantly within the last 10 years for cyclists (de Waard, Westerhuis, & Lewis-Evans, 2015). Furthermore, researchers have found that only 60% of cyclists in Denmark know about the laws against mobile phone use while cycling (Brandt, Haustein, Hagenzieker, & Møller, 2022), and this unawareness might further increase phone-use numbers. To better understand factors related to phone use by cyclists in Denmark, a comprehensive evidence-base of phone use prevalence is needed. Unfortunately, no automated system for the detection of phone use in the road system is available, preventing large scale data collection of this crucial variable. In this study, the feasibility of a state-of-the-art computer vision algorithm for the automated registration of cyclists' phone use is explored. Computer vision has successfully been applied for the application of other safety related variables, such as motorcycle helmet use (Siebert & Lin, 2020; Lin, Chen, & Siebert, 2021; Lin, Deng, Albers & Siebert, 2020), but has not been as frequently applied for the registration of safety-related behaviour of cyclists.

Method

Three observation sites in the greater Copenhagen area were chosen for video data collection. Sites were selected with two prerequisites in mind: a) existing video observation infrastructure and b) cycling density. As video observation in traffic is subject to considerations of road users' privacy rights, only sites were considered which already have video observation systems installed by the municipality. These sites have existing signage (Figure 1.), informing road users about the presence of road observation video cameras. The camera system consisted of a generic action camera powered by a 20,000mAH power bank, enclosed in a grey box. The box was attached to the roadside infrastructure, using a strapping system at a height of approximately 3 metres (Figure 1.). The videos were recorded with a resolution of 1920x1080 and a minimum frame rate of 30 fps and saved to a micro-sd card in the camera. After the data collection, videos were scaled down to a resolution of 960x540 and a frame rate of 5 fps. Each video was compressed using H265 compression. Adjacent areas in the camera's view were cropped out, such that the final video only captures the activity on the bike-path, not on the road and opposite side of the road.

Figure 1. Left: Signage for public video observation installation by the Copenhagen Municipality. Right: Camera system placement at two of the data collection sites (camera marked with a grey rectangle).

To select a subset of video data which contains cyclists, a pre-trained out of the box object-detection algorithm YOLO (Jocher et al., 2022; Redmon et al., 2016) was used to identify video frames, in which a bicycle (potentially without a rider) was present. The identified frames were then annotated in the following way: so called bounding boxes, i.e. rectangular boxes, were drawn around the active cyclists, encompassing the bicycle and the rider. In a second step, the phone use of the bicycle rider was annotated as either "yes" or "no". In total, 715 cyclists were identified in the video data, 12 of which were observed to use a phone (1.7%). The annotated data was split into three datasets: a training set (60%), a validation set (20%), and a test set (20%). One of the breakthroughs in computer vision is the concept of transfer learning (Weiss, K. et al. 2016), where large deep neural networks consisting of layers learn abstract feature representations of images in the first layers and it is only the last few layers convert the features into a classification or prediction. We therefore use the aforementioned object-detection algorithm, which already has a concept of humans and bicycles (potentially without a rider) and transfer it onto our new dataset for phone use detection by fine-tuning it on the training data set. The test set contains 126 total bicycles, with 124 instances where the rider does not use a phone, and 2 instances where the bicycle rider uses a phone.

Results

In order to assess the quality of our trained model, we will use the metrics used in object detection, namely, Precision, Recall and mAP@5. These evaluate the detection performance (*Recall*), the accuracy of the classifications (*Precision*), and the errors of the bounding boxes (*mAP@0.5*). The complete results for all the test set can be seen in Table 1. A visualisation of cyclist phone use detection is presented in Figure 2.

Table 1. Showing the performance for all the different classes in the test set. For all metrics (Precision, Recall and mAP@.5) the closer to 1 the better.

Class	Images	Target	Precision	Recall	mAP@.5
Cyclist	146	126	0.86	0.94	0.97
Cyclist with Phone	146	2	0.5	1	0.67

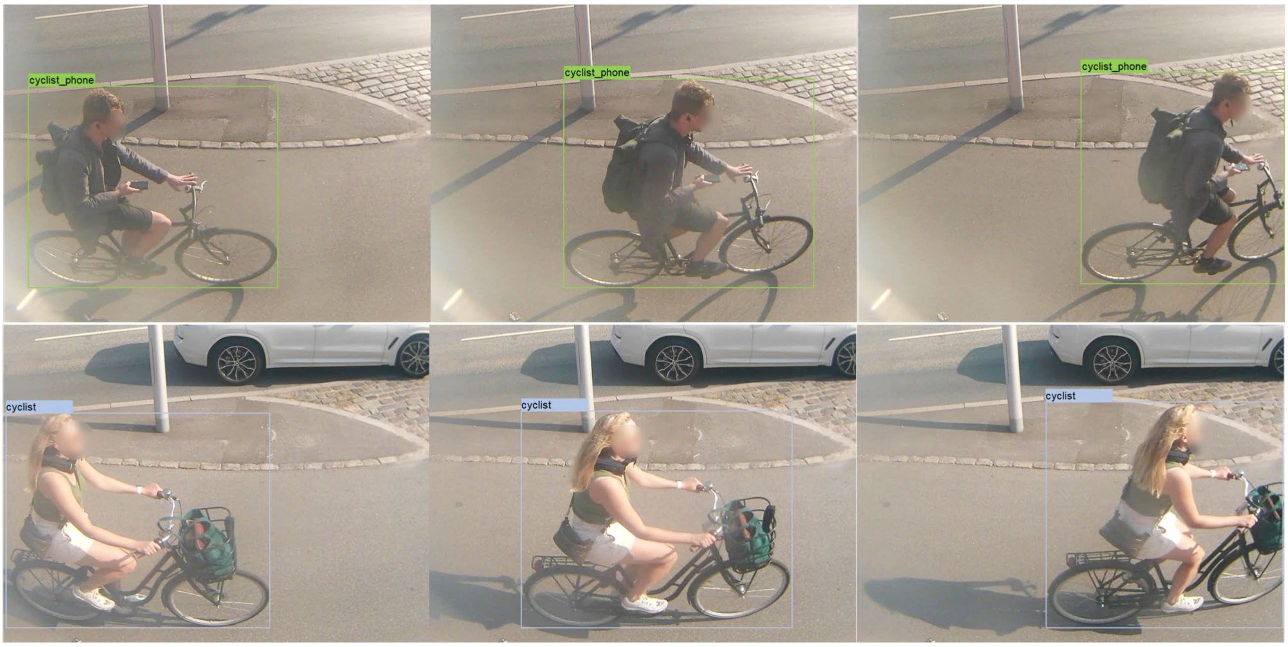

Figure 2. Visualization of the cyclist and phone use detection from a bicycle lane in Copenhagen.

Overall, the algorithm detected 89.5% of the bicycles present in the test video data, i.e. only 15 out of 126 cyclists were not detected. This corresponds to a recall of 0.92. Of all detections, 14% were instances of false positives, i.e., cyclists were detected by the algorithm, even though they were not there. This corresponds to a precision of 0.86. Overall, we achieved an mAP@0.5 of 0.97 for the detection of cyclists in the test data set. Of the two cyclists in the test data set who used a phone, both were detected by the algorithm (recall=1.0). However, two additional instances of phone use were detected, although the detected cyclists did not use a phone. These false positives lower the precision rate considerably (precision=0.5). Overall, we achieved an mAP@0.5 of 0.67 for the detection of cyclists' phone use in the test data set.

Discussion

Denmark's cycling lanes are well advanced in comparison to other European countries, and they have a crucial role in the Danish mobility of the 21st century. At the same time, there is a lack of comprehensive data on road users' behaviour in the cycle lane, as data collection by human observers is time consuming and expensive. This has led to a general lack of detailed data, prohibiting targeted education and enforcement campaigns, and evidence-based adjustments to the traffic laws. In this study, the feasibility of using a computer vision detection algorithm for phone use of cyclists was explored.

Our results show that pretrained state-of-the-art object detection algorithms fine-tuned on a relatively small train data set, perform reasonably well for the detection of cyclists in video frames. The relatively high detection rate of bicyclists could allow for the automated pre-screening of cyclists in large video data sets, which could be followed up by human observer-based registration of certain types of behaviour.

Our results on phone use detection show the general feasibility of computer vision algorithms for specific detection tasks, as both instances of phone use in the test data set were correctly identified. Our results also show room for improvement, as further research is needed, to decrease the number of false positive phone use detections. Larger datasets will help to increase robustness of both bicycle, and phone use detection. In general, our study shows that computer vision-based data collection methods can be a useful tool for the automated detection of safety-related behaviour and can potentially support Denmark in collecting valuable data to achieve the VISION ZERO of zero fatalities in traffic, as the automated detection of safety-related behaviour has been identified as an essential tool to achieve this goal (European Commission, 2019; Mwakalonge, White, & Siuhi, 2014).

References

Brandt, R. K., Haustein, S., Hagenzieker, M., & Møller, M. (2022). Cyclists' handheld phone use and traffic rule knowledge. *Transportation Research Part F: Traffic Psychology and Behaviour*, 86, 121-130.

Buehler, R. & Pucher, J. (2021). COVID-19 Impacts on Cycling, 2019–2020, *Transport Reviews*, 41:4, 393-400.

De Waard, D., Lewis-Evans, B., Jelijs, B., Tucha, O., & Brookhuis, K. (2014). The effects of operating a touch screen smartphone and other common activities performed while bicycling on cycling behaviour. *Transportation research part F: traffic psychology and behaviour*, *22*, 196-206.

De Waard, D., Westerhuis, F., & Lewis-Evans, B. (2015). More screen operation than calling: the results of observing cyclists' behaviour while using mobile phones. *Accident Analysis & Prevention*, *76*, 42-48.

European Commission (2019). EU Road Safety Policy Framework 2021-2030 - Next steps towards "Vision Zero", Available at: https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/move-2019-01178-01-00-en-tra-00_3.pdf.

Gioldasis, C., Christoforou, Z., & Seidowsky, R. (2021). Risk-taking behaviors of e-scooter users: A survey in Paris. *Accident Analysis & Prevention*, *163*, 106427.

Jocher, G., Ayush Chaurasia, Stoken, A., Borovec, J., NanoCode012, Yonghye Kwon, TaoXie, Jiacong Fang, Imyhxy, Kalen Michael, , Lorna, Abhiram V, Montes, D., Jebastin Nadar, Laughing, Tkianai, YxNONG, Skalski, P., Zhiqiang Wang, … Minh, M. T. (2022). ultralytics/yolov5: v6.1 - TensorRT, TensorFlow Edge TPU and OpenVINO Export and Inference (v6.1) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.6222936.

Goldenbeld, C., Houtenbos, M., Ehlers, E., & De Waard, D. (2012). The use and risk of portable electronic devices while cycling among different age groups. *Journal of safety research*, *43*(1), 1-8.

Lin, H., Chen, G., & Siebert, F. W. (2021, September). Positional Encoding: Improving Class-Imbalanced Motorcycle Helmet use Classification. In *2021 IEEE International Conference on Image Processing (ICIP)* (pp. 1194-1198). IEEE.

Lin, H., Deng, J. D., Albers, D., & Siebert, F. W. (2020). Helmet use detection of tracked motorcycles using cnn-based multi-task learning. *IEEE Access*, *8*, 162073-162084.

Martínez-Ruiz, V., Jiménez-Mejías, E., de Dios Luna-del-Castillo, J., García-Martín, M., Jiménez-Moleón, J. J., & Lardelli-Claret, P. (2014). Association of cyclists' age and sex with risk of involvement in a crash before and after adjustment for cycling exposure. *Accident Analysis & Prevention*, *62*, 259-267.

Mwakalonge, J. L., White, J., & Siuhi, S. (2014). Distracted biking: a review of the current state-of-knowledge. *International Journal of Traffic and Transportation Engineering*, *3*(2), 42-51.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

Siebert, F. W., & Lin, H. (2020). Detecting motorcycle helmet use with deep learning. *Accident Analysis & Prevention*, *134*, 105319.

Stavrinos, D., Pope, C. N., Shen, J., & Schwebel, D. C. (2018). Distracted walking, bicycling, and driving: Systematic review and meta-analysis of mobile technology and youth crash risk. *Child development*, *89*(1), 118-128.

Useche, S. A., Alonso, F., Montoro, L., & Esteban, C. (2019). Explaining self-reported traffic crashes of cyclists: An empirical study based on age and road risky behaviors. *Safety science*, *113*, 105-114.

Weiss, K., Khoshgoftaar, T.M. & Wang, D. A survey of transfer learning. *J Big Data* 3, 9 (2016). https://doi.org/10.1186/s40537-016-0043-6.