

Retraction

Retracted: Multitarget Tracking Algorithm in Intelligent Analysis of Football Movement Training Stance

Security and Communication Networks

Received 13 September 2023; Accepted 13 September 2023; Published 14 September 2023

Copyright © 2023 Security and Communication Networks. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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

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

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

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

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

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

References

- [1] C. Li and Q. Peng, "Multitarget Tracking Algorithm in Intelligent Analysis of Football Movement Training Stance," *Security and Communication Networks*, vol. 2022, Article ID 6579066, 8 pages, 2022.

Research Article

Multitarget Tracking Algorithm in Intelligent Analysis of Football Movement Training Stance

Changrui Li¹ and Qiuping Peng² 

¹School of Physical Education and Health, Nanning Normal University, Nanning 530001, China

²School of Physical Education and Health Science, Guangxi University for Nationalities, Nanning 530006, China

Correspondence should be addressed to Qiuping Peng; pqp2022@126.com

Received 8 April 2022; Revised 16 June 2022; Accepted 3 July 2022; Published 4 August 2022

Academic Editor: Fang Liu

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

In recent years, with the continuous development of computer technology, deep learning has been widely applied to computer vision tasks and has achieved great success in areas such as visual detection and tracking. On this basis, making deep learning techniques truly accessible to people becomes the next objective. Target detection and tracking in football gesture training is a quite challenging task with great practical and commercial value. In traditional football training methods, target trajectories are often extracted by means of a recording chip carried by the player. However, the cost of this method is high and it is difficult to replicate in amateur stadiums. Some studies have also used only cameras to process targets in football videos. However, due to the similarity in appearance and frequent occlusion of targets in football videos, these methods often only segment targets such as players and balls in the image but do not allow them to be tracked. Target tracking techniques are of great importance in football training and are the basis for tasks such as player training analysis and match strategy development. In recent years, many excellent algorithms have emerged in the field of target tracking, mainly in the categories of correlation filtering and deep learning, but none of them are able to achieve high accuracy in player tracking for football training videos. After all, the problem of locating clips of interest to athletes from a full-length video is a pressing one. Traditional machine learning-based approaches to sports event detection have poor accuracy and are limited in the types of events they can detect. These traditional methods often rely on auxiliary information such as audio commentary and relevant text, which are less stable than video. In recent years, deep learning-based methods have made great progress in the detection of single-player video events and actions, but less so in the detection of sports video events. As a result, there are few sports video datasets that can be used for deep learning training. Based on research in computer vision and deep learning, this paper designs a multitarget tracking system for football training. To be specific, this algorithm uses multiple cameras for image acquisition in the stadium in order to accurately track multiple targets in the stadium over time. Furthermore, the framework for a single camera multitarget tracking approach has been designed based on deep learning-based visual detection methods and correlation filter-based tracking methods. This framework focuses on using data correlation algorithms to fuse the results of detectors and trackers so that multiple targets can be tracked accurately in a single camera. To sum up, this research allows for robust and real-time long-term accurate tracking of targets in football training videos through multitarget tracking algorithms and the intercorrection of multiple camera systems.

1. Introduction

The vast majority of information that humans obtain from the outside world comes from vision. As a result, vision becomes the most important form of information acquisition for humans [1]. In recent years, with the rapid development of computers and information technology, people

are increasingly using computers to participate in everyday production and life [2, 3]. To be specific, computers can significantly extend human brainpower and perception, and the use of computers to simulate human visual perception has led to the creation of computer vision [4]. Computer vision is the science of acquiring and processing visual information through cameras and computers, with the

ultimate goal of enabling computers to perform some of the functions of human vision in order to understand the three-dimensional world [5]. What is more, computer vision is a multidisciplinary intersection and combination of mathematics, image processing, biology, computer science, and so on [6, 7]. As one of the most important tasks in computer vision, target tracking technology has become a hot research topic in recent years due to its promising applications and market demand. The goal of computer vision research is to enable computers to adapt to their environment autonomously and to observe and understand the world as humans do [8]. With the development of deep learning theory in recent years, computer vision has made promising advances in a number of tasks such as classification, detection, recognition, and tracking. Target tracking is the process of tracking an object of interest in the first frame of a video sequence to obtain the motion parameters associated with the target, including position, velocity, and trajectory [9]. The object is then recognised and tracked and its behaviour understood to perform subsequent higher level vision tasks. Target tracking technology has been widely used in many fields such as unmanned vehicles [10], intelligent video surveillance [11], smart construction [12], military guidance [13], and smart medical care [14].

With the rapid growth and distribution of videos on the Internet, there is an urgent need to find the video clips that people need quickly. Sports videos are a very important part of our daily lives and play a large part in people's entertainment [15]. Football is one of the most popular sports around the world, and videos of popular football matches are widely distributed and watched. For fans, they want to be able to watch the full match directly from the event they are interested in. For video editors, they want to be able to quickly organise the video content they need from a large number of sources, or even automatically generate video summaries or highlight reels [16]. What is more, for athletes and coaches alike, the analysis of players' movement data from video footage of football matches is one of the most important tasks in order to train better [17]. After all, the coaches need some of the motion data in the video for their analysis and the players themselves want to be able to record their own motion data in real time, which needs to be extracted from the target motion traces in the video. For professional football, these needs can be met, as this is only possible with the support of costly human and material resources [18]. For amateur football matches, which are much larger in scale, a low-cost target detection and tracking solution is of great practical value and relevance. As a result, event detection for football training videos has been an important topic in multimedia research.

Target detection, which focuses on determining the size and location of a target in a video image, is a core problem in machine vision [19]. However, target detection has always been one of the most challenging problems in machine vision due to a number of factors such as variations in illumination, object occlusion, and complex background conditions. In recent years, target detection algorithms have become a hot issue in the field of machine vision [20]. Many efficient algorithms such as twin algorithms (Figure 1) have

emerged and are widely used in target detection and tracking. The target tracking technology plays an essential role in football video analysis [21]. As the most watched sport in the world, football is a rich and widespread sport with huge commercial value. When watching football videos, different actors have different requirements for football videos. The average viewer tends to focus on a particular player of interest when watching a video of a football match. As a result, there is a need to provide viewers with high-level semantic analysis of video summaries, highlights, player movement recognition, and so on. The target tracking technology is the basis for these high-level semantic functions [22]. For professionals such as team coaches, there is a need for detailed information on the field of play to analyse matches and develop training programmes and game strategies, such as player trajectories, distances travelled, running speeds, and other movement parameters. Then the target tracking technology can provide this data directly [23]. In addition, for the referee on the field of play, there is a need for a variety of auxiliary information to ensure fair play, such as precise positioning of players, football trajectory analysis, and foul play recognition, in order to avoid potentially controversial calls during the game due to fierce competition. To sum up, researching and analyzing player tracking in football video is a fundamental research task for many practical applications and has important theoretical and practical value.

With the rise of deep learning theory in recent years, deep learning-based target tracking algorithms have been gaining prominence in the relevant industry in recent years [24]. However, the application of deep learning to target tracking has not been smooth. Unlike vision fields such as classification and recognition, the main problem with deep learning in target tracking is the lack of training data. Deep models can perform well by learning effectively from large amounts of labelled data, whereas tracking tasks only provide the first labelled image as training data [25]. In this case, it is quite difficult to train a deep model from scratch. The main idea of current deep learning-based tracking algorithms is to pretrain the depth model offline using the marker data and finetune the pretrained tracker online with the current samples during tracking. This migration learning approach will reduce the need for training samples and improves the feasibility of deep learning-based tracking algorithms [26]. Target detection is one of the main areas of research in computer vision and is widely used in areas such as video surveillance, face recognition, defect detection, and robot navigation. The goal of target detection is to find the exact location of the target of interest in the image and to give an accurate classification of the target. Deep learning is a data-driven approach that learns the features of a target from a large amount of data, and the features obtained from deep learning are often better than those designed by hand [27]. In the field of image detection, deep learning has achieved the best results to date, and the reason for this is that the large amount of training data allows the network to learn abstract, high-level features that are highly expressive of the target.

In fact, the target tracking technology is based on image processing techniques and aims to obtain the true position of

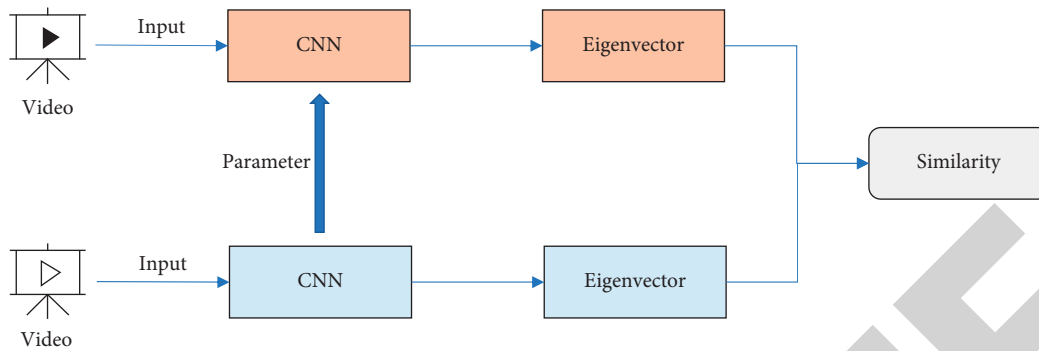


FIGURE 1: Basic framework of twin algorithm.

a target from a video sequence and to analyse it [28]. However, there are many difficulties in the practical application of video target tracking. Firstly, the similarity between the target and the background makes it difficult to capture the differences between the two. Secondly, the appearance of the target changes continuously over time. On the one hand, the object itself changes in shape, which is particularly noticeable in long-term tracking. On the other hand, external conditions such as lighting have an effect on the target. In addition to this, the position of the target in the video is constantly changing, and occlusion may occur. What is more, tracking has to balance the need for accuracy in target location with the need for real-time performance. As a result, there is a wide range of video target tracking algorithms available, but they lack generality. As shown in Figure 2, target tracking algorithms can be divided into different categories depending on the use of target tracking information.

The target tracking algorithm based on contrast analysis uses the difference between the target and the background to track the target and can be classified as form-focused tracking, centre-of-mass tracking, and edge tracking depending on the tracking reference point [29]. The algorithm does not work well in complex environments but is quite effective in situations where the difference between target and background is significant. What is more, the algorithm is computationally simple, responsive, and, in some cases, has great tracking accuracy. However, it is susceptible to external disturbances, has high random errors, and is poor for violent movements and occlusions. Matching-based tracking algorithms can distinguish between the attributes of a target and other things [30]. To be specific, a matching-based tracking algorithm begins by extracting the features of the target and matching them in each frame of the video sequence. The main features used in target tracking are feature points, contours, edges, colors, textures, etc. The main idea of motion detection-based tracking algorithms is to detect the difference between the motion of the target and the background in order to determine the position of the target, thus enabling the tracking of the target. This algorithm highlights the difference in time or space between the motion of the target and the background, avoiding the need for complex modelling of the target, and is now increasingly used. Finally, the detection-based tracking algorithm can essentially transform the

tracking problem into a binary classification problem, i.e., distinguishing between the target and the background [31]. These methods often train a detector that detects the location of the target frame by frame in the whole image and then connects it to the complete motion trajectory of the target. The biggest challenge for detection-based tracking algorithms is to train an effective detector and to ensure that the algorithm is real-time.

In recent years, deep learning techniques have been widely applied to computer vision tasks, with great success in areas such as vision detection and tracking [32]. Building on this foundation, the implementation of deep learning techniques is the next step and it even can be used in construction [33]. However, there are relatively few datasets for temporally bounded tagging of events in video. Some of the existing event annotation datasets are mostly mixed event types, and there are no datasets specifically for football video events. For deep learning methods, adequate labelled datasets are an important prerequisite for the development of related techniques. This study is based on football video, in which all players are tracked and the trajectory of each player in a normal match is extracted. Previously, this task has been carried out using a chip attached to the player to record the player's movement, but this is expensive and does not allow for real-time output of player movement data. The hardware in this research is able to accurately extract the player's entire movement data using only a camera. It can save costs, improve the user experience, and can be used in real life scenarios, which can become great reference for the football players as well as coaches.

2. Module Analysis

The video analysis system for target detection and tracking in football training is divided into four main modules: image acquisition module, storage system module, target detection, and tracking module, as well as VGA display module. The overall system block diagram is shown in Figure 3.

The main function of this system is to detect moving objects in the video image by means of an algorithm. In addition to this, the moving object can be extracted from the background image and then the background can be separated from the moving object according to the motion and stillness, and then the target can be tracked by the algorithm.

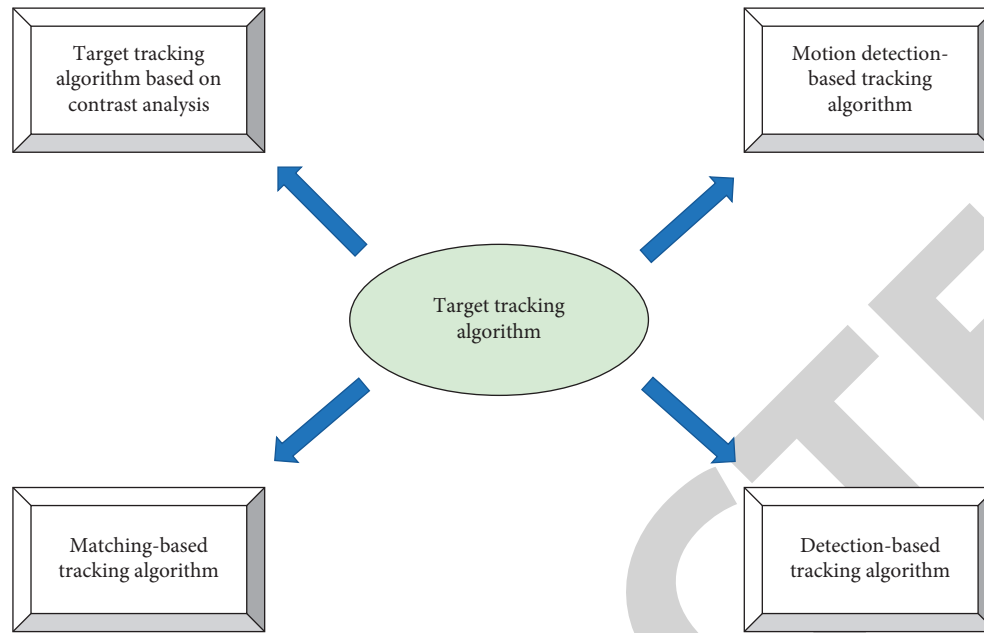


FIGURE 2: Various target tracking algorithms.

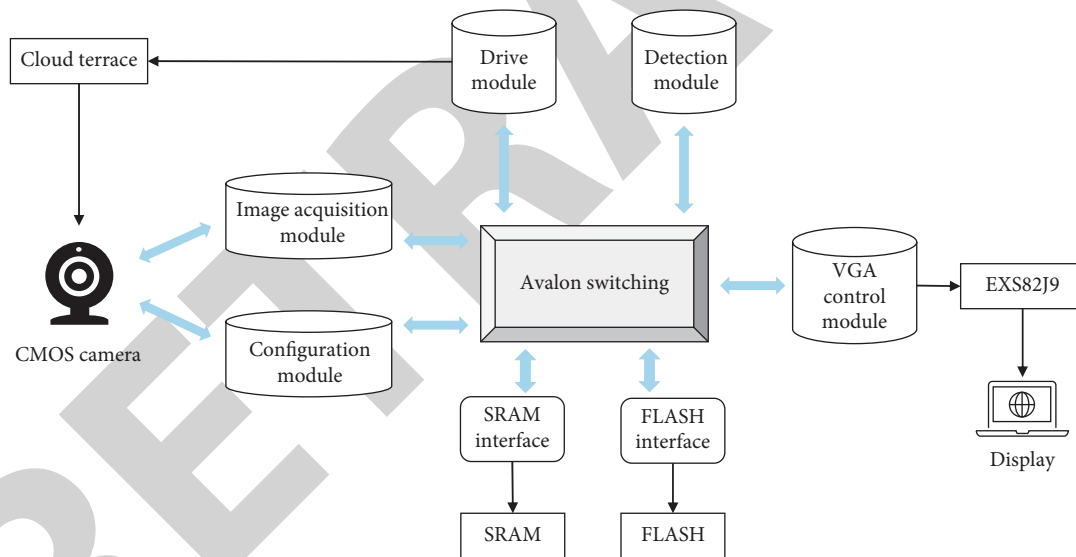


FIGURE 3: Overall system block diagram.

2.1. Image Acquisition Module. The video image capture module uses a CMOS video capture camera, the core of which is an internal large scale integrated circuit chip, a RAM chip that can be read and written. The core part of the video capture camera is the chip, which converts each pixel of the captured image into its own charge and voltage. The basic unit of the CMOS digital integrated circuit is a voltage-controlled amplifier with the advantages of high interference immunity and low-static power consumption.

2.2. Storage System Module. The storage module mainly applies a synchronous dynamic random storage memory as a picture memory. This memory receives a clock signal

before it receives a response, enabling it to be synchronised with the computer system bus.

2.3. Target Detection and Tracking Module. The internal structure of the FPGA chip consists of seven modules: programmable input and output units, configurable logic modules, digital clock management modules, embedded modules, wiring resources, underlying embedded functional units, and embedded dedicated cores. When the chip is powered up, the data is read into the on-chip programmable RAM memory and when the match is completed, the chip starts to operate and the logic inside returns to its original state when the power is removed. The chip can therefore be

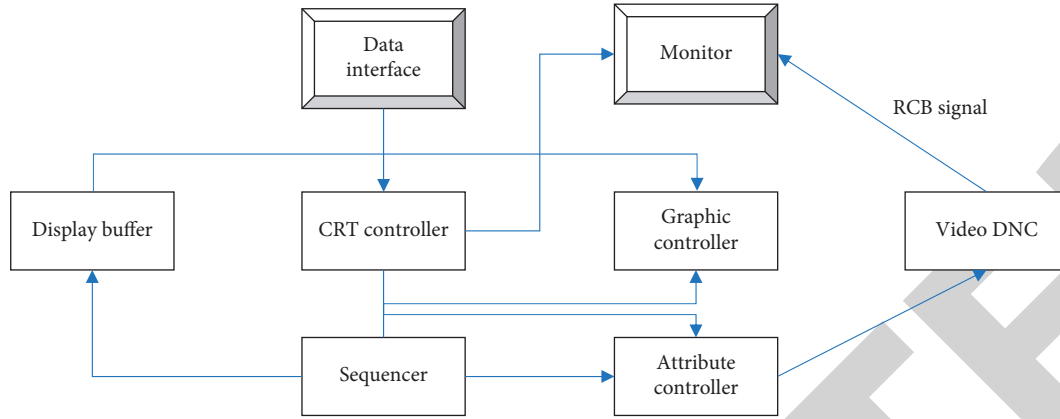


FIGURE 4: Control circuit of the VGA display module.

used repeatedly. When you want to change the function of the circuit, you can simply change the EPROM chip and write a different programming language, thus enabling different functions to be changed.

2.4. VGA Display Module. The VGA display system consists of three main modules: the control circuit, the display cache, and the video BIOS program. The control circuit is shown in Figure 4.

The control circuit is responsible for the chronological occurrence, display buffer data manipulation, master clock selection, and other functions. The display module enables the synchronisation of the display data update and the display.

3. Multitarget Tracking Algorithm

3.1. Feature Selection. The first step in the tracking task is to take the target features. To enhance the feature representation, conventional features and depth features are extracted separately. In terms of feature processing, discrete features are transformed into continuous features to enhance the feature representation and factorization is used to reduce the computational complexity. The output is corrected for tracking drift using prediction speed. Target feature extraction is an important part of the visual tracking task and the quality of the features directly affects the final tracking result. Target features can be divided into traditional manual features and depth features. Traditional manual features are an older form of feature representation and are usually represented using both global and local features, including colour and texture features. As one of the most widely used features, colour features are robust to angle changes and pose changes. Unlike the colour feature, the texture feature is not a pixel point-based feature. In particular, it requires a statistical calculation over a region containing multiple pixel points. Texture features have good nondeformation to illumination. In fact, conventional features rely heavily on the representation of the appearance of the object. However, due to the human factor of photography and various environmental factors, the actual images

obtained often show large variations in appearance. As a result, progress in traditional manual characterisation has been slow and it has been difficult to achieve significant breakthroughs in accuracy.

The football video tracking task mainly focuses on tracking the target player in the long shot. Firstly, the position of the target player is obtained in the initial frame and the position of the player is determined in the following frames. The rectangular frame corresponding to the player's position is denoted as (a, b, k, j) , where (a, b) represents the horizontal and vertical coordinates of the upper left corner of the rectangular frame, k represents the width of the rectangular frame, and j represents the height of the rectangular frame. The player's position is given in the initial frame as the target frame, and the images in the target frame are extracted and the tracker model is built. In subsequent frames, the images in the candidate frame are analysed and processed to determine the player's location. The candidate frame is the area where the target player is likely to be located, which is the search area for the tracking algorithm, as shown in Figure 5.

The candidate box area does not need to be the whole picture as the position of the player does not change much between the two frames. The candidate box is a square box centred on the target positioned in the previous frame, with sides as shown in the following equation:

$$L = \sqrt{[(\theta \times k) \times (\theta \times j)]}, \quad (1)$$

where L refers to the side of the candidate box and θ refers to the proportion of candidate areas.

3.2. Feature Extraction. Depth features and FHOG features are extracted separately from the feature representation of the target, and these two features are then fused to obtain the final feature representation. The depth model uses a pre-trained model trained offline on the ImageNet classification task and finetuned online during the target tracking phase. As the tracking task in this paper only requires feature extraction on the VGG model and not target classification, only the first 14 convolutional layers of the VGG model are



FIGURE 5: Candidate frame of feature selection.

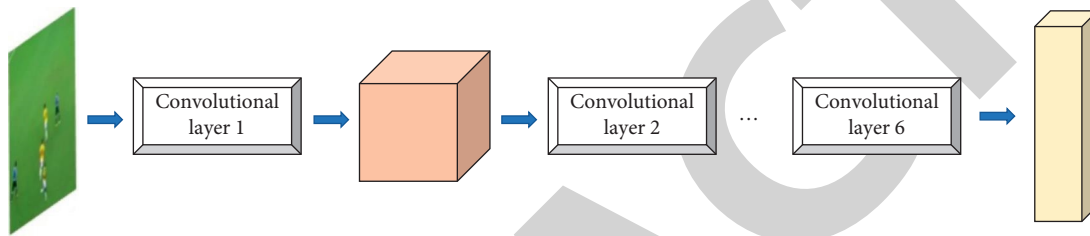


FIGURE 6: Feature output size.

used. The network model consists of 6 convolutional groups, with two to three convolutional layers within each convolutional group. The input to the network model is $64 \times 64 \times 3$ and the outputs of layers 3 and 7 are extracted as shallow and deep features, respectively. The feature output size of layer 3 is $16 \times 16 \times 68$ and that of layer 7 is $9 \times 9 \times 212$, as shown in Figure 6.

3.3. Feature Continuity. To address the need to improve the accuracy of the tracking results, an extension of the discrete features to a continuous approach is devised. This approach can not only improve the expressiveness of the features but also can result in a continuous response map. At the same time, precise subpixel positions can be achieved during target localization, thus providing further improvements in tracking accuracy. The conversion of discrete features to continuous features naturally involves interpolation, as shown in Figure 7. Three spline interpolations are performed on the extracted features. As the tracking is performed in the frequency domain, the feature interpolation operation is also an interpolation process in the frequency domain by first transferring the features from the spatial domain to the frequency domain via a Fourier transformation.

The basic process of cubic spline interpolation in one dimension is to fit a continuous function with a segmented set of cubic equations. The solution of this set of cubic equations requires four constraints, namely equal function values at the nodes, equal first-order derivatives at the nodes, equal second-order derivatives at the nodes, and zero second-order derivatives at the endpoints, to be solved,

respectively. To be specific, the implicit representation can be expressed by the following equation:

$$f(x, v) = \sum_{i=0}^{I-1} x(m) \times n \times \left(v - \frac{V}{I} \times i \right), \quad (2)$$

where v, i refers to the variables in the continuous and discrete domains, respectively, V, N indicates the boundary values in the continuous and discrete domains, respectively, $f(x, v)$ refers to the continuous function, x indicates the discrete function, and n is the interpolation function.

3.4. Multitarget Tracking. The multitarget tracker is based on a detector and a single-target tracker. The entire video is first decomposed into a number of tracking cycles, each containing a number of consecutive frames. In each tracking cycle, the first frame is detected and the next few frames are tracked. The results of the two adjacent tracking cycles are then fused by means of data correlation, as shown in Figure 8.

In each tracking cycle, the first frame will detect all the target positions in the image and then initialise the tracker with the initial position of each target. The tracker is then initialised with the initial position of each target. A short track is then performed over the remainder of the image sequence to obtain the trajectory of each target. The next step is to assign the corresponding numbers to all the tracks in the sequence, which is done by means of a data association algorithm. The data association algorithm matches the trajectories of each numbered target from the previous cycle with all the trajectories of the current cycle and then assigns the corresponding numbers to these trajectories and saves them.

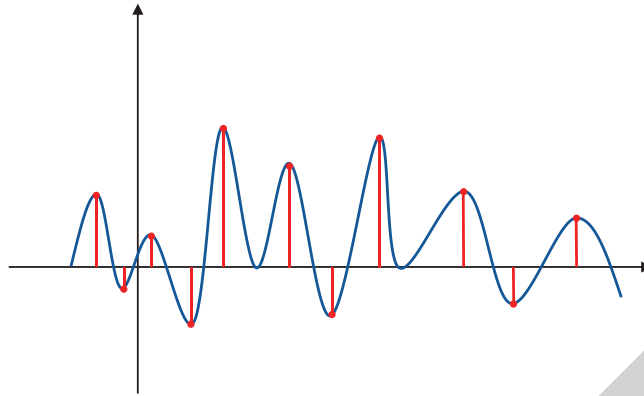


FIGURE 7: Conversion of discrete features to continuous features.

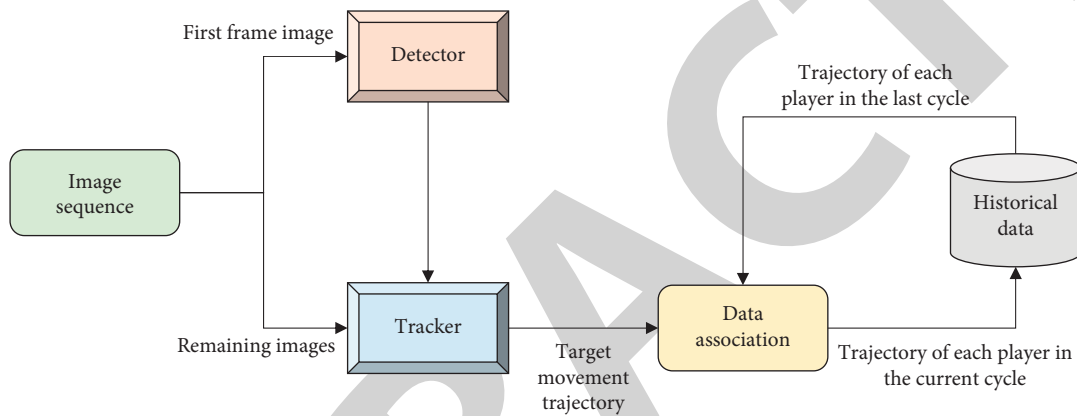


FIGURE 8: Framework of multitarget tracking algorithm.

4. Conclusion

In recent years, deep learning-based detection algorithms have been developed rapidly. On the basis of this, deep learning has continued to improve in terms of real time and accuracy. As a result, correlation filter-based trackers have been widely used due to their excellent tracking speed and accuracy. The target tracking technology plays an essential role in football match video analysis, with implications for player training, strategy development as well as football event detection. This research focuses on target detection and tracking algorithms in football video, and designs a multitarget tracking method based on a deep learning detector. First of all, this study designs a framework for multitarget tracking in football video and investigates the detection, tracking, and data association modules. This can enable the accurate tracking of multiple targets in a single camera. In the feature extraction process, a combination of traditional manual features and depth features was used to extract the FHOG features and the depth features of the VGG network model, respectively. The discrete features were transformed into continuous features in the frequency domain using cubic spline interpolation in order to improve the feature representation capability and the accuracy of the results. To speed up tracking and reduce the effects of overfitting, the dimensionality of the features is significantly reduced using factorization.

However, although the multitarget tracking algorithm designed in this research performs well in terms of accuracy, several problems still remain. Firstly, the predicted speed is not accurate enough. Although a moving target will maintain its inertia and move smoothly in a realistic scene, camera movement can cause the smoothness of the target's trajectory to be significantly reduced in the video frame. In fact, offsetting camera movement can significantly improve the accuracy of the predicted velocity. In addition, there is no dynamic adjustment for factorization dimensionality reduction. The dimensionality reduction matrix remains constant during the tracking process and cannot be targeted in subsequent features, which may reduce the accuracy of the tracking results to a certain extent.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The research was supported by the Education Research Project of Guangxi University for Nationalities (No. 2021XJGY22).

References

- [1] M. Deng and X. Gu, "Information acquisition, emotion experience and behaviour intention during online shopping: an eye-tracking study," *Behaviour & Information Technology*, vol. 40, no. 7, pp. 635–645, 2021.
- [2] B. Cheng, C. Fan, H. Fu, J. Huang, H. Chen, and X. Luo, "Measuring and computing cognitive statuses of construction workers based on electroencephalogram: a critical review," *IEEE Transactions on Computational Social Systems*, 2022.
- [3] B. Cheng, K. Lu, J. Li, H. Chen, X. Luo, and M. Shafique, "Comprehensive assessment of embodied environmental impacts of buildings using normalized environmental impact factors," *Journal of Cleaner Production*, vol. 334, Article ID 130083, 2022.
- [4] M. Ghobaei-Arani, A. Souri, and A. A. Rahmanian, "Resource management approaches in fog computing: a comprehensive review," *Journal of Grid Computing*, vol. 18, no. 1, pp. 1–42, 2020.
- [5] B. Zhang, L. Zhou, and J. Zhang, "A methodology for obtaining spatiotemporal information of the vehicles on bridges based on computer vision," *Computer-Aided Civil and Infrastructure Engineering*, vol. 34, no. 6, pp. 471–487, 2019.
- [6] B. G. Weinstein, "A computer vision for animal ecology," *Journal of Animal Ecology*, vol. 87, no. 3, pp. 533–545, 2018.
- [7] Y. Qian, S. Chen, J. Li et al., "A decision-making model using machine learning for improving dispatching efficiency in Chengdu Shuangliu airport," *Complexity*, vol. 2020, Article ID 6626937, 16 pages, 2020.
- [8] D. Yin, R. G. Lopes, J. Shlens, E. D. Cubuk, and J. Gilmer, "A fourier perspective on model robustness in computer vision," *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [9] L. Haibo, X. Lingyun, H. Bin, and C. Zheng, "Status and prospect of target tracking based on deep learning," *Infrared and Laser Engineering*, vol. 46, no. 5, Article ID 0502002, 2017.
- [10] Z. Yan, X. Liu, J. Zhou, and D. Wu, "Coordinated target tracking strategy for multiple unmanned underwater vehicles with time delays," *IEEE Access*, vol. 6, pp. 10348–10357, 2018.
- [11] G. Sreenu and M. A. S. Durai, "Intelligent video surveillance: a review through deep learning techniques for crowd analysis," *Journal of Big Data*, vol. 6, no. 1, pp. 48–27, 2019.
- [12] R. Yuan, F. Guo, Y. Qian et al., "A system dynamic model for simulating the potential of prefabrication on construction waste reduction," *Environmental Science and Pollution Research*, vol. 29, no. 9, pp. 12589–12600, 2022.
- [13] K. Wu, Z. Cai, J. Zhao, and Y. Wang, "Target tracking based on a nonsingular fast terminal sliding mode guidance law by fixed-wing UAV," *Applied Sciences*, vol. 7, no. 4, p. 333, 2017.
- [14] P. Sundaravadivel, E. Kougiannos, S. P. Mohanty, and M. K. Ganapathiraju, "Everything you wanted to know about smart health care: evaluating the different technologies and components of the internet of things for better health," *IEEE Consumer Electronics Magazine*, vol. 7, no. 1, pp. 18–28, 2018.
- [15] H. Morimitsu, I. Bloch, and R. M. Cesar-Jr, "Exploring structure for long-term tracking of multiple objects in sports videos," *Computer Vision and Image Understanding*, vol. 159, pp. 89–104, 2017.
- [16] A. A. Khan, J. Shao, W. Ali, and S. Tumrani, "Content-aware summarization of broadcast sports videos: an audio-visual feature extraction approach," *Neural Processing Letters*, vol. 52, no. 3, pp. 1945–1968, 2020.
- [17] B. Follmer, A. A. Varga, and E. P. Zehr, "Understanding concussion knowledge and behavior among mixed martial arts, boxing, kickboxing, and Muay Thai athletes and coaches," *The Physician and Sportsmedicine*, vol. 48, no. 4, pp. 417–423, 2020.
- [18] G. Russo and G. Ottoboni, "The perceptual-Cognitive skills of combat sports athletes: a systematic review," *Psychology of Sport and Exercise*, vol. 44, pp. 60–78, 2019.
- [19] Z. Liu, Y. He, C. Wang, and R. Song, "Analysis of the influence of foggy weather environment on the detection effect of machine vision obstacles," *Sensors*, vol. 20, no. 2, p. 349, 2020.
- [20] Z. Xiangyue, D. Qinghai, L. Haibo, H. Bin, C. Zheng, and Z. Junchao, "Infrared dim target detection algorithm based on improved LCM," *Infrared and Laser Engineering*, vol. 46, no. 7, Article ID 0726002, 2017.
- [21] Z. Ye and H. Ye, "Particle filter algorithm based spatial motion tracking of football landing location," *Multimedia Tools and Applications*, vol. 79, no. 7–8, pp. 5053–5063, 2020.
- [22] S. Pang, J. J. Coz, Z. Yu, O. Luaces, and J. Diez, "Deep learning to frame objects for visual target tracking," *Engineering Applications of Artificial Intelligence*, vol. 65, pp. 406–420, 2017.
- [23] T. Wang and C. Shi, "Basketball motion video target tracking algorithm based on improved gray neural network," *Neural Computing & Applications*, vol. 34, pp. 1–16, 2022.
- [24] C. Kwan, B. Chou, J. Yang et al., "Deep learning-based target tracking and classification for low quality videos using coded aperture cameras," *Sensors*, vol. 19, no. 17, p. 3702, 2019.
- [25] S. Hossain and D. J. Lee, "Deep learning-based real-time multiple-object detection and tracking from aerial imagery via a flying robot with GPU-based embedded devices," *Sensors*, vol. 19, no. 15, p. 3371, 2019.
- [26] Y. Wang, S. Cao, H. Ren et al., "Towards cost-effective service migration in mobile edge: a Q-learning approach," *Journal of Parallel and Distributed Computing*, vol. 146, pp. 175–188, 2020.
- [27] M. Reichstein, G. C. Valls, B. Stevens, M. Jung, J. Denzler, and N. Carvalhais, "Deep learning and process understanding for data-driven Earth system science," *Nature*, vol. 566, no. 7743, pp. 195–204, 2019.
- [28] Z. Pan, S. Liu, and W. Fu, "A review of visual moving target tracking," *Multimedia Tools and Applications*, vol. 76, no. 16, pp. 16989–17018, 2017.
- [29] J. Luo, Z. Zhang, C. Liu, and H. Luo, "Reliable and cooperative target tracking based on WSN and WiFi in indoor wireless networks," *IEEE Access*, vol. 6, pp. 24846–24855, 2018.
- [30] B. Yan, L. Xiao, H. Zhang et al., "An adaptive template matching-based single object tracking algorithm with parallel acceleration," *Journal of Visual Communication and Image Representation*, vol. 64, Article ID 102603, 2019.
- [31] Y. Wang, X. Luo, L. Ding, S. Fu, and X. Wei, "Detection based visual tracking with convolutional neural network," *Knowledge-Based Systems*, vol. 175, pp. 62–71, 2019.
- [32] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep learning techniques for medical image segmentation: achievements and challenges," *Journal of Digital Imaging*, vol. 32, no. 4, pp. 582–596, 2019.
- [33] Y. Peng, Y. Lin, C. Fan et al., "Passenger overall comfort in high-speed railway environments based on EEG: assessment and degradation mechanism," *Building and Environment*, vol. 210, Article ID 108711, 2022.