

Research Article

Analysis of the Occlusion Interference Problem in Target Tracking

Shuili Zhang ^{1,2}, Kexin Zheng,¹ and Sun Huaiyuan¹

¹College of Physics and Electronic Information, Yan'an University, Yan'an 716000, Shaanxi, China

²Shaanxi Key Laboratory of Intelligent Processing for Big Energy Data, Yan'an University, Yan'an 716000, Shaanxi, China

Correspondence should be addressed to Shuili Zhang; zhangshuili@yau.edu.cn

Received 16 August 2022; Accepted 7 September 2022; Published 26 September 2022

Academic Editor: Zaoli Yang

Copyright © 2022 Shuili Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As an indispensable part in the field of computer vision, target tracking has been widely used in intelligent transportation, missile guidance, unmanned aerial vehicle (UAV) tracking, and many other fields. It has become one of the hot directions in computer vision in recent years, while occlusion problem has always been a great difficulty and challenge in the process of target tracking. In this article, the problem of occlusion interference in target tracking is described, and the solution of occlusion problem is proposed based on different occlusion conditions. Due to the disadvantages of feature point center weighting, multiparticle template matching, and Kalman filter trajectory prediction algorithms in different cases, some algorithms with higher robustness and stability are developed to solve the occlusion problem. In the analysis of the anti-occlusion model, it is found that some tracking errors caused by occlusion can be solved by improving the quality of negative training samples and enriching the diversity of positive sample sets. According to the different training characteristics of online and offline tracking algorithms, the anti-occlusion model suitable for an active learning algorithm under different tracking conditions is found, and the tracking algorithm and characteristics of the active learning algorithm are listed, which is helpful to select the suitable tracking model in different scenarios. Finally, the future development of occlusion problem in target tracking is prospected.

1. Introduction

Computer vision can replace human brain perception and understanding of the external world through the simulation of human visual function, and image and video processing. In 1982, Marr [1] proposed a relatively complete visual theory system framework. This theory provides a relatively clear theoretical system for computer vision, which not only promotes the development of computer vision, but also lays a solid theoretical foundation for the wide application of computer vision technology in various fields [2]. Target tracking is one of the basic hot issues in the field of vision research and has been widely applied in the intelligent transportation system [3], advanced assisted driving [4], missile manufacturing [5], medical diagnosis [6], video surveillance, and other aspects [7] in recent years. The accuracy, robustness, and efficiency of target tracking directly affect the development of subsequent research work such as

target recognition and behavior analysis. Therefore, how to improve the accuracy of target tracking has become a hot topic for scholars at home and abroad. The main challenges of target tracking come from the interference of tracking scene and the uncertainty of target motion state. Target tracking is confronted with problems such as illumination change, internal and external rotation of plane, target deformation, scale change, background clutter, and occlusion, among which occlusion is one of the biggest difficulties and challenges of target tracking. Most of the tracking methods are easy to introduce background error results when the tracking target is blocked, which leads to the tracking contamination by the model and ultimately leads to tracking failure. Table 1 presents the data of the tracking algorithm competition in VOT (visual object-tracking challenge) platform [8] in recent years. As can be seen from Table 1, in the VOT competition from 2017 to 2019, the accuracy of both scale change and motion change increased from more than 40

TABLE 1: Average performance of the first three most challenging attributes of VOT.

Accuracy	VOT2017	VOT2018	VOT2019	VOT2020
Occlusion (OC)	0.39	0.40	0.44	0.45
Size change (SC)	0.41	0.43	0.48	0.51
Motion change (MC)	0.45	0.37	0.48	0.53

percent to more than 50 percent, and only the tracking failure caused by occlusion maintained an accuracy of about 40 percent. This shows that occlusion is a very challenging problem among various challenges that affect the tracking accuracy.

In the process of target tracking, the existence of various uncertain factors makes the tracking target blocked by unexpected objects. The tracking performance degradation caused by occlusion is not instantaneous, but a gradual process from entering occlusion state—continuous occlusion state—out of occlusion state. So far, the methods to solve the occlusion problem can be roughly divided into three cases, namely, target tracking algorithm based on predicted target trajectory [9], target tracking method based on space-time context [10], and target tracking method based on large-scale random detection [11]. With the promotion of deep learning in the field of target tracking, the training network is used to learn the general changing characteristics of the target, which provides a new idea to solve the problem of target occlusion [12].

In this article, the occlusion interference problem in target tracking is analyzed. In addition to the traditional methods, the solutions with higher robustness and accuracy, such as target tracking algorithm for predicting target trajectory, target tracking algorithm based on space-time context information, target tracking algorithm based on large-scale random detection, and target tracking algorithm based on deep learning, are analyzed and studied. By improving the quality of negative training samples and enriching the diversity of positive sample sets, an adaptive tracking anti-occlusion model based on active learning is found and its representative tracking algorithms and characteristics are listed.

2. Analysis of Occlusion Interference

Shielding interference is an important factor, which leads to the decrease in target tracking accuracy in the long term, and it is also an inevitable problem in the long-term tracking process. In different scenarios, different shielding interference problems have different influences on target positioning, so adopting different coping strategies for different shielding interference situations can reflect the stability of the tracking algorithm, and effectively classifying different jamming situations is also the prerequisite to solve the problem of shielding interference.

2.1. Tracing Scenarios. Simple scene refers to the single background of the tracking scene without too many texture features and similar targets, as shown in Figure 1.

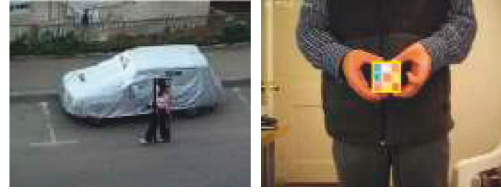


FIGURE 1: Simple scenario diagram.



FIGURE 2: Complex scene diagram.

In the simple background, the environment is relatively simple, and there is no similar target interference and no big deformation of the tracked target during target tracking, so the impact on target positioning accuracy is small, and the response value of the tracked target is large, which makes it easy to achieve target tracking and positioning. In this way, the tracking algorithm in simple scenes is suitable for verification and debugging of a certain component of the tracking algorithm because of its few external interference factors.

Complex scene refers to the background texture information of the tracking scene, which is rich, and there may be similar targets, background features similar to target features, targets may be blocked, targets may be deformed, or lighting changes, as shown in Figure 2.

In complex scenes, there are many background interference factors, so it is difficult to make prediction in advance to prevent interference, and the tracker needs to deal with the influence of interference in real time. This also tests the robustness and stability of the tracker algorithm. Therefore, robustness and stability tracking in complex background is an essential part of the tracking algorithm.

2.2. Shielding Duration. When the tracked target is in the process of occlusion within 20 frames, it is short-time occlusion, as shown in Figure 3.

In the case of short-time occlusion, the time from entering occlusion to leaving occlusion is short, so the error information introduced into the target feature template is less. Due to the short occlusion time and limited movement range of the tracking target, the tracking target is probably still within the location sampling range of the tracker. When the tracking target leaves the occlusion, the tracker will produce a drift for a short period of time near the position of the tracking target and finally recapture the target. However, such short-term drift may lead to a decrease in tracking accuracy.

In the tracking process, when the tracked target is in the occlusion process of more than 20 frames, it is regarded as long-term occlusion, as shown in Figure 4.



FIGURE 3: Short-time occlusion process.

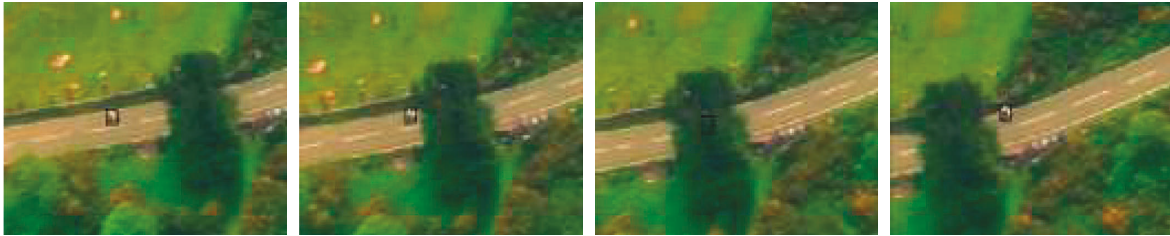


FIGURE 4: Long-term occlusion process.

In the process of target tracking, tracking the target model in each frame is updated in real time; therefore, under the shade for a long time, a lot of background information is introduced to the target feature template, which can cause the tracker serious pollution, the destruction of the tracking target model, and the low degree of confidence, and will eventually lead to the failure of tracking location drift.

2.3. Occlusion Range. According to the size of occlusion range, target tracking can be divided into partial occlusion and severe occlusion.

In target tracking, when the part of tracking target obscured by background is less than 40%, it is regarded as partial occlusion, as shown in Figure 5. Partial occlusion will cover up part of target features, leading to lower confidence of tracker. As the tracking feature template is being updated all the time, it will lead to the continuous introduction of wrong background information and gradually deepen the pollution degree of the tracking target model, and the tracking accuracy will decrease significantly in the following frames. To alleviate this problem, we must find a good feature extraction method to keep tracking. And different feature extraction methods will have some differences when applied to the same tracking algorithm.

When the tracking target is blocked by the background more than 40%, it is regarded as seriously blocked; when the target is completely blocked by the background, it is regarded as fully blocked, as shown in Figure 6.

In the case of severe occlusion or total occlusion, the tracker cannot correctly locate the tracking target because there is too much information for the target to be occluded and too little information for positioning. In addition, the sampling range of the tracker is limited, but the distance between its positioning coordinates and the position of the tracking target is too far. As a result, the model of the tracker will be damaged more and more seriously, resulting in tracking failure. Therefore, stopping the update of the target



FIGURE 5: Partial occlusion.

model when occlusion occurs, activating the detector to scan in a certain range until the target reappears, and continuing to keep tracking with the tracker is one of the effective methods to solve the problem of severe occlusion or full occlusion.

3. Solutions to Occlusion Problems

3.1. Traditional Solutions. The traditional methods to solve the occlusion problem mainly include the following four methods:

- (1) Target tracking algorithm based on feature point center weighting

This algorithm, represented by the mean-shift algorithm [13], firstly finds the distribution of the main feature points of the tracking target and then maximizes the weight of the feature points of the target center. The weight of the other feature points is carried out in inverse proportion to the distance from the target center. Finally, the tracking target is located through iterative search.

- (2) Template matching tracking algorithm based on multiparticle

The multiparticle template matching tracking algorithm, in essence, divides the tracking target into multiple tracking targets according to different regions for template matching and calculates the position of the tracking target through the position of

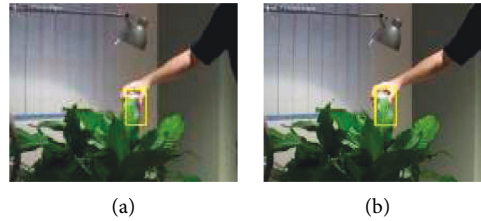


FIGURE 6: Severe occlusion and full occlusion. (a) Severe occlusion and (b) full occlusion.

multiple particles [14]. Particle filter has no special requirement for tracking model, can maintain the multimode distribution of state, and is not susceptible to clutter, so it has been greatly developed in the field of tracking. But the conventional particle filter tracking algorithm has some problems such as large computation and low sampling efficiency.

(3) Track tracking algorithm using the Kalman filter

When the tracking target is in the occlusion state, the Kalman filter tracking algorithm will predict the trajectory of the tracking target in the occlusion process according to the recorded coordinates of the previous target [15]. The Kalman filter has certain limitations on the tracking model, which can only deal with linear, Gaussian, and single-mode situations. However, in the application of image tracking, the actual tracking situation is often changeable and complex, which makes the application of the Kalman filter limited to a certain extent.

(4) Target tracking algorithm based on multi-algorithm fusion

For example, Kalman filter, mean-shift algorithm, and SVM classifier are fused in literature [16]. When the target is moving normally or partially occluded, the Kalman filter algorithm is used to predict the target position in the current frame according to the target position information of the previous frame, which is used as the iteration initial point of mean shift. In this way, the distance between the iteration initial point and the actual position of the target can be shortened, so as to reduce the number of iterations and improve the calculation of the algorithm. When the target is seriously blocked, an SVM detector is used to assist positioning, so that the re-detected target position is used as the starting point of the mean-shift algorithm to continue iterative tracking.

With the continuous deepening of domestic and foreign scholars' research, some algorithms with higher robustness, stability, and accuracy are developed based on the traditional algorithm to solve the occlusion problem.

3.2. Target Tracking Algorithm for Predicting Target Trajectory. The algorithm is mainly divided into three modules: image processing module, target automatic labeling and feature extraction module, and multitarget tracking module. The function of the image processing

module is to detect the moving target and eliminate the noise of the image to get the ideal binary image. The target automatic marking and feature extraction module is to label the binary image and calculate the target location of the center of mass, area, circumference, and other feature quantities; the multitarget tracking module is to feature track multiple targets in sequence images, and the segmentation measurement of short-range targets and target track crossing are processed in this module [17]. The basic flow of the algorithm is shown in Figure 7. The image preprocessing module detects moving objects and eliminates noise in the sequence images to obtain an ideal binary image. After that, the target is automatically marked to obtain the feature quantity of the image, and then, the multitarget tracking module tracks the features of the multiple targets in the sequence images. If the trajectory cross occurs between the target features, the Kalman trajectory is used to predict the target. If there is no cross between the target features, the Kalman filter is directly used to process the features for tracking.

3.3. Target Tracking Algorithm Based on Space-Time Context Information. In order to preserve the space-time context information, based on the Siam Mask algorithm, a short-term memory storage pool is introduced to store historical frame features, and an appearance distinctiveness feature enhancement module is also introduced. This not only enhances the salient features of the tracking target, but also suppresses the interference of the surrounding similar targets to the target, and finally achieves the purpose of improving the accuracy of the tracking algorithm.

Figure 8 shows the implementation process of the target tracking algorithm based on space-time context information [18]. This algorithm is mainly based on the Siam Mask algorithm [19] to construct the network system framework, which is divided into two branches, namely, twin network: the upper branch of the twin network takes 127×127 images as the input template and is mainly responsible for extracting the feature information of the video target; and the lower branch is modified to take 255×255 image as input template and is mainly responsible for extracting the feature information of the current frame of the video. The upper branch of the twin network only uses the shared weight ResNet-50 framework Φ to extract image information features, while the lower branch of the twin network also uses the short-term memory storage pool to retain the historical frame features of the video. Then, the lower branch captures the context information through the appearance distinctiveness feature enhancement module to enhance the distinctiveness

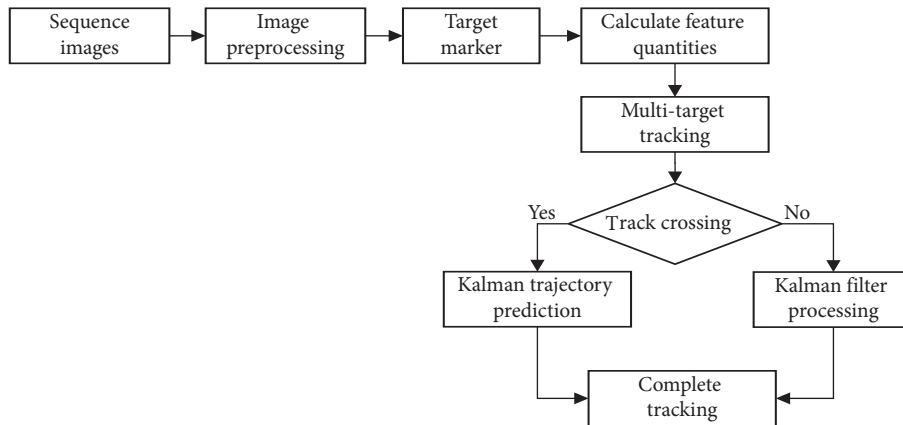


FIGURE 7: Target tracking algorithm flow for predicting target trajectory.

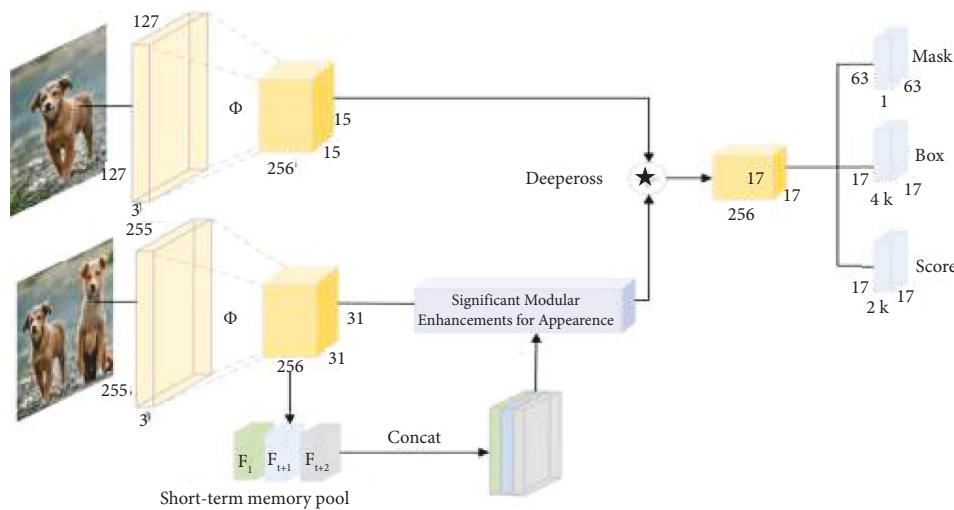


FIGURE 8: Target tracking algorithm flow based on space-time context information.

of the current frame feature and reduce the interference of similar objects in the environment. Next, cross-correlation is made between the feature image of the video target and the feature image of the current frame to get the response feature of the candidate region. Finally, the result of target tracking and segmentation is obtained by convolution activation.

3.4. Target Tracking Algorithm for Large-Scale Random Detection. On the basis of the tracking framework based on P-N learning, literature [20] proposes a target tracking algorithm adaptive to generate detection range. The Kalman filter is introduced to estimate the target position, scale, and their change speed, and the detection range is generated according to the estimated information before detection to improve detection efficiency.

The target tracking algorithm framework of large range random detection is mainly composed of four parts: tracker, detector, integrator, and learning module. The tracker is a mid-stream tracker that can detect tracking errors. When the target is visible, the tracker is responsible for tracking the target frame by frame and stops tracking when the error is greater than a certain value. The detector is a cascade

classifier that is responsible for detecting the target in each frame and restarting the tracker if necessary. The integrator evaluates the performance of the tracker and the detector according to the target model and fuses the results of both. The learning module generates training samples according to the results of the integrator and updates the detector to improve the performance of the detector in the next detection. It should be noted that the learning module only updates the detector, while the tracker remains fixed. The framework structure of the algorithm is shown in Figure 9. The specific implementation process is as follows: the estimated position and size of the target are filtered by the Kalman filter, the bounding box is initialized, and the following frame tracker and detector are processed together. The tracker estimates the state of the current target frame according to the target position and size of the previous frame. The detector scans the current frame completely in the sliding window mode and outputs one or several windows that may contain the target. The integrator processes the results of the tracker and the detector, and outputs the final tracking result of the current frame. Finally, the learning module updates the detector according to the result of the integrator to obtain the new target state, so as to realize tracking.

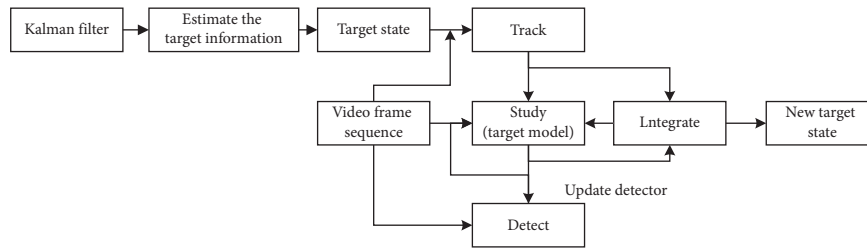


FIGURE 9: Target tracking algorithm framework of large-scale random detection.

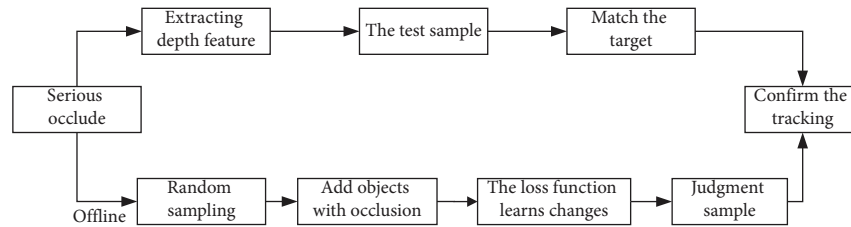


FIGURE 10: Deep learning flowchart.

3.5. Tracking Algorithm Based on Deep Learning. Due to its strong ability of tracking target recognition, the target tracking algorithm based on deep learning can recognize the tracking target well in the case of slight occlusion and partial occlusion. However, in the case of severe occlusion, tracking target information is seriously lost, and deep learning will adopt random sampling to detect target locations in a large range [21]. In the process of deep target tracking, there are two main methods to solve the problem of partial occlusion. One method is to add the target with occlusion in the training sample and learn the changes after the target is partially occluded through the loss function during offline training. Then, in the online tracking process, the network trained offline is used to judge the test sample so as to achieve accurate tracking. Another method is to divide the tracking target into several parts and extract the depth features of the target. When the target is blocked, the depth features of the test sample and the target template are matched. If the similarity degree is high, the sample is judged as the target. The deep learning flowchart is shown in Figure 10.

At present, the method to solve the problem of complete occlusion mainly relies on the full image search of the target detection. By matching the depth feature of the whole image with the feature of the target template, the test sample that may be the target can be found. However, due to the large size of the whole image and the large amount of data in the matching process, it is difficult to achieve accurate judgment of the target.

4. Anti-occlusion Model

In the process of target tracking, in view of the high difficulty and challenge of occlusion problem, scholars have classified and studied an effective anti-occlusion model from different aspects, which is of great significance to the construction of a high robustness tracking model.

4.1. Tracking Model Based on High-Quality Negative Training Samples. The current mainstream tracking models are driven by data, so improving the quality of training samples can significantly improve the performance of the tracker. By creating high-quality training samples, the tracking model can learn the discriminant information to distinguish the occluded target from the interfering object, so as to improve the anti-occlusion performance of the tracking algorithm. When the tracking target is occluded, the tracking model tends to drift to the occluded object. When the tracking target is severely occluded or fully occluded, the tracking model drifts to the semantic interference object, resulting in the tracking failure. In this case, by solving the problem of unbalanced distribution of training samples in the tracking task, the accuracy of the tracking model can be effectively improved. Therefore, sufficient negative training samples prone to false detection should be introduced or mined, so as to avoid a large number of false negative samples dominating the learning of the tracker, so that the tracking model can correctly distinguish the features of positive and negative semantic samples, and finally achieve the purpose of improving the anti-occlusion interference ability of the model. Table 2 shows three typical algorithm models of high-quality negative training samples and their respective strategies. These three algorithm models are the influence model of improving loss function to reduce invalid negative samples, the algorithm model of mining a few negative samples that are easy to misjudgment, and the high-quality negative sample model of real scenes, respectively.

4.2. Tracking Model Based on Effective Positive Sample Set.

When the target is occluded, the state of the target itself will change due to the interference of the outside world, and the occluded object will pollute the positive sample. The improvement of positive sample quality is beneficial to the feature discrimination of the tracker model and the ability to resist occlusion and background interference. The diversity

TABLE 2: Representative tracking model and strategy characteristics of high-quality negative training samples.

Methods	Algorithm	Negative training sample strategy	Algorithm is limited
The loss function is optimized to reduce invalid negative samples	VITAL [22]	The original cross entropy loss function is optimized into high order sensitive loss function	The improvement degree of cost-sensitive function is not high
	DSLTL [23]	The traditional ridge regression loss function is optimized into shrinkage loss function	The performance of the residual connection removal method does not decrease obviously
	DiMP [24]	The hinge-like loss was obtained by optimizing the focal loss function	Unable to handle large frame-to-frame shifts efficiently
Mining a few negative samples easy to misjudge	RPNT [25]	Candidate generation network generated background interference samples and distant similar appearance semantic interference samples	The features covered by semantic information are weak in discriminating between classes
	MDNet [26] SANet [27]	In the selection of small batch training samples, the samples with the highest positive score in the negative samples were selected to identify hard negative samples, and the degree of hard negative samples was increased in the process of iterative training tests	There is no optimization strategy for extracting potential targets, and it is difficult to apply panoramic scenes due to insensitive differences within classes, time-consuming of feature extraction and fusion modules, poor real-time performance, and limited estimation of target scale changes
	RT-MDNet [28]		
Introduce high-quality negative samples of real scenes	BACF [29]	The negative sample containing real background information is extracted from the most dense part of the image by clipping operation	Lack of temporal stability, fixed mask selection features have limitations
	DaSiamRPN [30]	Negative sample pairs, performing data enhancement such as translation, light transform, and blur are introduced from large-scale data sets	Candidate box extraction and region search take a long time, and the tracker speed decreases

TABLE 3: Representative models and characteristics based on valid positive sample sets.

Methods	Algorithm	Sample processing strategy and characteristics	Applicable scene
Improve sample quality	CFLB [31]	Binary mask matrix segmented cyclic shifted samples, increase the proportion of real training samples	Partial occlusion, scale illumination changes
	BACF [29]	The sampling area is enlarged to the whole image to extract high-quality positive samples, can remove interference information	Scale variations, background clutter, partial cover
	SRDCFdecon [32]	Dynamically manage training sample sets and capture valid information, error messages can be corrected	Background clutter, partial cover, in-plane rotation
Increase sample diversity	ECO [33]	Gaussian mixture sample generation model, the appearance difference between samples is large, redundant information can be removed	Serious barrier, in-plane rotation, be real-time
	TRACA [34]	Unsupervised clustering filtering reversal of context information, small cluster sample groups can be avoided	Occlusion completely, deformation, be real-time
	DaSiamRPN [30]	A positive sample pair is introduced to perform translation clipping transformation	Occlusion completely, out of sight, similar interference
	UPDT [35]	The wide Gaussian label function expands the area, rotates the blur operation, and captures semantic information	Occlusion completely, background clutter, out of sight, scale variations
	SINT++ [36]	Samples were collected by traversing the target manifold structure to generate objects that did not appear in the training data	Partial cover, deformation, elapsed time
Reference [37]	Generate adversarial network, and generate deformation, fuzzy samples, adaptive learning discrimination features	Deformation, motion blur, partial cover	

of samples can also enhance the generalization performance of the model and improve the ability of tracking deformation targets in occlusion scenarios. Table 3 analyzes the positive sample processing strategies and characteristics of different algorithms from two aspects of improving sample quality and increasing sample diversity, and presents the scenarios to which they are adapted.

4.3. Adaptive Tracking Algorithm Model Based on Active Learning. In the process of target tracking, the training characteristics of the tracking algorithm online and offline will be different. Although online training can better adapt to the appearance changes of the tracked target, there are few real labeled samples and the tracking target status is changeable, which makes it easy to over-fit sample errors

TABLE 4: Representative tracking algorithm and characteristics based on active learning strategy.

Methods	Algorithm	Algorithm model learning strategy	Applicable scene	
Construct the target model with high robustness	Feature fusion	SANet [27]	Structural perception learning, RNN self-modeling target structure, CNN, and RNN characteristics skip connection	The network structure is complex, time-consuming and difficult to track in real time
		SiamRPN++ [38]	Mitigating the impairment caused by network shift and removing the network depth limit, hierarchical aggregation, and depth cross-correlation	The linear weighting feature is simple and the model lacks self-adaptability
		UPDT [32]	Deep and shallow features are enhanced with different data, adaptive fusion strategy detects response graph	Feature extraction by deep convolution is time-consuming, and the algorithm is difficult to be applied to real scene
	Multidomain learning	MFT [39]	Gaussian windows with different bandwidths are used to train the corresponding characteristic filters independently, adaptive weight fusion multilayer response	The performance of adaptive weight and target offset estimation is low
		MDNet [26]	Sharing common characteristics independent of each domain, multiple branches capture target information under each domain attribute	The learning discriminant background information algorithm has limitations
		ANT [40]	The training samples in the data set are classified and tracked according to different attributes, Multibranch learning occlusion, and other multiclass tracking attributes	The calculation complexity of network parameters is high and difficult to implement
Generative adversarial networks	VITAL [22]	To generate the feature masks corresponding to the appearance changes of different types of targets, identify medium- and long-term robust features in the time domain to avoid fitting short-term discriminant feature information	Network prediction weight mask resolution is low, not suitable for low-pixel and large-scale deformation	
Improve the anti-interference ability of the model	Interferes with the ability to perceive objects	RT-MDNet [28]	Embedding loss function is introduced to increase class spacing in shared feature space, learn to identify instances of semantic objects across domains	The target positioning effect is limited in high precision area
		DaSiamPRN [30]	The reordering of semantic interference context information incremental learning target and interference perceptual candidate is introduced	The network depth is limited, and the features extracted are limited
		TADT [41]	Frame-by-frame regression loss function and rank loss function, select feature channels for instance tracking targets	There are algorithm limitations caused by fixed matching between templates
		MLT [42]	According to the gradient information, the feature space can be customized for each object to suppress the error response of background interference source	Meta-learning, migration, and merging are simple and low precision

and increase the tracking time in the complex scene covered. Offline training can better deal with occluded targets and avoid model contamination caused by learning wrong frames. However, in the real tracking process, there will be various complex problems, and it is often difficult to track the target accurately only by using the fixed appearance model of the offline training set. Therefore, in order to improve the performance of the tracker, active learning of the tracking model is required. The target model with high robustness is constructed through offline learning, and the

antifitting error damaged information is learned online to enhance the anti-interference of the model. Table 4 lists representative algorithms and characteristics of active learning strategies in tracking models.

5. Conclusions

In this article, the interference of occlusion problem in target tracking is analyzed. The most fundamental influence of occlusion on the tracking algorithm is that it destroys the

target location of the tracker. Some trackers with errors in sample analysis caused by occlusion can improve the performance of trackers by improving the quality of negative training samples and enriching the diversity of positive sample sets. According to the different training characteristics of online and offline tracking algorithms, anti-occlusion models suitable for active learning algorithms in different tracking situations are found, and tracking models with high robustness and high accuracy are selected in different situations, which helps to better complete the target tracking process. With the deepening of domestic and foreign scholars' research, the development of target tracking is becoming more and more perfect. Target occlusion is not only an important cause of tracking failure, but also a key problem of remote target tracking. The tracking task only tracks one target from beginning to end; once the target is blocked, the tracking accuracy will be greatly affected and even lead to tracking failure. Therefore, when facing the occlusion problem, the target tracking model is more strict. Research in the field of target tracking has been carried out for many years, and great progress has been made from the earliest generative algorithm to the algorithm based on correlation filtering and then to the current algorithm based on deep learning. Target tracking has also been widely used in many industries, and more and more people begin to pay attention to the development of target tracking. Therefore, to better solve the occlusion problem is still an important research direction in the future development. With adversarial generative learning matures, meta-learning, and other methods, the future tracker is expected to provide more prior multidimensional information for the tracker from capturing the target to tracking the scene and movement information of the target, to migrating to the long-term complex tracking task, so as to track the target more accurately. An algorithm with strong robustness, high accuracy, and high speed is still expected in the future, and it is believed that with the joint efforts of researchers, it will not take long to complete this task.

Data Availability

The authors confirm that the data supporting the findings of this study are available within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was approved by the Foundation of the Shaanxi Key Laboratory of Intelligent Processing for Big Energy Data (Grant no. IPBED17) and the Project of Shaanxi Provincial Department of Education, China (Grant no. 20JK983).

References

- [1] D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, The MIT Press USA, Cambridge, MA, USA, 2010.
- [2] W. Zhang, *Research on Spatio-Temporal Context Aware Target Tracking Method Based on Correlation Filtering*, Northwestern University, Evanston, Illinois, 2019.
- [3] D. W. Wu, "Research on target tracking algorithm based on adaptive deformation in intelligent traffic," *Shanxi Architecture*, vol. 46, no. 6, pp. 196–198, 2019.
- [4] J. N. Zhou, J. M. Duan, and G. Z. Yang, "Vehicle tracking method based on improved deterministic target association," *Pattern Recognition and Artificial Intelligence*, vol. 27, no. 1, pp. 89–95, 2014.
- [5] Q. C. Duan, T. F. Yuan, Y. Q. Wang, and P. P. Huangfu, "Target tracking system of UAV based on Kalman Filter.- Unmanned aerial vehicle target tracking system based on kalman filter," *Intelligent computers and applications*, vol. 10, no. 10, pp. 92–94, 2020.
- [6] M. B. Lu, X. F. Zhao, and G. Z. Liu, "Aerial high maneuvering target tracking algorithm. Modern defence technology," *Modern Defense Technology*, vol. 46, no. 2, pp. 45–50+172, 2018.
- [7] X. Chen, Y. F. Han, Y. F. Yan, D. L. Qi, and J. X. Shen, "A unified algorithm for object tracking and segmentation and its application on intelligent video surveillance for transformer substation," in *Proceedings of the CSEE*, vol. 40, no. 23, pp. 7578–7587, 2020.
- [8] K. W. Liu, *Research on Key Technology of Template Updating in Visual Target Tracking*, Shandong Technology and Business University, Yantai, China, 2021.
- [9] C. S. Hu, X. P. Yan, H. X. Wei, and G. L. Li, "Survey of target detection and trajectory prediction based on stereo vision," *Computer Engineering and Applications*, vol. 58, no. 03, pp. 50–65, 2020.
- [10] J. Wen and Q. Li, "Object tracking algorithm based on spatio-temporal context information enhancement," *Journal of Computer Applications*, vol. 41, no. 12, pp. 3565–3570, 2021.
- [11] Y. J. Huang, J. Y. Zhao, T. T. Cheng, and P. Q. Chen, "An adaptive approach for generating detective range in visual object tracking," *Journal of Ningbo University (Nsee)*, vol. 28, no. 04, pp. 36–41, 2015.
- [12] X. Li, Y. F. Zha, T. Z. Zhang et al., "Survey of visual object tracking algorithms based on deep learning," *Chinese Journal of Image and Graphics*, vol. 24, no. 12, pp. 2057–2080, 2019.
- [13] H. Shen, J. J. Zhuang, Q. Y. Zheng et al., "Design and implementation of target tracking system based on mean shift algorithm," *Electronic Measurement Technology*, vol. 41, no. 14, pp. 11–15, 2018.
- [14] B. Zhang, *Research on Visual Tracking Algorithm Based on Particle Filter*, Shanghai Jiaotong University, Shanghai, China, 2007.
- [15] X. B. Liang, *Research on On-Orbit Target Tracking Algorithm Based on Kalman Filter [D]*, University of Chinese Academy of Sciences, Beijing, China, 2017.
- [16] Y. Lin and W. B. Tang, "Anti-occlusion target tracking algorithm based on multi-algorithm fusion," *Modern Computer*, vol. 27, no. 31, pp. 40–45, 2021.
- [17] Y. Zhang, X. Y. Cai, J. H. Yan, K. Su, and K. Zhang, "Aerial multi-target detection based on spatial-temporal information and trajectory association," *Geomatics and Information Science of Wuhan University*, vol. 45, no. 10, pp. 1533–1540, 2020.
- [18] H. P. He, M. Zhao, and D. H. Sun, "Image Edge Detection Based on MeanShift Algorithm," *Journal of Chongqing Institute of Technology (Natural Science Edition)*, vol. 23, no. 08, pp. 122–125, 2009.

- [19] X. T. Cheng and Z. M. Tang, "Scale Adaptive Target Tracking Algorithm for Robot," *Computer Science*, vol. 41, no. 12, pp. 280–282+292, 2014.
- [20] K. B. Chu, L. Zhu, and J. Zhang, "Improved TLD target tracking algorithm based on KCF and HOG," *Journal of Changzhou University (Natural Science)*, vol. 34, no. 1, pp. 60–67, 2022.
- [21] H. R. Bai, "A review of depth single target tracking algorithm," *Modern Computer*, no. 09, pp. 113–116, 2021.
- [22] Y. Song, C. Ma, X. Wu et al., "Vital: Visual Tracking via Adversarial Learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8990–8999, Salt Lake City, UT, USA, June 2018.
- [23] X. Lu, C. Ma, B. Ni, X. Yang, I. Reid, and M.-H. Yang, "Deep regression tracking with shrinkage loss," *European conference on computer vision (ECCV)*, Springer, Cham, New York, NY, USA, 2018.
- [24] G. Bhat, M. Danelljan, L. Van Gool, and T. Radu, "Learning discriminative model prediction for tracking," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6182–6191, Seoul, Korea, October 2019.
- [25] G. Zhu, F. Porikli, and H. Li, "Robust visual tracking with deep convolutional neural network based object proposals on pets," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 26–33, Las Vegas, NV, USA, June 2016.
- [26] H. Nam and B. Han, "Learning multi-domain convolutional neural networks for visual tracking," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4293–4302, Las Vegas, NV, USA, June 2016.
- [27] H. Fan and L. H. Sanet, "Structure-aware network for visual tracking," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 42–49, 2017.
- [28] I. Jung, J. Son, M. Baek, and B. Han, "Real-time mdnet," in *Proceedings of the European conference on computer vision (ECCV)*, pp. 83–98, 2018.
- [29] H. Kiani Galoogahi, A. Fagg, and S. Lucey, "Learning background-aware correlation filters for visual tracking," in *Proceedings of the IEEE international conference on computer vision*, pp. 1135–1143, 2017.
- [30] B. Zhu, Z. Zhu, J. Wang et al., "Chinese woodchucks with different susceptibility to WHV infection differ in their genetic background exemplified by cytochrome B and MHC-DRB molecules," *Virology Journal*, vol. 15, no. 1, pp. 101–117, 2018.
- [31] H. Kiani Galoogahi, T. Sim, and S. Lucey, "Correlation Filters with Limited Boundaries," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4630–4638, Boston, MA, USA, June 2015.
- [32] J. Choi, H. J. Chang, T. Fischer et al., "Context-aware deep feature compression for high-speed visual tracking," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 479–488, Salt Lake City, UT, USA, June 2018.
- [33] G. Bhat, J. Johnander, M. Danelljan, F. S. Khan, and M. Felsberg, "Unveiling the power of deep tracking," *European Conference on Computer Vision (ECCV)*, Springer, Cham, New York, NY, USA, 2018.
- [34] M. Danelljan, G. Häger, F. S. Khan, and M. Felsberg, "Adaptive decontamination of the training set: a unified formulation for discriminative visual tracking," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1430–1438, Las Vegas, NV, USA, June 2016.
- [35] M. Danelljan, G. Bhat, F. S. Khan, and M. Felsberg, "Eco: efficient convolution operators for tracking," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 6638–6646, 2017.
- [36] X. Wang, C. Li, B. Luo, and J. Tang, "Sint++: robust visual tracking via adversarial positive instance generation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4864–4873, Salt Lake City, UT, USA, June 2018.
- [37] Y. Han, P. Zhang, W. Huang, Y. Zha, G. D. Cooper, and Y. Zhang, "Robust visual tracking based on adversarial unlabeled instance generation with label smoothing loss regularization," *Pattern Recognition*, vol. 97, Article ID 107027, 2020.
- [38] B. Li, W. Wu, Q. Wang, F. Zhang, J. Xing, and Y. Junjie, "Siamrpn++: evolution of siamese visual tracking with very deep networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4282–4291, Long Beach, CA, USA, June 2019.
- [39] S. Bai, Z. He, T.-B. Xu, Z. Zheng, D. Yuan, and H. Bai, "Multi-hierarchical independent correlation filters for visual tracking," in *Proceedings of the 2020 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, IEEE, London, UK, July 2020.
- [40] Y. Qi, S. Zhang, W. Zhang, L. Su, Q. Huang, and M.-H. Yang, "Learning attribute-specific representations for visual tracking," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 8835–8842, Honolulu, Hawaii, USA, January 2019.
- [41] X. Li, C. Ma, B. Wu, Z. He, and M. -H. Yang, "Target-aware deep tracking," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1369–1378, Long Beach, CA, USA, June 2019.
- [42] J. Choi, J. Kwon, and K. M. Lee, "Deep meta learning for real-time target-aware visual tracking," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 911–920, Seoul, Korea, October 2019.