Jörg Gamerdinger*, Sven Teufel, Georg Volk, Anna-Lisa Rüeck, and Oliver Bringmann

Analyzing Track Management Strategies for Multi Object Tracking in Cooperative Autonomous Driving Scenarios

Abstract: For autonomous driving to operate safely it is crucial to perceive surrounding objects correctly. Not only detection but also state estimation (track) of a perceived object is urgent. The state is required to enable a safe motion planning, since it allows to predict the future position of an object. To include only valid information, the state estimations must be maintained to determine which track is active and which is not. Mostly, a simple count-based approach is used. For this, we present an investigation of two common approaches from non-cooperative track management in comparison to two new management strategies to maintain tracks in a cooperative scenario. We evaluate them using three simulated scenarios with a varying rate of cooperative vehicles. A confidence-based approach was able to increase the average precision by up to 9 percentage points.

Keywords: Autonomous Driving, Cooperative Perception, Object Tracking, Track Management, V2X

Zusammenfassung: Für den sicheren Betrieb des autonomen Fahrens ist es von entscheidender Bedeutung, die Objekte in der Umgebung korrekt zu detektieren. Nicht nur die Detektion, sondern auch die Schätzung des Zustands (Track) eines wahrgenommenen Objekts ist dringend erforderlich. Der Zustand wird benötigt, um eine sichere Trajektorienplanung zu ermöglichen, da er es erlaubt, die zukünftige Position eines Objekts vorherzusagen. Um nur gültige Informationen einzubeziehen, müssen die Zustandsschätzungen verwaltet werden, um zu bestimmen, welcher Track aktiv ist und welcher nicht. Daher wird meist nur ein Anzahlbasierter Ansatz verwendet. In dieser Arbeit untersuchen wir zwei gängige Ansätze des nicht-kooperativen Trackmanagements im Vergleich zu zwei neuen Managementstrategien zur Verwaltung von Tracks in einem kooperativen Szenario. Wir evaluieren sie anhand von drei simulierten Szenarien mit einer variierenden Rate kooperativer Fahrzeuge. Ein Konfidenz-basierter Ansatz erreichte eine Steigerung der Average Precision um bis zu 9 Prozentpunkt.

Schlagwörter: Autonomes Fahren, Kooperative Wahrnehmung, Objekt Tracking, Track Management, V2X

***Corresponding author: Jörg Gamerdinger,** Sven Teufel, Georg Volk, Anna-Lisa Rüeck, Oliver Bringmann; University of Tübingen, Faculty of Science, Department of Computer Science, Embedded Systems Group. E-mail: {joerg.gamerdinger, sven.teufel, georg.volk, anna-lisa.rueeck, oliver.bringmann}@uni-tuebingen.de

1 Introduction

Interest in the development of autonomous vehicles has grown increasingly in recent years, both in research and industry. One of the most essential reasons for this is the promising contribution of automated vehicles to increasing road safety. The reason for 90 % of traffic accidents with fatalities is due to human error [1], therefore autonomous vehicles have the potential to make traffic as a whole safer. However, the realization of autonomous driving poses several challenges.

Autonomous vehicles must perceive their surroundings correctly and comprehensively. Therefore, these vehicles are equipped with various sensors like cameras, LiDAR, and RADAR. In real world driving the environment can be very challenging. The impact of adverse weather conditions on automotive surround sensors degenerates the perception capabilities [2, 3]. Hence, it is essential to develop perception algorithms that are both robust against these conditions and well tested in such situations. This can be achieved by training and testing the perception algorithm on diverse data sets extended by simulated weather effects [3, 4].

Besides the influences of weather on automotive sensors, the perception of self driving vehicles is also limited by the maximum sensor ranges and concealment by other road users or infrastructure [2, 5]. Therefore, a comprehensive perception is not possible in most cases for a single vehicle. Hence, cooperative perception using communication between vehicles or infrastructure is considered as a promising technology for realizing autonomous driving [5, 6]. Cooperative perception makes it possible to perceive the environment from different viewpoints, thus extending the locally perceived environment.

Based on the detection of an object, its state must be estimated using the object tracking. In the case of cooperative perception tracking comes with further challenges, like the trustworthiness estimation of perceived information [7]. Many publications consider different object tracking approaches but moreover, the management of tracks is an important step which is the main focus of this work. Track management includes the process of track creation, track deletion and the determination if a track is active based on current and previous perception iterations.

The goal of this work is to analyze the effects of different track management strategies on the perception performance in a cooperative driving context. Therefore, we present two common count-based approaches which are commonly used for local object tracking, one confidence-based method from Benbarka et al. [8] and furthermore, we propose a new management strategy combining the count and the confidence-based approaches in Sect. 3. Section 4 presents the conducted experiments and the obtained results. Finally, in Sect. 5 we conclude our work and give an outlook to further research.

2 Related Work

Object tracking describes the estimation of a vehicle's state based on its detection. The state can consist of position, the kinematic state with velocity, acceleration, orientation and turn-rate and information about size and shape of the object [9]. In contrast to image-based object detection, the object tracking uses multiple sequential images. Moreover, the object tracking can be divided into offline and online tracking. Here, offline tracking means that all images are taken into account and in the online tracking an actual image in combination with the previous images is used [10].

For object tracking various approaches such as Bayesian dynamic state estimation, the Gaussian sum filter or a sequential Monte Carlo method exist [9]. Furthermore, particle filters or the well known Kalman Filter with different movement models like constant acceleration or constant velocity can be employed. A detailed overview of different object tracking algorithms is presented by Vo et al. [11].

While different works consider combining these algorithms with various cost metrics for an improved tracking, the track maintenance is mostly neglected. To determine if a track is active mostly the count-based method is used [8]. The maintenance of tracks with a confidence-based approach was presented by Sun et al. [12]. Since they used an addition update function, the approach leads to a high number of false positives. A similar approach was presented by Benbarka et al. [8]. They improved the score update functions and achieved better results compared to the count-based approach of Sun et al.

To evaluate object tracking different metrics can be applied. As performance metrics precision and recall [13] are very well known in the field of object perception. These metrics describe the rate of perceived objects and the rate of correct track estimations. For this, tracking or detection information must be classified into true positive, false positive, false negative. Moreover, the average precision (AP) which is defined as the area of the corresponding precision recall curve is a common metric [14]. Alternatively, the interpolated average precision as introduced by Everingham et al. [15] can be used to reduce the influence of small variations. The interpolated AP can be calculated as shown in Eq. (1) where $p_{interp}(r)$ describes the interpolated precision score for a given recall r.

$$AP = \frac{1}{11} \sum_{r \in \{0, 0.1, \ldots, 1\}} p_{interp}(r) \tag{1}$$

We use the AP to incorporate both precision and recall. The AP averaged over all available classed is called mean average precision (mAP).

The more complex metrics Multiple Object Tracking Accuracy and Precision were presented by Stiefelhagen et al. [16]. These metrics are more meaningful due to the usage of the distance between tracking hypothesis and real position. The safety metric by Volk et al. [17] allows an evaluation of detection and tracking safety using common performance metrics combined with additional factors such as the object's relevance and velocities.

3 Track Management Strategies

As already introduced in Sect. 1, track management plays a crucial role for the perception quality. The track management is responsible for managing the lifetime of tracks including the creation and deletion. Before track management comes into play, detected objects need to be matched to already present tracks within the environmental model (EM). A detected object can then either be matched to an already existing track or will be a new, unmatched one. As last there also exist tracks with no assigned detected objects.

If a new, unmatched object is detected a new, yet inactive track gets created. When the track is e.g. observed

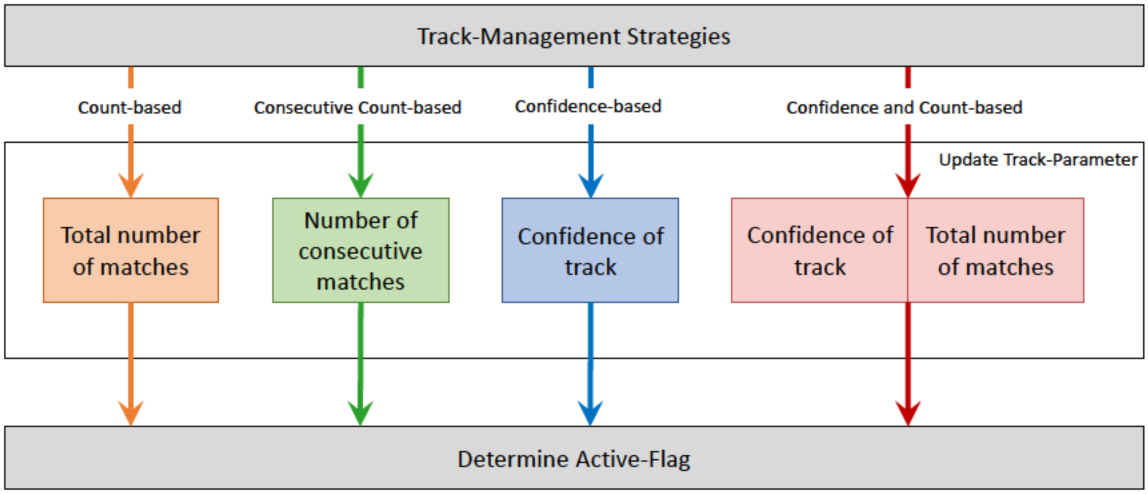

Fig. 1: Overview of investigated track management strategies and their relevant parameters

for at least a given observation time o_t, an observation count o_c or has reached a defined threshold of existence the corresponding track of an object gets valid and will be integrated into the EM. As long as the object is correctly perceived, the track stays active and remains in the EM. If an object is no longer observed, the track becomes inactive at first. Subsequently, the track management needs to decide whether a track has to be deleted if it is inactive for a given time span. Observation count-based methods alone are insufficient as additional parameters to e.g. o_c are necessary to decide if a track can be deleted. Hence, the time span without observation and the travelled distance of an object can be considered as criteria to mark the inactive track for deletion. Hence, the corresponding object will be deleted from the EM. Depending on the used track management strategy an object can on the one hand be deleted too early, too late or on the other hand be included too early or too late to the EM. This would either lead to a track id change or a false positive detection respectively. Neither one nor the other is desirable and results in a bad tracking performance.

Hence, this work evaluates different track management strategies for cooperative as well as vehicle-local tracking. The track management strategy will be applied similarly to the vehicle-local tracking and the cooperative tracking. Meaning that the track management is responsible for managing two different EMs one for local and the other one for cooperative perception. For cooperative perception the objects will be received via cooperative perception messages (CPMs) from other vehicles. These include an object list containing the locally perceived objects of cooperative vehicles.

As shown in Fig. 1, four different track management strategies will be evaluated. Starting with two simple count-based ones, followed by a more advanced confidence-based technique. Finally, the count and confidence-based techniques will be combined. In the following the four strategies will be explained in detail.

Count-based

The count-based track management decides based on the observation count o_c if a track is active. When a new track is created it is marked as inactive and will not be present in the EM. The track will be marked active if the cumulative number of observations during the lifespan of the track, respectively the observation count, is greater than a defined threshold th_c. It has to be stated that cooperatively perceived objects received via CPMs and locally perceived ones count equally.

Consecutive Count-based

The consecutive count-based track management is a specialization of the count-based track management. In contrast to the count-based track management the consecutive observation count o_{cc} will decide whether a track is active or not. The consecutive count will only be increased if the corresponding object o to track t_o is detected in every perception iteration. If o is not detected in a perception iteration and $o_{cc} < \mathrm{th}_{cc}$, where th_{cc} is the threshold for consecutive observations, we set $o_{cc} = 0$. Meaning that o_{cc} will be reset as soon as o is no longer detected. Similarly to the count-based method, t_o will become active as soon as th_{cc} is reached and will be incorporated into the EM.

Confidence-based

We use the confidence-based track management presented by Benbarka et al. [8]. This track management uses the confidence of a track to determine if the track is active or not. Therefore, three cases have to be considered. First, the matching determines objects without a given track, then a track t_o with the confidence conf_o from the detection of o is initialized, this confidence must exceed a given threshold for new tracks $\text{th}_{\text{conf-n}}$. The second case covers the scenario that no detection for an existing track occurred at a time i, here the confidence of the track $\text{conf}_{t,i}$ is decreased by a defined decay d as described in Eq. (2).

$$\text{conf}_{t,i} = \text{conf}_{t,i-1} - d \tag{2}$$

For an existing track with a match first Eq. (2) is applied, then a confidence update function (see Eq. (3)) is applied to determine the track confidence [8] .

$$\text{conf}_{t,i} = 1 - \frac{(1 - \text{conf}^*_{t,i}) \cdot (1 - \text{conf}_{o,i})}{(1 - \text{conf}^*_{t,i}) + (1 - \text{conf}_{o,i})} \tag{3}$$

$\text{conf}^*_{t,i}$ describes the track confidence after applying Eq. (2). Finally, a track is considered as active if $\text{conf}_{t,i}$ exceeds a defined confidence threshold th_{conf}.

Confidence and Count-based

We propose the confidence and count-based track management which combines two strategies. To be incorporated into the EM, a track must fulfill the requirements of count-based as well as confidence-based track management. A track is considered active if the observation count is greater or equal to the corresponding threshold ($o_c \geq \text{th}_c$) and additionally the confidence of the track must be at least equal to the confidence threshold ($\text{conf}_t \geq \text{th}_{\text{conf}}$).

4 Evaluation of Track Management Strategies

Three simulated scenarios are considered for the evaluation: Highway with dense traffic (60 vehicles/km), highway with sparse traffic (30 vehicles/km) and a road crossing. For the highway scenarios, the same highway section with a length of 660 m and three lanes per direction is used. The crossing has a size of about 120 m with two lanes per direction. CARLA [18] is used as simulation environment. As framework for implementing the different track management strategies the *RESIST* framework developed in our team by Müller et al. [19] with the improvements by Volk et al. [6, 20] is employed. The matching problem will be solved by the Hungarian algorithm [21]. For each scenario three equipment rates of cooperative vehicles are investigated: 5 % (ER5), 15 % (ER15) and 25 % (ER25).

To obtain the best perceptual performance with our track management system we conducted a parameter space exploration for the aforementioned thresholds of the proposed strategies in Sect. 3. An overview of the search space for the single thresholds and parameter is shown in Table 1. It must be stated that the given parameter search space for th_c and th_{cc} is rather small. However, using only one detection is not possible as the tracking requires at least two detections. Higher numbers would lead to a too late integration of the track into the EM. For each parameter we conducted 50 runs of each scenario and for each equipment rate of cooperative vehicles. Afterwards, we consider the minimum and maximum achieved AP overall equipment rates. Then, a sub-interval $[maxAP - 5\,\% \cdot i, maxAP]$ where $i \geq 1$ describes the step number, is constructed. We increase i while $maxAP - 5\,\% \cdot i > minAP$ holds and evaluate if a given parameter set achieves an AP within this interval for all three test scenarios. If multiple parameter sets satisfy the condition, the average AP overall scenarios is calculated and the one with the highest resulting AP is chosen. The resulting best parameter set overall test scenarios is listed in Table 1.

Parameter	Search space	Best Result
th_c	[2,3]	2
th_{cc}	[2,3]	2
th_{conf}	[0.5, 0.6,. . . , 0.8]	0.7
$\text{th}_{\text{conf-n}}$	[0.5, 0.6, 0.7]	0.6
d	[0.025, 0.05,. . . , 0.2]	0.075

Table 1: Parameter space for the exploration of the thresholds for the different track management strategies.

An overview of the results for the evaluated scenarios is shown in Fig. 2. In the highway scenario with dense traffic and ER5 the achieved AP was about 15.6 % for count-based as well as for the confidence and count-based strategy. Confidence-based performed best with 17.26 %, consecutive count-based performed worst with about 13 %. For ER15 and ER25, the count-based and confidence and count-based strategy achieved similar results with about 23.8 % and 32.5 % respectively. Equivalent to ER5,

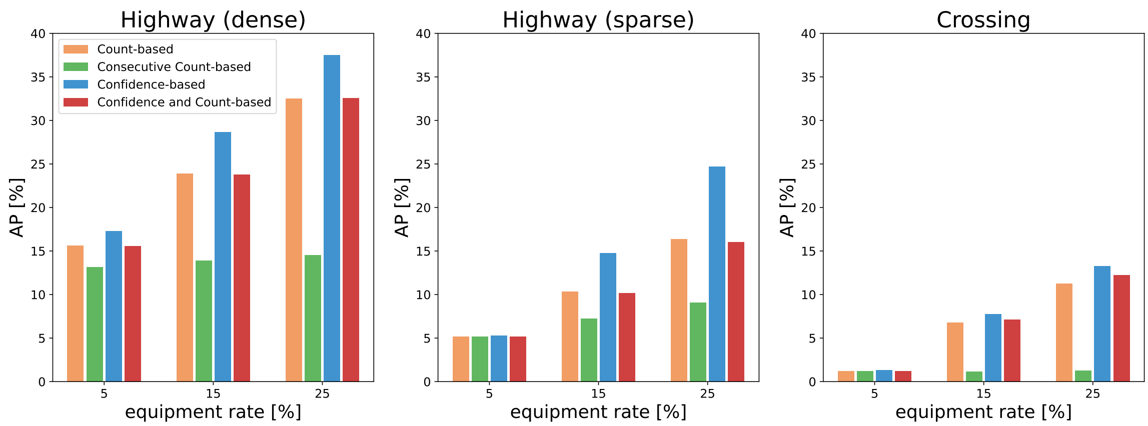

Fig. 2: Results of the AP metric on the evaluated scenarios with varying rates of cooperative vehicles. Left: Highway scenario with a high traffic density. Center: Highway scenario with a low traffic density. Right: Scenario at a road crossing

confidence-based performs best and the AP was increased by about 5 percentage points (p.p.) for both ER15 and ER25 compared to the count-based track management.

The highway scenario with sparse traffic and ER5 showed an AP of about 5 % for all strategies. Again, the confidence-based strategy performs best with up to 24.71 % for ER25 which corresponds to an improvement of 8.36 p.p. towards the count-based approach. The consecutive count-based strategy performed significantly worse and reached only up to 9 %.

For the crossing scenario the lowest AP scores appeared since this is the most complex scenario. For ER5 all strategies have an AP of about 1.2 %. The count-based strategy achieved an AP of 6.77 % (ER15) and 11.23 % (ER25). For the crossing scenario the confidence and count-based strategy showed slightly better results than the count-based with an AP increase of 0.35 p.p. and 1.01 p.p. for ER15 and ER25 respectively. For the consecutive count-based strategy the AP of ER15 and ER25 is similar to ER5. Comparable to the highway scenarios, the confidence-based strategy outperforms the other strategies and reached an AP of 13.29 % for ER25. It must be stated that, in this scenario, the difference between confidence-based and count-based is less than for the highway scenarios.

It must be stated that the overall AP is rather low which can be traced back to the scenarios in which lots of objects remain undetected due to occlusion. Moreover, the AP is not only affected by the given scenario but also the used sensor models (camera only) and algorithms have a significant influence on the performance which makes it difficult to compare the achieved results to other publications.

5 Conclusion and Outlook

In this paper we examined four strategies for track maintenance in different scenarios and with varying equipment rates of cooperative vehicles. A well performing track management is crucial to enable safety in autonomous driving. We have shown that the commonly used count-based approach performs well but is outperformed by the confidence-based approach by Benbarka et al. [8]. This strategy achieved the best results for all scenarios and equipment rates. Using a consecutive count-based approach showed a significant decrease in average precision and thus is not suitable for autonomous driving. A combined approach of the confidence-based and count-based strategy shows results similar to the common count-based strategy but with a little improvement in the crossing scenario. For further research combinations of the confidence-based as well as the confidence and count-based strategy with different cost metrics will be evaluated. For a more comprehensive and meaningful evaluation the safety metric by Volk et al. [17] will be used to determine not only the performance but also the safety of the different approaches. Using this metric to determine which approach shows the best results and should be used could increase the safety of an autonomous vehicle. Furthermore, machine learning approaches considering the semantics of the scene will be examined as track management strategy as using the semantics of the scene could help to improve the performance and safety in more complex scenarios. Moreover, higher rates of cooperative vehicles and static tracking sensors with an improved viewpoint on the scene are promising to improve the tracking quality in complex scenarios.

Acknowledgment: This work has been partially funded by the German Research Foundation (DFG) in the priority program 1835 under grant BR2321/5-2.

References

[1] De Jong Yeong, Gustavo Velasco-Hernandez, John Barry, and Joseph Walsh. Sensor and Sensor Fusion Fechnology in Autonomous Vehicles: A Review. *Sensors*, 21(6):2140, 2021.

[2] Joel Janai, Fatma Güney, Aseem Behl, Andreas Geiger, et al. Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art. *Foundations and Trends® in Computer Graphics and Vision*, 12(1–3):1–308, 2020.

[3] Georg Volk, Stefan Müller, Alexander Von Bernuth, Dennis Hospach, and Oliver Bringmann. Towards Robust CNN-based Object Detection Through Augmentation With Synthetic Rain Variations. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 285–292. IEEE, 2019.

[4] Sven Teufel, Georg Volk, Alexander Von Bernuth, and Oliver Bringmann. Simulating Realistic Rain, Snow, and Fog Variations For Comprehensive Performance Characterization of LiDAR Perception. In *2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring)*, pages 1–7. IEEE, 2022.

[5] Florian A Schiegg, Ignacio Llatser, Daniel Bischoff, and Georg Volk. Collective Perception: A Safety Perspective. *Sensors*, 21(1):159, 2020.

[6] Georg Volk, Alexander von Bemuth, and Oliver Bringmann. Environment-aware Development of Robust Vision-based Cooperative Perception Systems. In *2019 IEEE Intelligent Vehicles Symposium (IV)*, pages 126–133, Paris, France, June 2019. IEEE. ISBN 978-1-72810-560-4. 10.1109/IVS.2019.8814148.

[7] Georg Volk, Jörg Gamerdinger, Alexander von Bernuth, Sven Teufel, and Oliver Bringmann. Environment-aware Optimization of Track-to-Track Fusion for Collective Perception. In *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, pages 2385–2392. IEEE, 2022.

[8] Nuri Benbarka, Jona Schröder, and Andreas Zell. Score Refinement for Confidence-Based 3D Multi-Object Tracking. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, page 8083–8090. IEEE Press, 2021. 10.1109/IROS51168.2021.9636032. URL https://doi.org/10.1109/IROS51168.2021.9636032.

[9] Karl Granström, Marcus Baum, and Stephan Reuter. Extended Object Tracking: Introduction, Overview, and Applications. *Journal of Advances in Information Fusion*, 12, 12 2017.

[10] Joel Janai, Fatma Güney, Aseem Behl, and Andreas Geiger. Computer Vision for Autonomous Vehicles: Problems, Datasets and State-of-the-Art. *CoRR*, abs/1704.05519, 2017. URL http://arxiv.org/abs/1704.05519.

[11] Ba-ngu Vo, Mahendra Mallick, Yaakov Bar-Shalom, Stefano Coraluppi, Richard Osborne, Ronald Mahler, and Ba-tuong Vo. Multitarget tracking. *Wiley encyclopedia of electrical and electronics engineering*, (2015), 2015.

[12] Pei Sun, Henrik Kretzschmar, Xerxes Dotiwalla, Aurélien Chouard, Vijaysai Patnaik, Paul Tsui, James Guo, Yin Zhou, Yuning Chai, Benjamin Caine, Vijay Vasudevan, Wei Han, Jiquan Ngiam, Hang Zhao, Aleksei Timofeev, Scott Ettinger, Maxim Krivokon, Amy Gao, Aditya Joshi, Yu Zhang, Jonathon Shlens, Zhifeng Chen, and Dragomir Anguelov. Scalability in Perception for Autonomous Driving: Waymo Open Dataset. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2443–2451, 2020. 10.1109/CVPR42600.2020.00252.

[13] David Powers. Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation. *Journal of Machine Learning Technologies*, 2(1):37–63, 2011.

[14] Kendrick Boyd, Kevin H. Eng, and C. David Page. Area under the precision-recall curve: Point estimates and confidence intervals. In Hendrik Blockeel, Kristian Kersting, Siegfried Nijssen, and Filip Železný, editors, *Machine Learning and Knowledge Discovery in Databases*, pages 451–466, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[15] Mark Everingham, Luc Van Gool, Christopher Williams, John Winn, and Andrew Zisserman. The Pascal Visual Object Classes (VOC) challenge. *International Journal of Computer Vision*, 88(2):303–338, 2010.

[16] Rainer Stiefelhagen, Keni Bernardin, Rachel Bowers, John Garofolo, Djamel Mostefa, and Padmanabhan Soundararajan. The CLEAR 2006 Evaluation. In *CLEAR'06: Proceedings of the 1st International Evaluation Conference on Classification of Events, Activities and Relationships*, page 1–44, Apr 2006.

[17] Georg Volk, Jörg Gamerdinger, Alexander von Bernuth, and Oliver Bringmann. A Comprehensive Safety Metric to Evaluate Perception in Autonomous Systems. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–8. IEEE, 2020.

[18] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. CARLA: An Open Urban Driving Simulator. In *Conference on robot learning*, pages 1–16. PMLR, 2017.

[19] Stefan Mueller, Dennis Hospach, Joachim Gerlach, Oliver Bringmann, and Wolfgang Rosenstiel. Framework for Varied Sensor Perception in Virtual Prototypes. In *MBMV*, pages 145–154, 2015.

[20] Georg Volk, Quentin Delooz, Florian A. Schiegg, Alexander Von Bernuth, Andreas Festag, and Oliver Bringmann. Towards Realistic Evaluation of Collective Perception for Connected and Automated Driving. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pages 1049–1056, 2021. 10.1109/ITSC48978.2021.9564783.

[21] H. W. Kuhn. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97, 1955.