Target Detection in Martial Arts Competition Video using Kalman Filter Algorithm Based on Multi target Tracking

Zhiguo Xin

Department of PE, Wuxi Vocational Institute of Arts and Technology, Wuxi, 214200, China

Abstract—To solve the low accuracy and poor stability in traditional object tracking methods for martial arts competition videos, a Kalman filtering algorithm based on feature matching and multi object tracking is proposed for object detection in martial arts competition videos. Firstly, feature matching in multi target tracking is studied. Then, based on target feature matching, the Kalman filtering algorithm is fused to construct a target detection model in martial arts videos. Finally, simulation experiments are conducted to verify the performance and application effectiveness of the model. The results showed that the average tracking errors of the model on the X and Y axes were 3.86% and 3.38%, respectively. At the same time, the average accuracy and recall rate in the video target tracking process were 93.64% and 95.48%, respectively. After 100 iterations, the results gradually stabilized. This indicated that the constructed model could accurately detect targets in martial arts competition videos. It had high tracking accuracy and robustness. Compared with traditional object detection methods, this algorithm has better performance and effectiveness. The Kalman filter algorithm based on feature matching and multi target tracking has broad application prospects and research value in target detection in martial arts competition videos.

Keywords—*Multi target tracking; Kalman filtering algorithm; martial arts competition videos; target detection; feature matching*

I. INTRODUCTION

Object detection is an important research direction in the computer vision, which has broad application value. In martial arts competition videos, object detection can help referees judge the scores of players in real-time, improving the fairness and accuracy of the competition. However, due to the rapid and complex movements in martial arts competitions, traditional object detection methods face a series of challenges in this scenario, such as complex backgrounds and object occlusion [1-3]. In martial arts competitions, multi-target tracking and target detection are key aspects in analyzing the competition process, evaluating athletes' performance, and performing technical and tactical analysis. However, this task is very challenging due to the large number of targets, fast movement speed and frequent occlusion in the scene. Through a large number of studies, it has been found that the utilization of multi-target detection techniques can significantly improve the precision and accuracy of target tracking in martial arts competitions. Therefore, in order to solve the above problems, the study proposes a Kalman filter algorithm based on feature matching and multi-target tracking, which is used for target detection in martial arts competition videos. To address the

above issues, a Kalman filtering algorithm based on feature matching and multi target tracking is proposed for object detection in martial arts competition videos. This algorithm combines two technologies, feature matching and multi target tracking. The target position is accurately located through feature matching. The Kalman filtering algorithm is used to track targets to improve the accuracy and robustness of target detection [4-6]. Firstly, the study focuses on feature matching in multi target tracking. Then, based on target feature matching, the Kalman filtering algorithm is fused to construct a target detection model in martial arts videos. Finally, simulation experiments are conducted to verify the performance and application effectiveness of the model. This model locates the target position through feature matching. The Kalman filtering algorithm is used to track targets to improve the accuracy and robustness of target detection. The research expects to utilize the multi-target tracking Kalman filter algorithm to effectively solve the problems of low target detection accuracy and weak reliability in martial arts competition videos. The contribution of the research is reflected in the utilization of deep learning techniques for feature extraction, which effectively captures the nuances and dynamic changes of targets and improves the accuracy of target detection. Meanwhile, combining the Kalman filter algorithm to predict and correct the target trajectory effectively handles the tracking difficulties caused by target occlusion and fast movement, and enhances the robustness of tracking. By fusing feature extraction, multi-target tracking and Kalman filtering algorithms to construct the model, it can not only focus on the detection and tracking of single targets, but also analyze the interaction and collaborative behaviors of multiple targets, which provides a new perspective for the technical and tactical analysis of martial arts competitions. Compared with the existing techniques, the difference of the research is the organic combination of feature matching and Kalman filtering algorithm. While traditional methods tend to focus only on feature extraction or filter tracking, the research complements the advantages of both to improve the accuracy and robustness of target detection. At the same time, the study also fully considered the characteristics and practical needs of martial arts competitions, making the proposed method more practical and relevant.

Section II is about the related works. In Section III, based on the feature matching algorithm, the Kalman filtering algorithm is fused with it to construct a multi-objective tracking model. Section IV verifies the performance of the

constructed model for comment classification through simulation experiments and practical applications. Finally discussion and conclusion is given in Section V and VI respectively.

II. RELATED WORKS

Object detection in videos is an important research direction in the computer vision. The main goal is to automatically recognize and track specific target objects in video sequences. This technology has wide applications in many fields, such as racing videos, autonomous driving, security monitoring, medical image analysis, etc. Llano C R et al. State space tracking method based on particle filters for video object tracking. Through experiments, this method had strong performance in tracking objects/people in videos, including foreground/background separation for object movement detection [7]. Lu S and other scholars have investigated the accuracy of real-time video target detection algorithm based on YOLO network for network video image detection. The target information is obtained through image preprocessing and background elimination. Then the convolution operations are applied to reduce the parameters and shorten target detection time. The results showed that this algorithm could significantly shorten the real-time object detection time in videos [8]. Fang Y et al. Multi-intelligent body perception and trajectory prediction method based on spatio-temporal semantics and interaction graph aggregation for effective prediction of spatio-temporal allegories of images and interaction graph aggregation for scene perception and trajectory. The iterative aggregation network was used as background information. Then the trident encoder was decoded and finally detected using prediction methods. The results indicated that this method achieved significant improvements in scene perception and trajectory prediction [9]. Qiu, Ji et al. effectively classified and investigated the performance of small-scale pedestrian detection based on scale prediction method. This method could eliminate the anchor boxes set by most existing detectors. The pixel coordinates of pedestrians at a given center position were predicted. The comprehensive experiments on two real datasets demonstrated that the proposed method achieved excellent performance through [10]. Lyu Y et al. analyzed and studied to improve the detection performance of image detectors in classes without video labels based on agnostic convolutional regression tracker. The performance of the image detector was enhanced through this tracker. This tracker mainly utilized the features of reused image object detectors to learn and track objects. The results indicated that the image detector trained with this tracker could improve accuracy by 5% [11].

Chen Z et al. based on online multi-target tracking algorithm with Kalman filtering and multi-information fusion, conducted a study on leakage detection, false detection and target occlusion during online multi-target tracking. This method utilized Kalman filtering for modeling, and then combined target information with features. The results showed that this method could effectively solve the tracking drift problem caused by target interleaving and occlusion. The main tracking performance parameters were significantly improved [12]. Chen H et al. analyzed and studied the improvement of image target tracking capability based on distributed diffusion traceless Kalman algorithm with covariance intersection strategy. This method could diffuse policy information. Then the adjacent information was fused using a diffusion framework. The results showed that this method could significantly improve the tracking ability of image targets, while also reducing the impact of noise [13]. Liu S et al. based on the improvement algorithm of occlusion prediction tracking based on Kalman filter and spatio-temporal map, the target occlusion, drift and interleaving problems in the target tracking process have been effectively handled and studied. This method could distinguish different images using color histograms and color spatial distribution. The results indicated that the average tracking accuracy of this method was 34.1%. The proposed algorithm improved the performance of multi target tracking process [14].

In summary, the Kalman filtering algorithm is of great significance in the multi target tracking in video images. Based on the Kalman filtering algorithm in image target tracking and detection, the Kalman filtering algorithm can achieve real-time tracking and state estimation of targets in video image target tracking and detection, improving the accuracy and efficiency of target tracking. The research aims to provide a new method for multi object tracking and detection in martial arts competition videos.

III. DESIGN OF A MARTIAL ARTS COMPETITION VIDEO OBJECT DETECTION MODEL BASED ON FEATURE MATCHING AND KALMAN FILTERING ALGORITHM

Martial arts competition video object detection can provide real-time and accurate object tracking. Through feature matching algorithms, feature extraction and matching can be performed on the characters in the video, thereby accurately tracking the contestants in the competition.

A. Multiple Target Tracking Algorithm Based on Feature Matching

Athletes' movements in martial arts competitions are fast and complex. Traditional object detection algorithms may not be able to accurately track athletes. The multi target tracking algorithm based on feature matching can effectively analyze and track multiple targets to improve the analysis results of competition videos. In martial arts competitions, players have extremely fast movements. When tracking targets, feature comparison and target matching are performed between the current image and the previous frame image to obtain the correlation of target motion. In the tracking of multi-objective videos, feature extraction and matching of moving targets are required to complete the target tracking. The feature matching of moving targets directly affects the effectiveness of target tracking. Therefore, during feature selection, feature matching is performed on the tracked target to achieve target tracking [15-16]. The selected target matching indicators are the position and height to width ratio of the participants in the rectangular box, as well as the color value of the image. The factors that affect target feature extraction include target area, color, position, and the ratio of height to width. In the detecting the participants, the previous and second frames of images need to be collected. They are first grayed out, and

then subtracted before and after. The difference is binarized before edge detection. During the detection process, pixels are used as the corresponding pixels of the moving target, which represents the position occupied by the target. In the video object tracking, the color of the moving object itself can also serve as a feature matching element. The average color value of the moving target itself is used as a feature for matching. The process of matching and tracking moving targets is shown in Fig. 1.

Based on Fig. 1, to utilize the target feature matching indicators mentioned above, a feature vector is defined to match the target features. This vector can be defined using Eq. (1).

$$a_{n,i} = (s_{n,i}, r_{n,i}, g_{n,i}, b_{n,i}, x_{n,i}, y_{n,i}, rate_{n,i}) \tag{1}$$

In Eq. (1), $a_{n,i}$ represents the feature vector, which is defined as the i-th feature vector in the n-th frame image. $s_{n,i}$ represents the area occupied by the moving target in the selected image. $r_{n,i}$ represents the average value of red pixels. $g_{n,i}$ represents the average value of green pixels. $b_{n,i}$ represents the average value of blue pixels. $x_{n,i}$ represents the abscissa in the matrix. $y_{n,i}$ represents the ordinate in the matrix. $rate_{n,i}$ represents the ratio of matrix height to width. The variation of the moving target between two frames of images is very small, which makes the image have obvious continuity. It is used as a feature flux to define the similarity function of a target image. Then it is used for feature matching work. The similarity function can be represented by Eq. (2).

$$\Delta s_{i,j} = \frac{\left| s_{n,i} - s_{n-1,j} \right|}{s_{n,j}} \tag{2}$$

In Eq. (2), $\Delta s_{i,j}$ represents the similarity function. $s_{n-1,j}$ represents the area of the j-th target in the $n-1$-th frame image. $s_{n,j}$ represents the j-th feature vector in the frame image. To determine the color mean of the target, the similarity function between the three colors in the previous image and the current image is defined. The similarity function of the three colors can be represented by Eq. (3).

$$\begin{cases} \Delta r_{i,j} = \dfrac{\left| r_{n,i} - r_{n-1,j} \right|}{r_{n,i}} \\[3mm] \Delta g_{i,j} = \dfrac{\left| g_{n,i} - g_{n-1,j} \right|}{r_{n,i}} \\[3mm] \Delta b_{i,j} = \dfrac{\left| b_{n,i} - b_{n-1,j} \right|}{b_{n,i}} \end{cases} \tag{3}$$

In Eq. (3), $r_{n-1,j}$ represents the area of the j-th target in frame $n-1$ of the red image. $g_{n-1,j}$ represents the area of the j-th target in frame $n-1$ of the green image. $b_{n-1,j}$ represents the area of the j-th target in frame $n-1$ of the blue image. For the center of the moving target matrix, the similarity function between the current i-th target and the j-th target in the previous frame image is shown in Eq. (4).

$$\begin{cases} \Delta x_{i,j} = \dfrac{\left| x_{n,i} - x_{n-1,j} \right|}{x_{n,i}} \\[3mm] \Delta y_{i,j} = \dfrac{\left| y_{n,i} - y_{n-1,j} \right|}{y_{n,i}} \end{cases} \tag{4}$$

In Eq. (4), $x_{n-1,j}$ represents the center position of the j-th rectangular box in the $n-1$-th frame of the image in the x-axis direction. $y_{n-1,j}$ represents the center position of the j-th rectangular box in the $n-1$-th frame of the image in the y-axis direction. For the height to width ratio feature of the bounding rectangle of the moving object in the video, the similarity function is represented by Eq. (5).

$$\Delta rate_{i,j} = \frac{\left| rate_{n,j} - rate_{n-1,j} \right|}{rate_{n,i}} \tag{5}$$
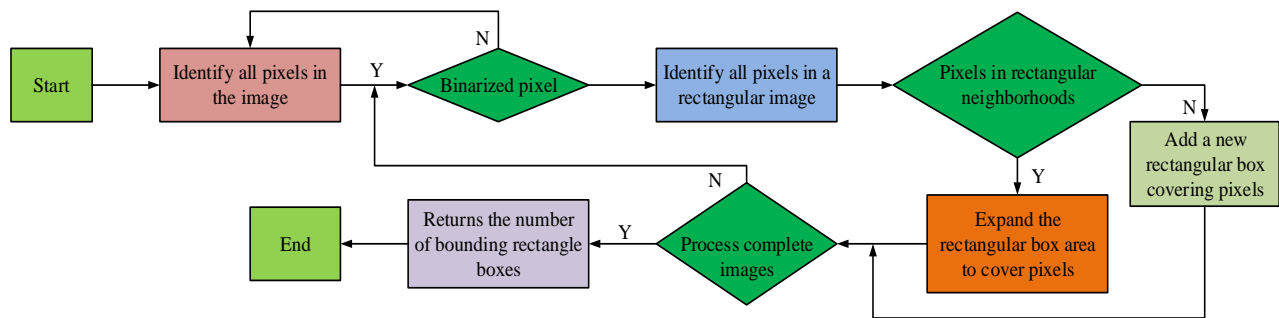

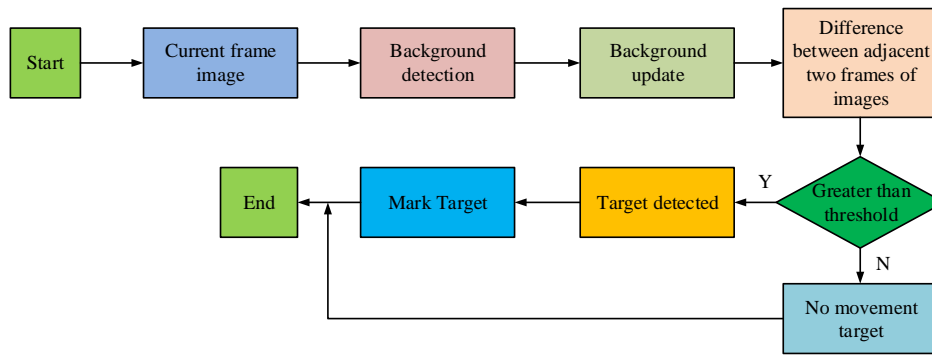Fig. 1. Flowchart of moving target matching and tracking.

Fig. 2. Flow chart of multi target tracking algorithm based on feature matching.

In Eq. (5), $rate_{n,j}$ represents the ratio of the height to width of the j-th target in the n-th frame image. $rate_{n-1,j}$ represents the ratio of the height to width of the j-th target rectangle in $n-1$-th frame image. $rate_{n,i}$ represents the ratio of the height to width of the i-th target rectangle in the n-th frame image. To fuse the four features used for target matching mentioned above together, a metric function is introduced into feature fusion. The definition process of the degree function can be represented by Eq. (6).

In Eq. (6), a is the target feature weighting coefficient. β is the color weighting coefficient. χ represents the x-axis area weighting coefficient. γ represents the y-axis target feature weighting coefficient.

$$\Delta a_{i,j} = a(\Delta s_{i,j})^2 + \beta(\Delta r_{i,j})^2 + (\Delta g_{i,j})^2 + \chi(\Delta x_{i,j})^2 + (\Delta y_{i,j})^2 + \gamma(\Delta rate_{i,j})^2 \quad (6)$$

Combined with the analysis of Fig. 2, it can be seen that the feature matching method is used for effective target detection of the features of athletes in the video of the game, and the feature matching process is completed by judging whether it meets the requirements of feature matching by the presence and absence of athletes as well as the size of the threshold value.

B. Construction of a Multi Target Tracking and Detection Model Based on Feature Matching and Kalman Filtering Algorithm

The feature matching algorithm can effectively track the target, but the tracking accuracy is significantly affected if the target is occluded. To address the impact of target occlusion, the Kalman filtering algorithm is introduced on the basis of the feature matching algorithm. The two are fused for the construction of a multi target tracking model. Under the principle of minimum mean square error, a Kalman filter is used to iterate the elements, thereby completing the entire tracking state [17-18]. The fused Kalman filtering algorithm can estimate the past and future states of moving targets based on their current states. The flowchart of the Kalman filtering algorithm after fusion feature matching is shown in Fig. 3.
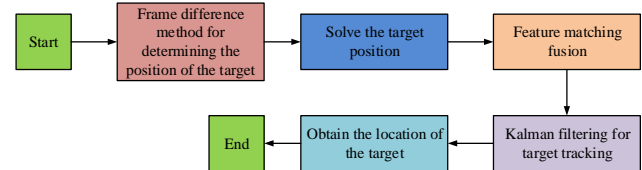

Fig. 3. Flowchart of Kalman filtering algorithm.

The tracking of moving target states using Kalman filtering algorithm is affected by random noise. Therefore, the tracking status is first determined. The tracking status is shown in Eq. (7).

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \quad (7)$$

In Eq. (7), A represents the state transition matrix. x_{k-1} represents the state x noise value of the $k-1$-th target in the state transition matrix. B represents the state control matrix. u_{k-1} represents the noise value of state u for the $k-1$-th target in the state control matrix. w_{k-1} represents the random noise value. After determining the tracking state, the storage capacity of the tracking state can be found through feature extraction methods. The observation formula is shown in Eq. (8).

$$z_k = Hx_k + v_k \quad (8)$$

In Eq. (8), H represents the observation matrix of the state. v_k represents observation noise. The determination of multi-objective states requires a significant amount of time. Therefore, to simplify the process, the covariance of state noise and observation noise is utilized to reflect the tracking effect by estimating the error in step k of the tracking process. It can be defined by Eq. (9).

$$\begin{cases} \bar{x} = A\bar{x}_{k-1} + Bu_{k-1} \\ p_{\bar{k}} = AP_{k-1} + Q \end{cases} \quad (9)$$

In Eq. (9), $\bar{x}_{k-1}$ represents the tracking result of the previous observation. $\bar{x}$ represents the observation results at the corresponding time. P_{k-1} represents the previous prediction result. Q represents the covariance difference of state noise. After obtaining the tracking status of multiple targets, the observation results are used to determine whether there is an error between the tracking status and the actual observed values. Furthermore, the revised state estimation values and noise values are obtained. It is the process of using the Kalman filtering algorithm to filter the noise. The flowchart of this process is shown in Fig. 4.
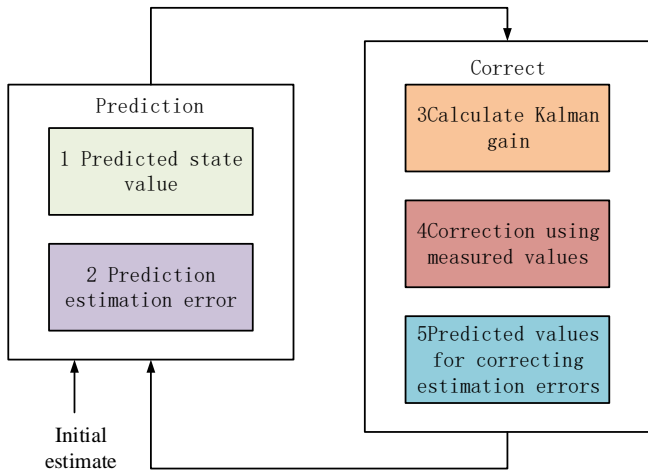

Fig. 4. Operation flowchart of Kalman filtering algorithm for noise filtering.

The target motion state in the competition video image is either high-speed or irregular. The common first-order motion model cannot complete the observation of the entire state. Therefore, a second-order motion model is introduced. The Kalman filtering algorithm is used to predict the targets in the second-order motion model to obtain relevant motion features, which are effectively fused with feature extraction algorithms. Assuming that at a certain moment in tracking, the tracked moving target is in a moving rectangle. The velocity of the tracked moving target in the vertical and horizontal directions is uniform motion. Then the motion state needs to meet the uniform motion, as shown in Eq. (10).

$$\begin{cases} v_x(t) = v_x(t-1) \\ v_y(t) = v_y(t-1) \end{cases} \tag{10}$$

In Eq. (10), $v_x(t)$ represents the uniform motion velocity on the x-axis at time t. $v_x(t-1)$ represents the uniform motion velocity on the x-axis at time $t-1$. $v_y(t)$ represents the uniform motion velocity on the x-axis at time t. $v_y(t-1)$ represents the uniform motion velocity on the y-axis at time $t-1$. The state transition of Kalman filtering can be

determined based on the uniform motion during tracking, as shown in Eq. (11).

$$x_z = Ax_{k-1} + w_{k-1} \tag{11}$$

In Eq. (11), x_z represents the transition state value of the Kalman filter. At this point, the corresponding state transition matrix is shown in Eq. (12).

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{12}$$

The state transition formula can only be used as a directly measured value if it satisfies the matrix. The measurement is shown in Eq. (13).

$$z_c = Hx_k + v_k \tag{13}$$

In Eq. (13), z_c represents the measured value obtained. The corresponding state observation matrix is shown in Eq. (14).

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{14}$$

After analyzing the competition video, there is a significant occlusion of objects in the martial arts competition video. This is mainly caused by the arena, participating athletes, and judges. It is inevitable in the real environment. This occlusion has a significant impact on multi target tracking in competitions. Through the above research, when analyzing the state of moving targets, if occlusion occurs, the image will disappear, and even all images will disappear. If the occluded image merges with the target after a period of time, it can be determined that the image has completely disappeared. If the occluded image appears again in the video rectangle after separating from other targets, it will be used as a new tracking target for tracking and recognition, and matched with a new feature quantity [19-20]. On the basis of image feature matching, the Kalman filtering algorithm is introduced to fuse the two for predicting the motion of moving targets in martial arts competition videos. By utilizing the characteristics of both, the position information of occluded targets is predicted to meet the real-time tracking of targets. The multi target tracking flowchart of this model is shown in Fig. 5.
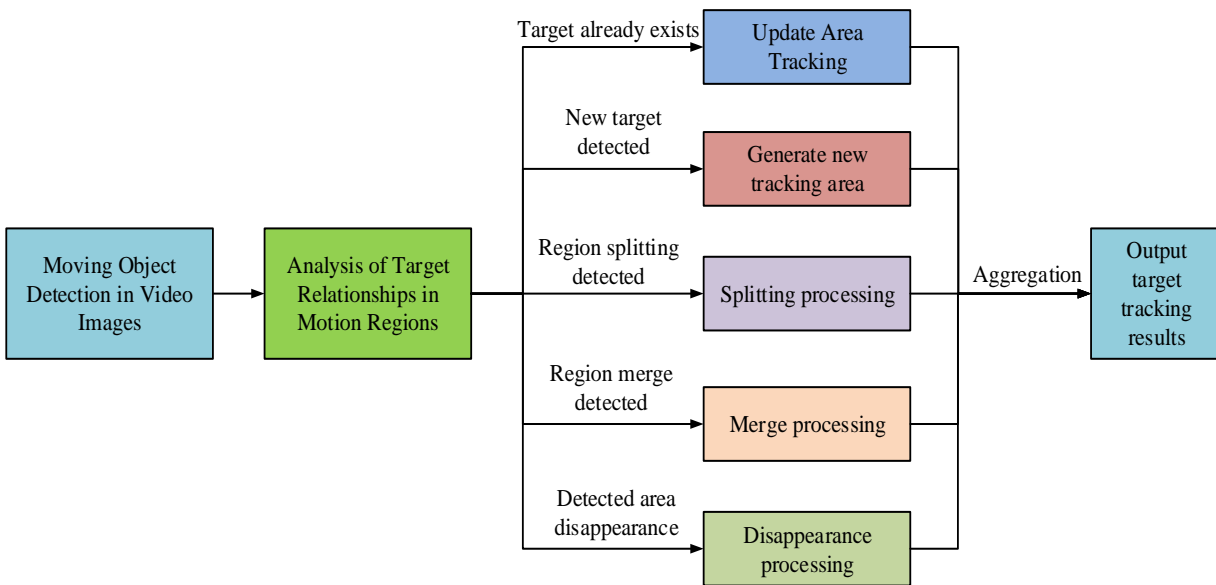
Fig. 5.   Flow chart of multiple targets tracking in martial arts competition video.

IV. Performance Analysis of Multi Target Tracking and Detection Models

To verify the performance of the multi-objective tracking model, 58 videos of different martial arts competitions are obtained through an authorized platform. Each is five minutes. The number of targets in the video varies, ranging from 1 to 3 people. These 58 videos are constructed into a multi target tracking dataset to validate the application performance of the model.

A. Performance Analysis of Multi Target Tracking and Detection Models

To analyze the performance of multi target tracking models, the Kalman filtering algorithm and the Minimum Output Sum Square Error (MOSSE) algorithm were compared with the propose method. The error comparison results of the three methods on the X and Y axes of the image are shown in Fig. 6.
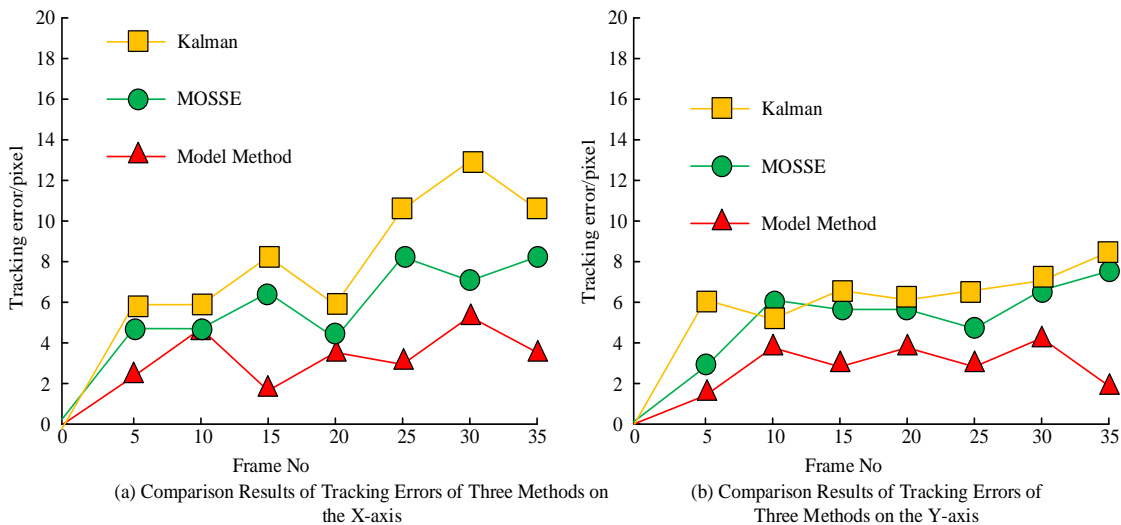
In Fig. 6 (a), there was a certain difference in the tracking error of the three methods on the X-axis. The average tracking error of the proposed model was 3.86%. The tracking errors of MOSSE and Kalman methods were 5.94% and 8.05%, respectively. In Fig. 6 (b), the tracking error of the three methods on the Y-axis was smaller than that on the X-axis. The tracking error of the proposed method was 3.38%. The tracking errors of MOSSE and Kalman were 5.17% and 6.23%, respectively. All errors did not exceed 10%. This indicated that the method used to construct the model had high robustness in identifying image targets. To verify the accuracy and recall of the model method in tracking targets, the ratio of the identified targets to the actual targets was used as an evaluation indicator. The comparison results of accuracy and recall were shown in Fig. 7.

(a) Comparison Results of Tracking Errors of Three Methods on the X-axis

(b) Comparison Results of Tracking Errors of Three Methods on the Y-axis

Fig. 6.   Error comparison results of three methods on the X and Y axis of images.

From Fig. 7 (a), all three methods had certain effects in the video target tracking process. The average accuracy of the proposed method in the video target tracking process was 93.64%. The average accuracy of MOSSE and Kalman methods was 81.09% and 78.16%, respectively. Compared to this method, it was 12.55% and 15.48% higher. In Fig. 7 (b), there was also a certain gap in the recall rate of video tracking data among the three methods. The proposed method had a recall rate of 95.48%. The recall rate of MOSSE method was

89.07%. The recall rate of the Kalman method was 83.47%. Recall rate refers to the proportion of correctly predicted positive samples to all actual positive samples. A high recall rate indicates that the model has a stronger ability to correctly predict positive examples. To further validate the multi target tracking performance of the model, three methods were applied to the training and validation sets for comparison. The comparison results were shown in Fig. 8.
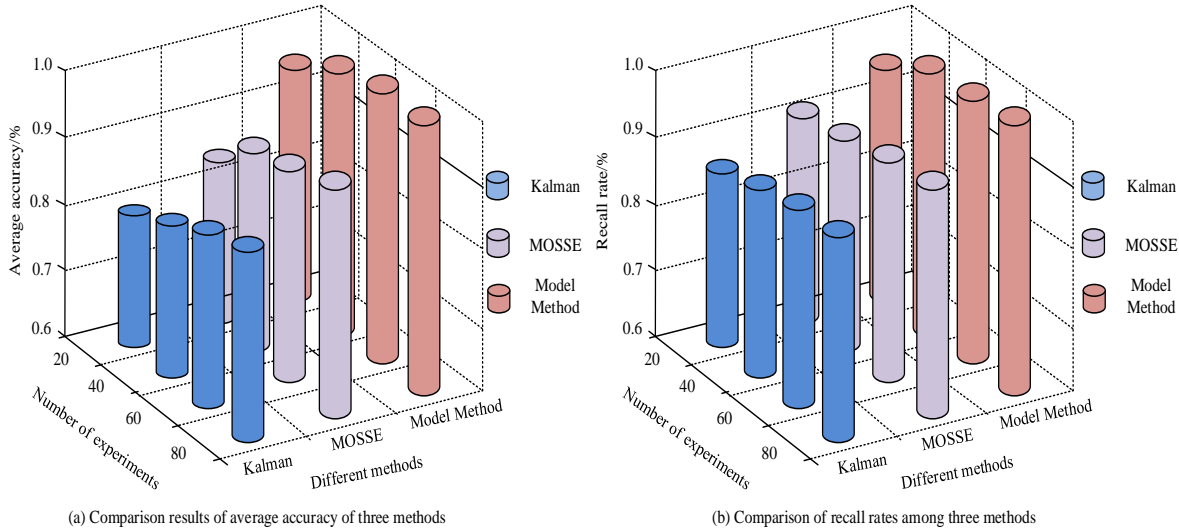

(a) Comparison results of average accuracy of three methods

(b) Comparison of recall rates among three methods

Fig. 7. Comparison results of target tracking accuracy and recall rate in videos using three methods.

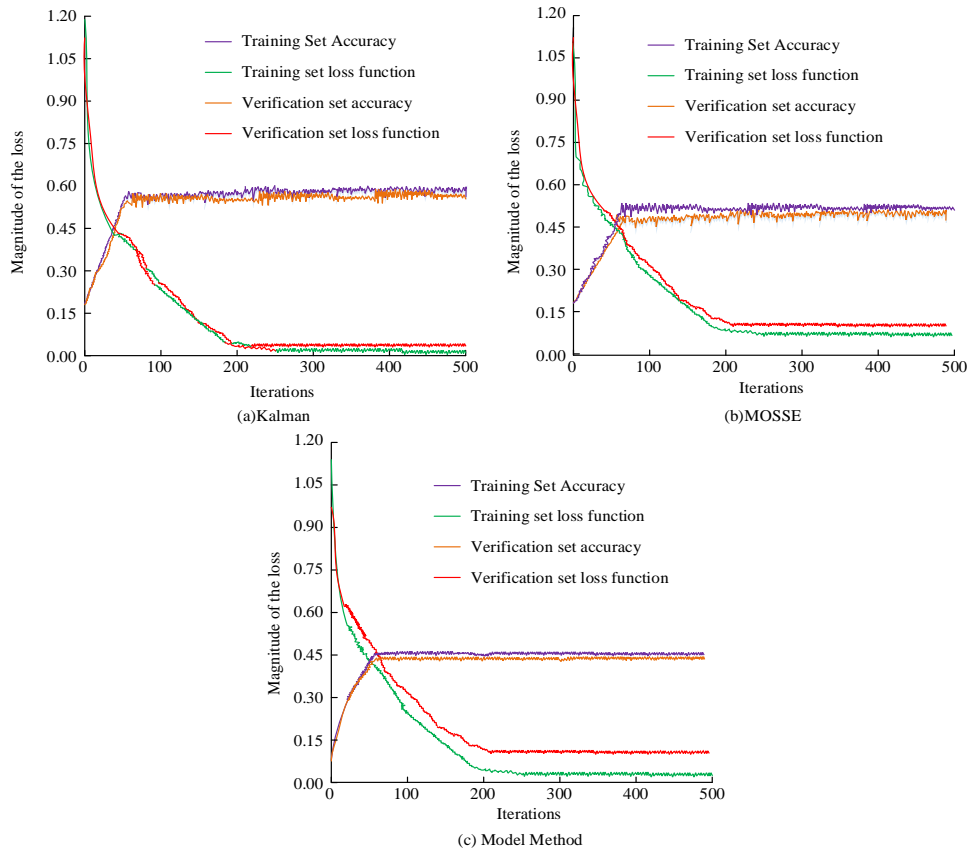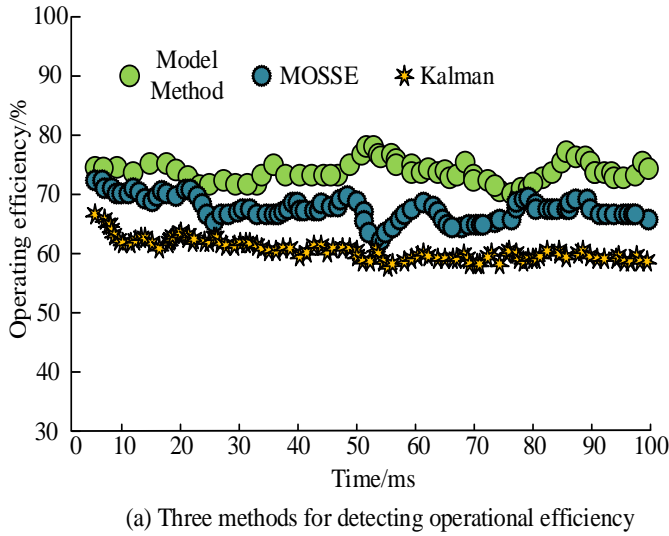(a)Kalman

(b)MOSSE

(c) Model Method

Fig. 8. Comparison results of loss values between three methods in training and validation sets.

From the comparative analysis in Fig. 8, the proposed method had a faster convergence speed at runtime compared to MOSSE and Kalman methods. After 100 iterations, the results gradually stabilized. At this point, the accuracy difference between the training set and the validation set was very small. The value of the loss function decreased faster and the value was also smaller. This indicated that the proposed method had better stability in the target tracking process compared to the comparison method. Compared to MOSSE and Kalman methods, this method could converge to the exact position of the target faster. The model could learn the features of the target faster, so that it could match more

accurately when the target reappeared. On the training and validation sets, the research model could better fit the features of the target with high accuracy.

B. The Application Effect of the Multi Target Tracking and Detection Model

To verify the effectiveness of the multi target tracking and detection model in practical applications, the real-time performance of target tracking and the complexity of algorithm operation were used as indicators for verification. The operational efficiency and computational complexity of the three methods were shown in Fig. 9.

(a) Three methods for detecting operational efficiency
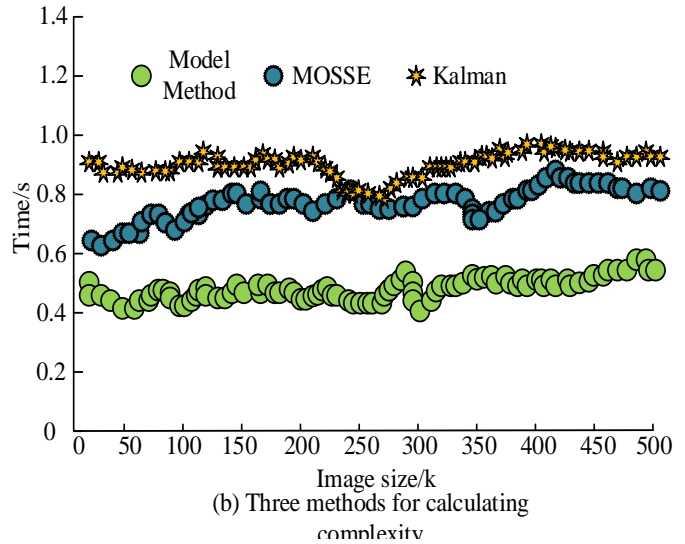
(b) Three methods for calculating complexity

Fig. 9. Comparison of efficiency and computational complexity of three methods for target checking.

In Fig. 9 (a), the operational efficiency of video object detection could reflect the speed of the method in practical applications. The operational efficiency of the proposed method in video object tracking was 77.48%. The operational efficiency of the MOSSE method was 67.34%. The operational efficiency of the Kalman method was 62.55%. In Fig. 9 (b), there were significant differences in the complexity of the three methods in actual operations. When using the proposed method to process images of the same size, the three methods also showed an increasing trend in time consumption as the image size increased. The average time consumption of the proposed method was 0.59ms. The average time consumption of the MOSSE method was 0.85ms. The average time consumption of the Kalman method was 0.93ms. The computational complexity represents the time and space resources required for algorithm execution. Low computational complexity results in less runtime and resources, greatly improving computational efficiency. To further validate the performance of the proposed method, the tracked predicted values were compared with the actual values. The comparison results were shown in Fig. 10.

In Fig. 10, the pixel error range of the true value was between [4.9-13.6]. The pixel error range of the predicted value was between [5.3-12.1]. The difference between the maximum and minimum predicted values and the true values was 1.5 and 0.4. By combining the pixel error trend, the predicted value was basically consistent with the actual trend.

This indicated that the model method had strong applicability in multi target tracking videos. To verify the tracking effect of the model method in multi target video orientation in martial arts competition videos, it was converted into a coordinate system for trajectory prediction. The results of trajectory prediction were shown in Fig. 11.
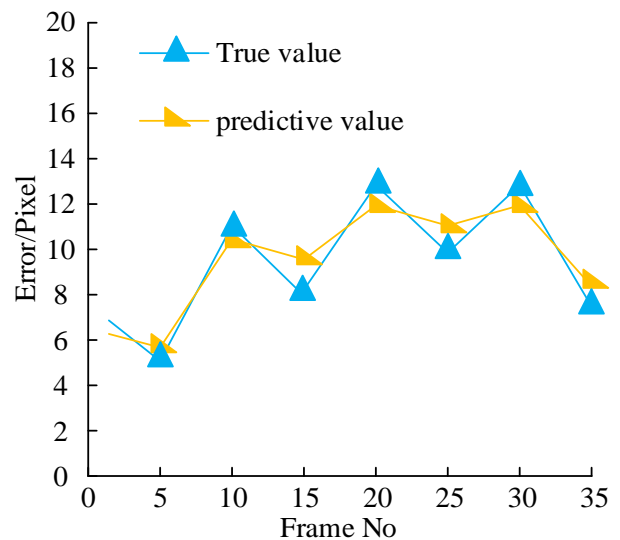

Fig. 10. Comparison results between video tracking predicted values and true values.
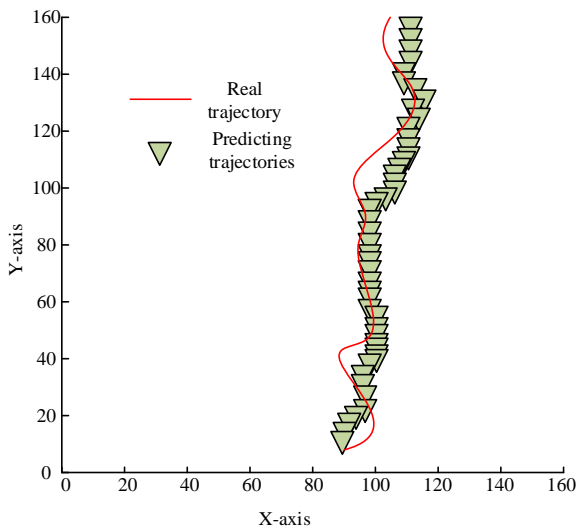
Fig. 11. Comparison of multiple target tracking effects of model methods in coordinate systems.

In Fig. 11, there may be some deviation in the predicted trajectory compared to the actual trajectory. However, the

overall trend gap was not significant. Especially in the Y-axis direction, there was a high degree of overlap between the predicted trajectory and the actual trajectory. This may be related to tracking target movement. Despite certain deviations, the predicted trajectory could still roughly reflect the movement trend of the target. This meant that the model could capture the motion patterns of the target and make relatively accurate predictions. To verify the target detection efficiency and multi target tracking ability of the model method, the average detection time and the ability to process multiple targets were used as validation indicators, as shown in Fig. 12.

In Fig. 12 (a), as the number of videos increased, the time required for the proposed method to track the target also showed an upward trend. But the rising speed was not fast, maintaining around 1ms. This indicated that the model method had strong adaptability, which could be well used for target tracking. In Fig. 12 (b), there were a total of 75 moving targets in the entire video image. The proposed method detected a total of 68 moving targets with a detection rate of 91.67%. This indicated that the tracking accuracy and stability of the model method in martial arts competition videos met the design requirements.
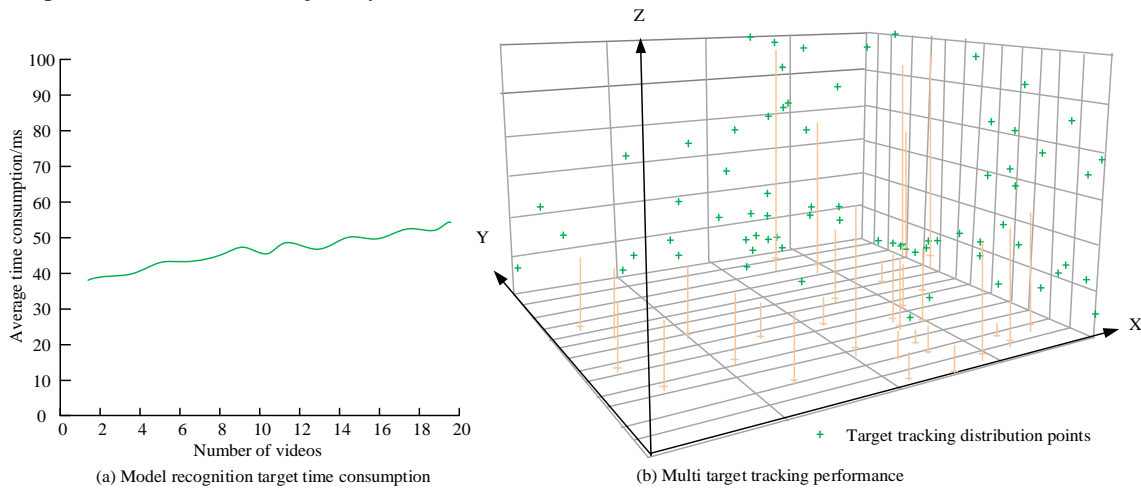

(a) Model recognition target time consumption

(b) Multi target tracking performance

Fig. 12. Model method average detection time and multiple target tracking ability results.

V. DISCUSSION

By analyzing the effective research of multi-target tracking based on feature matching with Kalman filter algorithm for target detection in martial arts competition videos, the method still has some limitations and challenges in performing target detection. First, the data volume and labeling problem is one of the key factors affecting the experimental effect. Due to the relatively small dataset of martial arts competition videos, the training of the deep learning model may not be sufficient, thus affecting the accuracy of target detection. Meanwhile, for the multi-target tracking task, labeling the trajectory of each target is a time-consuming and complex task, especially when there are frequent occlusions and interactions between the targets, and the difficulty of labeling will further increase. Second, the problem of fast target movement and occlusion is also one of the frequently encountered problems in the experiment. In martial arts competitions, the frequent rapid movement of targets with frequent occlusions leads to an increase in the

difficulty of feature extraction and also affects the accurate prediction and correction of target trajectories by Kalman filtering. To solve this problem, more advanced target detection algorithms can be tried to improve the accuracy and speed of target detection. In addition, the model generalization ability is also one of the issues that need to be paid attention to in the experiment. Although this study is optimized for martial arts competition videos, the generalization ability of the model may be limited when facing other types of videos or real application scenarios. In order to improve the generalization ability of the model, techniques such as migration learning can be tried to extract more representative features from large-scale datasets to enhance the generalization ability of the model. Finally, real-time performance requirements are also one of the factors to be considered in experiments. For practical applications, such as real-time match analysis or referee assistance systems, the real-time performance of the algorithm is very critical. Therefore, how to improve the running speed of the algorithm while ensuring accuracy is a

problem to be solved. Techniques such as parallel computing can be tried to accelerate the running speed of the algorithm to meet the real-time performance requirements.

In summary, the research on multi-target tracking based on feature matching and Kalman filtering algorithms for target detection in martial arts competition videos faces limitations and challenges in terms of the amount of data, labeling, target motion characteristics, and scene complexity. In order to solve these problems, the performance of the algorithm needs to be further investigated and improved to enhance its target detection accuracy and robustness in martial arts competition videos.

VI. CONCLUSION

In response to the low accuracy and poor stability in traditional target tracking methods for martial arts competition videos, a multi target tracking and detection model for martial arts competition videos is constructed by integrating feature matching and Kalman filtering algorithms. The results showed that the operational efficiency of the model method in video object tracking was 77.48%, with an average time of 0.59ms. The maximum and minimum predicted values of the proposed method differed from the true values by 1.5 and 0.4, respectively. In the entire video image, there were a total of 75 moving targets. The proposed method detected a total of 68 moving targets with a detection rate of 91.67%. The model performs well in object detection in martial arts competition videos. This model can accurately detect targets in videos under different scenes and lighting conditions. It has high stability and robustness. In addition, the model can also handle the multiple targets to ensure that each target is correctly detected and tracked. Overall, the proposed Kalman filter algorithm based on feature matching and multi target tracking has high accuracy and stability in the target detection model of martial arts competition videos. It can effectively handle the multiple targets, providing an effective technical means for real-time scoring of martial arts competitions. However, there are still shortcomings in the research. There are many complex backgrounds, lighting, etc. in martial arts competitions. In special environments, the tracking and detection capabilities of the proposed method still need to be further improved. Meanwhile, there are deficiencies in the research of multi-target tracking based on feature matching and Kalman filtering algorithm for target detection in martial arts competition videos in terms of data volume and annotation, model generalization ability, and real-time performance requirements. Future research should further expand the dataset, introduce advanced techniques, optimize the algorithm performance, and focus on the requirements of real-time applications to improve the accuracy and robustness of target detection. Interdisciplinary cooperation and communication will also bring new ideas and methods for research in this field.

REFERENCES

[1] J. Wąsik, D. Mosler, D. Ortenburger, T. Gora, "Stereophotogrammetry measurement of kinematic target effect as speed accuracy benchmark indicator for kicking performance in martial arts," Acta of bioengineering and biomechanics, vol. 23, pp. 1509-1514, 2021.

[2] C. Guo, S. You, Y. Luo, J. Hianlang, W. Dong, "A fast moving target detection tracking and trajectory prediction system for binocular vision," Wuhan University Journal of Natural Sciences, vol. 26, pp. 69-80, 2021.

[3] V. H. Le, "Deep learning-based for human segmentation and tracking, 3D human pose estimation and action recognition on monocular video of MADS dataset," Multimedia Tools and Applications, vol. 82, pp. 20771-20818, 2023.

[4] S. Yang, "Face feature tracking algorithm of aerobics athletes based on Kalman filter and mean shift," International Journal of Biometrics, vol. 14, pp. 394-407, 2022.

[5] R. Golash, Y. K. Jain, "Real-time automatic tracking of hand motion in RGB videos using local feature SIFT." International Journal of Intelligent Systems Design and Computing, vol. 3, pp. 161-177, 2020.

[6] P. Sharma, M. Gangadharappa, "Abnormal behavior detection of stationary objects in surveillance videos with visualization and classification," Concurrency and Computation: Practice and Experience, vol. 34, Pp. 56-70, 2022.

[7] C. R. Llano, Y. Ren, N. I. Shaikh, "Object Detection and Tracking in Real Time Videos," International Journal of Information Systems in the Service Sector (IJISSS), vol. 11, pp. 1-17, 2019.

[8] S. Lu, B. Wang, H. Wang, L. Chen, X. Zhang, "A real-time object detection algorithm for video," Computers & Electrical Engineering, vol. 77, pp. 398-408, 2019.

[9] Y. Fang, B. Luo, T. Zhao, H. Dong, B. B. Jiang, Q. Liu, "ST-SIGMA:Spatio-temporal semantics and interaction graph aggregation for multi-agent perception and trajectory forecasting," CAAI Transactions on Intelligence Technology, vol. 7, pp. 744-757, 2022.

[10] Qiu, Ji, L. Wang, Y. Hu, Y. Wang, "Effective object proposals: size prediction for pedestrian detection in surveillance videos," Electronics Letters 56, vol. 14, pp. 706-709.

[11] Y. Lyu, M. Y. Yang, G. Vosselman, G. S. Xia, "Video object detection with a convolutional regression tracke," ISPRS journal of photogrammetry and remote sensing, vol. 176, pp. 139-150, 2021.

[12] Z. Chen, L. Huang, "Online multi-target tracking algorithm based on Kalman filter and multiple information fusion," Information/Communication, vol. 17, pp. 35-38, 2019.

[13] H. Chen, J. Wang, C. Wang, D. Wang, M. Xin, "Distributed Diffusion Unscented Kalman Filtering Algorithm with Application to Object Tracking," IFAC-PapersOnLine, vol. 53, pp. 3577-3582, 2020.

[14] S. Liu, "The research of multi-object tracking algorithm using Kalman filtering method," International Journal of Innovative Computing and Applications, vol. 10, pp. 107-114, 2019.

[15] Y. Cai, D. Li, X. Wu, R. Song, Y. Wang, "A TLD target tracking algorithm combining CAMSHIFT and Kalman filtering," Computer applications and software, vol. 36, pp. 211-215, 2019.

[16] Y. Lei, "Research on micro video character perception and recognition based on target detection technology," Journal of Computational and Cognitive Engineering, vol. 1, pp. 83-87, 2022.

[17] B. B. Benuwa, B. Ghansah, "Locality-Sensitive Non-Linear Kalman Filter for Target Tracking," International Journal of Distributed Artificial Intelligence (IJDAI), vol. 13, pp. 36-57, 2021.

[18] H. G. Asl, A. D. Firouzabadi, "Integration of the inertial navigation system with consecutive images of a camera by relative position and attitude updating," Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, vol. 233, pp. 5592-5605, 2019.

[19] H. Li, J. Zhu, "Target tracking algorithm based on mean shift and Kalman filter," Journal of Beijing Institute of Technology, vol. 28, pp. 365-370, 2019.

[20] C. Lu, "Kalman tracking algorithm of ping-pong robot based on fuzzy real-time image," Journal of Intelligent & Fuzzy Systems, vol. 38, pp. 3585-3594, 2020.