

Research Article

Research on Design of Intelligent Background Differential Model for Training Target Monitoring

Ya Liu ¹, Fusheng Jiang ¹, Yuhui Wang ², Lu OuYang ², Bo Gao ², Jinling Jiang ²,
and Bo Zhang ²

¹Shenyang Sport University, Physical Education College, Shenyang 110102, Liaoning, China

²Shenyang Sport University, Graduate Department, Shenyang 110102, Liaoning, China

Correspondence should be addressed to Fusheng Jiang; jiangfusheng@syty.edu.cn

Received 22 February 2021; Revised 20 April 2021; Accepted 13 May 2021; Published 24 May 2021

Academic Editor: Wei Wang

Copyright © 2021 Ya Liu et al. 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.

The detection of moving targets is to detect the change area in a sequence of images and extract the moving targets from the background image. It is the basis. Whether the moving targets can be correctly detected and segmented has a huge impact on the subsequent work. Aiming at the problem of high failure rate in the detection of sports targets under complex backgrounds, this paper proposes a research on the design of an intelligent background differential model for training target monitoring. This paper proposes a background difference method based on RGB colour separation. The colour image is separated into independent RGB three-channel images, and the corresponding channels are subjected to the background difference operation to obtain the foreground image of each channel. In order to retain the difference of each channel, the information of the foreground images of the three channels is fused to obtain a complete foreground image. The feature of the edge detection is not affected by light; the foreground image is corrected. From the experimental results, the ordinary background difference method uses grey value processing, and some parts of the target with different colours but similar grey levels to the background cannot be extracted. However, the method in this paper can better solve the defect of misdetection. At the same time, compared with traditional methods, it also has a higher detection efficiency.

1. Introduction

The video surveillance system is mainly composed of moving target detection and tracking, target behaviour analysis, understanding of target appearance scenes, and subsequent behaviour analysis. With the smooth development of the safe city construction, a variety of cameras has begun to be distributed to all corners of the world. According to the various information collected by the cameras, the security is solved. Most of the problems in the industrial field are precisely because of the massive monitoring information provided by these large numbers of cameras that improve people's security. The working mode of the traditional monitoring system first obtains the monitoring information through the camera, then assigns the relevant operators to observe the video information with the naked eye, which conducts manual analysis of the observed information one

by one, and finally makes a human judgment. This manual operation mode has many shortcomings [1, 2]. If a small number of personnel are put into the monitoring system for management, the analysis results will be poor due to objective factors such as mental fatigue and visual fatigue. If a large number of staff are put into management, this results in a large amount of additional management costs and divergence of opinions on the results of the video information due to the analysis and judgment of multiple people, which will increase the economic cost and time cost at the same time. The best-expected effect will not be obtained. In short, this kind of traditional manual video surveillance system only realises the function of video storage of surveillance pictures [3]. To achieve the purpose of video surveillance and supervision, it is necessary to add human observation and subsequent processing [4]. It takes up human resources at the same time. It cannot effectively meet the safety

precaution effect of people's advance warning, so this traditional monitoring mode can no longer meet people's growing needs.

Moving target detection is a branch of image processing and computer vision. It has great significance in theory and practice and has been paid attention to by scholars for a long time. In practice, using a camera to monitor a specific area is a meticulous and continuous process, which can be completed by humans, but it is not reliable for humans to perform such long-term boring routine monitoring, and the cost is also high. Therefore, it is necessary to introduce exercise monitoring. The background difference method is currently the most commonly used method in motion detection. It uses the difference between the current image and the background image to detect the motion area. It can generally provide the most complete feature data, but it is particularly sensitive to changes in dynamic scenes, such as interference from lighting and extraneous events. This algorithm first selects the average of one or several images in the background as the background image and then subtracts the current frame of the subsequent sequence image from the background image to eliminate the background. If the number of pixels obtained is greater than a certain threshold, it is determined that there is a moving object in the monitored scene, and the moving target is obtained.

Computer intelligent video surveillance is an emerging application direction and a cutting-edge topic in the field of computer vision; it is placed in various important facilities, such as museums, airports, subway entrances and exits, or outdoor traffic intersections. A large number of cameras monitor the lens to achieve the purpose of controlling and recording accidents, but at the same time this kind of monitoring can only record the accidents that have occurred. For example, an ordinary camera placed in a parking lot does not have an automatic alarm function when the security personnel are not in front of the surveillance camera and theft occurs. This requires the security personnel to exist in front of the surveillance camera at all times to judge and analyse possible criminal acts at the same time. We also know that it is impossible for security personnel to concentrate on guarding in front of the camera during the hours; even if the group is replaced on duty, it is impossible to be in front of dozens of surveillance cameras, and every surveillance camera will not be missed every minute. In response to this situation, it is necessary to combine computers to help security personnel to achieve the purpose of intelligent surveillance. Computer vision and applied research scholars put forward the concept of a new generation of video surveillance in a timely manner. Intelligent video surveillance uses computer vision and video analysis methods without human intervention. Computer video surveillance system not only conforms to the future development trend of the information industry, but also represents the future development direction of the surveillance industry. It contains huge business opportunities and economic benefits and is highly valued by academia, industry, and management.

In order to solve the shortcomings of traditional video surveillance systems, researchers began to try to combine

computer vision related technologies with video surveillance systems. By using cameras instead of human eyes for monitoring and computers instead of human brains for observation and analysis, a complete set was formed. The so-called intelligent video surveillance system refers to the intelligent processing of the video information obtained by the camera through computer vision related technology, then the moving target detection of the obtained image information is used to extract useful information, and then the target information undergoes behaviour analysis, and finally a corresponding response is made according to the preset rules. In response to the above problems, this paper proposes a background differential target detection method based on RGB colour separation for training target monitoring, which improves the effectiveness of the colour separation method.

2. Related Work

The background difference algorithm is a commonly used video target detection algorithm and an important part of modern video surveillance applications. The background difference algorithm usually requires the background to remain stable without drastic changes, so it is mainly suitable for foreground segmentation in the video image collected by a fixed camera. Because of its fast calculation speed and high segmentation accuracy, it has been widely used in video surveillance fields such as road monitoring and regional security. This section mainly introduces the basic principles and classification of background difference, as well as the problems and difficulties that need to be solved in the algorithm.

The first step in the process of video vehicle detection is to take a video image through the digital image acquisition system and save it in the memory or frame buffer of the hardware device, which performs preprocessing on the digital video image to reduce the impact of small interference noise on the quality of image detection results. It has important applications in video surveillance, human-machine interface, video coding, and other fields. At present, the most researched motion detection methods include frame difference method, background subtraction method, and optical flow method [5]. Although the frame difference method can effectively remove the static background, the extracted target is often rough and larger than the actual moving target contour, and there will be holes in the target. The background subtraction method requires high background modelling and background update. Once the background structure is completed, the coarsely segmented moving area can be finely segmented by the background subtraction method to detect the precise moving target. Some scholars combine colour and shape information to detect circular areas in complex scenes. The literature [6] adopts the method of colour separation to detect the direction of traffic and pedestrian movement at traffic intersections. The literature [7] combines contour information and colour information to detect targets with specific contours. Traffic vehicle detection carried out according to a certain method so that the information of road vehicles can

be extracted from the image area is an important issue discussed in the current vehicle detection research [8]. Among the mentioned methods, the background difference is a current method in video vehicle detection. The background difference method needs to compare the real-time input frame image and the current background model. The accuracy of the background model will directly affect the overall detection effect, and it needs to make adjustment in time to update the background according to actual changes. Therefore, it is necessary to use a background update and extraction method that not only can meet the real-time requirements of vehicle detection, but also has good robustness under the interference of various situations. Therefore, an effective and stable background model is already the most effective method of background difference.

3. Difference Method to Construct Moving Target Detection

3.1. Background Difference Algorithm. The background difference algorithm continuously detects the difference between the current video image and the reference image. The reference image here is what we usually call the “background image” or “background model.” In an ideal state, both the surveillance camera and the background remain stationary, and the difference between the current image and the background model is close to zero. However, in reality, because the scene is constantly changing, it is necessary to preset a threshold as the criterion for judgment [9]. The area is determined to be the foreground; otherwise, it is the background. Assume that the input video image sequence is $I_i^v(x, y)$.

$$Dif_i^v(x, y) = I_{i-1}^v(Vx, Vy) - B_{i-1}^v(Vx, Vy). \quad (1)$$

Here, $I_i^v(Vx, Vy)$ represents the current video image, $B_i^v(Vx, Vy)$ represents the background model corresponding to the current image, and $Dif_i^v(x, y)$ is the difference result. Perform a binary operation on $Dif_i^v(x, y)$; that is, judge and classify each pixel in the image according to the threshold value, and the binarisation result of the current image can be obtained, as shown in

$$R_k^{\text{Binarization}} = \begin{cases} 1, & Dif(x, y) > T, \\ 0, & \text{others,} \end{cases} \quad (2)$$

where $R_k^{\text{Binarization}}$ is the result of binarisation, $R_k^{\text{Binarization}} = 1$ represents the foreground target, and T is the critical threshold for judging whether the pixel belongs to the foreground target, which can be preset or used in the monitoring process for continuous adjustment through algorithms. Finally, the size of the foreground area is judged [10]. If the area of the target area is larger than a certain value, it is the real foreground target; otherwise, it is an interference item.

Background difference algorithms are very active in the field of moving target detection, and many algorithms have been proposed in recent years [11]. The specific processes of these algorithms are roughly the same, which can be specifically represented in Figure 1, where T represents the

current moment, N is the number of image frames used to initialise the background model, and $I_i^v(t)$ and $B_i^v(t-1)$, respectively, represent the video image obtained by the algorithm at time T and its corresponding background image. The detailed description of the basic flow of the background difference algorithm is as follows.

3.1.1. Pretreatment. In the application of intelligent video surveillance, since the video image is interfered by various noises during the collection process of the surveillance camera, the quality of the obtained image information is low, which is not conducive to the subsequent work. Therefore, it is necessary to use the corresponding image processing technology to denoise the original data to obtain a more pure and high-quality image [12]. Commonly used image preprocessing methods include median filtering and RGB difference filtering.

3.1.2. Background Modelling. Background modelling is the most important link in the background difference algorithm, and the quality of the model has the most direct impact on the effect of target detection. The simplest method of background modelling is to use a video image that does not contain any foreground objects as the background. This method is less computationally intensive and suitable for real-time video target segmentation, but the disadvantages are also obvious. In reality, the background is always complex and diverse. Constantly changing, a single background image cannot be guaranteed to be suitable for all video images. Therefore, in most cases, a specific background description method is used to construct a background model, such as a RGB difference background model based on statistical probability and a background model based on principal component analysis. In background modelling, more features are used to describe the background, and the adaptability to the monitoring environment is stronger.

3.1.3. Foreground Detection. The background difference algorithm usually obtains the foreground target by continuously detecting the difference between the current video image and the background model, which is the so-called foreground detection stage [13]. At this stage, the focus of our attention is how to judge whether the current pixel matches the background model. Normally, we set the corresponding threshold according to the built background model. The size of the threshold can be set in advance based on experience or can be automatically calculated by an algorithm according to different scenarios. The former is simpler but not very adaptable to the scene. Although the latter increases the time complexity of the algorithm, it has better adaptability to dynamically changing scenes.

3.1.4. Postprocessing. The postprocessing is to refine the results obtained from the foreground detection; the purpose is to obtain a more accurate and complete target area. Because the image noise cannot be completely removed in the image preprocessing stage, there are often many noises in the

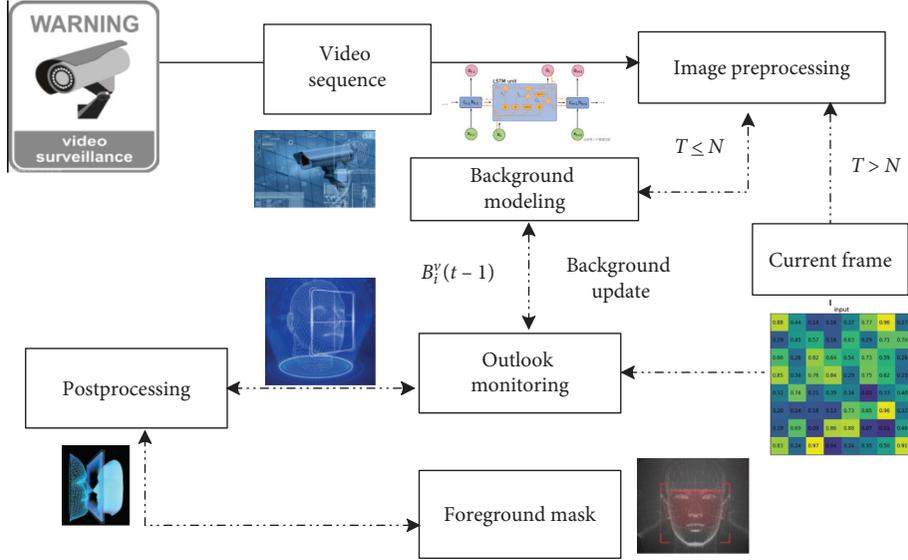


FIGURE 1: The structure of the scene difference algorithm.

results of foreground detection, and some characteristics of the current scene target are similar to the background, if the background difference algorithm cannot distinguish between them. It will produce holes in the foreground target, resulting in incomplete target detection. In the post-processing, we usually use image morphology methods, hoping to eliminate the remaining noise in the detection results, smooth the edges of the image, and at the same time fill in these void areas to obtain relatively complete results.

3.1.5. Background Update. During the video surveillance process, since the scene may change gradually, the algorithm needs to update the parameters of the background model according to these changes to adapt to the changing surveillance scene. It can update part of the background based on the result of foreground detection or update the background of the entire image based on statistical principles. The background update rate is a parameter index used to measure the background update speed. If the update rate is too large, it is easy to incorporate unstable factors such as noise in the scene into the background model. If the update rate is too small, it cannot keep up with the real-time changes of the monitoring scene. Therefore, the selection of the background update rate is very critical. An appropriate update rate can not only adapt to the real-time changes of the scene, but also effectively prevent some nonbackground targets from being updated to the background and thus affecting the performance of the algorithm.

3.2. Training Target Detection. Usually, the colour feature and location area of the target are used to achieve target detection. However, because the image background is complex, the target and the background are interlaced, and the target may be moving or stationary; the background may

also change. The flowchart of the training target detection is shown in Figure 2.

- (1) Extract a video obtained by monitoring, and process this video according to the background difference algorithm [14]. First, establish a corresponding model for the background. Here, we build a model for each pixel through the statistics of the acquired image sequence. Differentiate the current frame image and the previously constructed background image, and set $B_k^v(Vx, Vy)$ as the background model at time $F_k^v(Vx, Vy)$ as the video image at time as $Dif_k^v(x, y)$. The target area is $Dif_k^v(x, y) = F_k^v(Vx, Vy) - B_k^v(Vx, Vy)$.

Set a threshold T . When $Dif_k^v(x, y)$ is greater than T , it is determined as a moving area, when $Dif_k^v(x, y)$ is smaller than T , it is determined as a background area; that is, there is no moving target in this area. In order to better distinguish the background image and the moving area, the pixels of the moving area are assigned a value of 0, and the pixels of the background area are assigned a value of 1, which is a binary image.

- (2) The acquired video sequence is processed according to the algorithm of multiframe difference. $F_k^v(Vx, Vy)$, $F_{k-1}^v(Vx, Vy)$, and $F_{k+1}^v(Vx, Vy)$ are, respectively, k time, $k-1$ time, and T time. For the video image at time 1, apply the two-frame difference algorithm H to these three frames, respectively, to obtain the difference images $Dif_k^v(x, y)$ and $Dif_{k+1}^v(x, y)$, as follows:

$$\begin{aligned}
 Dif_{k-1}^v(x, y) &= |F_k^v(Vx, Vy) - F_{k-1}^v(Vx, Vy)|, \\
 Dif_k^v(x, y) &= |F_{k+1}^v(Vx, Vy) - F_k^v(Vx, Vy)|, \\
 d_k^v(x, y) &= |Dif_{k-1}^v(x, y) \cup Dif_k^v(x, y)|.
 \end{aligned} \tag{3}$$

This paper sets a threshold in the same way, which binarises the image to obtain a binary image.

- (3) Combine the two binary images obtained previously and use the following operations: $d_k^v(x, y) \cup Dif_k^v(x, y)$. After the previous series of processing, we generally cannot get a complete picture and will not destroy its original shape [15]. Morphology includes a series of algebraic calculations to deal with it accordingly. The core theory of mathematical morphology is set theory. In an image processed by mathematical morphology, the parts irrelevant to the target can be removed without destroying the original shape of the image. Morphology includes a series of algebraic operators, of which the four most basic are opening, expansion, closing, and corrosion. It is used to process binary images and can be used to process greyscale images. After the image is processed by mathematical morphology, some isolated noise points no longer exist, and small holes disappear [16]. However, some large holes still cannot be eliminated. Here, we need to use connectivity area detection. After this processing, the holes have disappeared, and the moving target we obtained is complete.

4. Framework of the Monitoring System for Training Targets

This article proposes a training target-oriented RGB colour separation of background difference target detection, which improves the effectiveness of the colour separation method. The concept of the training target monitoring system is to accurately segment and extract the training targets in the monitoring area by setting up a monitoring camera above the training targets and count their number and area for subsequent tracking and recognition processing, regardless of whether the area needs cleanup judgment or decision-making [17, 18]. The hardware components of the system include network cameras, routers, display terminals, storage servers, and other equipment to complete video collection, transmission, data processing, distribution, decision-making and execution, and archiving. The system hardware architecture is shown in Figure 3.

From a system engineering perspective, this system consists of a data collection layer, a data processing layer, and a data display layer. Subdivided from the perspective of specific functional modules, a complete intelligent video surveillance system can be divided into the following modules: front-end video acquisition module, image preprocessing module, background differential target detection module, target tracking and counting module, and display terminal and data log module. The overall design framework of the system is shown in Figure 4.

This section will give a detailed introduction to the functions of each module.

4.1. Front-End Video Capture Module. This module is responsible for acquiring image data from cameras or video files. Video data acquisition is mainly through network cameras. The acquired images are video-encoded and transmitted to the processing host via the network. Due to the uncertainty of the transmission network, there may be frame breaks in the video due to network delays. A video buffer is added to this module, and 30 frames of video images are taken as the size of the buffer. While taking into account the real-time requirements of the system, solutions are provided for possible network problems.

4.2. Video Preprocessing Module. This module first performs greyscale transformation on the acquired image to reduce the calculation amount of the algorithm and at the same time uses RGB difference filtering to blur the image to eliminate the noise caused by various reasons and improve the image quality. Subsequently, the video image and the background model are based on the RGB difference algorithm, so that the background difference algorithm can be used to segment the training target.

4.3. Background Differential Target Detection Module. This module is mainly responsible for judging whether the training target appears in the monitoring area, and if so, it needs to be accurately located and segmented. This module mainly includes the following aspects: First, it establishes a background model based on the RGD colour difference algorithm and completes it. It performs preliminary detection of the training target, then removes the interference of the light and shadow area on the target in the HSV space, and smooths the target contour through the image morphology method. Finally, it completes accurate segmentation of the training target for later statistics and tracking.

4.4. Target Tracking and Counting Module. When the training targets are successfully detected and divided, this module will track them and count their number and area so that the monitoring personnel can make judgments. Regarding the tracking of training targets, there are already relatively mature algorithms. Based on the background difference algorithm for foreground detection, this article is supplemented by the *Kalman* filter algorithm that comes with OpenCV to realise the tracking and statistics of training targets.

4.5. Display Terminal and Data Log Module. The display terminal must display both the current video image information and the statistical results of the training targets. When the number of training targets exceeds a certain threshold, an alarm is issued to the monitoring personnel and the information at this time is written into the monitoring log.

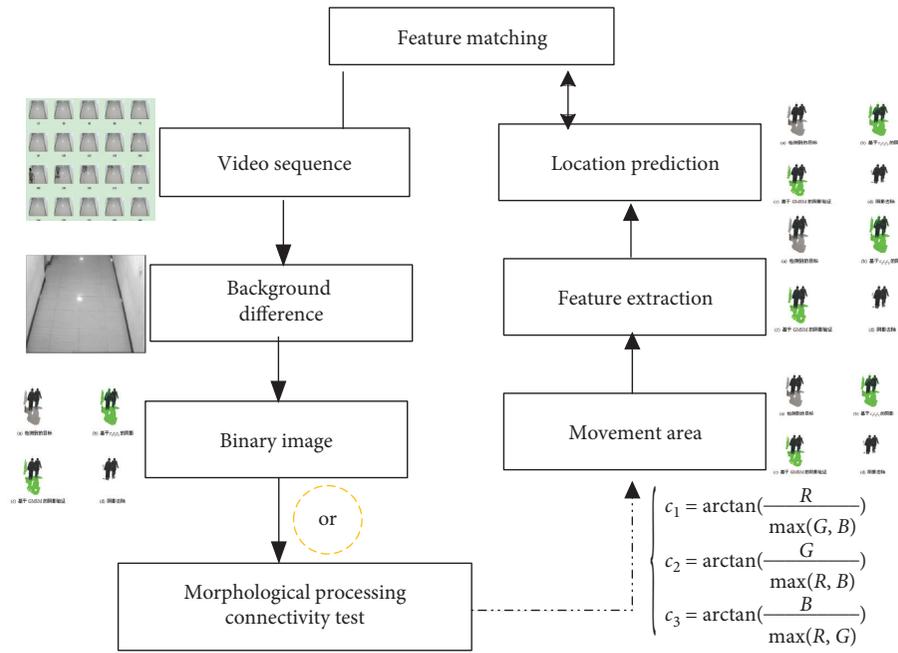


FIGURE 2: The flowchart of the training target detection.

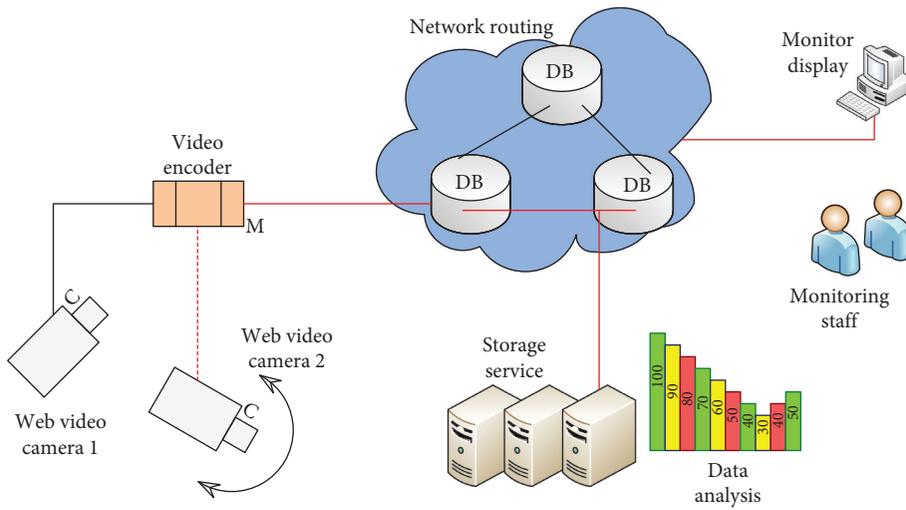


FIGURE 3: The system hardware structure.

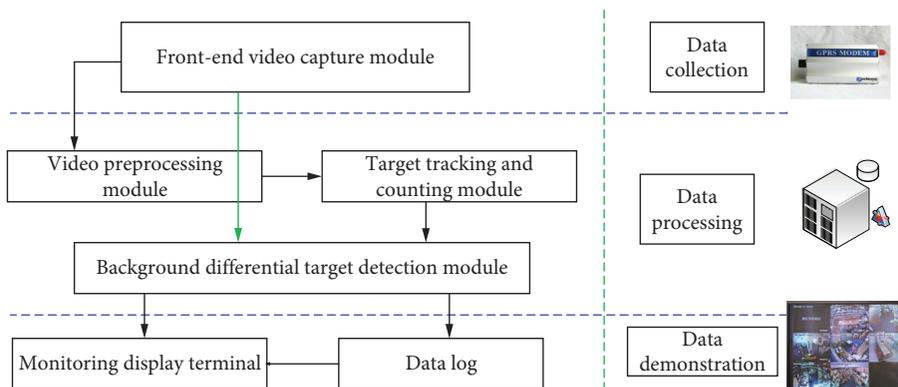


FIGURE 4: The overall design framework of the monitoring system for training targets.

5. RGB Colour Separation Background Differential Target Detection Method

The traditional background difference method is to convert the colour image into a greyscale image, using the foreground image to get the target method, such as (1). After the differential image is obtained, in order to eliminate the noise, the image needs to be binarised according to the set threshold, as shown in (2). The colour information in the image will be greatly reduced. It can be seen from (3) that different colour component combinations may convert the same grey value. If the threshold is considered, the point on the remote target must be judged once it has a background color [19]. If the converted grey value is the same as the background or within the range allowed by the threshold, it is impossible to detect the difference of the point in Figure 5.

Assume the following situation: a threshold value has been given $\theta(\theta > 3\xi(x, y))$ for a certain point (x, y) , $I_r^v(x, y) = a$, and $\theta > a > \theta_r$, $I_G^v(x, y) = 0$, $I_B^v(x, y) = 0$. However, at the same time, the noise at this point is very small or 0; that is, $\xi(x, y)$ is 0 or small and can be ignored; then, according to (7), $Dif_k^v(x, y) = a$. Because the threshold is 0 and so $I_r^v(x, y) = 0$, this point will be mistaken for noise and eliminated. In fact, the difference has been detected under the R primary colour [20]. According to this method, this difference cannot be reflected in the result, and the result is no longer complete.

$$\begin{aligned} Dif^v(x, y) &= Dif_R^v(x, y) + Dif_G^v(x, y) + Dif_B^v(x, y), \\ Dif^v(x, y) &= I_R^v(x, y) + I_G^v(x, y) + I_B^v(x, y) + 3\xi(x, y). \end{aligned} \quad (4)$$

In order to explain the problem more clearly, the above situation is expressed in a matrix form, as shown in Figure 6.

It can be seen from Table 1 that when the traditional method takes a smaller threshold, it does not have a good suppression effect on noise. Under the same threshold, the method in this paper has a more effective suppression effect on noise. When the threshold of the traditional method is very large, the suppression effect on noise is obvious, and the suppression effect on the target pixel will be great, resulting in incomplete detection. The threshold of the traditional method has increased to 120, and the threshold of the method in this paper is unchanged. The two methods are tested again, and the number of detected target pixels is used to evaluate the effect of moving target detection.

It can be seen from the experimental results that when the threshold is too high, the detection result of the traditional method is compared with the result of the method in this paper, many pixels are reduced, and the result is obviously incomplete.

6. Experiment and Result Analysis

In this paper, the theoretical knowledge of video background extraction algorithm and foreground detection algorithm is improved, and the experimental effect before background difference is compared with the improved effect. The evaluation of any kind of algorithm must have actual data to

illustrate the effect of the improvement. The experimental environment and platform used in this article; the parameters selected in the modelling process; the model's objective evaluation index; the algorithm accuracy, recall rate, etc. indicators; the traditional target detection algorithm, and the improved RGB separation algorithm in this paper are used for experimental simulation and data comparison analysis.

7. Experimental System and Platform

The experimental environment of this article is Visual Studio 2018 Professional Edition with OpenCV 6.2 library, the hardware platform is Intel® Core™ i7-6730, the main frequency is 4.75 GHz, the memory is 16G, and the operating system is Windows 64 bit. OpenCV is an open source platform of computer vision library that can run on the operating system. The library contains a large number of image processing algorithms and related data structures. The algorithms inside have a wide range of applications and are highly portable. This article uses the public dataset CDnet 2017, which includes a series of test videos in different scenes, such as the background in a still scene. There is no foreground target in the first frame of the scene. In addition to the static background, there are dynamic interference changes in the scene [21]. This article selects some suitable scenes for experimental simulation. During the experiment in this paper, the values of some parameters used are as follows. When extracting the video background image, the grey value interval coefficients are all 0.48, 0.31, and 0.2. During the modelling process, the number of pixels in background model N is 20, the random update factor is 16, the adaptive increase factor a is 0.7, the adaptive reduction factor is β 0.3, and the initial value of the matching radius of the front background judgment is $R = 20$.

8. Performance Evaluation Standards

In order to verify that the improved algorithm in this paper solves the smear problem in the detection result, the foreground target is present in the first frame of a video and continues to exist until the last frame. The entire Road video has 100 frames. Figure 7 shows the original video image, as well as the result of foreground target detection by RGB difference mixture model, the original foreground target detection result based on visual background modelling, and the improved extraction result of the background image initialisation background model proposed in this paper.

It can be seen intuitively from the above experimental results that there are foreground targets in the first frame of the Road video. In the RGB difference mixture model, when the shaking leaves pixels change greatly, such pixels will be detected as foreground targets. As a result, there will be a large number of false detection events in the detection result, which contains more noise points, and the detection result is not accurate. When the background model is initialised based on the visual background modelling method, the foreground target area in the first frame image is modelled as the background area, so the subsequent background matching process of the image to be tested will cause

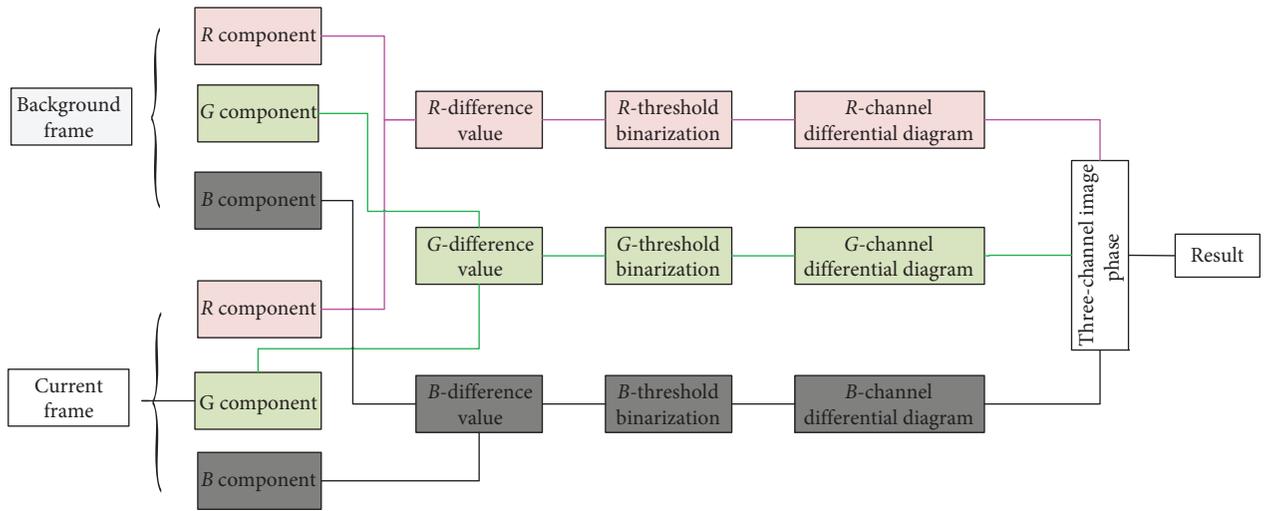


FIGURE 5: The overall design framework of the monitoring system for training targets.

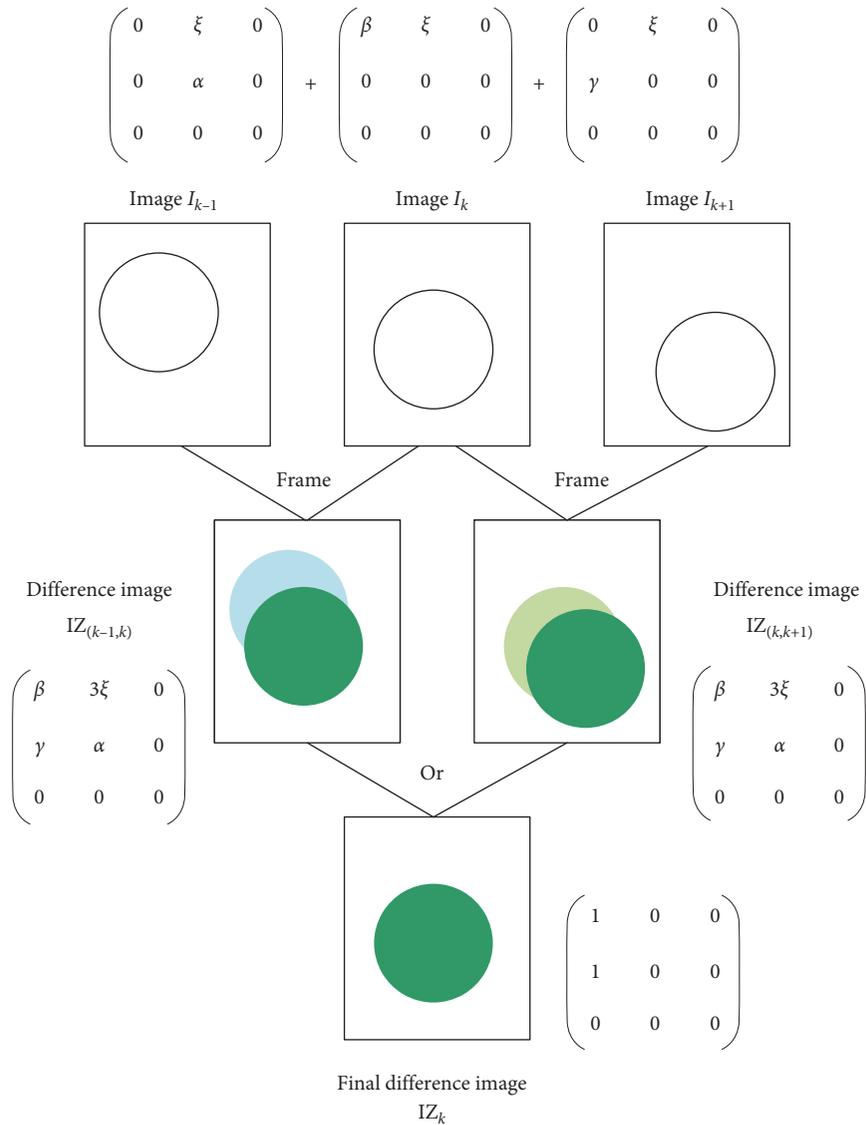


FIGURE 6: The schematic diagram of matrix transformation effect.

TABLE 1: Number of noise pixels under different thresholds.

Threshold Θ	30	60	90	120
Scene 1	380	10	5	0
Scene 2	3450	2800	350	0
Scene 3	760	250	20	2
Mean value	1539.3	1068.9	125	0.7

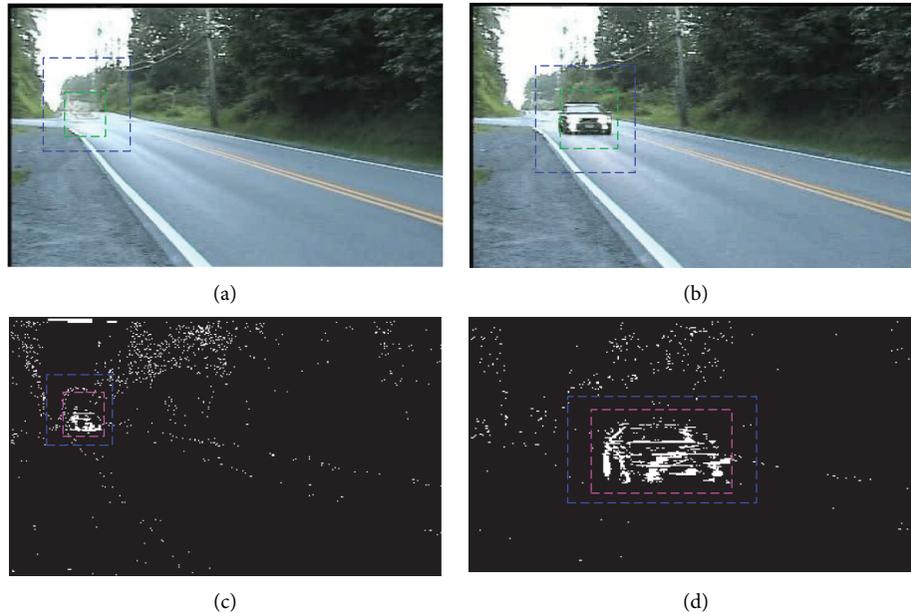


FIGURE 7: The experiment results of Road scene: (a) background image; (b) frame 50 image; (c) binarised image; (d) background differential binary image.

misjudgment. The smear phenomenon in the figure makes the detection result inaccurate. Although the improved visual background modelling method has a partial improvement in the accuracy of the detection result compared with the visual background modelling method, there is still smear overall.

The above is an experimental comparison of target detection when the first frame of video contains foreground targets. The actual situation is often target detection in complex scenes. It can reflect the efficiency and robustness of the algorithm itself. This paper selects three different types of video sequences under complex background scenes for experimental simulation and comparative analysis, including changes in light illumination, water wave shaking under sunlight, and leaf fluttering.

The first category includes a complex background video sequence with light changes. The video Office has 80 frames, there is a moving foreground target in the first frame of the video, and the foreground target contains two people. Figure 8 shows the experimental simulation results of the video sequence in the RGB difference mixture model, the original visual background model algorithm, and the improved visual background model. According to the experimental results, when the RGB difference mixture model processes the video in this scene, the detection result in the first frame is poor, the foreground target has not been

detected at all, and the detection result of the 42nd frame contains many lights. Noise points and the ghost image of the 1st frame are detected at the same time, which will make the detection accuracy lower. In the detection result of the original visual background model algorithm, the foreground target in the first frame is incomplete and contains holes; the detection result shows ghosting phenomenon and a large number of light noise points. It can be seen that the algorithm has greatly reduced the impact of flickering lights, but the processing is not thorough enough, and the improved visual background model algorithm does not remove ghost images. We can initialise the background model and add an adaptive threshold mechanism in the background matching process to distinguish background pixels, light changes pixel points, and foreground target pixels, thereby improving the detection accuracy. A relatively complete foreground target can be extracted from the processing result of the first frame, the 42nd frame does not contain light noise points and ghost phenomenon, and a complete foreground target can be obtained.

The detection effect based on RGB difference mixture modelling has been reduced when entering the foreground target detection in the complex dynamic scenes containing swaying leaves, swaying water ripples, and changes in the intensity of sunlight in the video scene. The outline of the foreground target is detected, but there are some small false

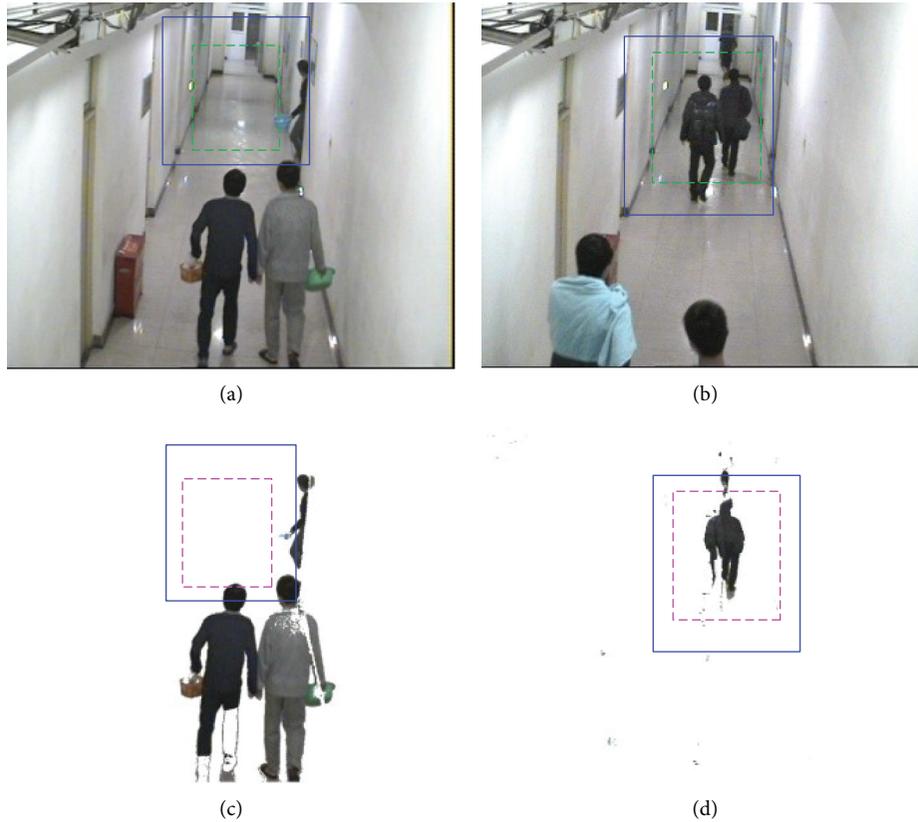


FIGURE 8: The experiment results of complex background scene: (a) first frame image; (b) 50th frame image; (c) binarized image; (d) background differential binary image.

phenomena in the target, and the original extraction result based on visual background modelling contains a lot of noise. It is obvious that the dynamic water ripples in the scene are misdetected as the foreground target, because the background modelling algorithm sets the matching radius threshold to a global fixed value, which is no longer applicable in this dynamic scene. The improved algorithm of matching threshold can effectively identify the dynamic background in the scene and detect it as the background area, which greatly reduces the noise interference. When there is a foreground target in the first frame of the video, it can effectively avoid dragging in the extraction result. It can be seen from the experimental results that the algorithm can detect a relatively complete foreground target in a complex environment and improve the detection performance.

9. Comparison and Analysis of Experimental Results

According to the video selected in this article, foreground target detection is performed based on the existing improved visual background modelling method. According to the quantitative index precision rate, recall rate, and comprehensive evaluation proposed in the previous section, from Table 2, it is known that in terms of accuracy indicators, RGB difference modelling has the lowest accuracy among the three videos of Road, Office, and Riverside, because the

original background pixels in the video have been detected as foreground pixels, resulting in the smallest precision value, while the visual background modelling method in Canoe video has the lowest precision. A large number of shaking leaves pixels in this method will be mistakenly detected as foreground pixels, resulting in minimal precision values. Compared with the previous two methods, the detection result of the improved visual background modelling method is slightly higher in accuracy, because it adds a scintillation point detection mechanism, which to a certain extent avoids the false detection of the flicker point as the foreground target pixel. This paper extracts the pure background as the initial value of the background model and adds an improved algorithm based on adaptive matching threshold to identify the dynamic background in the scene.

From Table 3, it is known that the recall rate of RGB difference modelling method in the four videos is the lowest among the four algorithms. Because the pixels that were originally foreground targets are misdetected as background pixels during the processing of RGB difference modelling method, this leads to the detection result, the foreground target has a hole phenomenon. Although the visual background extraction modelling method has reduced the hole phenomenon, the recall rate index has been improved. The improved visual background extraction modelling method is still in the first frame of the video sequence containing the foreground target. There is a void phenomenon in the foreground target, so the recall rate cannot be fully

TABLE 2: Accuracy comparison.

	Traditional method	RGB background difference
Scene 1	0.87	0.92
Scene 2	0.79	0.88
Scene 3	0.86	0.92

TABLE 3: Recall rate comparison.

	Traditional method	RGB background difference
Scene 1	0.64	0.82
Scene 2	0.59	0.89
Scene 3	0.76	0.86

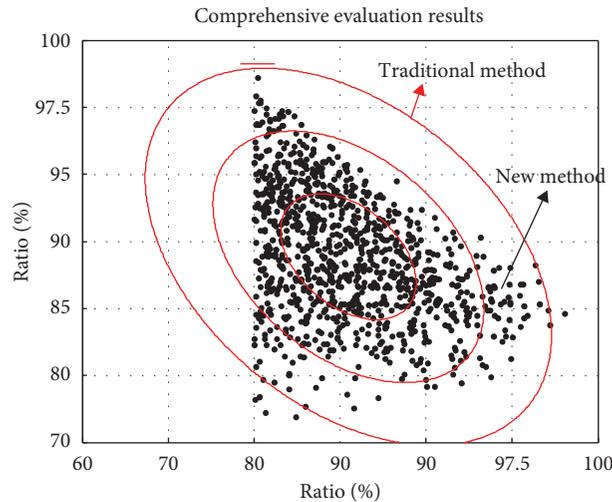


FIGURE 9: The experiment results of complex background scene.

improved. While the algorithm in this paper extracts an image that does not contain the foreground target, the final extracted foreground target avoids the void phenomenon and effectively improves the call rate index.

The comprehensive evaluation index is one of the important indexes to measure the overall performance of the algorithm experiment results. This article calculates the comprehensive evaluation index for the four algorithms through the processing results of the three videos, as shown in Figure 9.

The comprehensive evaluation index is the harmonic average of accuracy and recall. It can be seen from Figure 9 that the comprehensive index is the smallest among the four algorithms, which shows that the RGB difference modelling method has the worst effect on the foreground target extraction of the four videos. Compared with the RGB difference modelling method, the comprehensive index of the visual background modelling method has partially improved, but the improved visual background modelling method improves the performance by adding the scintillation point mechanism and increasing the update factor and further improves the comprehensive index. The improved algorithm in this paper optimises the model by extracting pure background and adding adaptive threshold matching to detect more accurate foreground targets. Through the

calculation of four video comprehensive indicators, the comprehensive indicators of the algorithm in this paper are higher than those of the other three algorithms. Through the comparison of accuracy, recall rate, and comprehensive index, this algorithm has obvious advantages compared with RGB difference modelling method, visual background modelling method, and improved visual background modelling method, and the overall performance is better than that of the other three algorithms.

10. Conclusion

This paper proposes an intelligent background differential model design method for training target monitoring, which overcomes the shortcomings of colour separation and differential extraction of targets that are not refined, which solves the difficulty of background structure update in the background subtraction method. This method can accurately detect motion in the surveillance scene, which is insensitive to light changes and background interference. At the same time, it can be seen from the analysis of the algorithm process that although the detection is divided into two steps: coarse segmentation and fine segmentation, the fine segmentation is only performed on some areas, so time consumption is small. By studying the principle of the

background difference method, the background construction principle based on RGB colour separation is analysed. The *OpenCV* programming tool is used to realise the capture and labelling of the moving target in the training target monitoring video. It provides a better detection method for the motion targets in dangerous areas and personnel in key production areas. This article has conducted in-depth research on several key technologies in intelligent video surveillance and has achieved certain results. In fact, training target information in video images is a powerful basis for detection. Combining motion information for detection will undoubtedly be beneficial to target judgment and improve the robustness of the detection algorithm.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] J. Xu and N. Liu, "Research on closed loop supply chain with reference price effect," *Journal of Intelligent Manufacturing*, vol. 28, no. 1, pp. 51–64, 2017.
- [2] M. S. Rahman, A. K. Manna, A. A. Shaikh et al., "An application of interval differential equation on a production inventory model with interval-valued demand via center-radius optimization technique and particle swarm optimization," *International Journal of Intelligent Systems*, vol. 35, no. 8, pp. 1280–1326, 2020.
- [3] A. O. Akmandor, Y. I. N. Hongxu, and N. K. Jha, "Smart, secure, yet energy-efficient, internet-of-things sensors," *IEEE Transactions on Multi-Scale Computing Systems*, vol. 4, no. 4, pp. 914–930, 2018.
- [4] M. Devarajan, V. Subramaniaswamy, V. Vijayakumar et al., "Fog-assisted personalized healthcare-support system for remote patients with diabetes," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 10, pp. 3747–3760, 2019.
- [5] M. Chen, Y. Zhang, M. Qiu et al., "Smart personal health advisor based on deep analytics," *IEEE Communications Magazine*, vol. 56, no. 3, pp. 164–169, 2018.
- [6] R. Punj and R. Kumar, "Technological aspects of WBANs for health monitoring: a comprehensive review," *Wireless Networks*, vol. 25, no. 3, pp. 1125–1157, 2019.
- [7] L. Liu, Y. Jiao, and F. Meng, "Key algorithm for human motion recognition in virtual reality video sequences based on hidden markov model," *IEEE Access*, vol. 8, no. 1, pp. 159705–159717, 2020.
- [8] J. Pagan, R. Fallahzadeh, M. Pedram et al., "Toward ultra-low-power remote health monitoring: an optimal and adaptive compressed sensing framework for activity recognition," *IEEE Transactions on Mobile Computing*, vol. 18, no. 3, pp. 658–673, 2018.
- [9] A. A. Uglu, O. Yilmaz, and A. R. J. Almusawi, "Development and control of shaped metal deposition process using tungsten inert gas arc heat source in additive layered manufacturing," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 232, no. 9, pp. 1628–1641, 2018.
- [10] L. X. Yan, "Research on impact factors extraction for driving mode of intelligent vehicle," *Zhongguo Gonglu Xuebao/China Journal of Highway and Transport*, vol. 31, no. 1, pp. 120–127, 2018.
- [11] P. Elamparithi and K. R. Soundar, "Trusted sensing model for mobile ad hoc network using differential evolution algorithm," *Information Technology and Control*, vol. 49, no. 4, pp. 556–563, 2020.
- [12] Z. Zhong, "A general moving detection method using dual-target nonparametric background model," *Knowledge Based Systems*, vol. 164, no. 15, pp. 85–95, 2019.
- [13] N. V. Munts and S. S. Kumkov, "A numerical method for solving time-optimal differential games with a lifeline," *Automation and Remote Control*, vol. 81, no. 8, pp. 1545–1561, 2020.
- [14] T. Saheb, "Big data analytics in the context of internet of things and the emergence of real-time systems: a systematic literature review1," *International Journal of High Performance Systems Architecture*, vol. 8, no. 1-2, pp. 34–50, 2018.
- [15] F. Jungang, M. Xiangyang, and J. Xiangfu, "Double differential cross sections for ionization of H by 75 keV proton impact: assessing the role of correlated wave functions," *Chinese Physics B*, vol. 29, no. 12, pp. 93–102, 2020.
- [16] U. Claudio, F. Garrido, and J. Kern, "Design and implementation of intelligent agent training systems for virtual vehicles," *Sensors*, vol. 21, no. 2, pp. 1–25, 2021.
- [17] D. Mišić, M. Zdravković, M. Mitković et al., "Real-time monitoring of bone fracture recovery by using aware, sensing, smart, and active orthopedic devices," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4466–4473, 2018.
- [18] W. Yang, "Research on quantitative evaluation of innovation capability of intelligent grid industry cluster based on BP NNA-taking Jiangsu power grid industry cluster as an example," *Revista de la Facultad de Ingenieria*, vol. 32, no. 8, pp. 726–734, 2017.
- [19] S. Kurokawa, T. Nagatsuma, T. Harada et al., "Image processor and image processing method," *Journal of the Acoustical Society of America*, vol. 29, no. 05, pp. 56–61, 2017.
- [20] T. Brunye Tad, S. B. Martis, H. Carlene et al., "Visual saliency and biological motion interact to determine camouflaged target detectability," *Applied Ergonomics*, vol. 73, no. 2, pp. 1–6, 2018.
- [21] C. Keyang, S. Wenxi, Z. Bowen et al., "Robust KCF pedestrian tracking method with complex scene," *Journal Of Nanjing University Of Aeronautics & Astronautics*, vol. 51, no. 5, pp. 625–635, 2019.