Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 1013714, 7 pages
https://doi.org/10.1155/2022/1013714

Research Article

RBF-Based 3D Visual Detection Method for Chinese Martial Art Wrong Movements

Xi Wang,[1] Yi-Hsiang Pan,[1] Zongbai Li,[2] and Bing Li[3]

[1] *Graduate Institute of Physical Education, National Taiwan Sport University, Tao Yuan, 333325 Taipei, China*
[2] *College of History and Culture, Hunan Normal University, Changsha, 410081 Hunan, China*
[3] *College of art, Design, & Physical Education, Chosun University, Gwangju 61452, Republic of Korea*

Correspondence should be addressed to Bing Li; libingwushu@poers.edu.pl

Received 6 March 2022; Revised 6 April 2022; Accepted 15 April 2022; Published 30 April 2022

Academic Editor: Jun Ye

The accuracy of action detection is limited by the extracted action, and there are problems of high processing complexity and low efficiency. Therefore, a three-dimensional visual detection method of martial art wrong action based on RBF is proposed. After noise reduction and weighting processing of martial art action video images, a martial art action 3D visual transformation model is established. According to the 3D visual model, C3D features are used to represent martial art actions. The video is segmented using sparse coding to determine the detection range. RBF neural network model is established, and the combination of the above 3D visual model and network parameters is obtained by sample training to detect martial art wrong actions. The test method of the experimental results shows the detection of the research under the condition of different degrees of precision, an average of at least 5%, and the method of detection of high efficiency and stability.

1. Introduction

Martial art is an ancient science in China's traditional sports, with attack action as the main content, routines, and combat as the main movement form, paying attention to the internal and external repairing of traditional ethnic sports [1]. As an excellent national traditional culture in China, Chinese martial art has formed its own unique expression and development means in the process of development and derivation for thousands of years. At present, the teaching of Chinese martial art at home and abroad mostly stays in the traditional way of face-to-face teaching or practitioners only follow video learning [2]. On the one hand, the inheritors of Chinese martial art are scarce, and it is difficult for people to get access to authentic face-to-face teaching; on the other hand, there are problems such as poor intuition and low efficiency in learning skillful movements only by following videos, and practitioners are very likely to do wrong movements and cause muscle damage. Therefore, an effective human movement posture detection method can play a role in correcting wrong movements for athletes' regular training. Action detection is widely used in industrial production, daily safety behavior monitoring, social operation management, and other work areas, and scholars at home and abroad have made some research results. Ohl and Rolfs use causality in the human visual system to adapt to show the causal relationship linked between action directions. It is used as a key low-level feature of visual events to detect motion direction [3]. Reference [4] combines perceptual learning and statistical learning to improve information acquired through experience and achieve action detection by statistical cooccurrence between environmental features. Reference [5] uses the frame difference method to subtract the background to achieve effective detection of the slightest motion. Reference [6] uses the YOLOv4 deep-learning motion target detection algorithm to achieve localization and recognition of motion targets. Real-time image detection of pictures, videos, and cameras is achieved by identifying and tagging the location and type of objects contained in the image, which improves the accuracy and speed of detection. NagiReddy et al. proposed a novel background modeling mechanism using a bias illumination field fuzzy C-means algorithm to separate nonstationary pixels from stationary ones by background subtraction. Feature extraction under the condition of noise and illumination changes is completed by using the fuzzy C-means method of biased lighting field, and the detection accuracy is improved through clustering [7]. The above methods can achieve high-quality motion detection in different environments, but the overall detection performance of the method is degraded due to the strong coherence of martial art actions and the influence of the accuracy of extracting wrong action features when wrong actions occur.

The neural network can simulate some systems that cannot be described by mathematical models, and it has strong learning and adaptability, but it also has obvious nonlinear characteristics. The radial basis function (RBF) neural network belongs to the forward neural network; this kind of network to the structure of the multilayer forward network control is similar, and it is a kind of forward neural network with three layers structure [8]. The transformation function of neurons in the hidden layer refers to the radial basis function and is aimed at the center of radial symmetry and nonlinear function with an attenuation trend [9]. In order to better detect the wrong movements of Wushu, in future research, we should conduct in-depth research on the martial art movements with rapid changes in the movement connection, so as to improve the detection accuracy of the detection method. On the basis of the above analysis, this paper will study the three-dimensional visual detection method of martial art wrong action based on RBF, combined with the advantages of the RBF neural network, combined with the three-dimensional visual model to realize martial art wrong action detection and test the comprehensive performance of this method.

2. Research on RBF-Based 3D Visual Detection Method for Martial Art Error Movements

2.1. Martial Art Action Video Image Processing

2.1.1. Video Image Preprocessing. The video image collection of Chinese martial art action is mainly completed by a high-resolution color camera. The image with a continuous signal taken in photogrammetry is the analog image, and its two-dimensional function is represented by $p(x, y)$. Any (x, y) in the image can be used as the two-dimensional coordinate point here. In the intake of the original image, there will be noise interference before certain processing, in the filtering process, and will lose part of the details of the picture, so in the process of noise; at the same time, we need to do our best to ensure the quality of the original picture. In this study, the commonly used mean filter and median filter are used to process Chinese martial art action video images to make the images smooth [10].

In the actual mean filtering, a filtering template is also set based on a point pixel (x, y), which is composed of the remaining surrounding pixels except for this pixel. The average value can be calculated by using the pixels in this template to replace the value of this point pixel (x_i, y_i). The gray value $h(x, y)$ corresponding to this point in the digital image can be obtained. In this way, the pixel value of each position in the original image can be replaced by the mean value solved by the filter template, which is the basic principle in the application of mean filtering. The formula is as follows [11]:

$$h(x, y) = \frac{1}{n}\sum_{i=1}^{n} p_i(x_i, y_i). \tag{1}$$

In the above formula, n is the total number of pixels in the filtering template after the target pixels.

Median filtering was adopted while on noise suppression to eliminate the nonlinear smoothing technique, which uses the principle of order statistics, based on the different pixels to build a template. The selected pixel values in the template, sorted somewhere in the middle of pixel values to replace the target pixel gray value, thus reduce the image noise pixel gray value. In practical operation, the neighborhood around the target pixel is required to conduct size-sorting statistics of some columns according to gray value, and the median of the two-dimensional sequence is filtered, which is expressed as follows [12]:

$$h'(x, y) = Med\{h(x - q, y - w), (q, w) \in M\}. \tag{2}$$

In the above formula, M is defined as a two-dimensional template. The function Med is used to get the median of the entire two-dimensional sequence; $h'(x, y)$ represents the gray value of the image after processing, and $h(x, y)$ represents the gray value of the image after the previous step of mean filtering.

Using the method of spatial domain and frequency domain method, the two methods are used to describe the image-enhancement processing. In the spatial domain method, the gray value of each image point is directly processed to ensure the enhancement effect of the image. Specifically, taking a target pixel and its adjacent pixels as a whole, the pixels excluding the target pixel can be set as a template, so that the original gray value can be represented by the average value solved by the filtering template. The processing of spatial domain method is as follows [13]:

$$H(x, y) = H_T\left[h'(x, y)\right]. \tag{3}$$

In the above formula, H_T is a spatial operation with respect to h'; $H(x, y)$ is the Chinese martial art action video image after enhanced processing; $h'(x, y)$ is the Chinese martial art action video image after smooth noise reduction.

2.1.2. The 3D Visual Transformation Model of Chinese Martial Art Action Is Established. The difference between Chinese martial art action and the human body's routine action is that Chinese martial art action has certain obvious changes in three-dimensional space, and the detection of Chinese martial art wrong action also needs to be discriminated from the perspective of three-dimensional space. Therefore, combined with the collection of Chinese martial art movements, this paper will establish a three-dimensional visual transformation model of Chinese martial art movements. The 3D visual data of Chinese martial art movement is accomplished by using the 3D visual collector of line structured light. Based on the principle of optical triangulation, the optical projector projects the structured light onto the surface of the human body, and the camera captures the information of the light strip, so as to obtain the two-dimensional distorted image of the light strip. The degree to which the strip varies depends on the position of the light plane relative to the camera and the surface shape of the object. Since the brightness of the strip is obviously different from that of the unilluminated region, the two-dimensional coordinates of the strip in the camera image can be obtained by using a specific image processing method. The mathematical model of line structured light sensor is used to establish the mapping between the image plane coordinate system and world coordinate system. According to this model, the coordinates of points on the strip can be calculated according to the pixel coordinates of the image. The camera model, the basic imaging model, often referred to as the basic pinhole model, is given by a central projection transformation from three-dimensional space to plane. The relation diagram of the Chinese martial art action acquisition object in the coordinate system of the online structured light image collector is shown in Figure 1 [14, 15].

If the coordinate of the Chinese martial art movement acquisition object is $R = (x_s, y_s, z_s)$ in the spatial coordinate system, and the equation $\alpha x_s + \beta y_s + \gamma z_s + \theta = 0$ of the spatial plane where the imaging object is located in the collector coordinate system is known, then the spatial linear equation between the collector imaging spot center and the measured object is as follows:

$$\frac{x_s}{x_s - X_s} = \frac{y_s}{y_s - Y_s} = \frac{z_s}{z_s - Z_s}. \tag{4}$$

In the above formula, (X_s, Y_s, Z_s) is the coordinate position of the measured object on the collector imaging plane. By putting the equation of the space plane of the imaging object in the collector coordinate system into formula (3), the three-dimensional space coordinates of any point on the light plane in the camera coordinate system can be worked out, and the three-dimensional visual transformation model $H'(x, y, z) = \mathrm{new}(X_s, Y_s, Z_s)$ of martial art action can be established.

2.2. Generate Chinese Martial Art Error Action Fragments to Be Detected. In this paper, C3D features are used for video motion representation, and C3D features show excellent performance in motion recognition tasks. C3D features are generated by a 3D-CNN deep network. Compared with traditional features, C3D features can better represent the characteristics of action videos in time and space. Compared with 2D-CNN, 3D-CNN can better extract timing features of videos, which is very suitable for motion detection tasks. The C3D network consists of 8 convolution layers, 5 maximum pooling layers, 2 full connection layers, and one softening output layer. All the convolution layers use $3 \times 3 \times 3$ 3D convolution cores, the first pooling layer uses $1 \times 2 \times 2$ pooling cores, and the other pooling layers use $2 \times 2 \times 2$ pooling cores [16, 17].

After feature extraction in the C3D network, the visual dictionary needs to be established when sparse coding is used to generate action fragments to be detected. The establishment of the visual dictionary is based on the sample video set. For each video sample, the run-time space point of the interest detection algorithm is used to extract 3D HOG features at each detected point of interest location to obtain a feature vector. The feature vector cannot be directly used as a visual word due to its high dimension and large variance, so it needs to be quantitatively processed. Moreover, as the feature quantity of the whole sample set is very large, the calculation is usually carried out on its subset [18]. All feature vectors extracted from all video sets are taken as a set, and a subset is obtained by random sampling. The clustering algorithm is performed on the subset to obtain K categories. The center of each category is computed as the visual words of that category, which constitute the dictionary of visual features on this data set.

The traditional sparse coding method of multiple dictionaries is used, and the basic sparse dictionary learning method is used to learn each dictionary. T represents the feature of the action fragment used to train the dictionary, and ZD represents the dictionary to be learned. Dictionary learning for each category uses the following formula [19, 20]:

$$(ZD, B) = \arg\min_{ZD,B} \frac{1}{b} |T - ZDB|^2 + \lambda B_t^2, \tag{5}$$

where B is the sparse representation coefficient and λ is the length of the sparse window. The learning process of a dictionary is the same as that of using a dictionary. In each iteration, the dictionary is updated by the fixed coefficient matrix first, then the coefficient matrix is updated by the fixed dictionary, and the result of minimizing formula (5) is finally obtained. Each dictionary learned was used to encode the candidate fragments. Formula (6) was used to calculate the reconstruction error of each dictionary, and the corresponding fragment score of each dictionary was calculated using the normalized formula. At this point, each candidate fragment has different scores from each category dictionary, and the final score of the candidate fragment can be obtained by calculating these scores [21].

$$B_m = \arg\min_{b_m} \frac{1}{b_m} |U_m - ZDB|^2 + \lambda B_t^2. \tag{6}$$

According to the above-obtained clip scores, the video clips that may contain wrong moves are selected using the correlation coefficient coding between Chinese martial art moves. After selecting the video clips that may contain wrong moves, the Chinese martial art move features in the images are extracted by combining the Chinese martial art

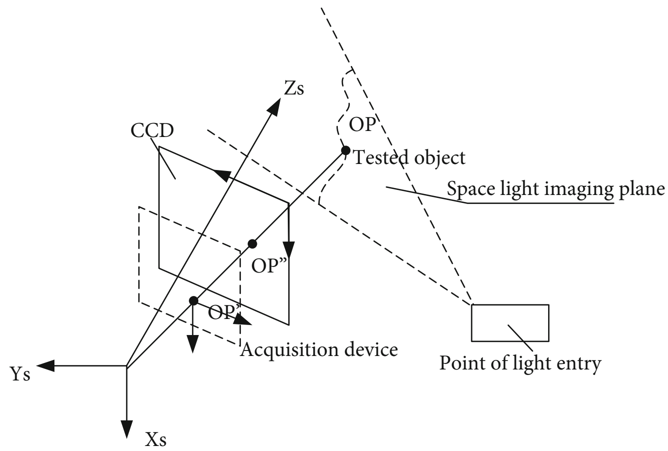

FIGURE 1: Coordinate diagram of the three-dimensional visual-spatial relationship of the detected object.

move 3D visual transformation model, and the Chinese martial art wrong moves are detected using RBF.

2.3. The RBF Model Is Used to Detect the Wrong Action of Chinese Martial Art. Since there are differences in the initial position relative to the camera and the body orientation when people perform actions, this has a significant impact on the description of human posture and action recognition. Therefore, it is necessary to first regularize the coordinated system so that the initial position and orientation of the human skeleton after coordinated transformation are the same. Since the set of 3D trajectories of all skeletal joint points contains the full information of the complete action, the original action data can be reconstructed from three projections, i.e., by a 2D sequence of motion units. 2D human joint point detection is performed using a cascade pyramid network (CPN) to determine the relative position between the human joint points corresponding to each frame of the Chinese martial art action in the video. Use the k-means clustering algorithm to extract the Chinese martial art movements and the Chinese martial art action characteristics, and use the convolution network fusion [22].

In the feature fusion structure, multiple video segments are fed into the network structure at the same time, but in this paper, only the same network model is used, and these segments fed into the network at the same time will share all the parameters of the convolutional layer and some of the parameters of the fully connected layer in the network. More specifically, for a given video, the video will be segmented for the first time into multiple nonoverlapping video segments of the same duration, and then a sequence of images will be obtained in each video segment using a certain sampling strategy [23]. In the proposed framework, the extracted image sequences will be fed into the 3D convolutional neural network, and each image sequence will be given a corresponding spatiotemporal feature. These features will be merged in the training phase and the resulting features will be considered the spatiotemporal features of the whole video and will be used in the subsequent optimization process. In this way, during the whole learning process, the target of optimization becomes the loss of the whole video, rather than the loss of a video segment or slice [24].

After extracting and fusing the Chinese martial art action features from the video clip, an RBF neural network model is built to detect the wrong Chinese martial art action. The radial basis function often used in RBF neural networks is a Gaussian function, from which the activation function involved in the RBF neural network can be represented by the following equation [25].

$$R_{bf}\left(x_r - c_g\right) = \exp\left(-\frac{1}{2\sigma^2}\left\|x_r - c_g\right\|^2\right), \tag{7}$$

where $\|x_r - c_g\|$ is a Euclidean norm, c_g represents the center of the Gaussian function, and σ is the variance of the Gaussian. Thus, the relationship between processing output O_i and input I_j of the RBF neural network is as follows:

$$O_i = \sum_{i=1}^{n} Q_{ij} R_{bf}\left(x_r - c_g\right)_{I_j}. \tag{8}$$

Based on the above analysis, this paper designs a stereo matching reconstruction model including four input nodes and three output nodes in the constructed RBF network model. The input nodes take the pixel values of the standard martial art action video image and the action video image to be detected in turn, and the output node is the three-dimensional coordinates of the corresponding points. The RBF network is trained with the training sample set, and the network parameters that minimize the output error are selected as the parameters of the final detection model. Input the processed standard Wushu action video image into the RBF network model, and the processed output result vector is the Wushu action detection result. Compare it with the standard Wushu action to find out whether there are wrong Wushu actions. Above, the research of three-dimensional visual detection of Chinese martial art wrong movements is realized by using radial basis function neural network technology.

Table 1: Comparison of average detection accuracy and time consumption between processing algorithm and model.

Data set	YOLOv algorithm		Fuzzy C-means algorithm		RBF neural network	
	Accuracy, %	Time consumed, ms	Accuracy, %	Time consumed, ms	Accuracy, %	Time consumed, ms
KTH	97.6	151.5	97.4	149.8	98.8	102.3
UCF101	96.5	232.6	95.7	241.7	98.0	137.8
HMDB51	96.9	298.1	93.2	301.5	97.6	151.4
Kinetics	95.7	364.3	90.3	344.3	97.1	156.9

3. Experimental Study

In order to verify the feasibility of the above theoretical design and the performance of the proposed detection method, experimental research on the detection method will be conducted in this section. According to the final experimental data analysis, the comprehensive performance and practical application of the proposed motion detection method are evaluated.

3.1. Experiment Content. In this experiment, various experimental standards were calibrated for a single user test environment for subsequent threshold settings. In the experiment, all experimental factors were consistent except for experimental control variables. The motion detection method based on the YOLOv algorithm in reference [6] and the motion detection method based on the fuzzy C-means algorithm in reference [7] are selected as comparison method 1 and comparison method 2, respectively. The two comparison methods were applied to the same experimental background and comprehensively compared with the 3D vision detection method based on RBF proposed in this paper. In order to verify the stability of this method, this experiment simulates the method test in different background environments and selects several subjects to test the method in the simple background and complex background, respectively.

3.2. Experimental Data and Preparation. Considering that martial art actions include kicking, hitting, falling, holding, falling, hitting, splitting, stabbing, and other actions, this paper uses KTH, UCF101, HMDB51, and dynamics data sets as the test data sets of algorithm performance in the experiment. The above four data sets contain different movements similar to Wushu movements. Among them, the contents of the KTH data set are simple six types of actions completed by 25 adults in four different scenes, including walking, jogging, running, boxing, hand waiting, and hand clipping, with a total of 2391 video samples. The fixed camera and single background used for image acquisition in this data set are not close to the objective and real scene performance. Part of the data set UCF101 comes from various sports samples collected by BBC/ESPN radio and television channels, and part comes from video samples downloaded from the Internet. The video website with the most sources is YouTube. UCF101 contains 13320 video samples, which are divided into 101 categories in total. Most of the data samples in the HMDB51 data set are collected from movies, which are more difficult to understand than videos in natural scenes. The data set has 6849 samples and 51 categories, and each category contains at least 101 data samples. The kinetics data set comes from YouTube and contains 400 kinds of actions, with a total of about 300000 videos. Professional practitioners in the martial art industry are invited to make a demonstration video and compare the video data as the detection standard sample of action detection methods. Using the known data set with parameters as above, the YOLOv algorithm, fuzzy C-means algorithm, and RBF neural network used in the three methods are tested.

3.3. Experimental Results. Table 1 shows the processing algorithms and models for the application of the detection method on the selection of test data set processing, and different data sets with corresponding processing compare the average detection accuracy and time.

Analyzing the data in Table 1, when processing the KTH data set, the detection accuracy of the three algorithm models is basically the same. With the increase in the complexity of samples in the data set, the accuracy of the three algorithms decreases, and the processing time increases rapidly. Among them, the processing accuracy of the RBF neural network for the four data sets is higher than 97%, and the processing time is 156.9 ms, which is far less than the other two algorithms, indicating that the performance of this algorithm is relatively better.

In order to further verify the effectiveness of the proposed method, the average accuracy of martial art error action detection under the use of the three methods is compared under the simple background and complex background. The results are shown in Figures 2 and 3. The simple background refers to the background of martial art error action, which is a single background, has low noise, and has stable illumination, and the complex background refers to the background with many interference factors, the background with more noise and unclear light and dark lines.

By comparing and analyzing the figures in Figures 2 and 3, it can be seen that in a simple scene, the detection accuracy of the three detection methods has little difference. The main reason is that the simple background is a single background, has low noise, and has stable illumination, which has little impact on the extraction action of the three methods. In complex scenes, the complexity of the scene gradually increases with the scene number, the interference factors such as background and noise in the scene increase resulting in the gradual decline of the detection accuracy of the three methods, and the decline of the detection accuracy

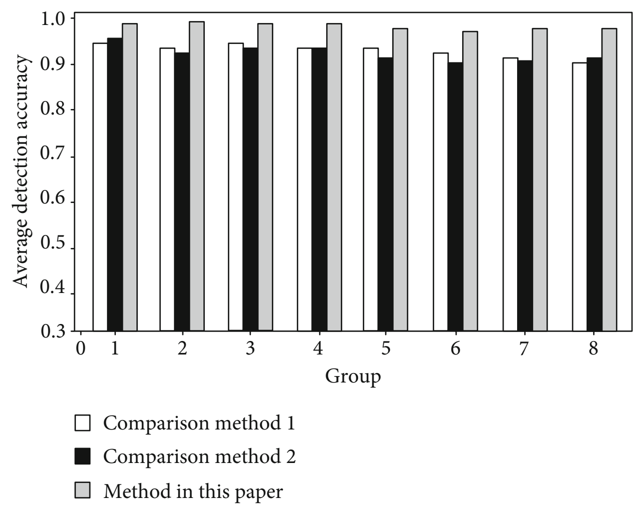

Figure 2: Comparison of detection accuracy in simple scenes.

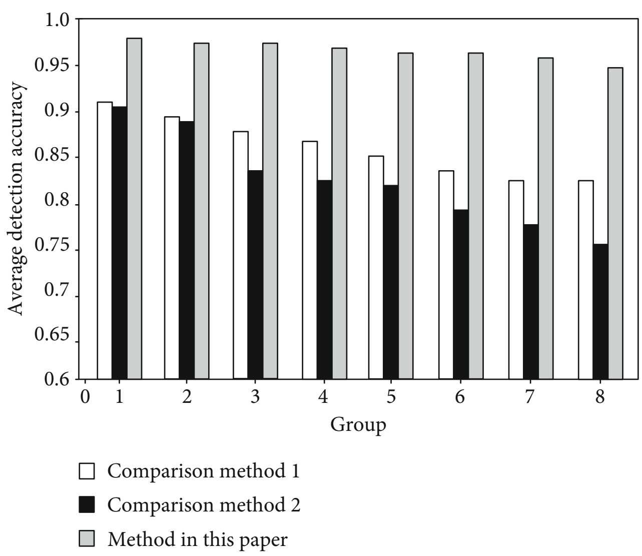

Figure 3: Comparison of detection accuracy in complex scenes.

of the comparison method is the largest. Based on the above analysis, in simple and complex detection scenarios, the detection accuracy of this method is higher than that of the two comparison methods. On average, the detection accuracy of this method is improved by at least 5%. Therefore, the detection efficiency of this method is higher, and the detection is less affected by the background environment. The stability of this detection method is good.

Summarizing the above test data, it can be seen that the three-dimensional visual detection method of martial art wrong action based on RBF proposed in this paper has high detection accuracy and sensitivity, and the application stability of the method is good, which is suitable for different action detection conditions. The comprehensive performance of this method is significantly improved and has higher practical application value and application effect. This method meets the research expectation.

4. Conclusion

With the continuous maturity of video image processing technology, the use of a computer to process images to detect martial art movements is gradually applied. However, in order to solve the problems of low detection accuracy, slow detection speed, and slow detection response, this paper proposes a three-dimensional visual detection method of martial art wrong action based on RBF. Using the characteristics of radial basis function neural network, this method realizes the accurate detection of martial art wrong actions. Experiments show that the proposed three-dimensional vision detection method has high precision, high efficiency, and good stability and can be used to effectively detect martial art wrong movements.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. Xu, J. Yiran, X. He, Y. Juntong, and Y. Gao, “Dynamic analysis of the complex motion of three-section cudgel in Wushu sports,” *Applied Sciences*, vol. 11, no. 21, pp. 10407–10407, 2021.

[2] L. Wenbo and D. Guobin, “The inheritance and dissemination of Wushu culture in the global era,” *The International Journal of the History of Sport*, vol. 38, no. 7, pp. 768–778, 2021.

[3] S. Ohl and M. Rolfs, “Causality detection in the visual system is tuned to motion direction,” *Journal of Vision*, vol. 21, no. 9, pp. 1946–1946, 2021.

[4] A. Phillips, G. Erlikhman, and P. J. Kellman, “On the relationship between perceptual learning and statistical learning: evidence from coherent motion detection,” *Journal of Vision*, vol. 21, no. 9, pp. 2871–2871, 2021.

[5] N. Pooja, B. S. Bhaskar, J. Aman, and C. Sarika, “Motion detection of webcam using frame differencing method,” *Research Journal of Engineering and Technology*, vol. 12, no. 2, pp. 32–38, 2021.

[6] H. Yu and W. Chen, “Motion target detection and recognition based on YOLOv4 algorithm,” *Journal of Physics: Conference Series*, vol. 2025, no. 1, article 012053, 2021.

[7] K. S. NagiReddy, T. Suresh, A. Prasanth, T. Muthumanickam, and K. Mohanram, “An effective motion object detection using adaptive background modeling mechanism in video surveillance system,” *Journal of Intelligent & Fuzzy Systems*, vol. 41, no. 1, pp. 1777–1789, 2021.

[8] S. Nabil, A. Ahmed, J. I. Muhammad, M. Muhammad, and O. Pablo, “Multi-sensor fusion for underwater vehicle localization by augmentation of RBF neural network and error-state Kalman filter,” *Sensors*, vol. 21, no. 4, pp. 1149–1149, 2021.

[9] I. Anantraj, B. Umarani, B. P. Kumar, B. R. Palaniyappan, and R. Ravishankar, “Feature space replica and locomote dimensional detection using Gaussian RBF,” *Journal of Critical Reviews*, vol. 7, no. 4, pp. 1019–1024, 2020.

[10] K. B. Gyu, "Digital signal, image and video processing for emerging multimedia technology," *Electronics*, vol. 9, no. 12, pp. 2012–2012, 2020.

[11] S. Takeda, M. Isogai, S. Shimizu, and H. Kimata, "Local Riesz pyramid for faster phase-based video magnification," *IEICE Transactions on Information and Systems*, vol. E103.D, no. 10, pp. 2036–2046, 2020.

[12] P. Shamsolmoali, M. Emre Celebi, and R. Wang, "Advances in deep learning for real-time image and video reconstruction and processing," *Journal of Real-Time Image Processing*, vol. 17, no. 6, pp. 1883-1884, 2020.

[13] A. Khandual, T. Grover, Y. Luximon, N. Rout, and R. Nayak, "Instrumentation and objective evaluation of flammability of textiles by video image processing," *The Journal of The Textile Institute*, vol. 111, no. 8, pp. 1176–1183, 2020.

[14] L. Santeri, N. Longchuan, H. Lionel, N. Jouni, and M. Jouni, "Autonomous robotic rock breaking using a real-time 3D visual perception system," *Journal of Field Robotics*, vol. 38, no. 7, pp. 980–1006, 2021.

[15] M. Kyritsis, S. R. Gulliver, and E. Feredoes, "Visual search fixation strategies in a 3D image set: an eye-tracking study," *Interacting with Computers*, vol. 32, no. 3, pp. 246–256, 2020.

[16] K. Kanagaraj and G. G. Priya, "A new 3D convolutional neural network (3D-CNN) framework for multimedia event detection," *Signal, Image and Video Processing*, vol. 15, no. 4, pp. 779–787, 2021.

[17] R. Alfaifi and A. M. Artoli, "Human action prediction with 3D-CNN," *SN Computer Science*, vol. 1, no. 5, pp. 689–704, 2020.

[18] D. Wang, "Simulation research on safety detection of pattern rope jumping motion based on large data background," *Connection Science*, vol. 33, no. 4, pp. 1047–1059, 2021.

[19] M. Marc, D. Glen, D. Jesse et al., "Motion sensor-based detection of outlier days supporting continuous health assessment for single older adults," *Sensors*, vol. 21, no. 18, pp. 6080–6080, 2021.

[20] Z. Jing, W. Wang, Z. Yang, and Z. Jiwen, "Robust high precision multi-frame motion detection for PMLSMs' mover based on local upsampling moving least square method," *Mechanical Systems and Signal Processing*, vol. 159, no. 6, article 107803, 2021.

[21] I. Lorato, S. Stuijk, M. Meftah et al., "Automatic separation of respiratory flow from motion in thermal videos for infant apnea detection," *Sensors*, vol. 21, no. 18, pp. 6306–6306, 2021.

[22] H. K. Azeem, P. Alain, L. Marcus, S. Didier, and A. M. Zeshan, "CasTabDetectoRS: cascade network for table detection in document images with recursive feature pyramid and switchable atrous convolution," *Journal of Imaging*, vol. 7, no. 10, pp. 214–214, 2021.

[23] C. Beltran-Perez, H.-L. Wei, and A. Rubio-Solis, "Generalized multiscale RBF networks and the DCT for breast cancer detection," *International Journal of Automation and Computing*, vol. 17, no. 1, pp. 55–70, 2020.

[24] K. Z. Ahmad, K. Laiq, A. Saghir, M. Sidra, J. Muhammad, and K. Qudrat, "RBF neural network based backstepping terminal sliding mode MPPT control technique for PV system," *PLoS One*, vol. 16, no. 4, pp. e0249705–e0249705, 2021.

[25] A. D. Izzat, P. Dimitrios, A. A. Kakei et al., "Adaptive robust controller design-based RBF neural network for aerial robot arm model," *Electronics*, vol. 10, no. 7, pp. 831–831, 2021.