

Retraction

Retracted: Basketball Object Extraction Method Based on Image Segmentation Algorithm

Security and Communication Networks

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 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This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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] H. Zhu and L. Liu, "Basketball Object Extraction Method Based on Image Segmentation Algorithm," *Security and Communication Networks*, vol. 2022, Article ID 3021682, 10 pages, 2022.

Research Article

Basketball Object Extraction Method Based on Image Segmentation Algorithm

Huachen Zhu¹ and Long Liu² 

¹College of Physical Education, Jilin University, Changchun 130012, Jilin, China

²Chongqing Preschool Education College, Wanzhou 404100, Chongqing, China

Correspondence should be addressed to Long Liu; liulong@aynu.edu.cn

Received 20 May 2022; Revised 7 July 2022; Accepted 26 July 2022; Published 29 August 2022

Academic Editor: Panagiotis D. Diamantoulakis

Copyright © 2022 Huachen Zhu and Long Liu. 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.

Finding your favorite videos from massive sports video data has become a big demand for users, accurate sports videos can better help people learn sports content, and the traditional data management and retrieval methods using text identifiers are difficult to meet the needs of users, so the research on the extraction of sports objects in sports videos is of great significance. This paper mainly studies and proposes the basketball object extraction method based on image segmentation algorithm and can accurately analyze the trajectory of the basketball target. By modeling the video frame of basketball game, the basketball object is selected for segmentation and extraction. The extracted basketball object can be used for tracking the target in the basketball video clip retrieval system. At the same time, the segmentation and extraction of the basketball object are also the core part in the basketball video clip retrieval framework. Combined with the characteristics of basketball video images in the database, the algorithm extracts the image block variance and contrast to form the training feature vector, and the correct segmentation rate on the database is higher than 95.2%. The results show that this method has a good effect on the segmentation and extraction of basketball objects in basketball videos.

1. Introduction

1.1. Background and Significance. The key technologies of robots include target recognition and tracking, path planning, and multisensor information fusion. They are widely used in hotel service robots, autonomous driving, drone and missile interception, and other application scenarios. Among them, the identification and tracking of fast flying objects have great research and application value in the identification and tracking of suspicious targets in sports competition monitoring and adjudication and military confrontation. With the development of digital video technology, Internet technology, and broadband multimedia services, video has gradually become one of the mainstream carriers of information dissemination and a commonly used tool for people to obtain and record information; in this case, various new video application methods have also continuously emerged, and there has been an explosive and rapid

increase in the volume of video data. People have put forward new and higher requirements for the storage, retrieval, and processing of video data. In order to make full use of video information, people want to consume video media in an adaptive manner, that is, to consume video media any way, at any time, and in any place. This consumption is not a passive consumption, and consumers of video media are also producer of video media. At the same time, due to the rapid development of communication technology, including the development of network technology, there are massive amounts of video on the Internet. Therefore, how to self-determine the outside world, interactively process information, and understand the content of information, how to obtain video information distributed on the global Internet according to needs, how to find video information of interest in massive videos, and how to organize knowledge through automatic learning and constitute a colorful interactive video content service have become the goal pursued by

people in the digital age. This makes the development of related technologies such as intelligent processing, effective organization, retrieval, and management of videos more urgent. Image segmentation technology can achieve accurate tracking of image targets and has a wide range of applications in motion capture and so on.

1.2. Research Status at Home and Abroad. In recent years, image segmentation methods based on graph theory have attracted a lot of interest from many scholars, and many researchers have focused on in-depth research on them. Image segmentation methods based on graph theory have a perfect theoretical background, and the segmented images have good results, which have also been widely used in image segmentation technology. After continuous research and efforts of researchers, image segmentation methods based on graph theory have developed today, and many improved methods have also been produced, which have made many contributions to image segmentation methods.

In 2004, someone proposed an isoperimetric cut method, whose idea was inspired by the classic isoperimetric problem [1]. To segment a good-quality image, the following conditions are met: the acreage of the area needs to be large, and the circumference of the area needs to be small [2]. For isolated points, the perimeter is very small, and at the same time the area is also small, so the cost of the cut set will not be the smallest, which effectively prevents the situation where the segmentation tends to be small regions or isolated points [3]. Isocyclic cut sets transform the feature vector problem of graph theory into a linear system problem of images, which improves the stability of the algorithm and runs faster [4]. When the image has multiple targets, the target needs to be iterated in depth, and the algorithm requires a large amount of calculation [5].

Someone trains a generative probability model, which adopts the method of unsupervised learning object class segmentation, in which Learning Object Classes with Unsupervised Segmentation is referred to as LOCUS [6]. The color and texture characteristics of different entities of the same type of objects are unstable, but their shape characteristics are relatively stable; in addition, the color and texture characteristics of individual entities are relatively stable, which is the premise of LOCUS assumptions [7, 8]. LOCUS combines low-level features such as colors, textures, edges, and top-level features such as shapes and postures through generative probabilistic models to simultaneously achieve object positioning, segmentation, and posture estimation [9, 10]. The biggest feature of this model is the ability of allowing large differences in appearance between different entities of the object. In order to avoid the hard decision of the object's pose, size, and segmentation during the calculation process, LOCUS iteratively updated its trust degree to achieve the optimization of these three aspects at any time [11, 12]. LOCUS trains the model by selecting a small number of unsegmented and labeled images and segmenting by inference. The purpose is to reduce human participation [13].

For the first time, someone used CRF in image pixel labeling. Pixel labeling refers to classifying pixels in an image according to a preset classification [14]. Both the input image segmentation and the label classification of the area can be achieved by pixel labeling [15, 16]. The use of context information to mark pixels is based on the principle that there is a great correlation between adjacent pixels. And it has different tag information at different image levels [17, 18]. For example, color and texture features are the basis for classifying pixels on local images; color and texture are commonly used for pixel classification; meanwhile, contour and shape features are the basis for classifying regions in the image [19]. In addition, according to the preset, the local characteristics between pixels classified into different categories will also be similar; for example, the water surface and the sky are blue; they have the same color characteristics; at the same time, noise easily affects the local features [20, 21]. Therefore, classification based on local features is not persuasive. At this time, the geometric relationship between the objects, the position of the object in the image, and other context information can be considered [22, 23]. The information labels on partial images have "ambiguity." It is obvious that the context information at a higher level is conducive to the elimination of this "ambiguity" [24]. Through the above investigations, the author describes the characteristics of the image at different levels, using the multiscale CRF method and further statistics of the degree of influence of their mutual relationship on pixel classification. The author uses the maximum likelihood to estimate the model parameters in the training process by increasing the gradient [25, 26]. At the same time, the maximum posterior probability is used to estimate the posterior distribution of the classification labeling conditions in the image segmentation process by Gibbs sampling, which is simple and easy to converge [27].

A variety of information in the active contour model is used to define a new energy function for image segmentation. These models aim to extract all potential objects from the background but at the same time also obtain nontarget objects and noise. Liu uses a sparse representation method to extract the target object. The original indicator function (binary function) relative to the level setting function is used to represent the foreground (value 1) and background (values 0). From another perspective, the indicator function can be represented by a linear combination of a set of basic functions. First, the label operator uses the indicator function in each iteration, and each connected region is represented by a basis function. Second, the linear combination of these basis functions is used to represent the object. Finally, through sparse constraints of basis function coefficients, object extraction is regarded as a sparse representation problem. At the same time, the corresponding improved orthogonal matching tracking algorithm is designed to obtain the ideal result [28]. His method is too complicated to solve the problem of efficiency in practical applications. Sun has proposed an improved ViBe algorithm to extract moving objects for the effects of ghosting in the visual background extractor (ViBe) and dynamic background. He improved the way by which the background was acquired during modeling to eliminate

ghosting. Detect the saliency of the previous M frames and synthesize a relatively real background. He improved the selection of thresholds in the model to reduce the impact of dynamic background. Adaptively, adjust the threshold according to the background complexity. In addition, finding the internal contour of the extracted object to be filled makes the detected target more complete [29]. His research method does not consider the problem handling when the moving object is blocked, and it is not comprehensive. In order to evaluate the accuracy of object extraction, Liping proposed several novelty measures that differ from the standard. First, based on the confusion matrix, he gave measures for the accuracy assessment based on the area and the number of objects. Secondly, he combines the similarity of multiple features to provide different accuracy evaluation measures. Third, in order to improve the reliability of the target extraction accuracy evaluation results, he designed two accuracy evaluation measures based on the differences in target details. Compared with existing methods, this method synthesizes feature similarity and distance difference, which greatly improves the reliability of target extraction evaluation [30]. Although the reliability of his method is very good, the processing time for the problem is too long to achieve the purpose of practical application.

The main work of this article is to describe the basketball video information description model and basketball video object extraction algorithm and discuss its application research. On the basis of reviewing and summarizing the existing results, a multilevel is proposed in the field of basketball video object description. Based on the sports video semantic object description model and combined with the video semantic information hierarchical description model, a multilevel basketball video object extraction algorithm is proposed.

2. Image Segmentation Algorithm and Object Extraction Method

2.1. Contour Extraction. The Snake model deforms the initial curve to obtain the target contour. From the mechanical point of view, when an object deforms, it must be subjected to force. The basic idea of the Snake model is to transform the curve drawing according to the image information into solving the optimal route problem. Therefore, the evolution of the curve in the Snake model must also be affected by different forces, that is, the internal force and the image force. The steps of the traditional Snake model are the following: first, manually point out some points in the image, which will be connected to each other to form an initial curve; then, the energy in the functional is minimized; that is, this process makes the initial curve like a target and contours are close. When the traditional Snake model forms the initial curve, the points in the image are $A(x) = [B(x), C(x)]x \in [0, 1]$. Among them, $B(x)$ and $C(x)$ represent the coordinates of each point when it is pointed out in the image where x is the independent variable. Snake uses these points to define the energy function and then relates the energy change to the deformation of the contour curve, which is defined as follows:

$$E_{\text{total}} = \int_s \left(\alpha \left| \frac{\partial}{\partial s} \bar{v} \right|^2 + \beta \left| \frac{\alpha^2}{\alpha s^2} \bar{v} \right|^2 + E_{\text{ext}}((\bar{v})(s)) \right). \quad (1)$$

Among them, $\alpha |\partial/\partial s \bar{v}|^2$ changes the shape of the curve, and $\beta |\alpha^2/\alpha s^2 \bar{v}|^2$ bends the curve, which is a force that changes the initial curve. Extract Snake model by bending curve. What $E_{\text{ext}}((\bar{v})(s))$ can do in the traditional Snake model is only the characteristics of the area near the curve, such as the gradient in the image area, that is, the figure force:

$$E_{\text{ext}}(\bar{v}(s)) = P(\bar{v}(s)) = -|\nabla I(v)|^2, \quad (2)$$

where the force is the vertically acting force. When the curve evolves near the edge of the target shape, the gradient value in the outline of the curve will become larger and then lead to the minimum energy in the formula. According to the defined curve evolution equation, it can be understood that the point at this time will be stationary, not changes. Eventually, the curve will stop at the edge of the target shape, and of course the target contour you want to get is obtained. Of course, to achieve this goal, an important requirement is needed; that is, the target in the image can have more obvious edge features to prevent the curve from overflowing.

The various energies that change the shape of the curve in the Snake model are called internal force, and they can operate the various changes of the curve while ensuring the continuity and smoothness of the curve. And $E_{\text{ext}}(\bar{v}(s))$ is also called external force, which represents the change of the curve in the same area as the characteristics of its vicinity. The values of α and β in the formula also have a great effect. They respectively determine that the curve can be changed within a certain range of length and round plaque.

When the curve is initialized in the image, the points used are because of its randomness, so solving the minimum value of $E_{\text{total}}(V)$ is equivalent to the variational method in mathematics. In this stochastic situation, the final solution can be made through the Euler equations. Because the functional extremum condition of the differential expression is equivalent to the Euler equation, the solution of the Euler equation, which is to solve the extremum of the functional, goes a step further, making the deep variational problem more understandable and simple differential problem.

$$-\alpha' \bar{v}' - (\alpha - \beta) \bar{v}'' + 2\beta' \bar{v}' + 2\beta \bar{v}''' = -\nabla P(\bar{v}). \quad (3)$$

Among them, $F = -\nabla P$ is external force. Through the discretization of the above formula, $B(x)$ and $C(x)$ are changed into linear systems with five diagonal matrices, and then they are solved. Of course, in actual applications, the initial curve of the Snake model is usually and manually calibrated around the target to form the initial curve, followed by the continuous solution of the energy function. Finally, the target contour curve is obtained as the target contour and can realize the detection and analysis of moving targets.

However, the Snake model only plays a pioneering role and proposes a new research direction for obtaining the target contour; that is, the curve is evolved into the target contour through the capability functional method. However,

it also has a lot of areas for improvement. First, the initial curve requires rigorous design and planning; otherwise it will definitely cause a great change in the final output. If the curve contains a concave shape, it will be difficult to divide, and it will easily converge to the local pole. If the topology of the curve changes, it will not be able to handle flexibly; secondly, the Snake model has poor robustness, it is easily affected by noise, and the accuracy of the obtained results is not high.

The parameterization of the line is a very difficult problem. For its serious shortcomings, after continuous research and development, the level set method based on the active contour model method was born. The curve evolution process of this method can be automatically split and merged. It has greater advantages and practical value than the Snake model.

2.2. Region-Based Object Identification Features

2.2.1. Fourier Painter. A series of boundary points in the image object area are regarded as a complex sequence. The Fourier transform of this complex sequence is called Fourier description sub-FD (Fourier description). Fourier transform is a linear combination of trigonometric functions or their integrals that can express a function that satisfies certain conditions. For any object, after normalization, it can avoid the dependence of the feature on position, size, and direction. However, the use of descriptors to represent the boundaries of objects, the amount of calculation, is relatively large; it is difficult to meet the real-time needs of actual needs.

2.2.2. Euler Number. Topological properties can be used to describe the shape of a planar object area, and Euler's number is one of the object's topological properties. The Euler number is a characterization of a vector bundle. If the number of holes in the object area in the image is and the connected part is, the Euler number is defined as

$$E = C - H. \quad (4)$$

Euler's number is invariant to rotation, scale, and translation.

2.2.3. Area Projection. In object shape analysis, object area projection is a very effective method. The projection of a two-dimensional image is a one-dimensional waveform whose value is the sum of the values of pixels along a specific direction. Taking character recognition as an example, the general equations for 0° , 90° , 45° , and 135° projection are

$$\begin{aligned} P(0^\circ, t) &= \int_R f(x, t) dx, \\ P(90^\circ, t) &= \int_R f(t, y) dy, \\ P(45^\circ, t) &= \int_R f(x, x - \sqrt{2}t) dx, \\ P(135^\circ, t) &= \int_R f(x, \sqrt{2}t - x) dx. \end{aligned} \quad (5)$$

In the formula, $p(0^\circ, t)$, $p(90^\circ, t)$, $p(45^\circ, t)$, and $p(135^\circ, t)$ are the projections of the object areas in the above four directions, respectively, and are the pixel values.

A lot of information about the rough appearance of the characters is included in the projection. The projection can detect strokes in a specific direction, and some characters can be strictly recognized by their projections, such as 1, 2, 5, and 7. Using projection technology, you can also detect other simple two-dimensional target parameters. For example, using the projection of the vehicle image, you can calculate the relative physical length. Using the projection of the edge image of the license plate image, you can locate the position of the license plate. The projection method is not invariant to object rotation.

2.3. Description of the Image. A graph uses topology to represent the association between edges and vertices and can also be defined as a set of edges and vertices. Geometry defines a graph as a collection of points (vertices) in space and lines (edges) connecting these points. Graph theory defines a graph as a pair $G = (V, E)$, where V represents the set of vertices and E represents the set of edges. Thus, Figure 1 can be expressed as follows.

$$V = (v_1, v_2, v_3, v_4), E = (E_1, E_2, E_3, E_4). \quad (6)$$

Among them, the edge can also be represented by two vertices. If the two endpoints of edge e are v_i and v_j , then edge e can be written as $e = (v_i, v_j)$ and (v_j, v_i) represents the disordered pair of v_i and v_j . That is, both (v_i, v_j) and (v_j, v_i) express undirected edges with v_i and v_j as end points. This can be rewritten as

$$\begin{aligned} G &= (V, E), V = (v_1, v_2, v_3, v_4), \\ E &= \{(v_1, v_2), (v_1, v_3), (v_1, v_4), (v_2, v_3), (v_2, v_4), (v_3, v_4)\}. \end{aligned} \quad (7)$$

2.4. Calculation of Machine Vision. As shown in Figure 2, it is a schematic diagram of two methods of machine vision. The color representation of an image is represented by grayscale values. The gray value of the image represents the performance of the complex factors of various aspects in the natural world scene. These complex factors that affect the gray value of the image are the geometry of the real object, the reflectivity of the surface, the light environment at that time, and the direction and distance of the viewer from the object. When any of the above complex factors change, the image will be changed, because these factors determine the grayscale of the image. The nature of objects in the natural world has nothing to do with the above factors. No matter how far away the observer is from the object, the size of the object observed will change, but the size of the object itself will not change. Viewing the object from different directions will get different image, but the shape of the object will not change. The characteristics of objects that exist in objective facts are unchanged. These images are imaged on our retina and then input to the brain, so people feel that the objects are changing. The

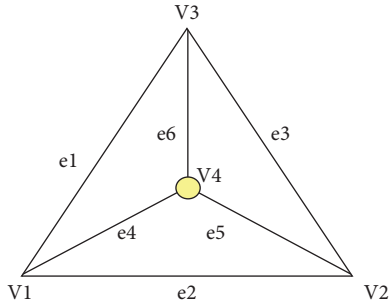


FIGURE 1: Undirected graph.

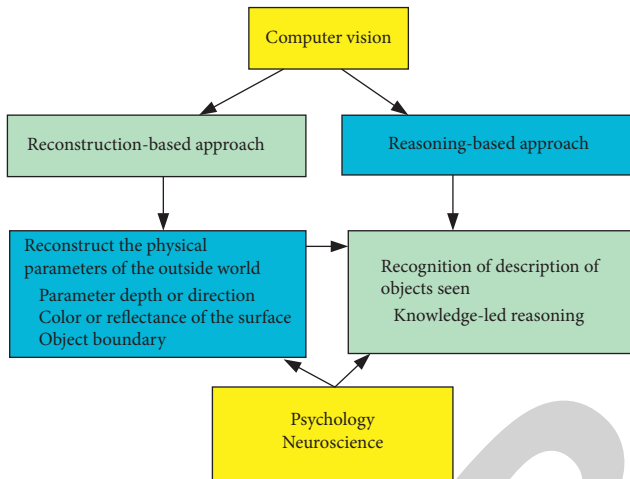


FIGURE 2: Two methods in computer vision.

external natural real world forms an image on the retina of the human eye, which is actually human perception of the natural world. The image obtained by this perception is constructed by gathering point data. Another point that is worth analyzing is that the characteristics of objects that people feel in the brain will not change. It is speculated here that human brain cells will converge many point-shaped data information into a complete body and restore the characteristics of the object, isolate the complex factors that affect image imaging, and obtain pure and clean data information that is only the most essential of the object. This purely clean data information that only belongs to the most essential objects will not be affected by factors such as the light environment, the distance and direction between the viewer and the object, and the reflectivity of the object surface, which is called constancy. In short, the visual brain nerve will not be based solely on the image formed on the retina in the natural real world but will identify objects based on the point data aggregation process and by separating those complex factors that affect image imaging.

2.5. Optical Flow Method. Optical flow method is to project each point on a moving object in three-dimensional space to the observation plane and express the motion information of the original object through the instantaneous velocity of each pixel on the observation plane. Mainly, used in

computer vision and other image processing fields, it is very useful for motion detection, object cutting, calculation of collision time and object expansion, etc. A pixel in the time domain has a position shift at the next moment (or in the next frame), which can be understood as the instantaneous velocity (including size and direction) of the pixel, and the entire image is the frame image optical flow field. Reducing the optical flow field to three-dimensional space is the motion field of the physical objects in three-dimensional space. The establishment of optical flow method is based on two conditions: color consistency (luminance constancy assumption) and small motion (spatial smoothness assumption). For a grayscale image, consistent colors can be understood as consistent brightness values; small motion means that no large motion shift occurs for each pixel. As shown in Figure 3, it is a schematic diagram of the principle of optical flow method.

3. Experimental Design of Basketball Object Extraction

3.1. Experimental Environment Settings. Image segmentation is to extract areas of interest. In order to objectively evaluate the advantages and disadvantages of the method, this paper compares the method with traditional methods in terms of algorithm running time and segmentation accuracy. The experimental operating environment constructed in this paper is shown in Table 1.

3.2. Experimental Procedure. After the conditional random field optimization process, the final semantic segmentation map is obtained at this time, but it cannot be used to extract the contour because we know that the semantic segmentation map obtained at this time has 300 candidate regions, in which there may be one target, or there may be multiple targets. Therefore, they need to be further processed; existing image segmentation methods include threshold-based, region-based, and edge-based segmentation.

Then, we can know the specific pixel value of each target. For example, if the target is basketball, then it has pixel values of $R = 128$, $G = 128$, and $B = 128$. Different targets correspond to different pixels. First, we need to distinguish between different types of targets. Among them, pixel maps that contain two or more pixels are also classified as one, with the single same pixel value or different pixel values. The goals are grouped together for later processing. Through the upsampling process in the above section, there is another important position information that is also particularly important; because the pixel map at this time corresponds to the position in the original image, we can use the position information for preliminary processing.

In the same and single pixel image class, if the IOU threshold of the two pixel images is greater than 0.8, that is to say, the two are related or partially overlapped, then they are fused on the new background image processing. If pixels at the same location are same, the pixel will not be put into the corresponding position. If one is there and the other is not, it

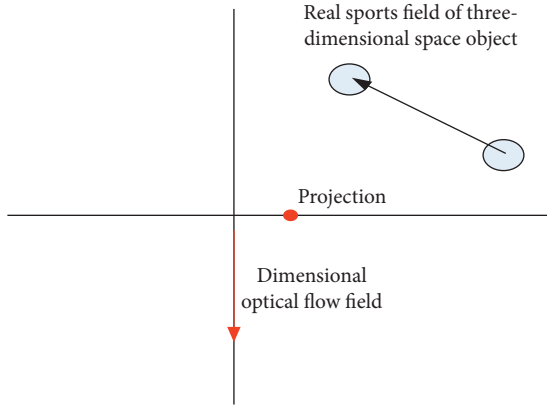


FIGURE 3: Schematic diagram of optical flow method.

TABLE 1: Experimental environment configuration table.

| Lab environment | Configuration |
|------------------|----------------------------------|
| CPU | Intel core i5 3.0 GHz |
| RAM | 4 GB |
| Operating system | Windows 7 Ultimate |
| Open environment | Visual Studio 2013; MATLAB 2014b |

will be filled in and so on until all the segmentation targets are executed. If the IOU value is less than the threshold of 0.8, the two are judged as different targets.

Pixel differentiation: the pixels at the same position of the two are unchanged, but, for pixels at different positions, it is necessary to fill the pixel values at the positions that have pixel values at different positions, so the target semantic segmentation after further fusion is divided into pixels. Further supplementary processing was carried out on the same time, and the final processing results were obtained at the same time.

4. Experimental Analysis of Basketball Object Extraction

4.1. Comparative Analysis of Accuracy Test. In this paper, the results of testing on VOC-2007 and 20012 are obtained by using PASCAL VOC-2012 as a training set and using the PASCAL VOC evaluation server. As shown in Table 2, it is clear that the method adopted in this paper achieves better results on the basis of DeepLab and the segmentation accuracy shown on the VOC-2012 test set is 9.5% higher than DeepLab.

In order to further illustrate the detection effect of small targets in the image, this article compares the results of the specific detection accuracy of various targets in the image on the data set with the results obtained by other methods, as shown in Figure 4. Obviously a detection and positioning effect based on the outline of the target is more accurate than that using a rectangular frame to detect the target.

4.2. Classification Analysis of Basketball Video Shots. The shots of the basketball video can be divided into the shots of the court area and the shots of the off-field area according to

TABLE 2: Comparison of accuracy test on data set.

| Method | Test accuracy (%) |
|----------------------|-------------------|
| FCN-8s | 63.7 |
| CRF-RNN | 66.5 |
| DeepLap-CRF | 66.8 |
| VGG19-FCN | 68.6 |
| Deep squee Net-CRF | 72.4 |
| This research method | 75.6 |

the areas presented therein. The field lens can be divided into far lens, medium lens, and close lens. The telephoto is a global perspective, which reflects the progress of each player's macrogame such as standing and running. The middle shot is basically a picture of the whole body of one or a few players, that is, the details of the player's collision and contention in the game. Close-up refers to the close-up of the athlete, that is, the close-up of the player's expression or physical contact. Figure 5 shows the action of the player when he shoots. Shots from outside the field include coaches, spectators, and referees.

Since the close-up lens and the off-field lens are very similar, they can be judged by human semantics, and it is more difficult to distinguish them, and the meaning is not much different, so this article classifies these two kinds of shots into the same category. Since the video frames in the same video sequence will contain multiple shot types, it is impossible to determine the type of a shot based on a certain frame. In order to improve the accuracy of judgment and reduce errors, this paper uses the frame with the largest proportion to determine the type of shot. The purpose of image segmentation is to segment the lens image and analyze each frame to improve the accuracy of image analysis. As shown in Figure 6, the close-up shots of the players are shot.

Using the judgment standard proposed in this paper, the motion information difference curve constraint is introduced for extraction, 25 key frames are obtained, and the redundant information is reduced by 30%. As shown in Figure 7, an example of key frames extracted using this algorithm and a motion information difference map of the test video are given.

4.3. Image Processing Performance Analysis of Occlusion. It can be seen from the experimental results that the target tracking algorithm based on multifeature fusion can overcome the influence of morphological changes during target movement when tracking human targets and can reliably track targets in normal motion in complex scenes. During the tracking process, the relevant information in the object collection is updated in time to ensure the correct matching under the smooth change of the target feature. And, because the state information about the target is stored in the object set, even if the target disappears after reappearing temporarily, it can still be correctly identified and label-tracked.

For the occlusion problem, the occlusion determination criterion in this paper can give mutual occlusion targets and occlusion types, which is convenient for the next occlusion

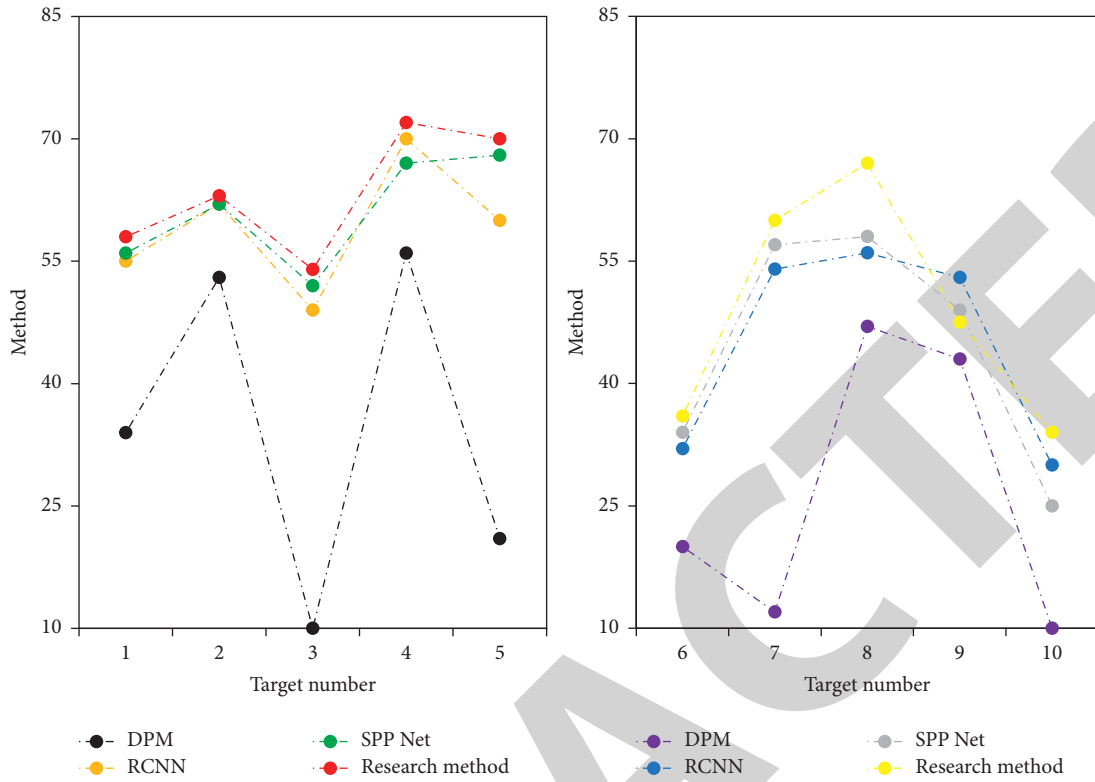


FIGURE 4: Various types of target detection effects.

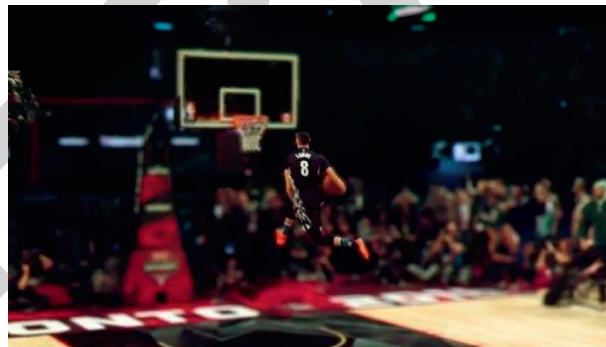


FIGURE 5: Player shooting action.



FIGURE 6: A close-up shot of the player's shooting action.

processing. In terms of occlusion processing, the algorithm in this paper changes the method of dealing with the occlusion problem in a predictive manner in the previous

algorithm, while, combining the direction of the target's movement, it makes full use of the effective brightness information of the exposed part of the occlusion target to



FIGURE 7: Basketball video key frames.

TABLE 3: Algorithm of performance comparison.

| Image under basketball video | Threshold pair | Calculation time/s |
|-------------------------------------|----------------|--------------------|
| Classic two-dimensional Otsu method | (126, 126) | 0.14318 |
| Algorithm for this study | (126, 126) | 0.13714 |
| MQPSO | (73, 100) | 0.12811 |

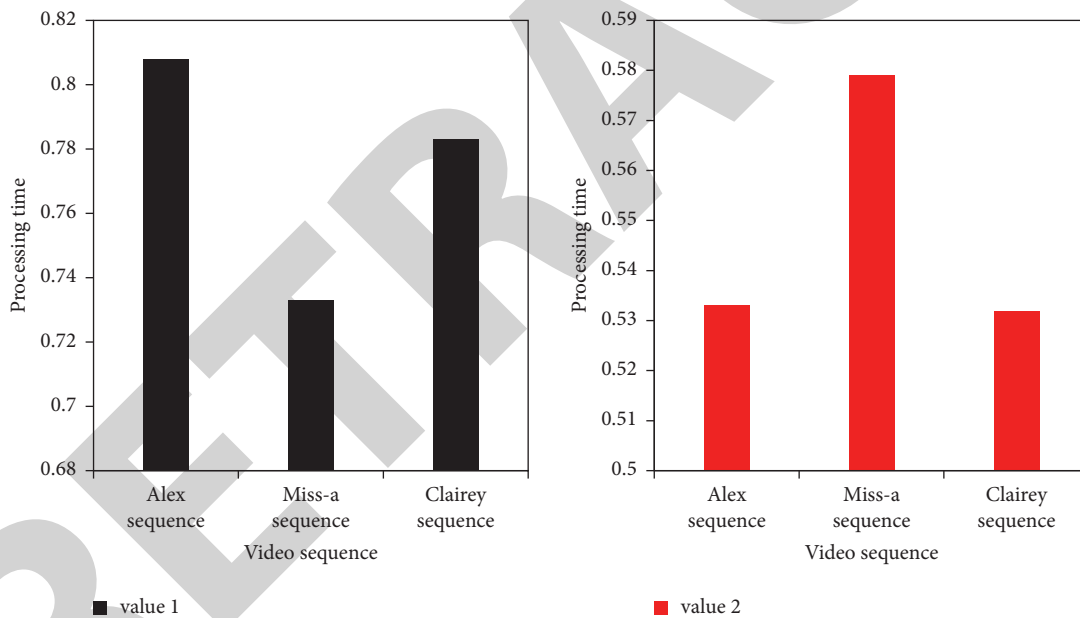


FIGURE 8: Relationship between cumulative frame difference and the processing time.

reduce the matching calculation of the occlusion part. The specific gravity can effectively deal with the problem of slight occlusion. For the occlusion between targets, by predicting the number of occlusion frames, the weighting formula can be automatically reversed in the second half of occlusion, which ensures the rationality of the calculation.

As shown in Table 3, for the images under the basketball video, the segmentation speed obtained by the algorithm proposed in this paper has been further improved. As shown in Figure 8, the algorithm uses the relationship between the time and the number of frames for the cumulative frame difference calculation of the three image sequences: Clairey,

Miss-a, and Alex. The calculation time of the research algorithm is 0.13714s, which is shorter than other algorithms. Taking the Clairey sequence as an example, the number of areas obtained by using this algorithm to find the motion area is 28, which means that the subsequent processing will mainly be based on 28.

As shown in Figure 9, in terms of segmentation accuracy and running time, although this algorithm has no obvious advantage in time, from the point of view of the accuracy of the segmentation result, this algorithm has obvious advantages. And the threshold pair of this algorithm is higher and can accommodate more image information. Therefore,

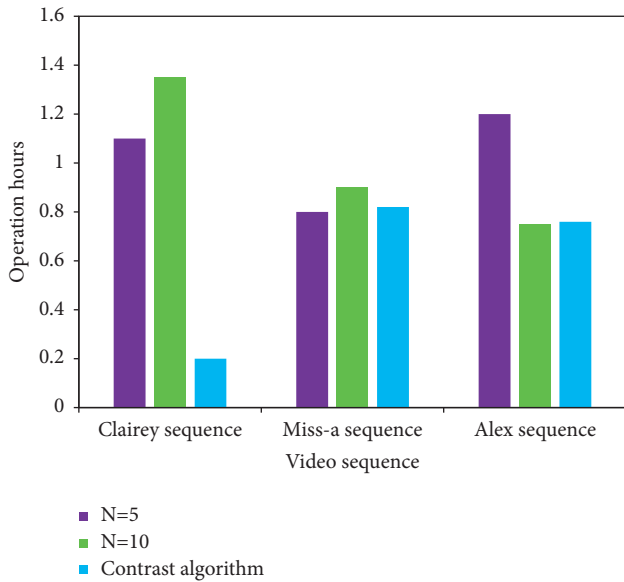


FIGURE 9: Running time comparison.

in general, this algorithm has better advantages than the comparison algorithm in the segmentation of moving objects in low bit rate video image sequences.

5. Conclusions

With the rapid development of multimedia technology, people's interest in the interactive function of video content has gradually increased. This paper analyzes the significance, background, and current situation of basketball image segmentation and recognition and based on the difficult environment of basketball video image classification analyzes the deficiencies and difficulties encountered in existing traditional image segmentation and recognition algorithms, aiming at studying these problems, and corresponding solutions were proposed.

This paper proposes a method for extracting basketball objects from basketball game video based on image segmentation algorithm. Through contour extraction and other techniques, each frame of the image is analyzed, and then the basketball target in the image is tracked. Experiments show that this method has the advantages of less interactive operations and high segmentation accuracy when segmenting and extracting basketball objects and it can be applied to segmentation and extraction of other ball objects.

This paper further improves the accuracy of segmentation and extraction of the basketball object in the video frame image by this method and applies the extracted basketball object to the basketball video clip retrieval system as the target tracking object, which provides an effective method for basketball video clip retrieval.

Data Availability

No data were used to support this paper.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

2022 Projects of Science and Technology in Henan Province: Algorithm and Application of Movement Image Based on Convolutional Neural Network (Grant number: 222102320063).

References

- [1] A. Wilkowski, M. Stefańczyk, and W. Kasprzak, "Training data extraction and object detection in surveillance scenario," *Sensors*, vol. 20, no. 9, p. 2689, 2020.
- [2] X. Zhang, Y. Qi, and G. Lin, "Dynamic object feature extraction with rotating parts[J]," *Dianbo Kexue Xuebao/chinese Journal of Radio Science*, vol. 34, no. 1, pp. 65–69, 2019.
- [3] K. E. Ko, H. Shim, and K. B. Sim, "Development of behavior understanding system based on object-human interaction pattern extraction using deep learning model," *Journal of Institute of Control, Robotics and Systems*, vol. 24, no. 1, pp. 57–63, 2018.
- [4] W. Song, Y. Zhang, L. Gao, and Y. Zhou, "The extraction and classification of Liaoning northwest desertification information based on object-oriented and decision-making model[J]," *Liaoning Gongcheng Jishu Daxue Xuebao*, vol. 37, no. 3, pp. 595–601, 2018.
- [5] X. Li and W. Li, "A case study of a two-stage image segmentation algorithm," *Multimedia Tools and Applications*, vol. 78, no. 7, pp. 8197–8206, 2019.
- [6] L. Peng, "A brain nuclear magnetic resonance image segmentation algorithm based on non-rigid registration," *Traitement du Signal*, vol. 35, no. 3-4, pp. 317–330, 2018.
- [7] H. Chen, X. S. Yu, C. D. Wu, and P. SUN, "Fast image segmentation algorithm based on parametric level set active contour model[J]," *Dongbei Daxue Xuebao/Journal of Northeastern University*, vol. 40, no. 1, pp. 6–10, 2019.
- [8] Q. Qi, Y. Tian, and L. Han, "An improved image segmentation algorithm based on the maximum class variance method," *MATEC Web of Conferences*, vol. 309, no. 11, Article ID 03029, 2020.
- [9] K. Hachemi, S. Vaidyanathan, and B. Beddad, "Development and optimisation of image segmentation algorithm on an embedded DSP-platform," *International Journal of Computer Applications in Technology*, vol. 58, no. 3, p. 250, 2018.
- [10] L. R. Dung, Y. M. Yang, and Y. Y. Wu, "A hierarchical grab cut image segmentation algorithm," *Journal of Computer and Communications*, vol. 06, no. 02, pp. 48–55, 2018.
- [11] J. Hu and Y. Zhou, "Fuzzy clustering image segmentation algorithm based on new distance matrix variance[J]," *Journal of Quantitative Economics*, vol. 035, no. 003, pp. 286–293, 2018.
- [12] W. Hongmei and S. Jiayi, "SAR image segmentation algorithm based on Contourlet domain AFMRf model[J]," *IET Image Processing*, vol. 12, no. 7, pp. 1124–1130, 2018.
- [13] Z. Zhen, "An image segmentation algorithm based on combination of slope width reduction and cross cortical model," *International Journal of Computer Applications in Technology*, vol. 61, no. 1/2, p. 75, 2019.
- [14] E. Harkaitz, C. Oskar, and M. Iciar, "The shannon entropy trend of a fish system estimated by a machine vision approach

- seems to reflect the molar Se:Hg ratio of its feed[J],” *Entropy*, vol. 20, no. 2, p. 90, 2018.
- [15] F. P. An and J. E. Liu, “Medical image segmentation algorithm based on optimized convolutional neural network-adaptive dropout depth calculation,” *Complexity*, vol. 2020, no. 7, pp. 1–13, Article ID 1645479, 2020.
 - [16] Y. Qi, G. Zhang, and Y. Li, “An auto-segmentation algorithm for multi-label image based on graph cut,” *Sensing and Imaging*, vol. 19, no. 1, p. 13, 2018.
 - [17] B. Gao, X. Q. Li, W. L. Woo, and G. Y. Tian, “Physics-based image segmentation using first order statistical properties and genetic algorithm for inductive thermography imaging,” *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2160–2175, 2018.
 - [18] C. Yang, W. Huang, R. Zhang, and R. Kong, “Portrait extraction algorithm based on face detection and image segmentation,” *Computer and Information Science*, vol. 12, no. 2, pp. 1–13, 2019.
 - [19] A. Banerjee and P. Maji, “Spatially constrained student’s t-distribution based mixture model for robust image segmentation,” *Journal of Mathematical Imaging and Vision*, vol. 60, no. 3, pp. 355–381, 2018.
 - [20] K. Nongmeikapam, W. K. Kumar, and A. D. Singh, “Fast and automatically adjustable GRBF kernel based fuzzy C-means for cluster-wise coloured feature extraction and segmentation of MR images,” *IET Image Processing*, vol. 12, no. 4, pp. 513–524, 2018.
 - [21] X. Wu, X. Ma, J. Zhang, and Z. Jin, “Salient object detection via reliable boundary seeds and saliency refinement,” *IET Computer Vision*, vol. 13, no. 3, pp. 302–311, 2019.
 - [22] Y. Du, D. Wen, G. Liu, S. Qiu, and N. Li, “A characteristic extraction algorithm based on blocking star images[J],” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 33, no. 09, pp. 104640J–104783J, 2019.
 - [23] L. Zeng, Z. Jia, and Y. Wang, “Extraction algorithm for optimal coarse-grained networks on complex networks,” *International Journal of Modern Physics C*, vol. 30, no. 11, Article ID 1950081, 2019.
 - [24] C. W. Zhang, B. Y. Guo, M. Y. Yang, and J. Wen, “A study for moving object extraction method of intelligent vehicle omnidirectional lidar[J],” *Journal of Information Hiding and Multimedia Signal Processing*, vol. 9, no. 6, pp. 1380–1390, 2018.
 - [25] M. Xian, Y. Zhang, H. D. Cheng, F. Xu, B. Zhang, and J. Ding, “Automatic breast ultrasound image segmentation: a survey,” *Pattern Recognition*, vol. 79, no. 3, pp. 340–355, 2018.
 - [26] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.
 - [27] F. Zhang and X. H. Zhang, “Distance regularized level set image segmentation algorithm by means of dislocation theory [J],” *Zidonghua Xuebao/acta Automatica Sinica*, vol. 44, no. 5, pp. 943–952, 2018.
 - [28] G. Liu and J. Zou, “Level set evolution with sparsity constraint for object extraction,” *IET Image Processing*, vol. 12, no. 8, pp. 1413–1422, 2018.
 - [29] P. Y. Sun, L. r Lü, and J. Qin, “Moving object extraction based on saliency detection and adaptive background model,” *Optoelectronics Letters*, vol. 16, no. 1, pp. 59–64, 2020.
 - [30] C. Liping, S. Wenzhong, M. Zelang, and M. Hao, “Accuracy assessment measures for object extraction from remote sensing images[J],” *Remote Sensing*, vol. 10, no. 2, p. 303, 2018.