

# Challenge of UAV-Based Vehicle Re-Identification

Subjects: [Computer Science](#), [Artificial Intelligence](#)

Contributor: Wenji Yin , Yueping Peng , Zecong Ye , Wenchao Liu

Vehicle re-identification research under surveillance cameras has yielded impressive results. However, the challenge of unmanned aerial vehicle (UAV)-based vehicle re-identification (ReID) presents a high degree of flexibility, mainly due to complicated shooting angles, occlusions, low discrimination of top-down features, and significant changes in vehicle scales.

UAV re-identification

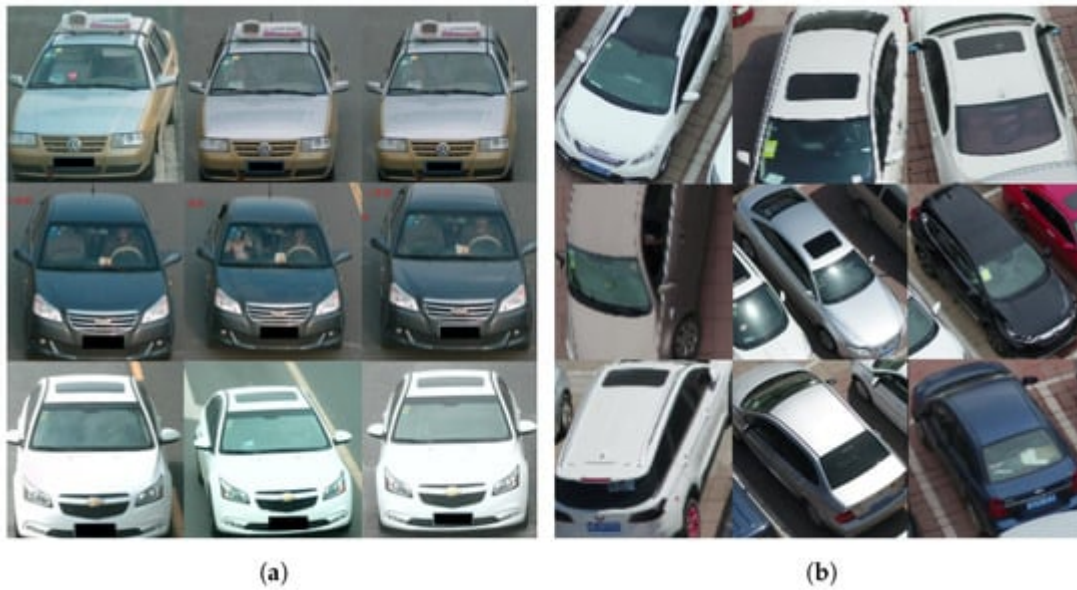
Vehicle re-identification (ReID)

UAV

## 1. Introduction

Vehicle re-identification (ReID) [\[1\]\[2\]\[3\]\[4\]](#) holds great importance in the realm of intelligent transportation systems (ITSs) in the context of smart cities. Vehicle ReID can be regarded as an image retrieval problem. Given a vehicle image, the similarity between each image and the image to be retrieved in the test set is calculated to determine whether the image to be retrieved is in the test set. Traditionally, license plate images have been employed for vehicle identification. However, obtaining clear license plate information can be challenging due to various external factors like obstructed license plates, obstacles, and image blurriness.

Thanks to the success of deep learning, the vehicle identification algorithm in the field of surveillance cameras again achieved impressive results [\[5\]\[6\]\[7\]\[8\]\[9\]](#). According to the idea of solving the vehicle ReID problem, the methods of vehicle ReID can be divided into a global feature-based method, local feature-based method, attention mechanism-based method, vehicle perspective-based method, and generative adversarial network-based method. Typically, these methods [\[6\]\[9\]\[10\]\[11\]\[12\]](#) employ a deep metric learning model that relies on feature extraction networks. The objective is to train the model to distinguish between vehicles with the same ID and those with different IDs to accomplish vehicle ReID. However, as shown in **Figure 1**, there are discernible disparities between vehicle images captured by UAV and those acquired through stationary cameras. The ReID challenge regarding UAV imagery introduces unique complexities stemming from intricate shooting angles, occlusions, limited discriminative power of top-down features, and substantial variations in vehicle scales.



**Figure 1. Comparison of two types of vehicle images.** There is a significant difference between the vehicle from the UAV perspective and the vehicle from the fixed camera. The vehicle under the fixed camera shooting angle is relatively fixed. In the view of UAVs, the shooting angles of cars are changeable, and there are many top-down shooting angles. **(a)** Surveillance cameras; **(b)** UAV cameras.

It is worth mentioning that traditional vehicle ReID methods, primarily designed for stationary cameras, face challenges in delivering optimal performance when adapted to the domain of UAV-based ReID. Firstly, the shooting angle of UAVs is complex. UAVs can shoot at different positions and angles, and the camera's viewpoint will change accordingly. This viewpoint change may cause the same object or scene to have different appearances and characteristics in different images. Second, the UAV can overlook or squint at a target or scene at different angles, resulting in viewpoint changes in the image. This viewpoint change may cause deformation or occlusion of the target shape, thus causing difficulties for feature extraction. To solve the above problems, it is necessary to add a mechanism [1][13][14][15] that can extract more detailed features when ReID extracts features to deal with the challenges brought by the drone perspective. The change in the UAV viewpoint makes the feature extraction algorithm need a certain robustness, which can correctly identify and describe the target in the case of significant changes in the viewpoint. The difference in the UAV view angle makes the feature extraction algorithm need to have the ability to adapt to shape changes and occlusions to improve the feature reliability and robustness in different views.

In recent years, the attention mechanism has gained significant popularity across multiple domains of deep convolutional neural networks. Its fundamental concept revolves around identifying the most crucial information for a given target task from a vast volume of available data. The attention mechanism selectively focuses on the image's different regions or feature channels to improve the model's attention and perception ability for crucial visual content. In the context of UAV-based vehicle ReID, the attention mechanism enables the model to enhance its perception capabilities by selectively highlighting the vehicle's specific regions or feature channels.

However, most attention mechanisms [15][16][17][18] focus on extracting features only from channels or spaces. The channel attention mechanism can effectively enhance essential channels, but it cannot deal with the problem of slight inter-class similarity. Spatial attention mechanisms can selectively amplify or suppress features in specific regions spatially, but they ignore the relationship between channels. To overcome the shortcomings of a single attention mechanism, recent studies have begun to combine channel and spatial attention [19][20][21]. Such a hybrid attention mechanism can consider the relationship between channel and space at the same time to better capture the critical information in the input feature tensor. By introducing multiple branches of the attention mechanism or fusing different attention weights, the interaction between features can be modeled more comprehensively. Shuffle attention (SA) [19] divides molecular channels to extract key channel features and local spatial fusion features, with each subchannel acquiring channel and spatial fusion attention. The bottleneck attention module (BAM) [20] is a technique that generates an attention map through two distinct pathways: channel and spatial. On the other hand, the dual attention network (DANet) [21] incorporates two different types of attention modules on dilated fully convolutional networks (FCNs). These attention modules effectively capture semantic dependencies in both spatial and channel dimensions.

## 2. Vehicle Re-Identification

The ReID problem [1][22] is first explored and applied to humans. Compared with pedestrian ReID, vehicle ReID is more challenging. Firstly, vehicles tend to have high similarity in appearance, especially in the case of the same brand, model, or color. A higher similarity makes vehicle re-identification more challenging because relatively few features may distinguish different vehicles, and there is little difference between features. Second, vehicle re-identification may face more significant pose variation than human re-identification. Vehicles may appear at different angles, positions, and rotations, resulting in changes in the geometry and appearance characteristics of the vehicle, which increases the difficulty of matching and alignment. Traditionally, vehicle Re-ID problems have been solved by combining sensor data with other clues [23][24][25][26][27][28], such as vehicle travel time [23] and wireless magnetic sensors [24]. Although the sensor technology can obtain better detection results, it cannot meet the needs of practical applications because of its high detection cost. In theory, based on the vehicle license plate number, feature recognition technology is the most reliable and most accurate again [25][26]. However, the camera's multi-angle, illumination, and resolution significantly influence license plate identification accuracy. Additionally, criminals block, decorate, forge, or remove license plates, making re-identifying vehicles only by license plate information less reliable. Accordingly, researchers have considered vehicle attributes and appearance characteristics, such as shape, color, and texture [27][28].

With the development of neural networks, deep learning-based approaches have outshone others [5][29]. Significant changes in camera angles can lead to substantial differences in local critical areas for vehicle re-identification, which leads to low precision. The hybrid pyramidal graph network (HPGN) [30] proposes a novel pyramid graph network, targeting features closely connected behind the backbone network to explore multi-scale spatial structural features. Zheng et al. [5] proposed the deep feature representations jointly guided by the meaningful attributes, including camera views, vehicle types and colors (DF-CVTC), a unified depth convolution framework for the joint

learning of depth feature representations guided by meaningful attributes, including camera view, vehicle type, and color of vehicle re-identification. Huang et al. [29] raised multi-granularity deep feature fusion with multiple granularity (DFFMG) methods for vehicle re-identification, which uses global and local feature fusion to segment vehicle images along two directions (i.e., vertical and horizontal), and integrates discriminant information of different granularity. Graph interactive transformer (GiT) [31] proposes a structure where charts and transformers constantly interact, enabling close collaboration between global and local features for vehicle re-identification. The efficient multiresolution network (EMRN) [32] proposes a multiresolution feature dimension uniform module to fix dimensional features from images of varying resolutions.

Although the current vehicle ReID method plays a specific role in the fixed camera perspective, the vehicle space photographed from the UAV perspective changes significantly, and extracting features from the top-down vertical angle is difficult. Moreover, the shooting angle of UAVs is complex. UAVs can overlook or squint at a target or scene at different angles, resulting in viewpoint changes in the image. The current method needs to be revised to solve the above problems well, and further research is required.

## References

1. Li, M.; Wei, M.; He, X.; Shen, F. Enhancing Part Features via Contrastive Attention Module for Vehicle Re-identification. In Proceedings of the 2022 IEEE International Conference on Image Processing (ICIP), Bordeaux, France, 16–19 October 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1816–1820.
2. Shen, F.; Peng, X.; Wang, L.; Zhang, X.; Shu, M.; Wang, Y. HSGM: A Hierarchical Similarity Graph Module for Object Re-identification. In Proceedings of the 2022 IEEE International Conference on Multimedia and Expo (ICME), Taipei, Taiwan, 18–22 July 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1–6.
3. Han, K.; Wang, Q.; Zhu, M.; Zhang, X. PVTReID: A Quick Person Re-Identification Based Pyramid Vision Transformer. *Appl. Sci.* 2023, 13, 9751.
4. Qiao, W.; Ren, W.; Zhao, L. Vehicle re-identification in aerial imagery based on normalized virtual Softmax loss. *Appl. Sci.* 2022, 12, 4731.
5. Li, H.; Lin, X.; Zheng, A.; Li, C.; Luo, B.; He, R.; Hussain, A. Attributes guided feature learning for vehicle re-identification. *IEEE Trans. Emerg. Top. Comput. Intell.* 2021, 6, 1211–1221.
6. Shen, F.; Wei, M.; Ren, J. HSGNet: Object Re-identification with Hierarchical Similarity Graph Network. *arXiv* 2022, arXiv:2211.05486.
7. Shen, F.; Shu, X.; Du, X.; Tang, J. Pedestrian-specific Bipartite-aware Similarity Learning for Text-based Person Retrieval. In Proceedings of the 31th ACM International Conference on Multimedia, Ottawa, ON, Canada, 29 October–3 November 2023.

8. Zhou, T.; Li, L.; Li, X.; Feng, C.M.; Li, J.; Shao, L. Group-wise learning for weakly supervised semantic segmentation. *IEEE Trans. Image Process.* 2021, 31, 799–811.
9. Zhou, T.; Qi, S.; Wang, W.; Shen, J.; Zhu, S.C. Cascaded parsing of human-object interaction recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 2021, 44, 2827–2840.
10. Wu, H.; Shen, F.; Zhu, J.; Zeng, H.; Zhu, X.; Lei, Z. A sample-proxy dual triplet loss function for object re-identification. *IET Image Process.* 2022, 16, 3781–3789.
11. Xie, Y.; Shen, F.; Zhu, J.; Zeng, H. Viewpoint robust knowledge distillation for accelerating vehicle re-identification. *EURASIP J. Adv. Signal Process.* 2021, 2021, 48.
12. Xu, R.; Shen, F.; Wu, H.; Zhu, J.; Zeng, H. Dual modal meta metric learning for attribute-image person re-identification. In *Proceedings of the 2021 IEEE International Conference on Networking, Sensing and Control (ICNSC)*, Xiamen, China, 3–5 December 2021; IEEE: Piscataway, NJ, USA, 2021; Volume 1, pp. 1–6.
13. Fu, X.; Shen, F.; Du, X.; Li, Z. Bag of Tricks for “Vision Meet Alage” Object Detection Challenge. In *Proceedings of the 2022 6th International Conference on Universal Village (UV)*, Boston, MA, USA, 22–25 October 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1–4.
14. Shen, F.; Wang, Z.; Wang, Z.; Fu, X.; Chen, J.; Du, X.; Tang, J. A Competitive Method for Dog Nose-print Re-identification. *arXiv* 2022, arXiv:2205.15934.
15. Qiao, C.; Shen, F.; Wang, X.; Wang, R.; Cao, F.; Zhao, S.; Li, C. A Novel Multi-Frequency Coordinated Module for SAR Ship Detection. In *Proceedings of the 2022 IEEE 34th International Conference on Tools with Artificial Intelligence (ICTAI)*, Macao, China, 31 October–2 November 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 804–811.
16. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
17. Li, X.; Wang, W.; Hu, X.; Yang, J. Selective kernel networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 15–20 June 2019; pp. 510–519.
18. Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Zuo, W.; Hu, Q. ECA-Net: Efficient channel attention for deep convolutional neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 13–19 June 2020; pp. 11534–11542.
19. Zhang, Q.L.; Yang, Y.B. Sa-net: Shuffle attention for deep convolutional neural networks. In *Proceedings of the ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, 6–11 June 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 2235–2239.

20. Park, J.; Woo, S.; Lee, J.Y.; Kweon, I.S. Bam: Bottleneck attention module. *arXiv* 2018, arXiv:1807.06514.
21. Fu, J.; Liu, J.; Tian, H.; Li, Y.; Bao, Y.; Fang, Z.; Lu, H. Dual attention network for scene segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 15–20 June 2019; pp. 3146–3154.
22. Shen, F.; Du, X.; Zhang, L.; Tang, J. Triplet Contrastive Learning for Unsupervised Vehicle Re-identification. *arXiv* 2023, arXiv:2301.09498.
23. Lin, W.H.; Tong, D. Vehicle re-identification with dynamic time windows for vehicle passage time estimation. *IEEE Trans. Intell. Transp. Syst.* 2011, 12, 1057–1063.
24. Kwong, K.; Kavalier, R.; Rajagopal, R.; Varaiya, P. Arterial travel time estimation based on vehicle re-identification using wireless magnetic sensors. *Transp. Res. Part C Emerg. Technol.* 2009, 17, 586–606.
25. Silva, S.M.; Jung, C.R. License plate detection and recognition in unconstrained scenarios. In *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, Germany, 8–14 September 2018; pp. 580–596.
26. Watcharapinchai, N.; Rujikietgumjorn, S. Approximate license plate string matching for vehicle re-identification. In *Proceedings of the 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Lecce, Italy, 29 August–1 September 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–6.
27. Feris, R.S.; Siddiquie, B.; Petterson, J.; Zhai, Y.; Datta, A.; Brown, L.M.; Pankanti, S. Large-scale vehicle detection, indexing, and search in urban surveillance videos. *IEEE Trans. Multimed.* 2011, 14, 28–42.
28. Matei, B.C.; Sawhney, H.S.; Samarasekera, S. Vehicle tracking across nonoverlapping cameras using joint kinematic and appearance features. In *Proceedings of the CVPR 2011*, Colorado Springs, CO, USA, 20–25 June 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 3465–3472.
29. Huang, P.; Huang, R.; Huang, J.; Yangchen, R.; He, Z.; Li, X.; Chen, J. Deep Feature Fusion with Multiple Granularity for Vehicle Re-identification. In *Proceedings of the CVPR Workshops*, Long Beach, CA, USA, 16–20 June 2019; pp. 80–88.
30. Shen, F.; Zhu, J.; Zhu, X.; Xie, Y.; Huang, J. Exploring spatial significance via hybrid pyramidal graph network for vehicle re-identification. *IEEE Trans. Intell. Transp. Syst.* 2021, 23, 8793–8804.
31. Shen, F.; Xie, Y.; Zhu, J.; Zhu, X.; Zeng, H. Git: Graph interactive transformer for vehicle re-identification. *IEEE Trans. Image Process.* 2023, 32, 1039–1051.
32. Shen, F.; Zhu, J.; Zhu, X.; Huang, J.; Zeng, H.; Lei, Z.; Cai, C. An Efficient Multiresolution Network for Vehicle Reidentification. *IEEE Internet Things J.* 2021, 9, 9049–9059.

Retrieved from <https://encyclopedia.pub/entry/history/show/115641>