

Retraction

Retracted: Target Tracking Algorithm Based on Adaptive Scale Detection Learning

Complexity

Received 23 January 2024; Accepted 23 January 2024; Published 24 January 2024

Copyright © 2024 Complexity. 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.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] D. Yang, "Target Tracking Algorithm Based on Adaptive Scale Detection Learning," *Complexity*, vol. 2021, Article ID 9033912, 11 pages, 2021.

Research Article

Target Tracking Algorithm Based on Adaptive Scale Detection Learning

Dawei Yang 

School of Information Technology and Engineering, Shenyang Ligong University, Liaoning, Shenyang 110159, China

Correspondence should be addressed to Dawei Yang; dwyang@sylu.edu.cn

Received 16 April 2021; Revised 17 May 2021; Accepted 29 May 2021; Published 7 June 2021

Academic Editor: Zhihan Lv

Copyright © 2021 Dawei Yang. 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.

In this paper, to better solve the problem of low tracking accuracy caused by the sudden change of target scale, we design and propose an adaptive scale mutation tracking algorithm using a deep learning network to detect the target first and then track it using the kernel correlation filtering method and verify the effectiveness of the model through experiments. The improvement point of this paper is to change the traditional kernel correlation filtering algorithm to detect and track at the same time and to combine deep learning with traditional kernel correlation filtering tracking to apply in the process of target tracking; the addition of deep learning network not only can learn more accurate feature representation but also can more effectively cope with the low resolution of video sequences, so that the algorithm in the case of scale mutation achieves more accurate target tracking in the case of scale mutation. To verify the effectiveness of this method in the case of scale mutation, four evaluation criteria, namely, average accuracy, cross-ratio accuracy, temporal robustness, and spatial robustness, are combined to demonstrate the effectiveness of the algorithm in the case of scale mutation. The experimental results verify that the joint detection strategy plays a good role in correcting the tracking drift caused by the subsequent abrupt change of the target scale and the effectiveness of the adaptive template update strategy. By adaptively changing the number of interval frames of neural network redetection to improve the tracking performance, the tracking speed is improved after the fusion of correlation filtering and neural network, and the combination of both is promoted for better application in target tracking tasks.

1. Introduction

More than 90% of the information that humans use to understand the world comes from vision, and the main goal of computer vision is to enable computers to see the world as humans do. The theory of visual computing has made great progress over the years, and in recent years, with advances in computer hardware and software equipment, computer vision technology has been widely used in various areas of life [1]. Computer vision is a discipline that spans several fields, including computer graphics, computer system architecture, information retrieval, neurological disciplines, and natural language processing. Computer vision is also one of the popular research areas for deep learning techniques. Computers have to acquire external information through auxiliary devices such as smart cameras, frame

receivers, and visual interfaces, etc. There are two main categories of image information that computers acquire, one of which is static images, mainly pictures, such as face recognition technology, and face recognition is a very popular application in the field of computer vision [2]. The other category is dynamic content, including video and three-dimensional capture, such as video analysis; through the video content analysis, retrieval, extraction of which has the use of information for further processing applications; video analysis is widely used in many fields such as security, transportation, and even retail. Target tracking as a branch under the field of computer vision has a wide range of applications [3]. Target tracking technology is mainly to standardize the information collected by the device, feature extraction, model building, and other operations to obtain the target size and location information, and feedback to the

system model to achieve the application in actual scene. Target tracking has important applications in several fields [4].

Due to the large and difficult to process video data, traditional algorithms have great limitations for video information analysis and processing. In recent years, with the advancement of the information age and the popularity of hardware, the efficiency of big data processing has been greatly improved, and video target tracking technology has been able to make a big impact. Visual target tracking technology is given the size and location of the target to be tracked in the initial frame of the video image, relying on algorithms to predict the target's position, direction, trajectory, and other motion information in subsequent frames to complete the analysis and understanding of the behavior of the moving target, the purpose of which is to use the computer instead of the human brain to complete time-consuming and laborious or high-risk and difficult specific tasks. Because visual target tracking technology has more outstanding advantages and wide applications than detection and recognition in the field of computer vision, it has attracted a wave of research from industry and scholars in recent years. However, many challenges in its research have plagued many researchers, such as illumination changes, scale changes, fast deformation, motion blur, background mottling, and target occlusion, which can affect tracker performance [5]. For target tracking systems, tracker performance is not only affected by the degree of model merit but is also related to the variation of the target itself and the environment. For tracking a moving target, the apparent state of the target and the background environment can change to different degrees, such as rotation, deformation, motion blur, etc., and target occlusion, illumination changes, background clutter, etc. These uncertainties can lead to inaccurate apparent modelling and frame by frame accumulation of errors, which eventually cause tracking "drift" or incorrect prediction of similar target position information [6].

The main research of this paper is to adopt the idea of detection before tracking for tracking, based on the kernel correlation filtering algorithm for the problem of the sudden change of target scale to analyze both template update and scale adaption and propose improvement methods to enhance the tracking accuracy and robustness of the algorithm. Section 1 is the introduction, which introduces the research background and significance of video tracking technology and also explains the research framework of this paper. Section 2 is related work, which mainly analyzes the current research status of target tracking techniques. Section 3 is the study of target tracking algorithms with adaptive scale detection learning. This chapter focuses on the improvement strategy of scale adaption. Firstly, the scale estimation of KCF is introduced, and then the scale estimation mechanism of this paper is introduced from three aspects: scale prediction, feature extraction, and position prediction. Section 4 is the result analysis, and this chapter mainly tests the algorithm proposed in this paper. Firstly, the evaluation index of target tracking is introduced, and then the more mainstream tracking algorithms are compared with the algorithm

of this paper, for the tracking process will occur due to the complete occlusion of the target after. Finally, this paper summarizes and analyzes the research content and results, analyzes the shortcomings of the algorithm, and provides an outlook on the research direction and further improvement direction.

2. Related Work

The research of target tracking technology can be traced back to the last century, and the early target tracking algorithms focus more on the changes of target feature points, and the optical flow method is a representative classical early tracking algorithm [7]. Abdhussain et al. proposed to track the target using optical flow points, mainly by extracting the pixel feature points for the target appearance model, then finding the optical flow matching points of the feature points in the adjacent frames, and finally using the pixel feature points; the unique changes of the pixel feature points are used to estimate the motion state of the target, thus realizing the tracking of the target. However, there are many limitations of target tracking using the optical flow method, so many improved algorithms based on the optical flow method have appeared [8]. Elisei-Iliescu et al. used Harris feature points to replace the pixel feature points in the optical flow method, which reduced the computational effort of the algorithm and improved the tracking speed [9]. Khan et al. used the idea of introducing foreground constraints on the tracking framework of the optical flow method to improve the matching accuracy and success rate of the algorithm. The improved algorithm based on the optical flow method improves the tracking effect to some extent, but such algorithms still have many drawbacks and high computational complexity [10]. Jumani et al. proposed a method of training continuous convolution filters in C-COT (Continuous Convolution Operators for Visual Tracking). In order to solve the problem of different resolutions of different convolution layers, a frequency domain implicit interpolation model was used to integrate the feature map. Interpolation into the continuous spatial domain facilitates the integration of multiresolution feature maps, so that feature maps of different resolutions can be input to the filter to estimate the target position [11].

Yuan et al. pointed out that feature extraction is a crucial part of target tracking, and the integrated model has strong generalization and effectiveness, and if the features are robust enough, the observation model is not so important [12]. Wang et al. proposed to extract the motion features of multiple videos by neural networks in MDNet (Multidomain Network) based on the idea of transfer learning and migrated the features of the target classification problem to the tracking field. Although the speed is difficult in real-time, the accuracy achieves state-of-the-art [13]. Luo et al. proposed HDT (Hedged Deep Tracking) to train filters using features of different depths of the VGG16 network while using the idea of adaptive integration learning to integrate multiple trackers into a more robust tracker [14]. With the application of deep networks, scholars have found that powerful network structures bring much less significant improvement

in the field of tracking than in other branches of computer vision; powerful network structures cannot be as powerful as they should be in the field of tracking [15].

In this paper, we analyze the theoretical and practical significance of target tracking technology, understand the challenges in the field of target tracking in detail, analyze the theory of deep learning technology, gain insight into the development process and technical advantages of deep learning technology, and propose the convolutional regression network algorithm in this paper based on the advantage of the speed of the correlation filtering algorithm after analyzing the traditional correlation filtering algorithm, and combine it with the advantage of the strong feature expression capability of deep learning to target. The traditional correlation filtering algorithm based on artificial features is less effective in complex environments such as out-of-field, low resolution, and motion blur [16]. The analysis of experimental results shows that the algorithm in this paper improves in all aspects. In the process of target tracking, the target model needs to be updated continuously to ensure the robustness of the target model for target tracking. However, the strategy of updating the model with a fixed learning rate frame by frame will accumulate errors in the process of updating and learning, which affects the tracking effect. Therefore, this paper proposes an adaptive model update method to achieve a targeted selection of learning rates for model updates [17]. In this paper, the model learning rate is divided into two parts: fixed learning rate and adaptive learning rate, and different learning rates are selected by judging whether the ratio between the confidence of the current frame tracking result and the maximum confidence of the historical frames is higher than the present update threshold.

3. Study of Target Tracking Algorithm with Adaptive Scale Detection Learning

3.1. Adaptive Scale Target Model. Occlusion redetection mechanism is used to deal with the situation when the target is occluded. Determine the reliability of the tracking result by calculating the difference between the credibility of the current frame tracking result and the historical credibility average and compare the difference with the present redetection threshold, and then determine whether the target needs to be redetected. If the difference is higher than the redetection threshold, it indicates that the peak response of the target area fluctuates sharply, and the target may be blocked or other conditions. Start the occlusion redetection model, expand the search area, and redetect the target position; otherwise, it indicates that the peak fluctuation range is relatively flat. The tracking result is highly reliable, and the model is updated normally without a redetection process. The method of adaptively updating the model realizes the targeted selection of the learning rate to update the model. This paper divides the model learning rate into two parts: a fixed learning rate and an adaptive learning rate. Different learning is selected by judging whether the ratio of the credibility of the current frame tracking result to the maximum credibility of the historical

frame is higher than the present update threshold. If it is higher than the update threshold, it means that the current frame information is more reliable, and a fixed initial learning rate is used to update the model; otherwise, it means that the current frame information is less reliable, and the model is updated with an adaptive learning rate. The adaptive update learning rate is determined by the average response of the current frame, the peak response of the current frame, and the average response of the previous frame, which effectively improves the anti-interference ability of the environment during the model update process.

In this paper, the target model is built based on the correlation filter and improved from the perspective of features, occlusion rechecking, and update strategies. This section introduces the correlation filter model, the color probability model, and the overall framework of the target model to be used in the target model building process and provides an overall visual overview of the algorithm proposed in this paper. The establishment of the target model in this paper draws on the idea of the correlation filter model in the FDSST algorithm [18].

The algorithm focuses on locating the position of the target in the current frame by training a position filter. The location filter mainly extracts the HOG features in the search region and transforms the extracted features into a response map using a correlation filter. In the filter response map, the peak represents the highest response score at the location, indicating that the location is the most likely to be the target location. The target model in this paper is based on this idea and aims to construct an optimal correlation filter L_i to establish the minimization cost function [19]. The formula is equation (1), $g(x)$ denotes the expected response value calculated from the Gaussian function, l represents the channel of the image feature, $g(i)$ is the feature extracted for the l th dimension, and the second term in the formula denotes the regularization term with the weight parameter β .

$$G(x) = \frac{\beta l(x)}{|w - \sum_{i=1}^m L_i v_i|^2 + \beta l(x)}. \quad (1)$$

To reduce the time consumed in the calculation process, equation (1) is transferred from the time domain to the frequency domain, and the corresponding transformation result is equation (2), where U , W , and V correspond to L_i , w , and v in equation (1) after Fourier transformation to the frequency domain, respectively. The w denotes the complex conjugate, and V is the optimal filter to be calculated. The filter model needs to be updated during the tracking process to ensure the accuracy of the filter model for the characterization of the target.

$$U(x) = \sum_{i=1}^d W_i V_i + \frac{\beta}{\beta + 1} W(x). \quad (2)$$

The model update strategy is equation (3), where Y and Z denote updates to the numerator and denominator parts of equation (2).

$$\begin{cases} Y = \frac{1-\varphi}{1+\varphi} WU_t^l, \\ Z = \frac{\varphi}{1+\varphi} 1, \end{cases} \quad l \subseteq [1, m]. \quad (3)$$

A single feature describes the target with limitations, and in the position filter, the response image is obtained by extracting HOG features for calculation. However, a single HOG feature cannot describe the information of the target comprehensively [20]. Therefore, this paper establishes a color probability model based on the processing method of color histogram features for extracting the color histogram features of the target to enhance the target description capability of the target model.

In this paper, by constructing mutually independent position filters and color probability models, the HOG features and color histogram features of the target are extracted in the search area, respectively, and the response images of the two features are computed, and the response results are analyzed and processed to obtain the position of the target in combination with the scale filters, and the model is updated in time. The tracking framework in this paper is mainly divided into three stages: modelling, tracking, and updating, and the overall framework is shown in Figure 1.

3.2. Adaptive Scale Target Tracking Algorithm. In practical target tracking application scenarios, there may be some differences in color between the target object and background information, so the color information of the target is also a very critical research point in some tracking algorithms, and it is especially important to select the color information with strong descriptive power [21–23]. The color features are features of global nature, which are usually based on the representation of pixel points and express the surface properties of the image or the corresponding object in the image region. That is, all pixel points in the image region contribute to the color features of the region, so the color features are not very sensitive to the changes of size and direction, etc. in the image region, and thus the color features are not very sensitive to the local changes of the target [24–27]. The response is not very sensitive to local changes in the target. Also, if only color features are used alone in the tracking process, they are easily associated with other unrelated images.

In this paper, the algorithm uses a modified SURF algorithm based on Krawtchouk moments, and when the calculated value is below the search range update threshold, the modified SURF algorithm is used to correct the target location to improve the accuracy of the algorithm. The improved algorithm uses the n th order Krawtchouk polynomial, normalized Krawtchouk polynomial as in equation (4), where $a = 0, 1, \dots, N$, $b > 0$, $E \in [0, 1]$.

$$KC_m(a, b, E) = kc(a, b, E)(\beta(a, b, E))^2. \quad (4)$$

The target state given in the initial frame in target tracking is shown in equation (5), where x_i, y_i is the upper left corner of the target Ground Truth, and w_i, h_i denote the width and height of the tracked target, respectively. The real value given in the initial frame is used as the initialized positive sample of the tracked target, while the negative sample is obtained by random sampling around the real position. Considering the limited number of positive samples, this paper expands them reasonably.

$$W_i = \{x_i, y_i, w_i, h_i\}, \quad i \in [0, n], n > 1. \quad (5)$$

After the training is completed, the target position estimation is started for the subsequent frames. The center of the predicted tracking result of the $m-1$ frame is taken as the center, and the search region is obtained by extending M pixel points in all directions I . The confidence A of each sample can be obtained by the sigmoid classification layer, and the sample with the largest confidence is the estimated result.

$$A_m = \ln(\max(\omega_i)) = \{x_m, y_m, w_m, h_m\}. \quad (6)$$

To reduce the influence of noise features on the model update, the algorithm in this paper adopts a predefined threshold update strategy to update the model adaptively. When the target location may be out of the previous local search range, then the search area needs to be moderately expanded to improve the accuracy of the algorithm, the conditions for expanding the search range are as follows:

$$\max(\omega_i) < \lambda_i + \beta \sum_{i=1}^m \omega_i. \quad (7)$$

The λ_i is the search range update threshold, and when equation (7) holds, the search area is expanded according to equation (8), where N is the initial search range and β is the search range increment.

$$M = \frac{M}{1+M} \left\| 1 - \sum_i^n \omega_i \right\|^{1/3} - 2\beta. \quad (8)$$

The main algorithm of the robust motion target tracking algorithm with depth feature fusion is described above. The algorithm flow chart of this paper is shown in Figure 2. In this paper, a 32×32 image is used as the input, so the input layer dimension is set to 1024. After that, the computational complexity and data loss are traded off, and the hidden nodes are halved layer by layer to achieve the purpose of data reduction and compression.

3.3. Target Tracking Evaluation Indicators. Target tracking effect evaluation generally has two metrics: one is used to evaluate the tracking accuracy; the other is used to evaluate the scale adaptive ability of the tracking algorithm. One of the commonly used metrics is the offset of the center coordinate position: the second commonly used metric is the size of the intersection ratio of two areas [28–30]. The average pixel error is based on the distance between the center position of the predicted target and the pixel of the

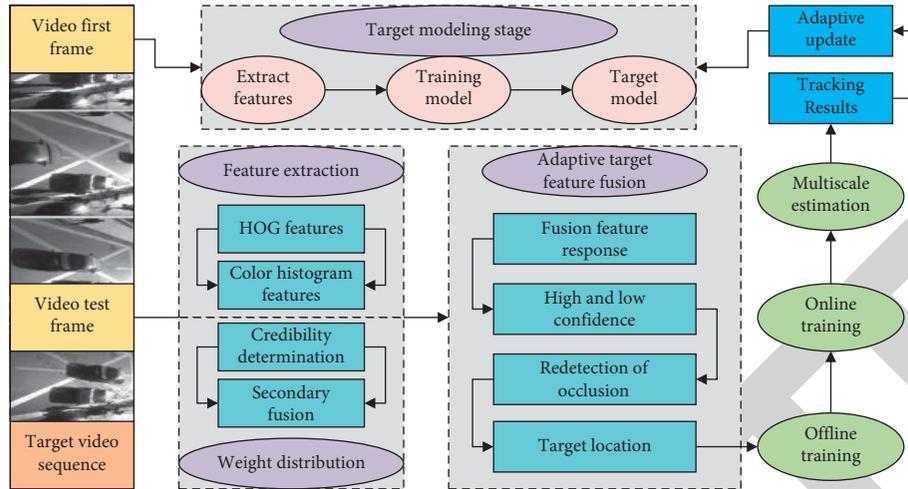


FIGURE 1: General framework diagram.

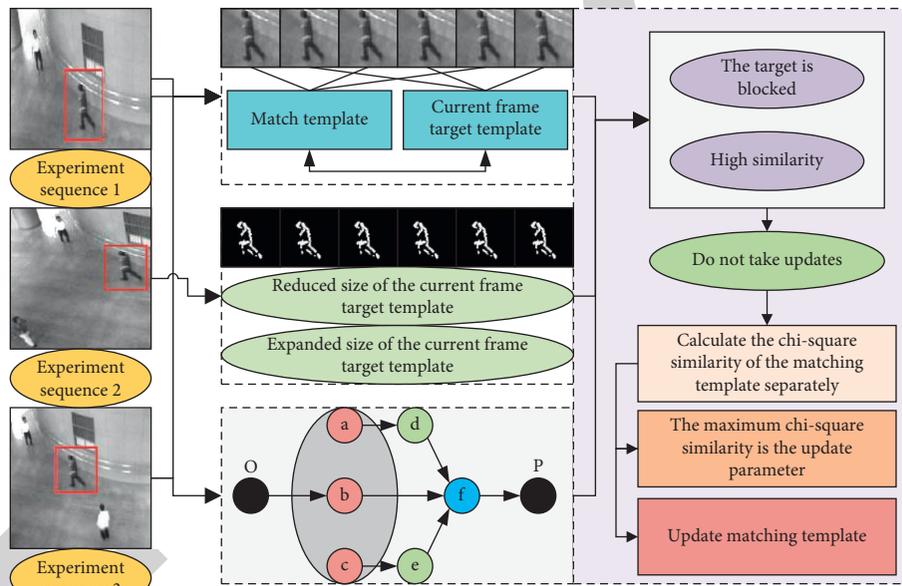


FIGURE 2: Algorithm flow chart.

artificially labelled real position, and the distance between them is judged as accurate tracking if it is less than the given threshold and as tracking error if it is greater than the given threshold, and the percentage of video frames less than the given threshold to the total number of frames is calculated, and the larger the value is, the larger the error is. The percentage obtained is different for different thresholds, and the general threshold is set to 20 pixels.

Since the resolution of the video applied by the target tracking algorithm is unstable, the robustness of the tracking algorithm has a large impact on the tracking effect, and the robustness of the algorithm can be evaluated in time and space. Temporal robustness is evaluated by taking different video frames as the starting point for tracking. The initial bounding box is the manually marked position of the corresponding frame. Finally, these results are averaged to

obtain a TRE score. The spatial robustness evaluation focuses on different bounding boxes, which are affected by the fact that some algorithms are more sensitive to a given bounding box at initialization since the real positions of the targets currently used for testing are manually labelled. To evaluate the sensitivity of the tracking algorithm to the bounding box at initialization, the real position is panned and a new bounding box is generated by scaling up and down, the size of the pan is generally 10% of the target size, and the scale is varied from 80% to 120% of the real target scale, increasing by 10% each time, and finally, these results are averaged to obtain the SRE score. The comparison of evaluation metrics is shown in Table 1.

In the tracking framework, the effectiveness of features is the primary consideration in building the target model, and the good or bad extraction of features directly affects the final

TABLE 1: Comparison of evaluation indicators.

Evaluation index	Index advantages	Indicator shortcomings
APE	Can reflect the change of target location	Inability to reflect changes in the size and scale of the target object
IOU	Can reflect changes in the size and scale of the target object	Cannot reflect changes in the target location
TRE	Can reflect the time robustness of the algorithm	Disrupt time
SRE	Can reflect the robustness of the algorithm space	Disrupt the space

tracking effect. The basic requirements for tracking features are, firstly, to be able to effectively describe the information of the target, and secondly, since different features perform differently in different scenes, the tracking features need to be adaptable to the changes in the scene with the ability to describe the target. Finally, considering the real-time requirements of tracking, the computational process of feature extraction should ensure a low complexity. The credibility strategy is used to compare the reliability of the differentiated features in the initial fusion stage, screen out the features with low credibility, calculate the weights of the remaining features based on the credibility and historical credibility information, and adaptively fuse different features based on the weights to form the final tracking features of the target. In this paper, after extensive experimental tests, three weight assignment methods with better performance are selected for fusion. The weight parameters are assigned in the way shown in Table 2.

4. Analysis of Results

4.1. Target Tracking Evaluation Analysis. Figure 3(a) is the horizontal coordinate overlap threshold that indicates the threshold value of the detected target and the real target: intersection area and the vertical coordinate Overlap Precision indicate the accuracy of the intersection area of the detected target and the real target, which reflects the tracking algorithm's spatial robustness of the tracking algorithm. The robustness of the algorithm can be compared by the size of the area enclosed by the curve and the x -axis. As can be seen from Figure 3(a), when IOU is less than 0.75, the method KCF_YOLO has a greater advantage than other tracking algorithms, and when IOU is greater than 0.75, the advantage is not obvious, which means that the method KCF_YOLO has certain robustness in space. Figure 3(b) is the horizontal coordinate Overlap Threshold that indicates the threshold value of the intersection area of the detected target and the real target, and the vertical coordinate Overlap Precision indicates the accuracy of the intersection area of the detected target and the real target, which can reflect the robustness of the tracking algorithm in time. As can be seen from Figure 3(b), when IOU is less than 0.75, the area enclosed by this paper's algorithm and c -axis is the largest, which indicates that this paper's algorithm has a greater advantage than other tracking algorithms and verifies that this paper's method KCF_YOLO has certain robustness in time.

The method in this paper is an improved tracking method based on KCF to achieve the effect of adaptive scale

TABLE 2: Integration weight assignment strategy.

Numbering	HOG weight	Color weight	Total weight
1	0.25	0.75	1.0
2	0.3	0.7	1.0
3	0.35	0.65	1.0
4	0.4	0.6	1.0
5	0.45	0.55	1.0
6	0.5	0.5	1.0

mutation. To illustrate the effectiveness of this method, the tracking method with scale adaptive capability is selected for comparison in the comparison experiment of this paper, where Precision is the error between the center point of the tracked target frame and the manually labelled target frame center point. The accuracy comparison of the six tracking algorithms is shown in Figure 4. As can be seen from Figure 4, KCF only updates the location of the target (x, y) , and the size of the target box remains the same, so the ability to cope with sudden changes in the target scale is poor; AMF improves on KCF by adding color features to the target feature extraction, i.e., HOG features and CN features. The combination of HOG features and CN features, the addition of 7 scales to the scale pool, and the cyclic selection of the DSST uses the scale and position filters to be independent of each other and can perform scale evaluation and target localization at once. FDSST accelerates DSST and performs PCA downscaling and QR decomposition on the position filter and scale filter, respectively, to reduce the computation and improve the computation speed; TLD adopts the strategy of learning while detecting for tracking; KCF_YOLO, the method of this paper, chooses the method of detecting first and then tracking and utilizes the detection method of neural network in template updating and scale adaption.

4.2. Performance Analysis of Target Tracking Algorithm. Tracking speed is an important indicator to evaluate whether the tracker can be widely used, and to some extent, it is the only criterion to evaluate the performance of the tracker. In this section, in addition to the tracking accuracy and success rate, the tracking speed is also compared, and the results are shown in Figure 5. Since the tracking framework is based on the idea of feature fusion, the tracking speed of this algorithm is not much different from that of the Staple algorithm, but the tracking robustness of this algorithm is better, and it can better adapt to the tracking of targets in various situations. The Staple algorithm is less stable in the tracking

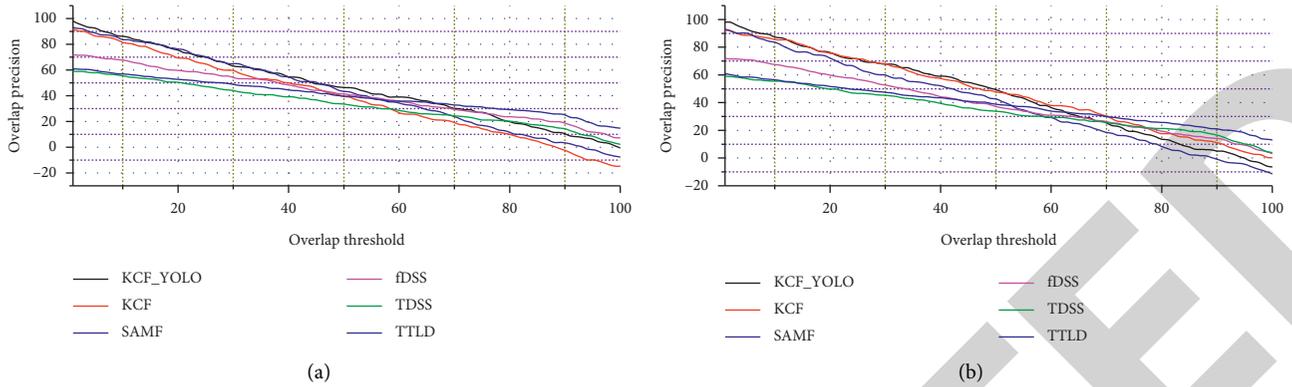


FIGURE 3: Robustness comparison of six trackers. (a) SRE, (b) TRE.

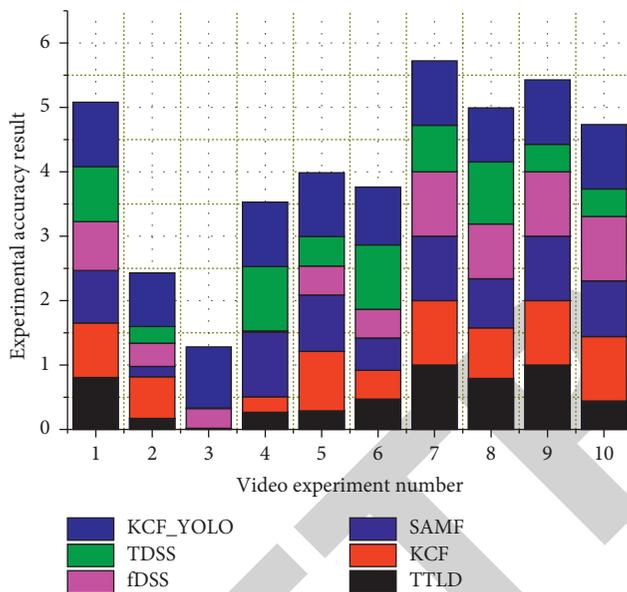


FIGURE 4: Comparison of the accuracy of six tracking algorithms.

process, especially in the case of motion blur. The tracking performance is poor. For the SAMF algorithm, due to the constraints of tracking framework design, the tracking speed does not exceed 16 frames/s, and the tracking performance is poor in real-time. Compared with the classical algorithm, the algorithm in this paper has better robustness and real-time performance in the tracking process by taking into account the tracking accuracy and tracking speed.

The tracking accuracy curves and tracking success rate curves of various algorithms on four sets of video sequences are plotted by the macro representation of the comparative effects of the algorithms in this paper and the classical algorithms, as shown in Figure 6. It can be seen from this Figure 6 that the tracking performance of this paper's algorithm is better, and the tracking accuracy and tracking success rate are better than the classical comparison algorithm compared with other algorithms. Compared with the SAMF algorithm, the average tracking accuracy of this algorithm is about 3% higher than that of this algorithm, but in the average tracking success rate, it is about 15.1% higher,

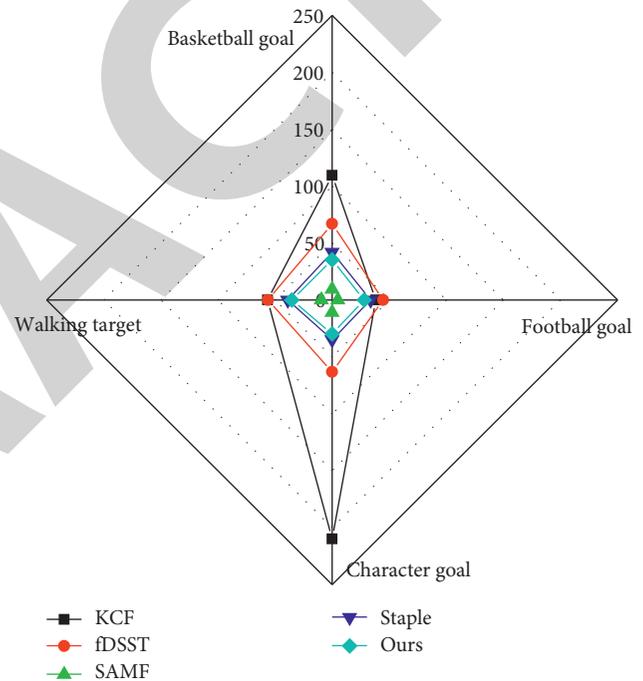


FIGURE 5: Algorithm tracking speed.

and the change of curve plunge is smoother, and it can always maintain a high tracking accuracy and success rate, which has a greater advantage. In terms of the average tracking success rate, the Staple algorithm, which also adopts the feature fusion idea, has the smallest difference with the average tracking success rate of this paper's algorithm, which illustrates the correctness and effectiveness of the feature fusion idea of this paper's algorithm, but in terms of the average tracking accuracy rate, there is a large gap between Staple algorithm and this paper's algorithm. For KCF and fDSS algorithms, this paper's algorithm has a large improvement in the average tracking accuracy and success rate. The reason for this improvement is that on the one hand, these two algorithms use a single feature in the process of tracking the target, while the algorithm in this paper uses the idea of feature fusion, which has a stronger feature characterization ability; on the other hand, the algorithm in this

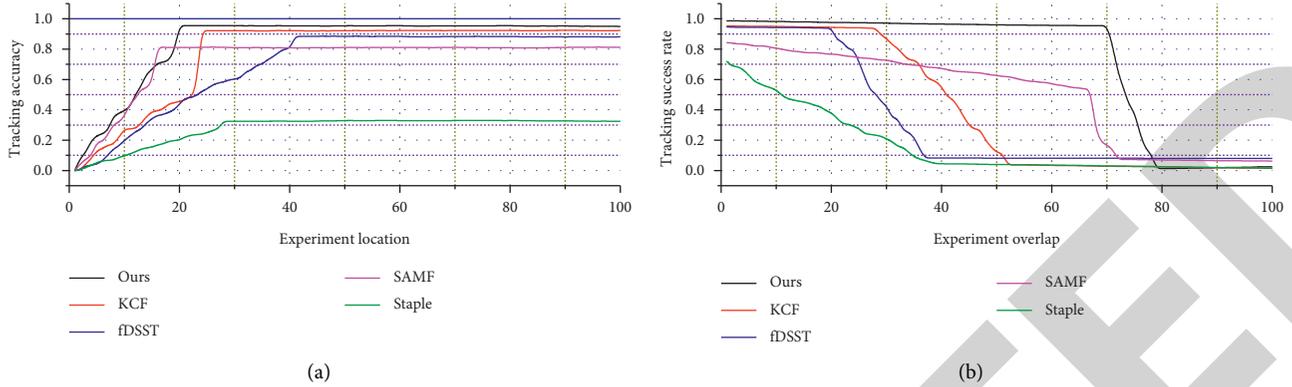


FIGURE 6: Accuracy and success curves of the algorithm on the experimental sequence. (a) Tracking accuracy curve. (b) Tracking success rate curve.

paper judges the reliability of the tracking results and obtains more reliable tracking results, while the KCF algorithm and the f DSST algorithm lack such a design in the tracking framework, so the tracking disadvantage is greater in comparison.

4.3. Target Tracking Effect Analysis. To verify the effectiveness of the detection system, the traditional convolutional regression network algorithm is first compared with the detection algorithm designed in this paper. As shown in Figure 7, the algorithm incorporating the detection mechanism achieves an 88.15% success rate of 66.81 in one pass evaluation, which is 2.15 percentage points higher in accuracy and 6.1 percentage points higher in success rate than the algorithm CRN without a detection mechanism. This indicates the effectiveness of the redetection mechanism.

As shown in Figure 8, the multilayer convolutional network fusion algorithm (MCT) achieves 78.21% accuracy on video sequences with fast motion attributes, which is 3.94 percentage points better than algorithm SiamFC, 12.41 percentage points better than algorithm CACF, 19.44 percentage points better than algorithm KCF, 23.51 percentage points better than algorithm CSK, 53.41 percentage points better than algorithm CT, 23.51 percentage points compared to algorithm CSK, and 53.41 percentage points compared to algorithm CT. Algorithm MCT achieves 77.1% accuracy on video sequences with motion blur attributes, which is 4.51 percentage points better than algorithm SiamFC, 6.83 percentage points better than algorithm CACF, 16.57 percentage points better than algorithm KCF, 42.94 percentage points better than algorithm CSK, and 49.82 percentage points better than algorithm CT. The improvement is 49.8 percentage points compared to algorithm CT. Algorithm MCT achieves 85.92% accuracy on video sequences with in-field rotation property, which is 9.9 percentage points better than SiamFC, 7.41 percentage points better than CACF, 16.1 percentage points better than KCF, 31.24 percentage points better than CSK, and 48.3 percentage points better than CT, and algorithm CT by 48.41 percentage points. Algorithm MCT achieves 89.81% accuracy on video sequences with low resolution attributes, which is 16.62 percentage points better

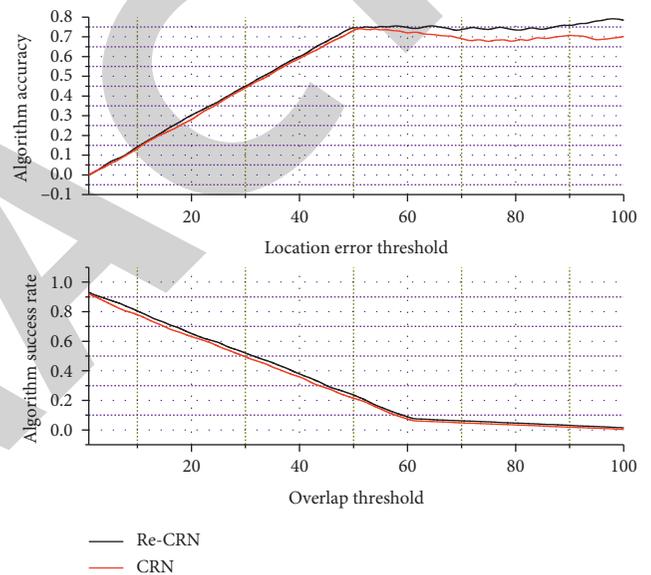


FIGURE 7: Comparison of algorithm accuracy and success rate.

than SiamFC, 49.6 percentage points better than CACF, 46.3 percentage points better than KCF, 45.62 percentage points better than CSK, and 74.34 percentage points better than CT.

The above comparison tracking results show that the depth feature-based target tracking algorithm is more accurate than the traditional correlation filtering algorithm based on artificial features in complex environments, especially in cases such as fast motion and low resolution. The algorithm in this paper achieves superior results than the traditional algorithm. In this paper, we improve the update mechanism of the kernel correlation filter tracking algorithm to avoid the model pollution that can be caused by occlusion and other situations in real applications, set a judgment threshold in advance, use this threshold to determine the trustworthiness of the tracking results in advance, and then decide the updated law of the model, and use such update mechanism in the tracking algorithm with adaptive scale change of multifeature fusion proposed in the last two chapters, to improve the tracking. The accuracy of the original algorithm is improved from 0.781 to 0.792 after the

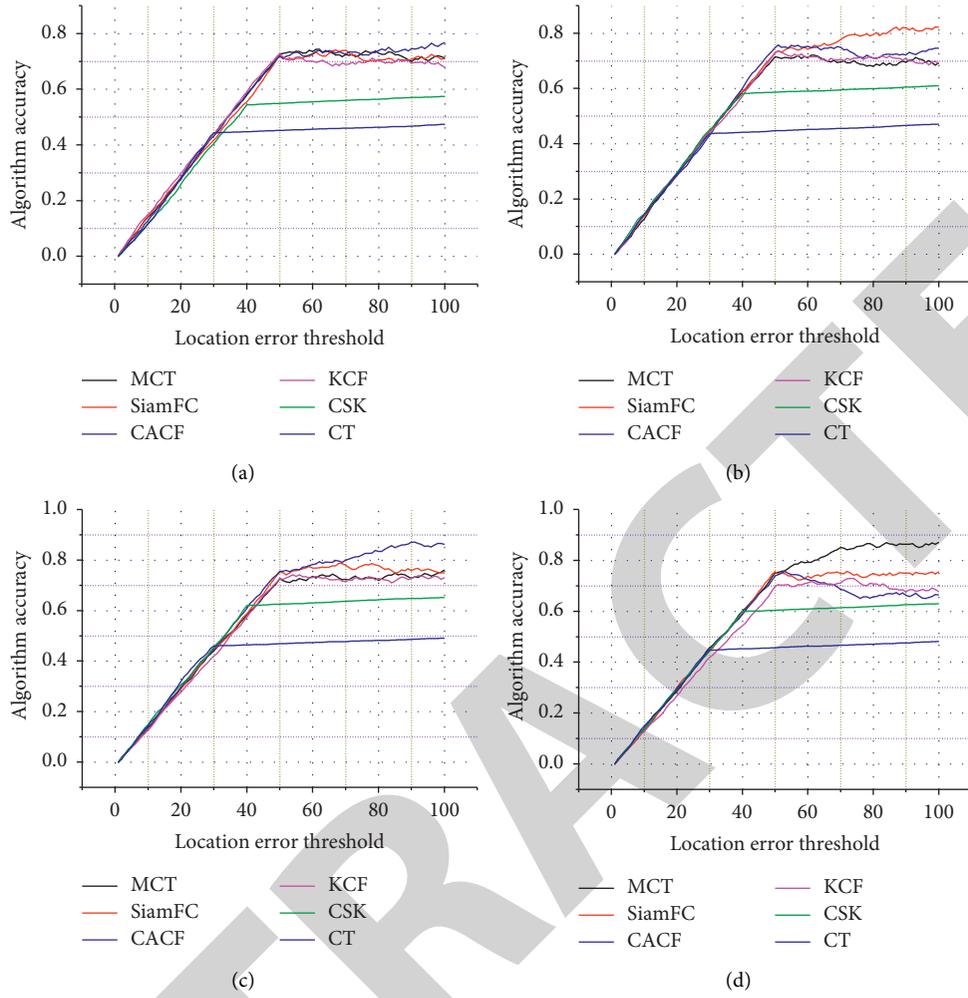


FIGURE 8: Accuracy plots of different algorithms on video sequences with different attributes. (a) Fast motion detection accuracy. (b) Motion blur detection accuracy. (c) Rotation detection accuracy in the field of view. (d) Low resolution detection accuracy.

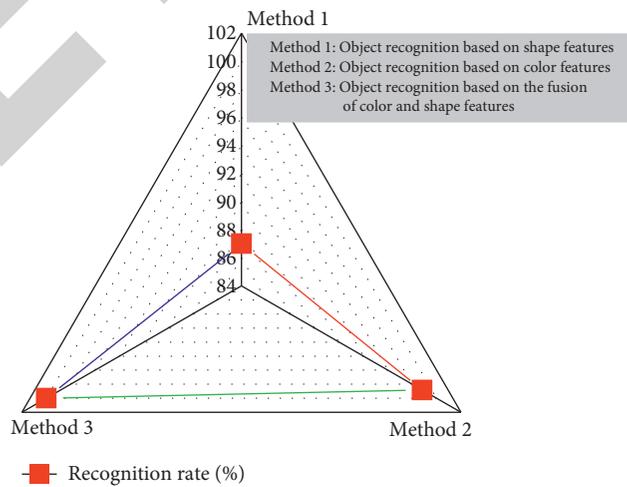


FIGURE 9: Experimental results of three methods based on color histogram feature processing.

optimization of the update mechanism, which shows that the algorithm improvement idea adopted in this paper has certain practicality and can improve the robustness of tracking well. Moreover, the experimental comparison with other classical algorithms also shows that the improved algorithm in this paper has better tracking accuracy under two types of challenges, namely, scale change and occlusion.

The validation is performed using the COIL-100 image dataset, which contains 100 object images, each containing 100 images whose 100 images were taken by rotating the object by 5° interval for one week. To illustrate that the method of fusing color and shape features for object recognition is superior to the method using color feature recognition alone or shape feature recognition alone, images taken at position 0 for each object are selected as samples. The test images at 15°, 20°, 55°, 305°, 370°, and 345° are used for object recognition based on shape features, object recognition based on color features, and object recognition based on color and shape features, respectively. Object recognition based on shape features, object recognition based on color features, and object recognition based on the fusion of color and shape features. The recognition rate is defined as follows: each test image is feature matched with the sample, sorted in the order of similarity from the largest to the smallest, and the ratio of the first similar bit to the number of test images and the number of samples is calculated as the recognition rate, and finally, the recognition rates of the six angle test images are summed up, and the average value is taken as the final recognition result. Figure 9 shows the experimental results of the three methods based on color histogram feature processing. The experimental results show that the object recognition method based on the fusion of color and shape features proposed in this paper has a good object recognition effect.

5. Conclusion

In this paper, we propose to use the KCF algorithm as the base algorithm for target tracking and add the redetection mechanism to update the target template and achieve the scale adaptive effect. Through the experimental analysis in Chapter 4, the idea of adding detection and then tracking in KCF can achieve real-time tracking although the tracking speed is reduced, and the tracking speed in this paper reaches 36 frames/second; the tracking progress is improved from the average tracking accuracy of KCJF of 0.564 to 0.956, which is improved by 39.35% percentage points. Through experiments, this paper shows that the video sequences with scale mutation on the OTB100 and VOT2016 datasets are compared with five mainstream algorithms, KCF, SAMF, FDSST, DSST, and TLD, and the comprehensive results are better than other algorithms, with an average improvement of 31.75% over the advanced trackers; tracking experiments are conducted after a long period of complete occlusion and a sudden change in scale and the effectiveness of the model proposed in this paper is verified based on the analysis of tracking results. In this paper, the correlation filter is designed as a convolutional regression network starting from a deep neural network, while a conventional

convolutional neural network is added for feature extraction to make it a whole, and end-to-end training is performed on a large-scale data set to design features that are more suitable for the target tracking task, and the resulting algorithm is evaluated better than the traditional artificial feature-based algorithm in terms of accuracy, temporal robustness, and spatial robustness. Considering that the process of feature fusion is to some extent at the expense of tracking speed, there is still room for further improvement in tracking speed, and the next research work is mainly focused on improving the scale adaptation capability and tracking speed, and on finding a method that can balance tracking accuracy and tracking speed.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that no conflicts of interest.

Acknowledgments

This work was supported by the Education Department of Liaoning Province: Research on Target Tracking Algorithm Based on Siamese Network (No. LG201915) and Shenyang Ligong University: Design and Implementation of Multi-Target Tracking Algorithm Based on Deep Learning.

References

- [1] J. Wen, J. Yang, B. Jiang et al., "Big data driven marine environment information forecasting: a time series prediction network," *IEEE Transactions on Fuzzy Systems*, vol. 29, 2020.
- [2] H. Yu, F. He, and Y. Pan, "A novel segmentation model for medical images with intensity inhomogeneity based on adaptive perturbation," *Multimedia Tools and Applications*, vol. 78, no. 9, pp. 11779–11798, 2019.
- [3] J. Yang, Y. Zhao, J. Liu et al., "No reference quality assessment for screen content images using stacked autoencoders in pictorial and textual regions," *IEEE Transactions on Cybernetics*, 2020.
- [4] A. Knutas, R. Van Roy, T. Hynninen, M. Granato, J. Kasurinen, and J. Ikonen, "A process for designing algorithm-based personalized gamification," *Multimedia Tools and Applications*, vol. 78, no. 10, pp. 13593–13612, 2019.
- [5] J. Wu, R. Nan, and L. Chen, "Improved salp swarm algorithm based on weight factor and adaptive mutation," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 31, no. 3, pp. 493–515, 2019.
- [6] T. Xu, Z.-H. Feng, X.-J. Wu, and J. Kittler, "Learning adaptive discriminative correlation filters via temporal consistency preserving spatial feature selection for robust visual object tracking," *IEEE Transactions on Image Processing*, vol. 28, no. 11, pp. 5596–5609, 2019.
- [7] H. Ullah, T. Saba, N. Islam et al., "An ensemble classification of exudates in color fundus images using an evolutionary algorithm based optimal features selection," *Microscopy Research and Technique*, vol. 82, no. 4, pp. 361–372, 2019.

- [8] S. H. Abdulhussain, A. R. Ramli, B. M. Mahmmud, M. I. Saripan, S. A. R. Al-Haddad, and W. A. Jassim, "Shot boundary detection based on orthogonal polynomial," *Multimedia Tools and Applications*, vol. 78, no. 14, pp. 20361–20382, 2019.
- [9] C. Elisei-Iliescu, C. Paleologu, J. Benesty, C. Stanciu, C. Anghel, and S. Ciochina, "Recursive least-squares algorithms for the identification of low-rank systems," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 5, pp. 903–918, 2019.
- [10] A. Khan, R. Shah, M. Imran, A. Khan, J. I. Bangash, and K. Shah, "An alternative approach to neural network training based on hybrid bio meta-heuristic algorithm," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 10, pp. 3821–3830, 2019.
- [11] T. A. Jumani, M. W. Mustafa, Z. Hussain et al., "Jaya optimization algorithm for transient response and stability enhancement of a fractional-order PID based automatic voltage regulator system," *Alexandria Engineering Journal*, vol. 59, no. 4, pp. 2429–2440, 2020.
- [12] D. Yuan, X. Zhang, J. Liu, and D. Li, "A multiple feature fused model for visual object tracking via correlation filters," *Multimedia Tools and Applications*, vol. 78, no. 19, pp. 27271–27290, 2019.
- [13] Y. Wang, W. Zhou, J. Luo, H. Yan, H. Pu, and Y. Peng, "Reliable intelligent path following control for a robotic airship against sensor faults," *IEEE/ASME Transactions on Mechatronics*, vol. 24, no. 6, pp. 2572–2582, 2019.
- [14] J. Luo and B. Shi, "A hybrid whale optimization algorithm based on modified differential evolution for global optimization problems," *Applied Intelligence*, vol. 49, no. 5, pp. 1982–2000, 2019.
- [15] B. Hekimoğlu, "Sine-cosine algorithm-based optimization for automatic voltage regulator system," *Transactions of the Institute of Measurement and Control*, vol. 41, no. 6, pp. 1761–1771, 2019.
- [16] S. Larabi Marie-Sainte and N. Alalyani, "Firefly algorithm based feature selection for Arabic text classification," *Journal of King Saud University-Computer and Information Sciences*, vol. 32, no. 3, pp. 320–328, 2020.
- [17] C. Kwan, B. Chou, J. Yang et al., "Target tracking and classification using compressive measurements of MWIR and LWIR coded aperture cameras," *Journal of Signal and Information Processing*, vol. 10, no. 3, pp. 73–95, 2019.
- [18] N. A. Gawande, J. A. Daily, C. Siegel, N. R. Tallent, and A. Vishnu, "Scaling deep learning workloads: NVIDIA DGX-1/pascal and intel knights landing," *Future Generation Computer Systems*, vol. 108, no. Jul, pp. 1162–1172, 2020.
- [19] U. Leicht-Deobald, T. Busch, C. Schank et al., "The challenges of algorithm-based HR decision-making for personal integrity," *Journal of Business Ethics*, vol. 160, no. 2, pp. 377–392, 2019.
- [20] X. Chen, C. Wu, T. Chen et al., "Age of information aware radio resource management in vehicular networks: a proactive deep reinforcement learning perspective," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2268–2281, 2020.
- [21] M. Zhang, X. Liu, D. Xu, Z. Cao, and J. Yu, "Vision-based target-following guider for mobile robot," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 12, pp. 9360–9371, 2019.
- [22] J. Wang, N. Cui, and C. Wei, "Optimal rocket landing guidance using convex optimization and model predictive control," *Journal of Guidance, Control, and Dynamics*, vol. 42, no. 5, pp. 1078–1092, 2019.
- [23] J. Qiu, K. Sun, T. Wang, and H. Gao, "Observer-based fuzzy adaptive event-triggered control for pure-feedback nonlinear systems with prescribed performance," *IEEE Transactions on Fuzzy Systems*, vol. 27, no. 11, pp. 2152–2162, 2019.
- [24] G. Sun, C. Li, and L. Deng, "An adaptive regeneration framework based on search space adjustment for differential evolution," *Neural Computing and Applications*, pp. 1–17, 2021.
- [25] L. Ding, S. Li, H. Gao et al., "Adaptive partial reinforcement learning neural network-based tracking control for wheeled mobile robotic systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 7, pp. 2512–2523, 2018.
- [26] J. Yang, S. Li, Z. Wang, H. Dong, J. Wang, and S. Tang, "Using deep learning to detect defects in manufacturing: a comprehensive survey and current challenges," *Materials*, vol. 13, no. 24, p. 5755, 2020.
- [27] S. Xia, D. Peng, D. Meng et al., "A fast adaptive k-means with no bounds," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.
- [28] G. Zhang, J. Yang, W. Wang et al., "Adaptive visual target tracking algorithm based on classified-patch kernel particle filter," *EURASIP Journal on Image and Video Processing*, vol. 2019, no. 1, pp. 1–12, 2019.
- [29] J. Yang, C. Wang, B. Jiang et al., "Visual perception enabled industry intelligence: state of the art, challenges and prospects," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2204–2219, 2020.
- [30] W. Wei, B. Zhou, D. Połap, and M. Woźniak, "A regional adaptive variational PDE model for computed tomography image reconstruction," *Pattern Recognition*, vol. 92, pp. 64–81, 2019.