Proceedings of Machine Learning Research vol 278:1–7, 2025 2025 2nd International Conference on Machine Learning and Intelligent Computing

# DCMTrack: Rethinking the Motion for Vehicle Multi-Object Tracking

**Peng Li** PENG52415@GMAIL.COM
*School of Information Science & Engineering, Yunnan University, Kunming, China*
**Jun Ni*** JUN.NI@YNU.EDU.CN
*School of Information Science & Engineering, Yunnan University, Kunming, China*
**Dapeng Tao** DPTAO@YNU.EDU.CN
*School of Information Science & Engineering, Yunnan University, Kunming, China*

**Editors:** Nianyin Zeng, Ram Bilas Pachori and Dongshu Wang

## Abstract

Vehicle multi-object tracking has significant applications in many fields. Existing methods struggle to address the challenges of nonlinear motion and prolonged occlusion in vehicle tracking. In this paper, an advanced tracker featuring a Nonlinear Noise Adaptive Unscented Kalman Filter, namely DCMTrack, is designed to finely adjust measurement noise and significantly enhance the accuracy of target motion state predictions. An Adaptive Direction and Confidence Cost Matrix is designed to more precisely calculate trajectory direction and confidence, enhancing the accuracy of target association. Ultimately, a Category-Aware Initialization Mechanism that integrates target category and environmental information is proposed to minimize false trajectories and optimize the overall trajectory management process. We conducted extensive experiments on the VehiclesMOT dataset, validating its effectiveness.

**Keywords:** Multi-object tracking, Vehicle tracking, Cost matrix

## 1. Introduction

Multi-Object Tracking (MOT) aims to track the positions of multiple targets in a video sequence while maintaining the continuity of their trajectories. In the context of vehicle tracking in traffic scenarios, existing methods (Bewley et al., 2016) typically assume that: (1) the tracked targets maintain a constant velocity over a given time interval, implying a linear motion pattern; (2) The targets are not subject to prolonged occlusion. However, in the realm of MOT for traffic vehicles, several challenges arise, including: (1) nonlinear motion due to perspective distortions in surveillance imagery. (2) Frequent and extended occlusions (Wang et al., 2022) between vehicles, often exacerbated by differences in shape and relatively low speeds. Figure 1 illustrates a failure case of current methods, highlighting these challenges.

To address these challenges, we propose three key innovations to enhance tracking performance. First, to address inaccurate trajectory prediction due to nonlinear motion, we introduce a nonlinear noise-adaptive Unscented Kalman Filter (UKF) that adjusts measurement noise based on detection confidence. Second, to overcome the limitations of the conventional Intersection over Union (IoU) cost matrix in handling extended occlusions, we propose an adaptive direction and confidence cost matrix that integrates direction and confidence information for improved tracking accuracy. Finally, we introduce a category-aware initialization mechanism that considers trajectory categories and environmental factors to minimize repeated detections and reduce false trajectories.

Taking the state-of-the-art (SOTA) tracking-by-detection algorithm OC-SORT (Cao et al., 2022) as our baseline, we have developed the Direction and Confidence Motion Track (DCMTrack) method.

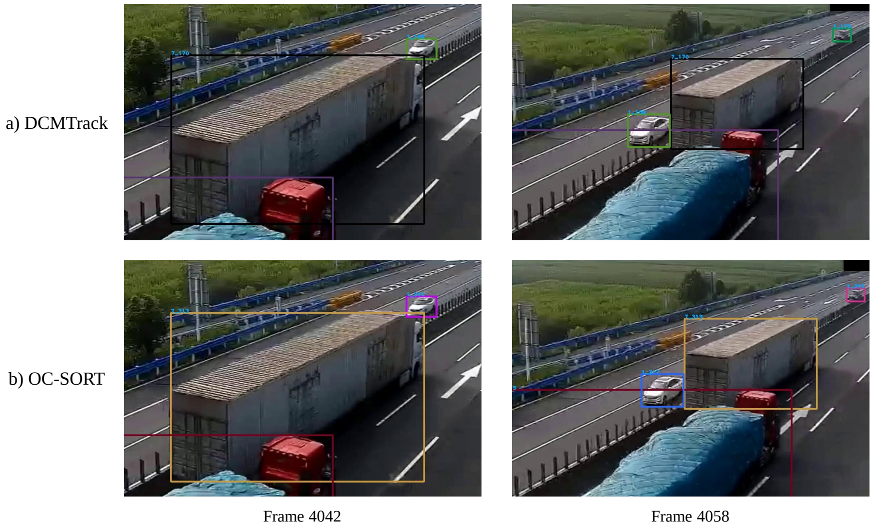


Figure 1: Samples from the results on VehiclesMOT. At frame 4058, OC-SORT experiences an ID switch for the occluded target, while DCMTrack consistently maintains its tracking.

Our contributions are summarized as: (1) DCMTrack is proposed in this paper. It is a concise, effective motion-based tracker for vehicle multi-object tracking without re-identification models. (2) A nonlinear noise-adaptive UKF is implemented to achieve a more accurate description of the target's motion state. (3) a trajectory direction and confidence cost matrix is proposed to enhance the accuracy of target association. (4) A category-aware initialization mechanism is proposed to mitigate the impact of duplicate detections by the detector.

## 2. Related Work

In recent years, the rapid advancement of object detection technology has driven the development of multi-object tracking. Tracking-by-detection trackers primarily focus on improving motion models and cost matrices.

**Motion Models.** Most tracking-by-detection algorithms rely on motion models, with many adopting the Kalman filter. Due to the Kalman filter's limitations, subsequent studies introduced improvements, such as EKF and UKF. Some methods adjusts Kalman filter measurement noise based on detection confidence. However, the relationship between confidence and detection box accuracy is nonlinear. In this paper, we propose a Nonlinear Noise-Adaptive Unscented Kalman Filter, which more accurately adjusts measurement noise using confidence levels.

**Similarity Metrics.** Most tracking-by-detection algorithms utilize a cost matrix to achieve the association between trajectories and detections. The SORT (Bewley et al., 2016) algorithm employs the IoU distance as the cost matrix for association. DeepSORT (Wojke et al., 2017) further incorporates deep visual features into the cost matrix. C-IoUTracker (Yang et al., 2022) replaces IoU with BIoU, expanding the matching space. OC-SORT (Cao et al., 2022) introduces trajectory direction consistency into the cost matrix.

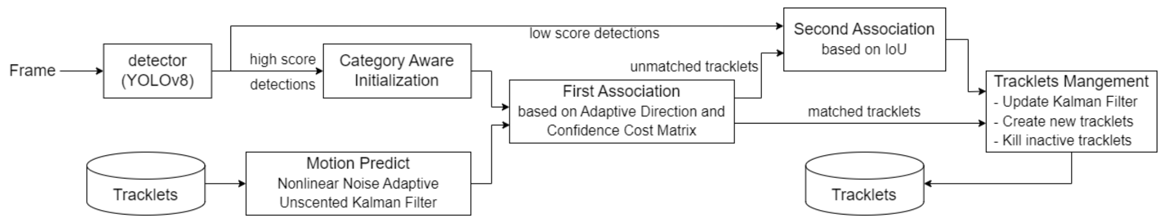


Figure 2: Overview of ours DCMTrack tracker pipeline.

## 3. Method

In this section, we present the three improvements in this paper. As show in Figure 2, the DCMTrack process primarily consists of three stages: Firstly, high-confidence objects are associated using the Adaptive Direction and Confidence Cost Matrix. Then, low-confidence objects are associated through the IoU cost matrix. Finally, the trajectories are updated based on the association results.

### 3.1. Nonlinear Noise Adaptive Unscented Kalman Filter

In SORT-based tracking algorithms, the conventional motion model assumes constant velocity for the target. However, in vehicle tracking scenarios, the motion within the frame often exhibits nonlinear characteristics. UKF is designed to address state estimation in nonlinear systems. Unlike traditional Kalman Filters, the UKF uses the Unscented Transform (UT) to handle state estimation and observation updates in nonlinear systems without relying on Taylor series expansion, thus reducing linearization errors. Recently, several studies have utilized detection confidence to linearly adjust the measurement noise $R$. Although this approach improves tracking accuracy, the relationship between the confidence score and bounding box accuracy is nonlinear. A slight decrease in confidence results in a rapid drop in positional accuracy, suggesting that the Kalman Filter fails to adjust observation noise promptly when bounding box accuracy declines.

To address this issue, we proposed a Nonlinear Noise Adaptive Unscented Kalman Filter (NNAUKF), which more accurately utilizes the confidence score to adjust the measurement noise. When the confidence score falls below a predefined threshold, the bounding box accuracy rapidly diminishes, leading to a sharp increase in $R$. The dynamic noise adjustment formula for the NNAUKF is expressed as follows:

$$\widetilde{R} = \frac{1}{1 + e^{\alpha(c-\beta)}} R \tag{1}$$

where, $c$ represents the confidence score, $\alpha = 30$ is the coefficient controlling the slope of the formula and $\beta = 0.8$ is the confidence threshold.

### 3.2. Adaptive Direction and Confidence Cost Matrix

During vehicle tracking, partial occlusion can cause the direction of the detection bounding box center to deviate from the true direction. We utilize the direction of the detection result vertices and the historical information of the trajectory direction to calculate the true motion direction (Li et al., 2024).

Specifically, the motion direction of each vertex in the trajectory detection box is first calculated. Then, based on the motion direction from the previous frame, the two vertices whose motion directions are closest to the target's motion direction in the previous frame are selected. Finally, the trajectory's motion direction is determined by calculating the average direction of these two vertices. Let the direction of vertex $i$ at time $t$ be denoted as $D_t^i$, the trajectory direction calculation formula proposed is as follows:

$$D_t^i = (v_t^i, u_t^i) \tag{2}$$

$$v_t^i = x_t^i - x_{t-\Delta t}^i \tag{3}$$

$$u_t^i = y_t^i - y_{t-\Delta t}^i \tag{4}$$

$$C_{dir} = 1 - arccos(\frac{v_1 v_2 + u_1 u_2}{\sqrt{(v_1^2 + u_1^2)(v_2^2 + u_2^2)}}) \tag{5}$$

$$C_{dir}(D_t^{min1}, D_{t-1}), C_{dir}(D_t^{min2}, D_{t-1}) = F_{min_2}(C_{dir}(D_t^i, D_{t-1})) \tag{6}$$

$$D_t = \frac{D_t^{min1} + D_t^{min2}}{2} \tag{7}$$

where $x_t^i$ and $y_t^i$ represent the positions of the detection box vertex at time $t$ and $i \in [1, 2, 3, 4]$ is the index of the vertex. $D_t^i$ denotes the motion direction of the detection box vertex and $C_{dir}$ represents the distance between the motion direction. $F_{min_2}$ refers to the selection of the two smallest values, $D_t^{min1}$ and $D_t^{min2}$ are the motion directions of the two vertices whose motion directions have the smallest distance to the target motion direction in the previous frame.

To enhance the accuracy of the cost matrix, we introduce trajectory confidence as a new metric. The confidence of a detection successfully associated with a trajectory is used as the trajectory's confidence. We utilize the detection confidence as the tracking trajectory confidence and employ linear prediction to estimate the current frame confidence for association. The trajectory confidence update formula and the confidence cost matrix formula are as follows:

$$\hat{conf} = \begin{cases} conf_{t-1} & \text{c}onf_{t-2} = None \\ conf_{t-1} - (conf_{t-2} - conf_{t-1}) & \text{e}lse \end{cases} \tag{8}$$

$$C_{conf} = |\hat{conf}_{track} - conf_{det}| \tag{9}$$

### 3.3. Category Aware Initialization

Repeated detections can result in false trajectories. To address this problem, we propose a category-aware trajectory initialization mechanism that incorporates both the category of the trajectory and the surrounding environmental information during initialization. This approach helps mitigate the impact of repeated detections and reduces false trajectories during the tracking process.

For a set of detection results, the IoU of detection boxes is calculated. Detections with IoU above a threshold are considered duplicates and merged, while retaining their individual results (box positions, categories, and confidence scores). During target association, the detection with the highest confidence from pending detections is used to associate with the tracking trajectory. If successful, the associated detection updates the trajectory. If no category match is found, the highest-confidence detection updates the trajectory and sets its category. After two association attempts, if no pending detections are successfully associated, the highest-confidence detection initializes the trajectory in the final tracking stage.

# 4. Experiments

## 4.1. Implementation Details

We evaluated DCMTrack on our custom dataset, VehiclesMOT, which is based on road surveillance images. The dataset consists of nine vehicle types, 30 video sequences, and a total of 15,222 frames, encompassing various environmental conditions.

To evaluate our method, we use common CLEAR metrics in MOT, including Multiple-Object Tracking Accuracy (MOTA), False Positive (FP), False Negative (FN), and ID Switch (IDSW). Additionally, we employ Higher Order Tracking Accuracy (HOTA) for a comprehensive assessment and use IDF1 to evaluate the association performance of the tracker. We selected YOLOv8s as the object detector with an input image size of 640. For the two-stage tracking algorithm, confidence thresholds are 0.1 and 0.6, and for the single-stage algorithm, it is 0.5. Lost tracklets are retained for 30 frames. The parameters of the compared methods were consistent with their original papers.

## 4.2. Experiment Analysis

We compared DCMTrack with several popular and advanced algorithms, including SORT (Bewley et al., 2016), ByteTrack (Zhang et al., 2021), BoTSORT (Aharon et al., 2022), OC-SORT (Cao et al., 2022), and C-BIoUTracker (Yang et al., 2022), while keeping the detector consistent across all methods. The test results on VehiclesMOT are presented in Table 1.

Table 1: Comparison of SOTA methods with the same detector on the VehiclesMOT dataset.

| Detector | Tracker | MOTA(↑) | IDF1(↑) | HOTA(↑) | FP(↓) | FN(↓) | IDSW(↓) |
|---|---|---|---|---|---|---|---|
| YOLOv8 | SORT | 82.06 | 86.07 | 85.84 | 4571 | 5366 | 146 |
| | BoTSORT | 85.94 | 92.45 | 90.63 | 4686 | 3164 | 50 |
| | ByteTrack | 84.26 | 90.83 | 90.08 | 4800 | 3994 | 50 |
| | OC-SORT | 86.28 | 90.47 | 88.73 | 4091 | 3453 | 167 |
| | C-BIoUTracker | 83.95 | 90.72 | 89.99 | 4623 | 4352 | 45 |
| | DCMTrack(Ours) | **87.32** | **92.59** | **91.04** | **4030** | **3043** | **42** |

As shown in Table 1, our method outperforms OC-SORT, improving MOTA by 1.04%, IDF1 by 2.12%, and HOTA by 2.31%, while also reducing FP, FN, and IDSW. These results highlight significant improvements in tracking accuracy and target association. Additionally, compared to the second-best method, our approach improves MOTA by 1.04%, IDF1 by 0.14%, and HOTA by 0.41%, with lower FP, FN, and IDSW, demonstrating its superiority over other methods.

## 4.3. Ablation Study

To further validate the effectiveness of the proposed improvements and their specific contributions to the algorithm's performance, we conducted systematic ablation experiments on VehiclesMOT. Our improvements are based on the OC-SORT method, where NNAUKF, ADCCM, and CAI represent Nonlinear Noise-Adaptive Unscented Kalman Filter, Adaptive Direction and Confidence Cost Matrix, and Category-Aware Initialization, respectively. The ablation experiment results are presented in Table 2.

Table 2: Components ablation on VehiclesMOT dataset. The best results are highlighted in **bold**.

| Tracker | | | | MOTA(↑) | IDF1(↑) | HOTA(↑) | IDSW(↓) |
|---|---|---|---|---|---|---|---|
| OC-SORT | | | | 86.28 | 90.47 | 88.73 | 167 |
| + NNAUKF | ✓ | | | 87.13 | 90.62 | 88.94 | 164 |
| + ADCCM | | ✓ | | 86.84 | 92.31 | 89.23 | 44 |
| + CAI | | | ✓ | 86.89 | 90.71 | 89.12 | 152 |
| DCMTrack(Ours) | ✓ | ✓ | ✓ | **87.32** | **92.59** | **91.04** | **42** |

Our proposed improvements consistently enhance tracking accuracy. Specifically, NNAUKF improves tracking precision, resulting in an increase in the MOTA metric. The Direction and Confidence Cost Matrix enhances object association, significantly reducing ID switches. Additionally, the Category-Aware Trajectory Initialization method reduces FP, leading to improvements in both MOTA and HOTA metrics.

## 5. Conclusion

We propose the DCMTrack algorithm to address non-uniform vehicle motion and long-term occlusion challenges. It incorporates a Nonlinear Noise-Adaptive Unscented Kalman Filter for adapting to non-uniform motion, a Trajectory Direction Confidence Cost Matrix to improve target association and recover trajectories after occlusion, and a Category-Aware Trajectory Initialization mechanism to reduce false trajectories and optimize management. DCMTrack demonstrates robust performance against occlusion and non-linear motion while maintaining simplicity, online operation, and real-time capability.

## Acknowledgments

This work was supported by the Yunnan Provincial Major Science and Technology Special Plan Projects under Grant $202202AD080003$ (was equally grateful for the vehicle dataset provided by this project), by the National Natural Science Foundation of China under Grant $62302429$, and by the Yunling ScholarTalent Program of Yunnan Province under Grant $K264202230207$.

## References

Nir Aharon, Roy Orfaig, and Ben-Zion Bobrovsky. Bot-sort: Robust associations multi-pedestrian tracking. *arXiv preprint arXiv:2206.14651*, 2022.

Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, and Ben Upcroft. Simple online and realtime tracking. In *2016 IEEE international conference on image processing (ICIP)*, pages 3464–3468. IEEE, 2016.

Jinkun Cao, Jiangmiao Pang, Xinshuo Weng, Rawal Khirodkar, and Kris Kitani. Observation-centric sort: Rethinking sort for robust multi-object tracking. *arXiv e-prints*, pages arXiv–2203, 2022.

Xiuwen Li, Yibing Zhan, Jun Ni, Fei Cai, Hua Jin, Xu Lin, Yufeng Zhang, and Dapeng Tao. Multimodal feature analysis for surgery duration predication. In *Proceedings of the 2024 16th International Conference on Bioinformatics and Biomedical Technology*, pages 188–195, 2024.

Shujuan Wang, Bochun Huang, Huafeng Li, Guanqiu Qi, Dapeng Tao, and Zhengtao Yu. Key point-aware occlusion suppression and semantic alignment for occluded person re-identification. *Information Sciences*, 606:669–687, 2022.

Nicolai Wojke, Alex Bewley, and Dietrich Paulus. Simple online and realtime tracking with a deep association metric. In *2017 IEEE international conference on image processing (ICIP)*, pages 3645–3649. IEEE, 2017.

Fan Yang, Shigeyuki Odashima, Shoichi Masui, and Shan Jiang. Hard to track objects with irregular motions and similar appearances? make it easier by buffering the matching space. *arXiv e-prints*, pages arXiv–2211, 2022.

Yifu Zhang, Peize Sun, Yi Jiang, Dongdong Yu, Fucheng Weng, Zehuan Yuan, Ping Luo, Wenyu Liu, and Xinggang Wang. Bytetrack: Multi-object tracking by associating every detection box. *arXiv e-prints*, pages arXiv–2110, 2021.