Heliyon 10 (2024) e38865

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Deep learning neural network-assisted badminton movement recognition and physical fitness training optimization

Chuanbao He [a,*], Min Zhang [b]

[a] Department of Physical Education, Tianjin Sino-German University of Applied Sciences, Tianjin, 300350, China
[b] Graduate School, Metharath University, Bangkok, 10400, Thailand

ARTICLE INFO

Keywords:
Deep learning neural networks
Badminton movement
Trajectory recognition and classification
Movement trajectory prediction

ABSTRACT

This work aims to solve the problem of low accuracy in recognizing the trajectory of badminton movement. This work focuses on the visual system in badminton robots and conducts side detection and tracking of flying badminton in two-dimensional image plane video streams. Then, the cropped video images are input into a convolutional neural network frame by frame. By adding an attention mechanism, it helps identify the badminton movement trajectory. Finally, to address the detection challenge of flying badminton as a small target in video streams, the deep learning one-stage detection network, Tiny YOLOv2, is improved from both the loss function and network structure perspectives. Moreover, it is combined with the Unscented Kalman Filter algorithm to predict the trajectory of badminton movement. Simulation results show that the improved algorithm performs excellently in tracking and predicting badminton trajectories compared with the existing algorithms. The average accuracy of the proposed method for tracking badminton trajectories is 91.40 %, and the recall rate is 84.60 %. The average precision, recall, and frame rate of the measured trajectories in four simple and complex scenarios of badminton flight video streams are 96.7 %, 95.7 %, and 29.2 frames/second, respectively. They are all superior to other classic algorithms. It is evident that the proposed method can provide powerful support for badminton trajectory recognition and help improve the accuracy of badminton movement recognition.

1. Introduction

Badminton, as an ancient and enduring sport, has garnered widespread attention and passion worldwide today [1–3]. Its rapid development and popularity have enriched people's sports and cultural lives and led to continuous improvement in professional athletic performance. However, in this fiercely competitive domain, athletes' technical and physical training increasingly demands more precise and personalized support [4]. Traditional coaching experience alone no longer suffices to meet the demands for sports data and performance analysis and the introduction of technology has become a vital avenue to enhance training efficacy [5–7]. With the rapid advancement of computer science and artificial intelligence (AI), breakthroughs in fields like computer vision and deep learning (DL) present new opportunities for the analysis and training of badminton [8–10]. Accurately capturing athletes' movements and trajectories is crucial for improving training effectiveness and optimizing competitive performance. However, traditional motion trajectory recognition methods fall short when confronted with the challenges posed by high-speed badminton flight and diverse

* Corresponding author.
E-mail address: hechuanbao@tsguas.edu.cn (C. He).

https://doi.org/10.1016/j.heliyon.2024.e38865
Received 3 June 2024; Received in revised form 27 September 2024; Accepted 1 October 2024
Available online 2 October 2024

motion patterns in badminton [11,12].

Several key issues arise in the field of badminton motion trajectory recognition. First, the high-speed flight of badminton results in significant trajectory changes within a short period, necessitating an efficient real-time recognition method [13–15]. Besides, conventional image processing and feature extraction techniques struggle to handle complex and variable motion scenarios, prompting the need for more advanced computer vision and DL methods [16]. Additionally, badminton in flight often appears as a small target in video streams, placing higher demands on detection and tracking algorithms [17–19]. In this context, this work primarily focuses on utilizing DL neural networks to assist in achieving more precise, real-time recognition of badminton movements and physical training optimization [20–22]. Specifically, the following core issues need to be addressed: How can an effective visual system be designed to achieve lateral detection and tracking of flying badminton? How can the convolutional neural network (CNN) and attention mechanisms be leveraged to extract and accurately identify badminton's motion trajectory from video streams?

In light of this, this work introduces a trajectory recognition method for badminton movement based on DL neural networks. It achieves precise tracking of flying badminton by crafting an efficient visual system. Additionally, it incorporates an attention mechanism to further enhance the accuracy of identifying badminton movement trajectories, which helps address complex and diverse motion patterns. The detection and tracking performance of flying badminton as the small target is optimized through improvements made to the Tiny YOLOv2 network from both network structure and loss function perspectives. Moreover, this work combines the Unscented Kalman Filter (UKF) algorithm to accurately predict the trajectory of badminton movement, providing reliable data support for optimizing physical training. With these concerted efforts, this work aims to bridge existing gaps in trajectory recognition research for badminton movement and offer more dependable support for athletes' physical training and competitive performance.

2. Literature review

In recent years, the field of motion analysis has achieved remarkable accomplishments with the rapid advancement of computer vision and DL technology [23]. In badminton sports, the application of visual systems has become one of the crucial means to enhance training effectiveness and competitive performance. In prior research, some classic motion trajectory recognition algorithms have been attempted in the context of badminton, such as methods based on feature extraction and traditional object detection algorithms. Siyal et al. (2021) [24] developed and tested a model to detect variations in badminton movements. Chen et al. (2021) [25] emphasized the importance of visual management in improving the accuracy of badminton movement recognition. Liu (2022) [26] utilized visual search technology and stereo vision cameras to record videos. Then, image stabilization technology was employed to perform histogram equalization on the recorded visual images, resulting in visual images after noise elimination and enhancement. The method also utilized the scale-invariant feature transform algorithm to capture image corner features. It combined with the support vector machine algorithm to build a feature recognition model, enabling the distinction of the third-arm movement of badminton players. Steels et al. (2020) [27] organized a data collection activity and designed a novel neural network using different frame sizes as inputs. The findings demonstrated that when utilizing accelerometer data alone, the new CNN could distinguish 99 types of badminton activities with 50 % accuracy using a sampling frequency of 86 Hz. Fang & Sun (2021) [28] employed body feature similarity based on css to construct an angle-adaptive and continuously scaled spatial template matching algorithm. This algorithm calculated the similarity between the horizontal board and detected image values, set a specific threshold, and matched the region of the human body in sports activities.

Cui & Zheng (2021) [29] developed a badminton recognition and tracking system using two high-speed cameras with 2,000,000 pixels each. This system captured images from the high-speed cameras through the Camera Link interface and processed all captured images in real-time with different regions of interest settings. Compared to traditional circular detectors and points in three-dimensional coordinates, the proposed elliptical detector in this method reduced the error by approximately 3 mm. Chen & Hu (2022) [30] employed a mobile neural network inference engine to create a new AI-powered intelligent training model for sports disciplines. This model could segment motions into a series of decomposed movements, which were individually recognized and assessed to measure the overall performance. Lin et al. (2023) [31] aimed to decrease coaching workload and enhance learners' badminton performance and designed a multi-feedback physical education model for badminton courses. This model provided learners with visual feedback, information feedback, and verbal guidance. Duncan et al. (2023) [32] examined the short-term (pre-competition) and long-term (up to 10 weeks post) effects of the World Badminton Federation's shuttle time program on the fundamental motor skills and physical health of Saudi boys and girls.

It suggests that recent research has made some progress in the field of badminton movement trajectory recognition. However, limitations still exist in accurately tracking and recognizing the trajectories of flying badminton. Some classic motion trajectory recognition algorithms have been attempted in the context of badminton, such as methods based on feature extraction and object detection. Traditional approaches struggle to meet the demands of high-speed motion and variable trajectories of flying badminton. However, existing DL methods still require small target detection and real-time performance improvements. This work aims to address the limitations above in the field of badminton movement trajectory recognition. This work presents a comprehensive and innovative method to enhance flying badminton trajectories' recognition and prediction accuracy. The approach encompasses various aspects, from visual system design, the application of CNN, the introduction of attention mechanisms, improvements to the Tiny YOLOv2 network, and integration with the UKF algorithm. By overcoming the shortcomings of existing methods, this work strives to provide more reliable and advanced support for the analysis, training, and competitive performance of badminton.

3. Research methodology

3.1. Badminton movement trajectory recognition with the introduction of attention mechanisms

In order to address the lateral detection and tracking of flying badminton, this work establishes a visual system aimed at accurately capturing the lateral information of flying badminton from video stream data and tracking their motion trajectory in real time. This process involves several key steps: segmentation and edge detection of the lateral images of the flying badminton, video image processing, application of CNN, and the introduction of an attention mechanism. First, image processing techniques, specifically edge detection and image segmentation, are employed to separate the badminton from the background, extracting the lateral view of the flying badminton from the video stream, thereby emphasizing its shape. A target-tracking algorithm based on Kalman filtering is utilized to achieve real-time badminton tracking. Kalman filtering predicts the current position of badminton based on previous state and observation data, and corrects it using new observation data. This enables real-time tracking of the badminton's movement trajectory. Fig. 1 illustrates the specific tracking process.

To improve the accuracy and robustness of badminton recognition, a CNN is introduced for further feature extraction and image classification. A dedicated CNN architecture is designed specifically for the lateral view images of badminton. This architecture comprises multiple convolutional layers and pooling layers, along with several fully connected layers. An activation function, such as ReLU, is introduced after each convolutional layer to enhance the network's ability for nonlinear modeling. CNN is trained using labeled lateral view images of badminton. During training, the stochastic gradient descent optimization algorithm is employed to fine-tune network parameters by minimizing the loss function. Dropout technology is also incorporated to prevent overfitting. The inclusion of attention mechanisms enhances the network's ability to focus more on important feature regions, resulting in stronger robustness against variations and noise. Attention modules are integrated at different levels of the CNN, enabling the adjustment of feature map weights based on the significance of image features. This allows the network to pay more attention to critical information of badminton during the recognition process, thereby enhancing sensitivity to movement trajectories. Fig. 2 depicts the specific network structure composition of the attention module.

Within the attention module, the process begins with the application of a convolutional layer to extract local features from the image. Subsequently, a global average pooling layer is adopted to compute the average for each channel within the feature map, resulting in weights of channel dimensions. Following this, the computed average weights are mapped to a weight coefficient through two fully connected layers, with the coefficient constrained between 0 and 1 (achieved through the sigmoid function). Finally, this weight coefficient is applied to the original feature map, resulting in a weighted feature map, namely the feature attention map. These

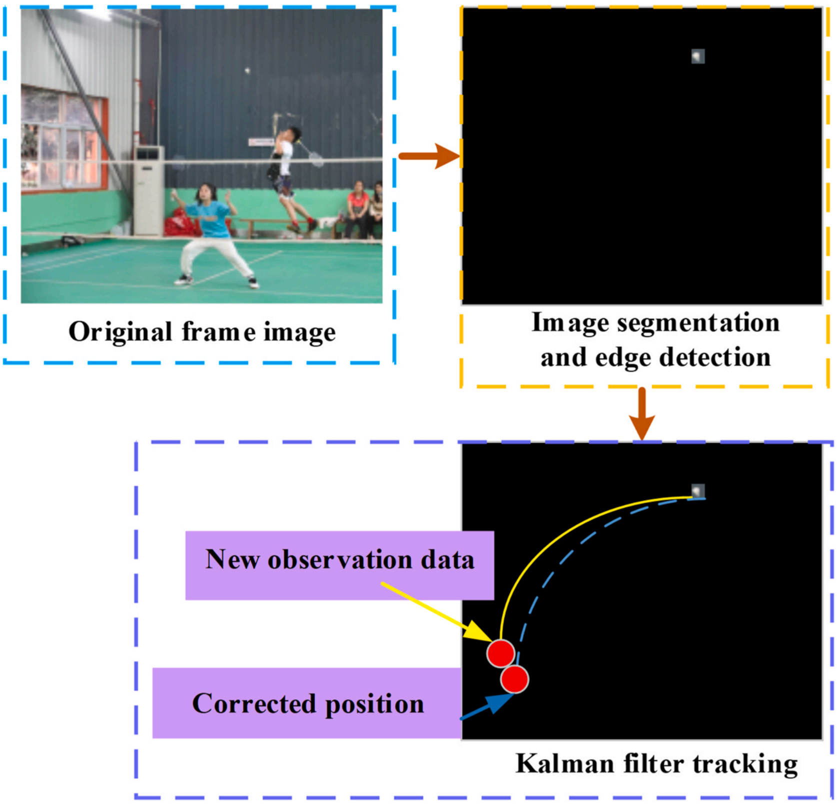


Fig. 1. Real-time tracking process of badminton movement trajectory.

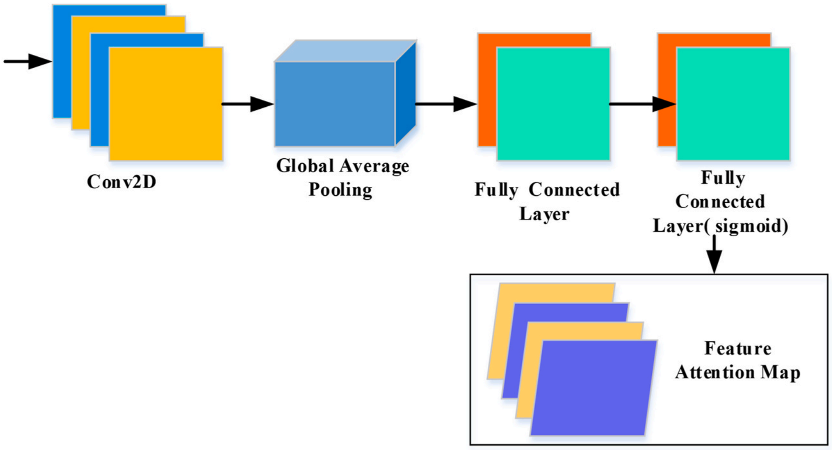


Fig. 2. Badminton trajectory recognition by introducing attention mechanism.

components collaborate synergistically, forming a comprehensive visual system that enables accurate recognition and prediction of flying badminton trajectories.

3.2. Enhancing the structure and loss function of the tiny YOLOv2 network

In video streams, badminton typically appears as small objects, necessitating an efficient real-time processing capability in the detection network. Tiny YOLOv2, a lightweight object detection network, maintains high accuracy while offering fast inference speed, making it particularly suitable for real-time applications. Additionally, considering the constraints of the competition environment and hardware resources, a lightweight model is required, and Tiny YOLOv2 meets this need effectively. By utilizing multi-scale anchor boxes, it can detect objects of various sizes, which is beneficial for capturing badminton from different angles and sizes. Moreover, Tiny YOLOv2 has optimized its network structure by adding convolutional layers, increasing the number of channels, and adjusting the receptive field size. This allows the network to more accurately capture detailed features of badminton, thereby improving the accuracy of trajectory detection. The introduction of residual connections has also deepened the network, effectively mitigating the vanishing gradient problem and enhancing training stability and performance. This improves the network's ability to model the trajectory of the badminton. Finally, to further enhance detection accuracy and robustness, this work also improves Tiny YOLOv2's loss function. By incorporating trajectory information such as the position and movement direction of the badminton into the loss function, the network is guided to focus more on the shuttle's trajectory, further improving detection accuracy. Additionally, the weights of the confidence loss and classification loss are adjusted to prioritize the accuracy of the shuttle's bounding box and trajectory, thereby enhancing detection precision.

Tiny YOLOv2 is a lightweight object detection network. While maintaining high accuracy, it offers faster inference speed, making it particularly suitable for real-time applications such as detecting and recognizing badminton movement trajectories. Given the constraints of the competition environment and hardware resources, a lightweight model is needed to meet these requirements, and Tiny YOLOv2 fits this need effectively. Zhan et al. (2024) [33] proposed an improved version of the YOLO series in their research, aiming to enhance the speed and accuracy of object detection. Tiny YOLOv2, a lightweight version of YOLO, achieves real-time detection by reducing the number of network layers and parameters. This design enables the model to perform rapid inference even with limited computational resources [33]. To further improve Tiny YOLOv2's performance in detecting badminton trajectories, this work introduces additional convolutional layers and increases the number of channels to enhance feature extraction capabilities. By adjusting the receptive field size, the network can more accurately capture the detailed features of the shuttle, thereby improving trajectory detection accuracy. Moreover, the introduction of residual connections helps alleviate the vanishing gradient problem, enhancing the network's training stability and performance. This approach has also been thoroughly discussed in the research by Yogeswararao et al. (2022) [34]. Utilizing multi-scale anchor boxes, Tiny YOLOv2 can detect targets of varying scales, which is advantageous for capturing different angles and sizes of flying badminton objects. In the context of badminton matches, where badminton sizes and motion trajectories vary significantly, Tiny YOLOv2 is well-suited for object detection. Improvements have been incorporated into the network structure to further enhance its performance in detecting badminton movement trajectories, combining several advantages to better meet the task demands.

Initially, the network is augmented by adding convolutional layers and increasing the number of channels to enhance feature extraction capabilities. By adjusting the receptive field size, the network can more accurately capture detailed features of the badminton object, thereby elevating the accuracy of trajectory detection. Subsequently, a residual connection structure is introduced to deepen the network and effectively alleviate the issue of gradient vanishing. This enhances both training stability and performance, bolstering the network's capacity to model badminton movement trajectories. Fig. 3 illustrates the improved network structure:

Lastly, additional trajectory information, encompassing the position and motion direction of the badminton object, is incorporated

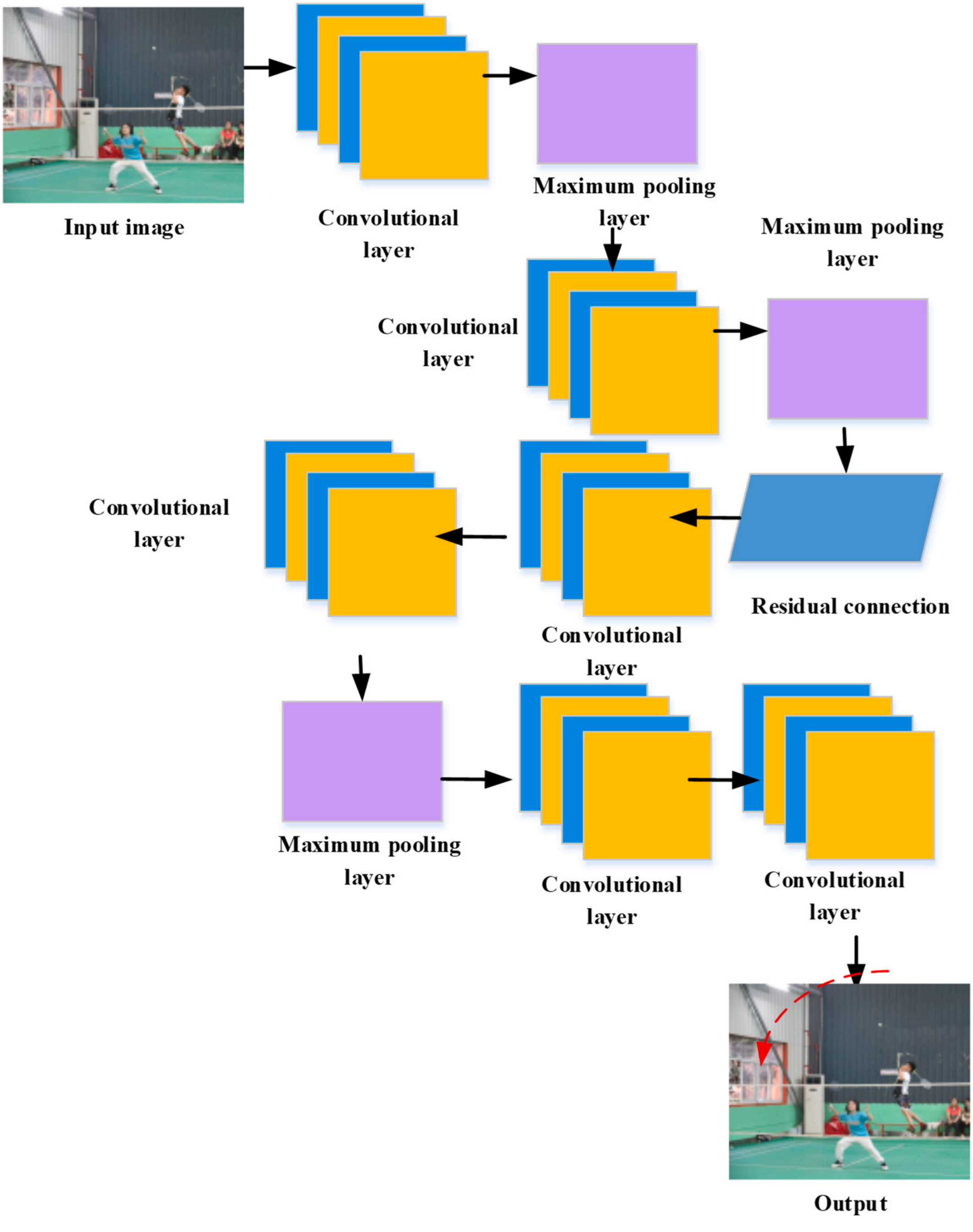


Fig. 3. Improved the Tiny YOLOv2 network structure.

into the loss function. This design directs the network's focus towards the movement trajectory of the badminton object, further elevating detection accuracy and robustness. Encouraging the network to learn trajectory information makes it more adept at precisely predicting the badminton object's movement trajectory. Furthermore, adjustments are made to the weightings of the confidence loss in the loss function, prioritizing the accuracy of the badminton object's bounding box and movement trajectory. This aids in enhancing detection precision. In object detection tasks, the design of the loss function is crucial for the model's training effectiveness. This work introduces additional trajectory information loss into the loss function to enhance the model's focus on the shuttle's movement trajectory. The residual learning framework proposed by Shafiq et al. (2022) provided significant insights for designing the loss function, especially when handling complex tasks [35]. By incorporating additional trajectory information such as position and movement direction into the loss function, the network can more accurately predict the shuttle's movement trajectory. Specifically, this work optimizes localization loss, confidence loss, classification loss, and trajectory information loss. The expression of the revised loss function reads:

$$L_{\text{total}} = L_{\text{loc}} + L_{\text{conf}} + L_{cls} + L_{\text{traj}} \tag{1}$$

L_{total} denotes the total loss, L_{loc} represents the localization loss, L_{conf} signifies the confidence loss, L_{cls} stands for the classification loss, and L_{traj} denotes the trajectory information loss. The localization loss L_{loc} is measured using the Mean Squared Error (MSE) to quantify the disparity between the target's predicted and actual bounding boxes. The computation is as follows:

$$L_{\text{loc}} = \sum_{i}^{N} MSE\left(\ominus_i^{box}, \ominus_i^{true}\right) \tag{2}$$

Here, $\ominus_i^{true}$ represents the true bounding box, and $\ominus_i^{box}$ represents the predicted bounding box. The confidence loss L_{conf} is evaluated using the Cross-Entropy Loss to measure the disparity between the predicted and actual values of target confidence. The computation is as follows:

$$L_{\text{conf}} = \sum_{i}^{N} \left(c_i^{ture} * \log\left(c_i^{pred}\right) + \left(1 - c_i^{ture} * \log\left(1 - c_i^{pred}\right)\right)\right) \tag{3}$$

c_i^{ture} represents true confidence, and c_i^{pred} represents the predicted confidence. The classification loss L_{cls} employs the Cross-Entropy Loss to measure the disparity between the predicted and actual target categories. The computation is as follows:

$$L_{cls} = \sum_{i}^{N} \left(p_i^{ture} * \log\left(p_i^{pred}\right) + \left(1 - p_i^{ture} * \log\left(1 - p_i^{pred}\right)\right)\right) \tag{4}$$

p_i^{ture} denotes the true category, and p_i^{pred} signifies the predicted category. The trajectory information loss L_{traj} is calculated as follows:

$$L_{cls} = \sum_{i}^{N} \left(MSE\left(x_i^{pred}, x_i^{ture}\right) + MSE\left(y_i^{pred}, y_i^{ture}\right) + MSE\left(\theta_i^{pred}, \theta_i^{ture}\right)\right) \tag{5}$$

N represents the number of anchor boxes, i stands for the index of the anchor box, $\ominus$ denotes the target bounding box, c signifies confidence, p represents the class probability, and x, y, and θ denote position and motion direction. By comprehensively considering the losses above, it is possible to more accurately optimize Tiny YOLOv2, enabling it to adapt to the task of detecting badminton movement trajectories.

3.3. Incorporating the UKF algorithm for badminton trajectory prediction

This work introduces the UKF algorithm to predict the trajectory of badminton, highlighting its importance and effectiveness compared to methods that do not use the UKF algorithm. First, the UKF algorithm is a nonlinear state estimation method that uses a series of unscented transformations to approximate the state distribution of nonlinear systems, achieving more accurate state estimation. In predicting badminton trajectories, the UKF algorithm combined with the improved Tiny YOLOv2 model can better predict the badminton's future positions and movement trajectories. Specifically, in the prediction step, the UKF algorithm first utilizes the improved Tiny YOLOv2 model to detect and track the current position of badminton. Then, it takes the detected bounding box information as the initial state, including position, speed, and direction. This method shows superiority in handling nonlinear systems, such as the movement trajectory of a badminton shuttle. By generating Sigma points and propagating them through the nonlinear function, the UKF algorithm can approximate the changes in state distribution, thus predicting the shuttle's state (position, speed, and direction). This process considers not only the badminton's speed and direction but also external influencing factors such as wind force. In contrast, methods that do not use the UKF algorithm may be affected by nonlinear factors when predicting shuttle trajectories, leading to reduced prediction accuracy. Traditional Kalman Filters might encounter issues with accurately approximating state distributions in nonlinear systems, whereas the UKF algorithm effectively addresses this problem through unscented transformations.

The UKF algorithm is a nonlinear state estimation method used for predicting the trajectory of a target. It approximates the state distribution of a nonlinear system through a series of unscented transforms, leading to more accurate state estimation. Liu et al. proposed the UKF algorithm to improve the application of traditional Kalman filters in nonlinear systems. Their research demonstrated the superiority of the UKF in handling nonlinear state estimation, particularly in approximating state distributions [36]. In the badminton trajectory prediction, the UKF algorithm is combined with the enhanced Tiny YOLOv2 model to better forecast the badminton's future position and movement trajectory.

1) First, the improved Tiny YOLOv2 model is utilized to detect and track the current position of badminton. The detected bounding box information is used as the initial state, including position, velocity, and direction. Diwan et al. pointed out that the application of YOLO series models in object detection could provide real-time and accurate object localization information [37].
2) Ampountolas further improved and implemented the UKF and proposed a more efficient unscented transformation method that could more accurately estimate the state distribution of nonlinear systems [38]. The UKF algorithm estimates the state distribution of nonlinear systems through a series of unscented transformations. Subsequently, the UKF algorithm is employed for the prediction

step, whereby the badminton's future position and motion state are forecasted based on the current state estimate and system model. Specifically, in the prediction step of the UKF, a sequence of unscented transforms is employed to estimate the state distribution of the nonlinear system. For the badminton's state (position, velocity, and direction), Sigma points are first generated based on the current state estimate and covariance matrix. These Sigma points are then propagated through nonlinear functions to approximate changes in the state distribution. For each Sigma point, the system model is utilized to predict the state, involving updates to the position based on the object's velocity and direction, and considering external influences such as wind forces. Finally, the predicted state and covariance matrix are computed based on the propagated Sigma points, which will be used for correction in the subsequent update step. Fig. 4(a and b) presents the process of the UKF prediction step. This work only illustrates a single Sigma point's propagation and prediction process for simplification. In reality, there will be more Sigma points involved in the calculations.

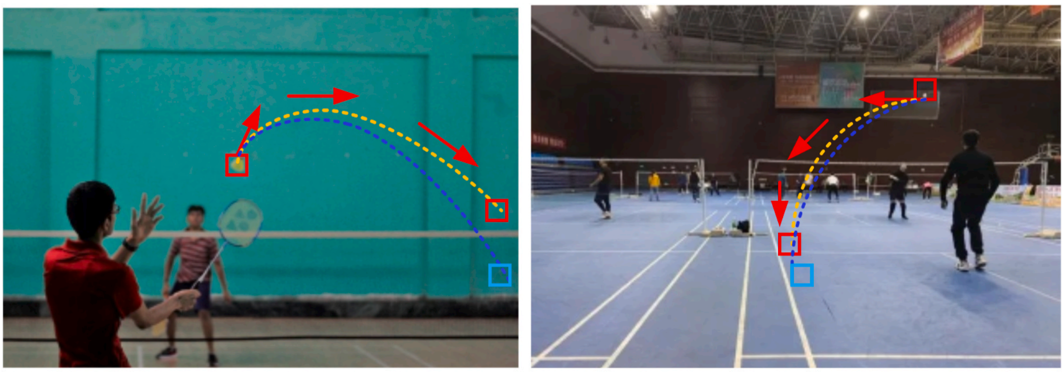
Through such prediction steps, it is possible to combine the UKF algorithm with an improved Tiny YOLOv2 model to more accurately forecast badminton's future position and motion trajectory, thereby enhancing the performance and robustness of sports trajectory recognition.

3) Finally, the UKF algorithm is used for the update step. At each time step, observational data are obtained from new image frames, which include the badminton's new position. These data are then compared with the predicted state, and this information is utilized to adjust the state estimation. The correction process consists of three main parts. Computing the observational model involves deriving the observed values for each predicted state based on the system model. This is akin to simulating the anticipated observational results using the predicted state. The update steps of the Gaussian Kalman Filter have been validated in many applications. For instance, Rigatos discussed the computation process and the correction steps of the Kalman gain, and how to adjust state estimates based on observational data [39]. Computing the Kalman gain entails determining the degree of correction by comparing the predicted state with actual observational data by calculating the Kalman gain. The Kalman gain considers the uncertainty of the state estimation and observational errors. Finally, the Kalman gain is applied to the predicted state to obtain the corrected state estimation. This step considers the observational data, making the state estimation closer to the actual situation. This yields the adjusted badminton state and covariance matrix.

3.4. Experiments and settings

Simulation experiments are conducted to validate the effectiveness of the proposed method. In order to meet the training requirements of the DLYOLO series detection network for badminton sample data, the dataset needed to possess conditions such as a large quantity, high quality, and diversity. The aim is to facilitate the network's better learning of data features and prevent model overfitting. A custom dataset is created and employed because the existing dataset can not fully satisfy these demands. In each image frame, the position of the badminton is manually annotated, indicating its pixel coordinates within the image. These annotations are used for subsequent target tracking and trajectory prediction. Through stereo vision cameras, 15 badminton flight video streams are captured from different scenarios (such as laboratory, parking lot, sports arena, forest, and roadside). Then, they are categorized into simple and complex scenarios based on factors like background, lighting, and interference. Fig. 5(a and b) illustrates some example scenarios from the badminton dataset.

To meet the training needs of the DL YOLO series detection networks for badminton sample data, this work creates and utilizes a custom dataset. The dataset includes 15 video streams of badminton flights captured from various scenes, such as laboratories, parking lots, sports fields, forests, and roadways. These scenes are categorized into simple and complex types based on factors like background, lighting, and interference. Each scene comprises approximately 100–150 consecutive image frames, covering different stages and trajectories of the badminton's aerial movement. The entire dataset contains around 1000 image frames, representing a variety of



(a) (b)

Fig. 4. Badminton trajectory prediction process based on UKF algorithm (a. Simple scenario; b. Complex scenario).


(a)

(b)

Fig. 5. Scenario examples of badminton dataset (a. Simple scenario; b. Complex scenario).

badminton movement scenarios. To improve model training and prevent overfitting, the dataset is split into a training set and a testing set in a 2:1 ratio. In each image frame, the badminton's position is manually annotated with its pixel coordinates in the image. These annotations are used for subsequent object tracking and trajectory prediction. The video streams are captured using stereo vision cameras to ensure the dataset's quality and diversity, facilitating better learning of data features by the network. When training the improved Tiny YOLOv2 model, the following parameters are set: learning rate: 0.001; batch size: 16; epochs: 50; activation function: ReLU; convolution kernel size: 3x3. These settings are aimed at optimizing the model's training effectiveness and convergence speed. During training, the stochastic gradient descent optimization algorithm is used to fine-tune the network parameters and minimize the loss function. Additionally, to prevent overfitting, Dropout techniques are incorporated. Table 1 outlines the parameter settings.

The model training and testing are conducted on high-performance computing hardware. A server equipped with an NVIDIA Tesla V100 GPU is used, enabling efficient processing of large-scale data and acceleration of model training. Additionally, all code is implemented on the TensorFlow DL framework and CUDA is utilized for accelerated computation, enhancing training speed and efficiency.

The algorithm performance evaluation metrics include recall rate (r_c), precision rate (r_p), F1 score, Intersection-Over-Union (IOU),

Table 1
Parameter settings.

Parameter type	Parameter name	Parameter value
Improved Tiny YOLOv2 model	Learning Rate	0.001
	Batch Size	16
	Epochs	50
	Activation	ReLU
	Kernel Size	3x3
UKF algorithm	Initial State Estimation	(x, y, vx, vy) = (0, 0, 2, 2)
	Initial Covariance Matrix	Identity matrix
	Observation Noise	0.5
	Process Noise	0.1

frames processed per second (FPS), and central location error (CLE). The calculation method for each metric is as follows:

$$r_c = \frac{N_{tp}}{N_{tp} + N_{fn}} \tag{6}$$

$$r_p = \frac{N_{tp}}{N_{tp} + N_{fp}} \tag{7}$$

N_{tp} represents the count of correctly detected badminton instances, N_{fp} indicates the count of erroneously detected badminton instances, and N_{fn} represents the count of missed badminton instances.

$$F_1 = \frac{2^*r_c{}^*r_p}{r_c + r_p} \tag{8}$$

$$IOU = \frac{predicted\_box \cap ground\_truth}{predicted\_box \cup ground\_truth} \tag{9}$$

predicted_box is the predicted box of badminton, and ground_truth represents the annotated box of badminton. The closer the IOU value is to 1, the better the algorithm performance is.

$$FPS = \frac{Frames}{T} \tag{10}$$

$$Avgerr = \frac{1}{N}\sum_{n-1}^{N} e_k(t) \tag{11}$$

$$e_k(t) = d(center_k(t), g_t(t)) \tag{12}$$

Frames represents the number of video frames, and T is the video processing time. $d()$ is the Euclidean distance, and $center_k(t)$ refers to the detected, tracked, or predicted badminton coordinate representation. $g_t(t)$ represents the center of the badminton annotation box, and N indicates the number of badminton detected, tracked, or predicted in each video stream.

4. Results and discussion

4.1. Quantitative data on trajectory recognition performance

This work employs an improved version of the YOLO series algorithm, specifically Tiny YOLOv2, for detecting and tracking badminton trajectories. The YOLO algorithm is a widely used real-time object detection method known for its speed and efficiency in computer vision applications. Tiny YOLOv2, as a lightweight version of the YOLO series, achieves faster inference speeds by reducing the number of network layers and parameters, making it particularly suitable for real-time scenarios. Here, the Tiny YOLOv2 network structure has been optimized by increasing the number of convolutional layers and channels, and adjusting the receptive field size. These modifications enable the network to more accurately capture detailed features of the shuttle, thereby improving trajectory detection accuracy. Additionally, residual connections have been introduced to deepen the network, effectively alleviating the vanishing gradient problem, and enhancing training stability and performance. This improvement enhances the network's ability to model

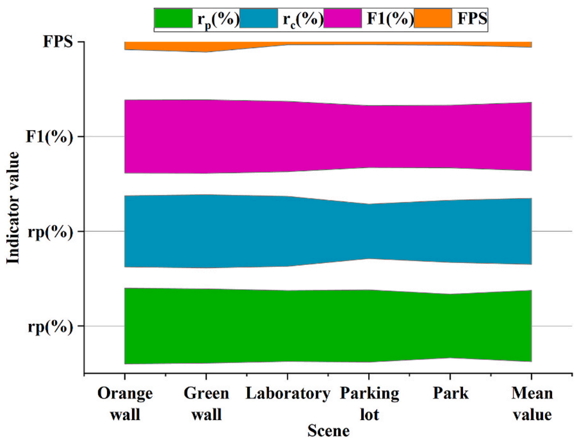


Fig. 6. Badminton tracking results of the algorithm proposed in multiple scenarios.

the badminton's movement trajectory. Furthermore, the loss function of Tiny YOLOv2 has been refined by incorporating additional trajectory information, such as the badminton's position and movement direction. This adjustment guides the network to focus more on the badminton's trajectory, further improving detection accuracy. To further enhance detection precision and robustness, additional trajectory information loss has been introduced into the loss function to emphasize the model's focus on badminton's movement trajectory. Through these improvements, this work aims to more accurately optimize Tiny YOLOv2 for better performance in badminton trajectory detection tasks. Experimental validation shows that the improved Tiny YOLOv2 performs excellently in various scenarios, achieving high accuracy and robustness in detecting badminton trajectories, regardless of whether the background is simple or complex.

Fig. 6 shows the badminton tracking results in multiple scenarios.

Fig. 6 depicts that in the orange wall scene, the model demonstrates very high precision and recall rates of 97.14 % and 90.91 %, respectively. This indicates that the model can detect and track badminton with high accuracy in this simple, low-interference background. The F1 score is 93.92 %, further confirming the model's overall performance in this scenario. The IoU value reaches 0.94, showing an excellent overlap between the predicted and ground truth bounding boxes. However, the frame rate is relatively low at 19.76FPS, which may limit the model's practicality in applications requiring high real-time performance. In the green wall scene, the precision and recall rates are slightly lower at 95.12 % and 93.67 %, respectively, but the F1 score remains high at 94.39 %. The IoU value also stays at 0.94, indicating that the model maintains high detection accuracy across different background colors. The frame rate improves to 26.3FPS, suggesting better real-time performance in this scenario. In the laboratory scene, the model's precision and recall rates are 90.91 % and 89.47 %, respectively, with an F1 score of 90.18 %. Although these metrics are slightly lower than in the previous two scenes, the IoU value remains at 0.89, demonstrating the model's detection capability in the complex laboratory environment. However, the frame rate significantly drops to 7.36FPS, likely due to the increased difficulty posed by the complex background and lighting variations in the laboratory setting. In the parking lot scene, the model achieves a high precision rate of 92.5 %, but the recall rate drops significantly to 69.64 %, resulting in an F1 score of 79.46 %. Despite the IoU value remaining at 0.94, the frame rate further decreases to 7.11FPS. This suggests that the model faces more challenges in detecting badminton in the complex parking lot environment. In the park scene, the precision and recall rates are 81.34 % and 79.45 %, respectively, with an F1 score of 80.47 %. The IoU value slightly decreases to 0.9, and the frame rate is 8.43FPS. These results indicate a drop in model performance in natural environments, although it remains within an acceptable range. Overall, the model's performance across different scenes demonstrates its adaptability and robustness in complex backgrounds. Despite lower frame rates in some scenarios, the generally high IoU values indicate good detection accuracy. On average, the model achieves a precision rate of 91.4 %, a recall rate of 84.6 %, an F1 score of 87.68 %, a frame rate of 13.792FPS, and an IoU value of 0.922 across all scenes. These results suggest that the model provides reliable performance for badminton trajectory recognition.

Fig. 7(a–d) provides the comparison results of badminton detection results using different YOLO series networks in various scenes.

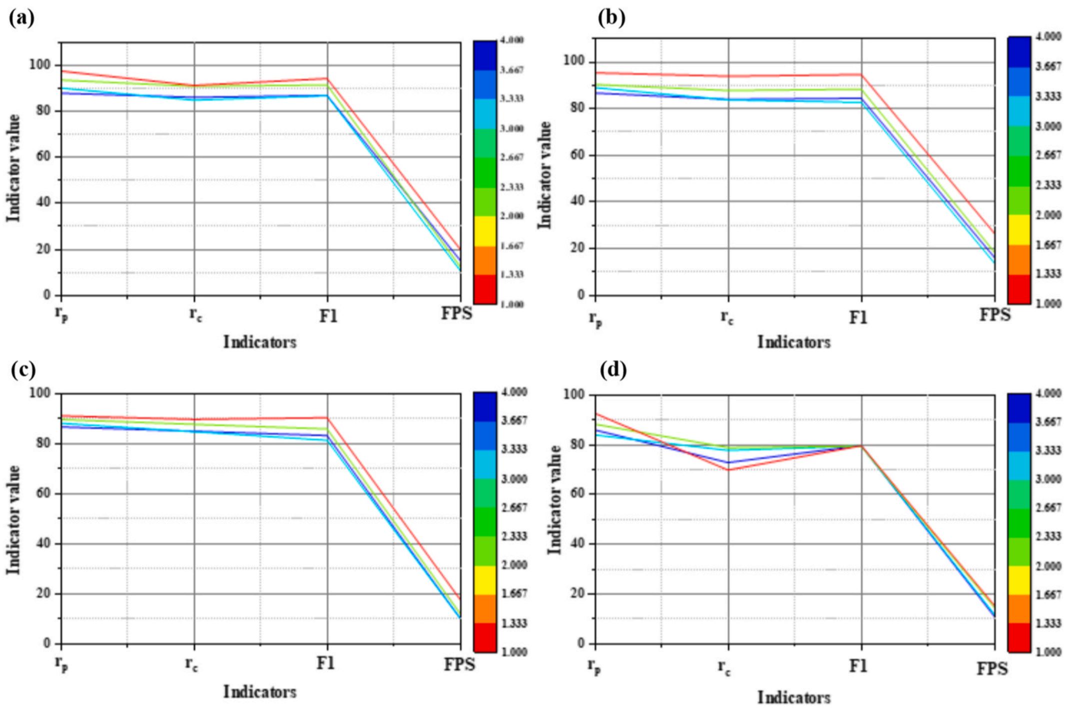


Fig. 7. Badminton tracking results in different scenarios using the YOLO series (a. Orange wall scenario; b. Green wall scenario; c. Laboratory scenario; d. Parking lot scenario).

Fig. 7 displays a comprehensive comparison of the results across the four mentioned scenarios. The improved Tiny YOLOv2 shows impressive performance in different scenarios. In the Orange Wall and Green Wall scenarios, it outperforms other networks significantly in terms of precision, recall, F1 score, and FPS, highlighting its robust capabilities in complex backgrounds. In the Laboratory scene, despite the lower FPS, the enhanced precision and recall still positively impact trajectory recognition performance. In the Parking Lot scenario, while the recall rate is lower, the relatively higher precision and well-balanced F1 score suggest that the improved Tiny YOLOv2 maintains a certain advantage in complex scenarios. In summary, the improved Tiny YOLOv2 exhibits favorable trajectory recognition performance across multiple scenes, displaying high precision and recall rates. Additionally, it achieves higher FPS in some scenarios. These results indicate the significant advantages and potential applications of the proposed approach in the domain of badminton trajectory recognition.

4.2. Comparison of trajectory prediction results

Fig. 8 depicts the badminton trajectory prediction results for different scenarios obtained with varying numbers of fitting points. The average predicted position error and the maximum predicted position error measure the accuracy of the predictions. The average predicted distance error and the maximum predicted distance error take into consideration the actual physical distances. Fig. 8 illustrates that the proposed method achieves fairly accurate prediction results across various scenarios, with an average predicted position error of around 3.76 pixels and an average predicted distance error of about 0.26 m. The maximum errors are around 8.80 pixels and 0.60 m, indicating that this method performs well in predicting badminton trajectories. Fig. 9 depicts the time consumption at different stages of the badminton trajectory prediction process across different scenarios.

Data loading, detection and feature extraction, and trajectory prediction represent the three main stages of the entire prediction process. The total time consumption is the cumulative value of these stages. Fig. 9 displays the time consumption variation for each stage across different scenarios. However, the average value is around 35.03 ms, indicating that this method has achieved a certain level of real-time performance.

It is evident that across various background scenarios, the proposed method has achieved significant improvements in both accuracy and robustness. The enhanced Tiny YOLOv2 network accurately detects and tracks badminton, leading to more precise trajectory prediction. Moreover, the introduced attention mechanism directs the network's focus towards crucial information about badminton, further enhancing trajectory prediction performance. These outcomes validate the efficacy of this approach in the domain of badminton trajectory prediction. Additionally, the proposed method excels across different background scenarios. The proposed method consistently achieves accurate trajectory prediction, whether it is the Orange Wall, Green Wall, laboratory, parking lot, or park scenarios. This highlights the method's adaptability to multiple scenarios, offering a reliable solution for diverse environments in practical applications.

4.3. Discussion

In the field of badminton trajectory prediction, various methods have been proposed to enhance the accuracy of trajectory recognition. This work compares existing methods and discusses their advantages and limitations. First, Li et al. (2022) proposed a badminton trajectory prediction method based on CNN, which improved trajectory precision through feature extraction and object detection techniques [40]. Their method leveraged deep convolutional networks to extract key features of badminton's movement, but accuracy decreased under complex backgrounds and lighting variations. In contrast, the improved Tiny YOLOv2 network proposed,

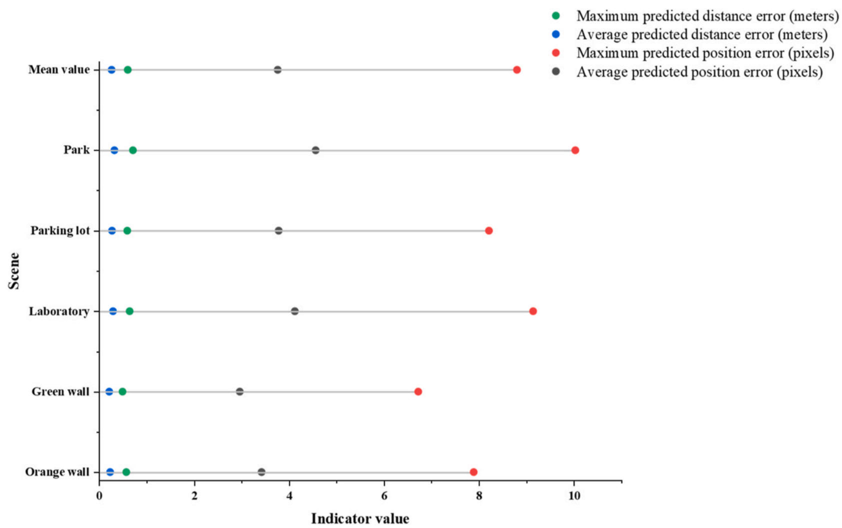


Fig. 8. Badminton trajectory prediction results of the proposed algorithm in different scenarios.

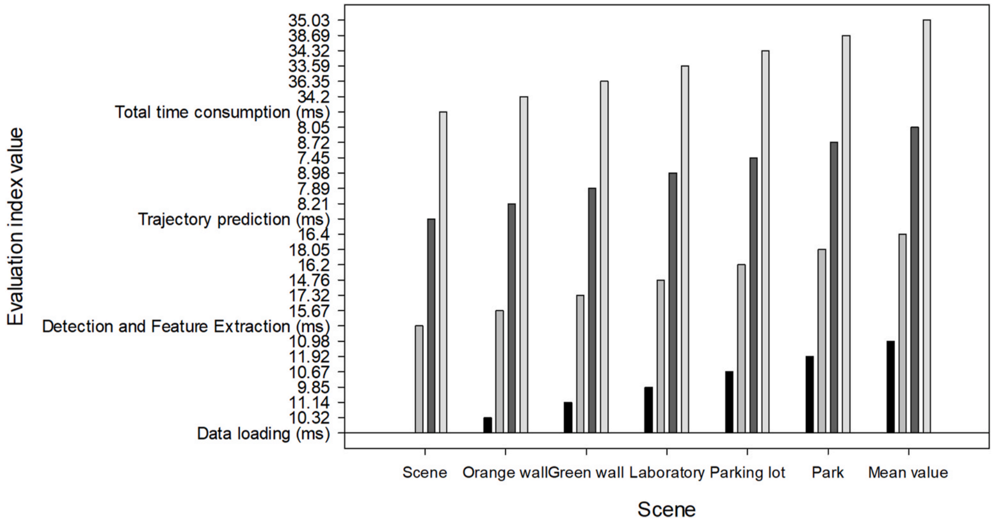


Fig. 9. Time consumption of badminton trajectory prediction in different scenarios.

combined with an enhanced loss function and network structure, significantly improves accuracy and robustness in complex environments. Notably, the introduced attention mechanism allows the model to better focus on critical areas of the badminton, resulting in an average precision of 91.40 %. This improvement markedly outperforms Li et al.'s results in complex scenarios. Moreover, Nokihara et al. (2023) utilized time series models for badminton trajectory recognition, and combined object detection and tracking algorithms to enhance overall system performance [41]. However, their method showed certain limitations when dealing with small targets in dynamic environments, especially in high-speed motion or occlusion scenarios. The approach proposed here overcomes these limitations by combining the improved Tiny YOLOv2 network with the UKF algorithm. In four different scenarios, the proposed method outperforms Li et al.'s research in detection and prediction accuracy, particularly in precise tracking of small targets. Lastly, Zhi et al. (2022) introduced an adaptive filtering-based trajectory prediction method, adjusting filter parameters in real-time to adapt to environmental changes [42]. However, their method still depended on filter parameter settings, affecting prediction accuracy under varying background conditions. The method proposed here, through the improved Tiny YOLOv2 and attention mechanism, reduces dependence on environmental parameters, enhancing prediction stability and accuracy. In simulation experiments, this approach achieves an average position prediction error of 3.76 pixels across different backgrounds, significantly lower than Wang et al.'s results. Overall, the proposed method demonstrates significant advantages in badminton trajectory recognition and prediction. By improving the Tiny YOLOv2 network structure and incorporating attention mechanisms, the model achieves high-precision trajectory recognition and prediction across various complex scenes, and excels in tracking small targets in dynamic environments. This result validates the effectiveness of the proposed method in the field of badminton trajectory recognition and showcases its potential in practical applications.

Despite the significant advantages of the proposed method, it does have some potential limitations. These limitations might affect the method's effectiveness and applicability in different scenarios. First, the improved Tiny YOLOv2 network combined with attention mechanisms better focuses on the shuttlecock's critical areas. However, in scenes with complex backgrounds and significant lighting changes, accuracy may decline if the background changes drastically or lighting conditions are extremely poor. Especially in outdoor scenes with frequent lighting changes, despite the partial mitigation provided by the attention mechanism, recognition errors may still occur. Besides, while the method excels in handling small targets in dynamic environments, its accuracy might be impacted in cases of high-speed motion or severe occlusion. For example, if the shuttlecock is occluded for a period during high-speed movement, although the UKF algorithm can predict and track to some extent, long-term occlusion may still lead to increased prediction errors. Moreover, although the proposed method demonstrates high accuracy and robustness across four different scenes, the varying complexity and characteristics of real-world scenarios may require further fine-tuning and optimization of the model for specific contexts. For instance, in laboratory settings, complex structures and variable lighting might necessitate more detailed preprocessing steps to enhance the model's robustness and accuracy. Finally, the proposed method relies on high-quality video data and substantial computational resources. For applications requiring high real-time performance, such as instant analysis during badminton matches, the proposed method achieves a degree of real-time capability. However, in practical applications, delays in data processing and model inference may still impact the effectiveness of real-time analysis. In summary, while the proposed method shows significant advantages in badminton trajectory recognition and prediction, and achieves good results in various complex scenarios, potential limitations in different application contexts should be considered. Future research could focus on further optimizing model structures and algorithms to improve adaptability and robustness in complex environments, and explore efficient implementations in resource-constrained settings to expand practical applications.

5. Conclusion

This work introduces an improved Tiny YOLOv2 network into the domain of badminton movement trajectory recognition, taking into full consideration the small target characteristics of badminton within video streams. Optimizing the network structure and loss function enhances the accuracy and robustness of badminton detection and tracking. Additionally, incorporating an attention mechanism heightens the network's focus on crucial badminton information, thereby further elevating trajectory recognition performance. The attention mechanism plays a crucial role in capturing essential features of badminton movement trajectories. This work combines the UKF algorithm to precisely predict and track badminton trajectories. UKF demonstrates excellent performance in integrating motion models and observation data, providing effective support for trajectory prediction. However, due to the utilization of complex algorithms such as DL networks and Kalman filtering, the computational complexity is relatively high, potentially posing challenges in terms of real-time performance. Further algorithm optimization and hardware support are needed to enhance real-time capabilities. Future research could consider the integration of multimodal information, such as sound and depth data, to further improve the performance and robustness of badminton trajectory recognition.

Data availability statements

All data generated or analysed during this study are included in this published article [and its supplementary information files].

CRediT authorship contribution statement

Chuanbao He: Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Min Zhang:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e38865.

References

[1] J. Xie, G. Chen, S. Liu, Intelligent badminton training robot in athlete injury prevention under machine learning, Front. Neurorob. 15 (2021) 621196.
[2] K.C. Lin, C.W. Wei, C.L. Lai, et al., Development of a badminton teaching system with wearable technology for improving students' badminton doubles skills, Educ. Technol. Res. Dev. 69 (2021) 945–969.
[3] G. Jiang, A study on the construction of the evaluation system of the teaching ability of students using pattern recognition for studying majoring in badminton in the mixed learning model of physical education majors and self-learning system, ACM Transactions on Asian and Low-Resource Language Information Processing 21 (6) (2022) 1–18.
[4] N. Gao, Construction and simulation of athlete's wrong action recognition model in sports training based on embedded wireless communication and computer vision, Wireless Commun. Mobile Comput. 2021 (2021) 1–11.
[5] I. Ghosh, S.R. Ramamurthy, A. Chakma, et al., Decoach: deep learning-based coaching for badminton player assessment, Pervasive Mob. Comput. 83 (2022) 101608.
[6] M.F. Tsai, S.H. Huang, Enhancing accuracy of human action Recognition System using Skeleton Point correction method, Multimed. Tool. Appl. 81 (5) (2022) 7439–7459.
[7] R. Gimazutdinov, Using foreign experiences in getting primary class pupils interested in badminton, Sci. Innovat. 2 (B3) (2023) 214–217.
[8] Y. Lyu, S. Zhang, Badminton path tracking algorithm based on computer vision and ball speed analysis, J. Sens. 2021 (2021) 1–13.
[9] B.T. Naik, M.F. Hashmi, N.D. Bokde, A comprehensive review of computer vision in sports: open issues, future trends and research directions, Appl. Sci. 12 (9) (2022) 4429.
[10] Z.Y. Yip, I.M. Khairuddin, W.H.M. Isa, et al., Badminton smashing recognition through video performance by using deep learning, MEKATRONIKA 4 (1) (2022) 70–79.
[11] Y.L. Chen, J.H. Hsu, D.H.L. Tai, et al., Training-associated superior visuomotor integration performance in elite badminton players after adjusting for cardiovascular fitness, Int. J. Environ. Res. Publ. Health 19 (1) (2022) 468.
[12] F. Liang, Auxiliary teaching of badminton basic movements based on wireless network communication and kinect, J. Sens. 2022 (2022) 1–12.
[13] J. Jaworski, G. Lech, M. Żak, et al., Relationships between selected indices of postural stability and sports performance in elite badminton players: pilot study, Front. Psychol. 14 (2023) 1019.
[14] X. Sha, G. Wei, X. Zhang, et al., Accurate recognition of player identity and stroke performance in table tennis using a smart wristband, IEEE Sensor. J. 21 (9) (2021) 10923–10932.
[15] J. Chen, R.D.J. Samuel, P. Poovendran, LSTM with bio inspired algorithm for action recognition in sports videos, Image Vis Comput. 112 (2021) 104214.
[16] G. Zhu, P. Ren, J. Hu, et al., Flexible and anisotropic strain sensors with the asymmetrical cross-conducting network for versatile bio-mechanical signal recognition, ACS Appl. Mater. Interfaces 13 (37) (2021) 44925–44934.

[17] Y. Zhang, J. Tang, H. Huang, Motion capture and intelligent correction method of badminton movement based on machine vision, Mobile Inf. Syst. 2021 (2021) 1–10.
[18] M.A. Haq, N. Tagawa, Improving badminton player detection using YOLOv3 with different training heuristic, JOIV: International Journal on Informatics Visualization 7 (2) (2023) 548–554.
[19] A. Wang, X. Yi, Attitude Perception of Badminton Players Based on Mobile Edge Computing, vol. 2021, Scientific Programming, 2021, pp. 1–12.
[20] L. Zhu, A prediction method for the service trajectory of badminton moving video based on fuzzy clustering algorithm, Int. J. Innovat. Comput. Appl. 12 (4) (2021) 216–223.
[21] L. Liu, Objects detection toward complicated high remote basketball sports by leveraging deep CNN architecture, Future Generat. Comput. Syst. 119 (2021) 31–36.
[22] M. Kopania, J. Nowisz, A. Przelaskowski, Automatic shuttlecock fall detection system in or out of a court in badminton games—challenges, problems, and solutions from a practical point of view, Sensors 22 (21) (2022) 8098.
[23] J.J. Rubio, Y. Pan, J. Pieper, et al., Advances in robots trajectories learning via fast neural networks, Front. Neurorob. 15 (2021) 671519.
[24] S. Siyal, M. Saeed, M.H. Pahi, et al., They can't treat you well under abusive supervision: investigating the impact of job satisfaction and extrinsic motivation on healthcare employees, Ration. Soc. 33 (4) (2021) 401–423.
[25] D. Chen, P. Wawrzynski, Z. Lv, Cyber security in smart cities: a review of deep learning-based applications and case studies, Sustain. Cities Soc. 66 (2021) 102655.
[26] L. Liu, Arm movement recognition of badminton players in the third hit based on visual search, Int. J. Biometrics. 14 (3–4) (2022) 239–252.
[27] T. Steels, B. Van Herbruggen, J. Fontaine, et al., Badminton activity recognition using accelerometer data, Sensors 20 (17) (2020) 4685.
[28] L. Fang, M. Sun, Motion recognition technology of badminton players in sports video images, Future Generat. Comput. Syst. 124 (2021) 381–389.
[29] Z. Cui, T. Zheng, Badminton recognition and tracking system, Int. J. Pattern Recogn. Artif. Intell. 35 (14) (2021) 2150039.
[30] Y. Chen, H. Hu, Design and research of the AI badminton model based on the deep learning neural network, J. Math. 2022 (2022) 1–10.
[31] K.C. Lin, H.C. Hung, N.S. Chen, The effect of wearable technology on badminton learning performance: a multiple feedback WISER model in physical education, Smart Learning Environments 10 (1) (2023) 28.
[32] M.J. Duncan, A. AlShabeb, K. Fitton Davies, et al., A 6-week badminton-based movement intervention enhances fundamental movement skills and physical fitness in Saudi boys and girls, Sports 11 (7) (2023) 132.
[33] J. Zhan, Y. Luo, C. Guo, et al., Anchor-free multi-task learning network for panoptic driving perception, Pattern Recogn 148 (2024) 110152.
[34] G. Yogeswararao, V. Naresh, R. Malmathanraj, et al., An efficient densely connected convolutional neural network for identification of plant diseases, Multimed. Tool. Appl. 81 (23) (2022) 32791–32816.
[35] M. Shafiq, Z. Gu, Deep residual learning for image recognition: a survey, Appl. Sci. 12 (18) (2022) 8972.
[36] X. Liu, C. Wen, X. Sun, Design method of high-order kalman filter for strong nonlinear system based on kronecker product transform, Sensors 22 (2) (2022) 653.
[37] T. Diwan, G. Anirudh, J.V. Tembhurne, Object detection using YOLO: challenges, architectural successors, datasets and applications, Multimed. Tool. Appl. 82 (6) (2023) 9243–9275.
[38] K. Ampountolas, The unscented kalman filter for nonlinear parameter identification of adaptive cruise control systems, IEEE Transactions on Intelligent Vehicles 8 (8) (2023) 4094–4104.
[39] G. Rigatos, Nonlinear optimal control for the translational oscillator with rotational actuator, Cybern. Syst. 53 (6) (2022) 541–562.
[40] H. Li, S.G. Ali, J. Zhang, et al., Video-based table tennis tracking and trajectory prediction using convolutional neural networks, Fractals 30 (5) (2022) 2240156.
[41] Y. Nokihara, R. Hachiuma, R. Hori, et al., Future prediction of shuttlecock trajectory in badminton using player's information, Journal of Imaging 9 (5) (2023) 99.
[42] J. Zhi, D. Luo, K. Li, et al., A novel method of shuttlecock trajectory tracking and prediction for a badminton robot, Robotica 40 (6) (2022) 1682–1694.