Cloud-Resilient Aircraft Tracking: An Antidrift Multifilter Approach for Remote Sensing video

[1]Ilwar Sruthi, [2]Dr.M. Dhanalakshmi
[1]MCA Student, Department of Information Technology, Jawaharlal Nehru Technological University, India
[2] Professor of IT, Department of Information Technology, Jawaharlal Nehru Technological University, India

Abstract: Aircraft tracking in satellite video data holds paramount importance in various domains such as military operations, airport management, and aircraft rescue missions. This paper introduces an innovative approach, combining correlation and Kalman filtering, to develop an antidrift multifilter tracker tailored for this purpose. We propose a novel temporal consistency-constrained background-aware correlation filter algorithm, integrating temporal regularization to combat model drift caused by cloud occlusion, thereby enhancing tracking accuracy. Our experimental evaluations demonstrate superior antidrift performance compared to contemporary methods, particularly in scenarios involving cloud occlusion, while maintaining stability in complex conditions. Additionally, we present an extension by incorporating diverse techniques including ADMFT and YOLO variants (v5, v6, v7, v8) for dataset analysis. Moreover, to facilitate user testing and validation, we propose integrating this solution into a frontend utilizing the Flask framework with authentication features. We anticipate that our model will offer valuable insights for researchers interested in satellite video object tracking, especially in mitigating challenges posed by cloud occlusion.

Index Terms - Cloudy conditions, model drift, object tracking, satellite videos.

1. INTRODUCTION

Many video satellite constellations have been successfully launched worldwide in recent years due to the ongoing advancements in video satellite technology. With the use of remote sensing technology, video satellites may monitor targets in real-time over an extended period of time by continuously observing changes in the Earth's surface dynamics. As of right now, there are 31 satellites in orbit in the Jilin-1 constellation, 12 of which are capable of taking pictures. The Jilin-1 SP-01, SP-02, and LQ satellites are examples of first-generation colour video satellites; on the other hand, the Jilin-1 SP-03 satellite is an example of second-generation colour video satellite. Third-generation dual-mode push-broom and gaze imaging video satellites are the Jilin-1 SP-04–SP-08 spacecraft. The Jilin-1 GF-03C01–GF-03C03 satellites are part of the fourth generation of small-batch television satellites. For a maximum of 180 seconds, these satellites can transmit colour videos at 10 frames per second (fps).with a spatial resolution of approximately one meter. These remote sensing videos form the basis for developing more diverse and convenient applications.

Contemporary monitoring techniques and technologies have been greatly enhanced and promoted by the introduction of high-resolution remote sensing video satellites. The qualities of the data determine whether satellite data is appropriate for use in monitoring and change detection applications. Oil and gas exploration [1], disaster monitoring [2], maritime monitoring [3], ecosystem changes and disturbances [4], traffic monitoring [5], change detection [6], and identifying and monitoring military objects [7], [8] are a few examples of applications for satellite video. One essential step in these applications of remote sensing data is object tracking. The detection and tracking of moving targets has been the main focus of research on satellite video tracking tools up to this point, such as the video background extractor algorithm [9]. These techniques use object identification modules that have been pretrained in order to recognise and follow targets in every frame, But it's still difficult to make such models able to discriminate between objects in the same class and accurately track moving targets.

Existing research has also investigated approaches based on deep learning [10], [11]. Nevertheless, the data processing times of these approaches are not fast enough for real-world applications. Existing research has offered certain correlation filtering based techniques for tracking remote sensing targets, which can be used as a guide. However, these approaches frequently fail to remain stable in real-world applications since they are customised for specific application contexts. Current techniques [7], [12], [13], and [14] generally concentrate on straightforward video scenarios and struggle to handle intricate target tracking circumstances like smoke, clouds, and light patches resulting from variations in lighting. Furthermore, these algorithms have trouble with slow-moving targets. Research on identifying and following moving targets has up to this point mostly concentrated on terrestrial vehicles [13], paying little attention to other types of vehicles, like ships and aeroplanes. This article concentrates on creating an algorithm that can monitor aeroplanes rapidly and correctly using correlation filters because of the importance of aircraft for both military and transportation applications.

The issue is caused by tracking algorithms that drift when evaluating satellite video data, especially when clouds obscure the view, which results in imprecise aircraft monitoring.

Accuracy is a challenge for traditional tracking techniques in difficult environmental situations like cloud cover.

Accurate aeroplane monitoring is essential for military operations, airport administration, and aircraft rescue missions in order to maintain operational safety and efficiency.

Ineffective airport administration, hampered security, and delayed emergency reaction times can all result from inaccurate tracking.

In order to improve tracking accuracy, we suggest combining a novel approach to counteract model drift brought on by cloud occlusion with correlation and Kalman filtering techniques in an antidrift multifilter tracker.

2. LITERATURE SURVEY

With the introduction of deep reinforcement learning, adaptive correlation filters, and other machine learning methods, object tracking has advanced significantly. Accurate object tracking in complex and dynamic situations is essential for many applications, such as autonomous navigation, remote sensing, and surveillance.

For UAV tracking, Yuan et al. (2022) suggested learning adaptive spatial-temporal context-aware correlation filters. This technology uses adaptive learning to take into account both temporal and spatial contexts, improving the resilience and accuracy of UAV tracking. The method greatly enhances tracking performance in difficult situations such sudden object movements and occlusions [8].

Cui et al. (2021) presented a deep reinforcement learning framework for object tracking under occlusion in the field of remote sensing. Their method makes use of reinforcement learning's capacity for decision-making to continue tracking even in situations where objects are entirely or partially obscured. When compared to conventional tracking algorithms, this approach performed better, particularly in situations when occlusions occur often [10].

A straightforward online and real-time tracking method utilising a deep association metric was provided by Wojke et al. (2017). This approach efficiently combines a strong association metric for object tracking in real-time with deep learning for feature extraction. When these elements are combined, a tracking system that is incredibly accurate and efficient is produced that can be used for a variety of purposes, such as autonomous systems and surveillance [11].

Xuan et al. (2020) used motion estimation methods with correlation filters to enhance object tracking in satellite footage. This technique tackles the particular difficulties presented by satellite photography, including intricate background clutter and large-scale changes. By adjusting to the object's motion patterns, the enhanced correlation filters improve tracking accuracy and offer a reliable solution for satellite-based tracking applications [12].

Zhang et al. (2018) investigated the use of correlation filters in conjunction with online learning for visual object tracking. Their method allows the system to adjust to changes in the object's appearance and its surroundings by dynamically updating the correlation filters throughout the tracking process. Because of its great tracking accuracy and resilience, this approach is appropriate for real-time applications [13].

Shi et al. (2020) employed better similarity metrics and normalised frame difference labelling to detect and track moving aircraft from space. The difficulties of identifying and following tiny, swiftly moving objects in satellite imagery are successfully addressed by this method. By precisely differentiating the target item from the background, the enhanced similarity measures improve tracking accuracy [14]

Adaptive correlation filters were first used for visual object tracking by Bolme et al. (2010). The groundwork for numerous later developments in correlation filter-based tracking was established by this early study. The system can continue to track objects well even when the object's appearance and surroundings change since the filters are adaptive [15].

With their work on using kernels to exploit the circulant structure of tracking-by-detection and kernelized correlation filters for high-speed tracking, Henriques et al. (2012, 2015) made a substantial contribution to the field. These techniques make use of circulant matrices' mathematical characteristics to provide incredibly accurate and successful tracking. Particularly, the kernelized correlation filters allow for real-time tracking with no computing overhead, which makes them appropriate for a variety of applications [16][17].

Danelljan et al. (2017) developed discriminative scale space tracking. Their method improves tracking resilience in situations with notable size variations by precisely estimating the target object's scale through the use of a multi-scale detection algorithm. This approach performed exceptionally well on a variety of difficult datasets [18].

Danelljan et al. (2014) concentrated on precise scale estimate for reliable visual tracking in their previous work. Through the introduction of algorithms for accurate scale estimate, this method set the foundation for further developments in scale-aware tracking. The method greatly increased the visual tracking systems' resilience and accuracy [19].

By learning continuous convolution operators, which may capture more intricate patterns and characteristics, this technique goes beyond conventional correlation filters. These operators' continuous nature makes tracking more accurate and versatile, especially in dynamic contexts [20].

The advancements in object tracking, particularly through the integration of machine learning techniques, have significantly improved the performance and robustness of tracking systems. Adaptive correlation filters, deep reinforcement learning, and advanced motion estimation techniques have each contributed to the ability to track objects accurately in complex scenarios. The continued development of these methods promises further enhancements in tracking accuracy, efficiency, and applicability across a wide range of domains.

3. METHODOLOGY

i) Proposed System:

This paper proposes an advanced aircraft tracking system for satellite video data, crucial for military operations, airport management, and rescue missions. The system leverages a novel combination of correlation and Kalman filtering to develop an antidrift multifilter tracker. A key component is the temporal consistency-constrained background-aware correlation filter algorithm, which incorporates temporal regularization to mitigate model drift caused by cloud occlusion, thereby enhancing tracking accuracy. Additionally, the system integrates multiple tracking techniques, including ADMFT and various YOLO versions (v5, v6, v7, v8), for comprehensive dataset analysis. Notably, YOLOv8 achieves the highest mAP50, outperforming its predecessors. To facilitate user testing and validation, the solution is implemented in a Flask-based frontend with authentication features. This proposed system aims to provide valuable insights and robust tracking capabilities, particularly in overcoming challenges posed by cloud occlusion in satellite video data.

ii) System Architecture:

The image depicts a diagram illustrating the process of image processing and data augmentation. Data augmentation involves creating modified versions of data to increase the size and diversity of a training set. This can improve the performance of machine learning algorithms, especially in computer vision tasks.

The left side of the diagram shows a dataset, which is a collection of images. The dataset is fed into the image processing and data augmentation module. This module modifies the images in various ways, such as rotating them, cropping them, or changing their brightness. The augmented images are then used to train a machine learning model.

The right side of the diagram shows a block labeled "Model building." This block represents the process of training a machine learning model on the augmented dataset. The text below this block lists different machine learning models, such as YOLOv5, YOLOv6, and YOLOv8. These are all object detection models that can be used to identify objects in images.

The bottom of the diagram shows a block labeled "Performance Evaluation." This block represents the process of evaluating the performance of the machine learning model on a test dataset. The test dataset is a collection of images that the model has not seen before. By evaluating the model's performance on the test dataset, we can get an idea of how well it will generalize to new data.

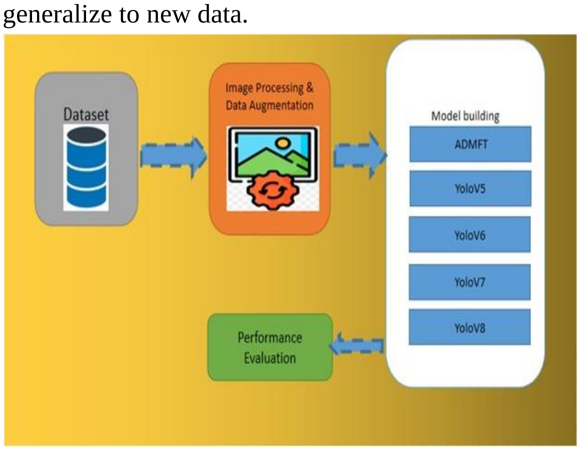

Fig 1 Proposed Architecture

iii) Image Processing:

In image processing, the workflow typically involves several key steps to prepare and augment data for various applications. Converting images to blob objects starts with segmenting them into coherent regions based on shared properties like color or intensity. Defining classes and declaring bounding boxes around objects of interest facilitates subsequent annotation and classification tasks, crucial for supervised learning models.

Next, converting image arrays to numpy arrays enhances computational efficiency and facilitates manipulation using array-based operations. The process continues with appending images to annotation files, ensuring alignment between visual data and metadata. Converting BGR (Blue-Green-Red) images to RGB format standardizes color representation across different platforms and applications.

Creating masks overlays binary images highlighting specific areas of interest, essential for tasks like semantic segmentation. Resizing images maintains consistency in dimensions, ensuring compatibility with model input requirements.

Data augmentation techniques, such as randomizing, rotating, and transforming images, introduce variability and robustness into training datasets. These operations simulate diverse real-world conditions, improving model generalization and performance across different scenarios. Collectively, these steps form a comprehensive pipeline for effective image processing and augmentation in machine learning and computer vision applications.

iv) Algorithms:

ADMFT (Antidrift Multifilter Tracker):

ADMFT is an advanced tracking algorithm designed to mitigate model drift in object tracking scenarios, especially in satellite video data affected by cloud occlusion. It integrates background-aware correlation filters with temporal consistency constraints to maintain tracking accuracy over time. In your project, ADMFT serves as a key component of the multifilter tracker, ensuring robust and stable object tracking performance in challenging environmental conditions.

YOLOv5:

YOLOv5 is a state-of-the-art object detection model known for its efficiency and accuracy. It operates on the principle of dividing an image into a grid and predicting bounding boxes and class probabilities for each grid cell simultaneously. YOLOv5 excels in real-time object detection tasks, making it suitable for applications where speed and accuracy are crucial, such as tracking moving airplanes in satellite videos. In your project, YOLOv5 can be utilized for initial object detection and localization.

YOLOv6:
YOLOv6 is an enhancement of the YOLO (You Only Look Once) series, incorporating improvements in architecture and training methodologies to further boost performance metrics such as accuracy and speed. It builds upon the strengths of YOLOv5 with additional optimizations and fine-tuning techniques, making it a compelling choice for object detection tasks in satellite video data. In your project, YOLOv6 can provide enhanced detection capabilities, especially in scenarios with complex backgrounds or variable lighting conditions.

YOLOv7:
YOLOv7 continues the evolution of the YOLO family, focusing on improving detection accuracy and efficiency through refined model architecture and training strategies. It leverages advancements in deep learning to achieve higher precision in object localization and classification, making it suitable for applications demanding superior performance in satellite video object tracking. In your project, YOLOv7 can contribute to more accurate and reliable detection of moving objects, crucial for tasks like aircraft tracking under varying atmospheric conditions.

YOLOv8:
YOLOv8 represents the latest iteration of the YOLO series, characterized by significant enhancements in model architecture and training methodologies. It achieves state-of-the-art performance in object detection tasks, particularly in challenging environments such as satellite video data with cloud occlusion. YOLOv8 incorporates advanced features and optimizations to deliver superior accuracy and robustness, making it ideal for your project's requirements. In satellite video object tracking, YOLOv8 can excel in detecting and tracking moving airplanes with high precision, even in adverse weather conditions or situations with partial occlusion.

4. EXPERIMENTAL RESULTS

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

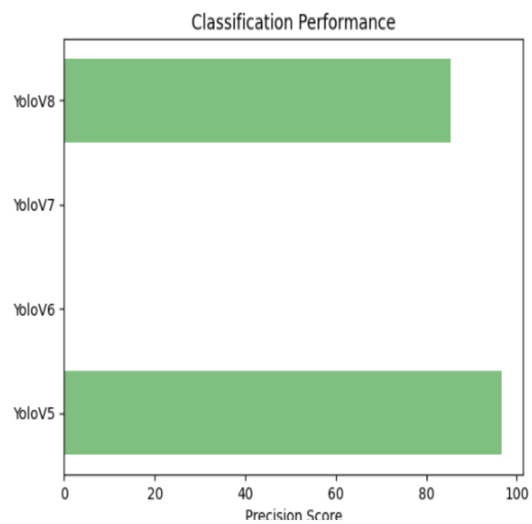

Fig 2 Precision Comparison Graph

Recall: Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a model's completeness in capturing instances of a given class.

$$Recall = \frac{TP}{TP + FN}$$

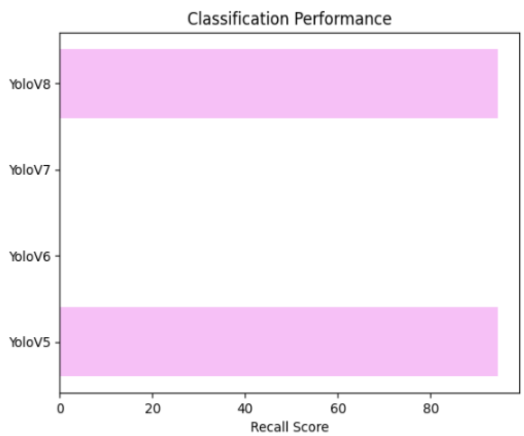

Fig 3 Recall Comparison Graph

mAP: Mean Average Precision (MAP) is a ranking quality metric. It considers the number of relevant recommendations and their position in the list. MAP at K is calculated as an arithmetic mean of the Average Precision (AP) at K across all users or queries.

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k$$

$$AP_k = \text{the AP of class } k$$

$$n = \text{the number of classes}$$

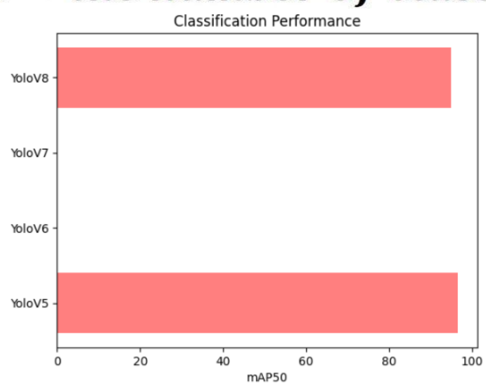

Fig 4 mAP50 Comparison Graph

Fig 5 Home Page

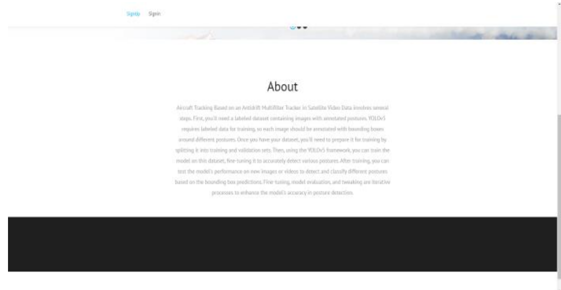

Fig 6 About page

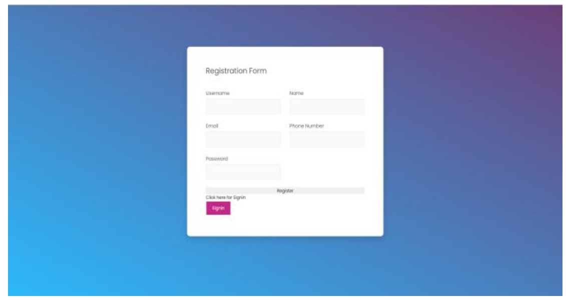

Fig 7 Registration Page

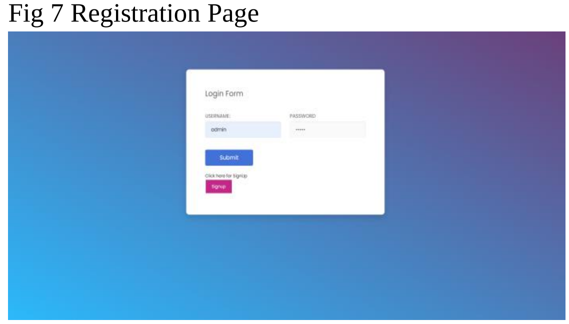

Fig 8 Login Page

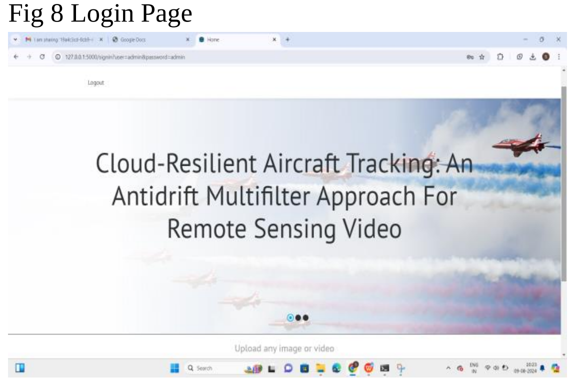

Fig 9 Main Page

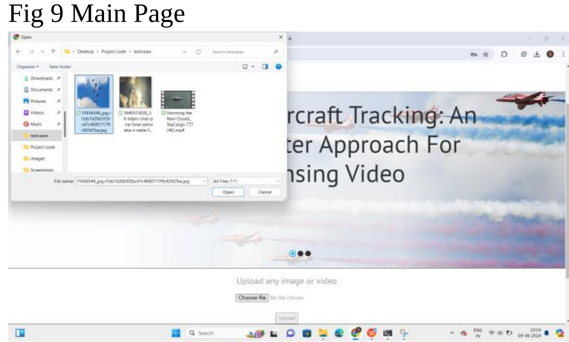

Fig 10 Upload Input Image

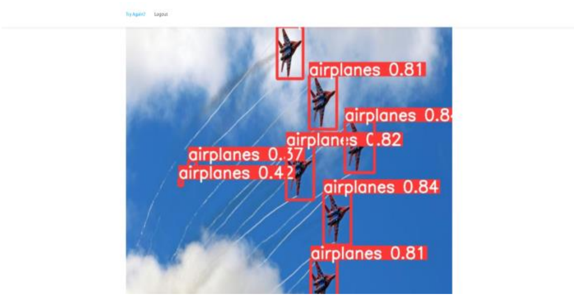

Fig 11 Final Outcome for Given Input Image

5. CONCLUSION

In conclusion, this paper introduces a sophisticated aircraft tracking system tailored for satellite video data, addressing critical needs in military, airport management, and rescue operations. By combining correlation and Kalman filtering in an antidrift multifilter tracker, the system effectively mitigates model drift exacerbated by cloud occlusion. The temporal consistency-constrained background-aware correlation filter algorithm significantly enhances tracking accuracy by maintaining object continuity through temporal regularization.

Integration of advanced tracking techniques such as ADMFT and multiple YOLO versions (v5, v6, v7, v8) enables comprehensive dataset analysis, with YOLOv8 demonstrating superior performance in mAP50 metrics compared to earlier iterations. This capability ensures robust detection and tracking of moving objects, crucial for real-time decision-making in dynamic environments.

The implementation of a Flask-based frontend with authentication features facilitates seamless user testing and validation, enhancing usability and reliability in operational settings. By addressing challenges associated with cloud occlusion in satellite video data, the proposed system not only offers valuable insights but also establishes a foundation for future advancements in satellite-based tracking technologies, potentially transforming how aerial surveillance and monitoring tasks are approached and executed.

6. FUTURE SCOPE

In future work, we aim to extend our tracking capabilities to encompass other remote sensing targets, including ships and ground vehicles. By adapting and refining our system's algorithms and techniques, we seek to enhance its versatility and applicability across diverse scenarios in remote sensing. This expansion will enable us to address broader operational needs, contributing to more comprehensive surveillance and monitoring solutions for various maritime and terrestrial applications.

REFERENCES

[1] C. Hu et al., “MODIS detects oil spills in Lake Maracaibo, Venezuela,” Eos Trans. Amer. Geophysical Union, vol. 84, no. 33, pp. 313–319, 2003.

[2] J. Yang et al., “Research on natural disaster emergency monitoring system,” Int. J. Sensor Netw., vol. 32, no. 4, pp. 218–229, 2020.

[3] R. Danovaro et al., “Implementing and innovating marine monitoring approaches for assessing marine environmental status,” Front. Mar. Sci., vol. 3, 2016, Art. no. 213.

[4] A. K. Skidmore et al., “Priority list of biodiversity metrics to observe from space,” Nature Ecol. Evol., vol. 5, no. 7, pp. 896–906, 2021.

[5] G. Kopsiaftis and K. Karantzalos, “Vehicle detection and traffic density monitoring from very high resolution satellite video data,” in Proc. IEEE Int. Geosci. Remote Sens. Symp., 2015, pp. 1881–1884.

[6] Y. K. Seong, Y.-H. Choi, J.-A. Park, and T. S. Choi, “Scene change detection in the hard disk drive embedded digital satellite receiver for video indexing,” in Proc. Dig. Tech. Papers Int. Conf. Consum. Electron., 2002, pp. 210–211.

[7] F. Shi et al., “A method to detect and track moving airplanes from a satellite video,” Remote Sens., vol. 12, no. 15, 2020, Art. no. 2390.

[8] D. Yuan et al., “Learning adaptive spatial-temporal context-aware correlation filters for UAV tracking,”ACM Trans. Multimedia Comput., Commun., Appl., vol. 18, no. 3, pp. 1–18, 2022.

[9] O. Barnich and M. Van Droogenbroeck, “ViBe: A powerful random technique to estimate the background in video sequences,” in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., 2009, pp. 945–948.

[10] Y. Cui, B. Hou, Q. Wu, B. Ren, S. Wang, and L. Jiao, “Remote sensing object tracking with deep reinforcement learning under occlusion,” IEEE Trans. Geosci. Remote Sens., vol. 60, pp. 1–13, 2021.

[11] N. Wojke, A. Bewley, and D. Paulus, “Simple online and realtime tracking with a deep association metric,” in Proc. IEEE Int. Conf. Image Process., 2017, pp. 3645–3649.
[12] S. Xuan, S. Li, M. Han, X. Wan, and G.-S. Xia, “Object tracking in satellite videos by improved correlation filters with motion estimations,” IEEE Trans. Geosci. Remote Sens., vol. 58, no. 2, pp. 1074–1086, Feb. 2020.
[13] X. Zhang et al., “Visual object tracking by correlation filters and online learning,” ISPRS J. Photogrammetry Remote Sens., vol. 140, pp. 77–89, 2018.
[14] F. Shi et al., “Detecting and tracking moving airplanes from space based on normalized frame difference labeling and improved similarity measures,” Remote Sens., vol. 12, no. 21, 2020, Art. no. 3589.
[15] D. S. Bolme, J. R. Beveridge, B. A. Draper, and Y. M. Lui, “Visual object tracking using adaptive correlation filters,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2010, pp. 2544–2550.
[16] J. F. Henriques et al., “Exploiting the circulant structure of trackingby-detection with kernels,” in Proc. Eur. Conf. Comput. Vis., 2012, pp. 702–715.
[17] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, “High-speed tracking with kernelized correlation filters,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 3, pp. 583–596, Mar. 2015.
[18] M. Danelljan, G. Häger, F. S. Khan, and M. Felsberg, “Discriminative scale space tracking,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 8, pp. 1561–1575, Aug. 2017.
[19] M. Danelljan et al., “Accurate scale estimation for robust visual tracking,” in Proc. Brit. Mach. Vis. Conf., 2014, pp. 1–5.
[20] M. Danelljan et al., “Beyond correlation filters: Learning continuous convolution operators for visual tracking,” in Proc. Eur. Conf. Comput. Vis., 2016, pp. 472–488.
[21] A. Lukezic et al., “Discriminative correlation filter with channel and spatial reliability,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 6309–6318.
[22] M. Tang and J. Feng, “Multi-kernel correlation filter for visual tracking,” in Proc. IEEE Int. Conf. Comput. Vis., 2015, pp. 3038–3046.
[23] M. Tang, B. Yu, F. Zhang, and J. Wang, “High-speed tracking with multikernel correlation filters,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2018, pp. 4874–4883.
[24] M. Danelljan, G. Häger, F. S. Khan, and M. Felsberg, “Learning spatially regularized correlation filters for visual tracking,” in Proc. IEEE Int. Conf. Comput. Vis., 2015, pp. 4310–4318.
[25] F. Li, C. Tian, W. Zuo, L. Zhang, and M.-H. Yang, “Learning spatialtemporal regularized correlation filters for visual tracking,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2018, pp. 4904–4913.