

Retraction

Retracted: Footprint Extraction and Sports Dance Action Recognition Method Based on Artificial Intelligence Distributed Edge Computing

Wireless Communications and Mobile Computing

Received 15 November 2022; Accepted 15 November 2022; Published 5 December 2022

Copyright © 2022 Wireless Communications and Mobile Computing. 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.

Wireless Communications and Mobile Computing has retracted the article titled “Footprint Extraction and Sports Dance Action Recognition Method Based on Artificial Intelligence Distributed Edge Computing” [1] due to concerns that the peer review process has been compromised.

Following an investigation conducted by the Hindawi Research Integrity team [2], significant concerns were identified with the peer reviewers assigned to this article; the investigation has concluded that the peer review process was compromised. We therefore can no longer trust the peer review process and the article is being retracted with the agreement of the Chief Editor.

References

- [1] X. Hu and X. Xu, “Footprint Extraction and Sports Dance Action Recognition Method Based on Artificial Intelligence Distributed Edge Computing,” *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 6246842, 12 pages, 2022.
- [2] L. Ferguson, “Advancing Research Integrity Collaboratively and with Vigour,” 2022, <https://www.hindawi.com/post/advancing-research-integrity-collaboratively-and-vigour/>.

Research Article

Footprint Extraction and Sports Dance Action Recognition Method Based on Artificial Intelligence Distributed Edge Computing

Xiaolin Hu¹ and Xiang Xu² 

¹Sichuan Technology and Business University, School of Music and Dance, Chengdu, 611745 Sichuan, China

²Military Ministry, Jiangxi Agricultural University, Nanchang, 330045 Jiangxi, China

Correspondence should be addressed to Xiang Xu; xuxiang@jxau.edu.cn

Received 9 March 2022; Revised 25 April 2022; Accepted 6 May 2022; Published 11 June 2022

Academic Editor: Rashid A Saeed

Copyright © 2022 Xiaolin Hu and Xiang Xu. 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.

Distributed edge computing technology for artificial intelligence refers to an emerging technology that integrates network core processing functions, computing functions, storage functions, etc. into one end source closer to objects or data on the basis of an open platform to optimize service quality. In this paper, distributed edge computing technology is applied to footprint extraction and sports dance action recognition, aimed at improving the recognition efficiency and recognition quality. Firstly, the overview of edge computing theory is introduced; these include edge computing concepts, edge computing characteristics, and edge computing platforms; and then, the classification of action recognition technology is described. Finally, the edge computing recognition technology and traditional recognition technology are compared and tested. The experimental results show that the average accuracy of edge computing technology for footprint extraction can reach 98.98, and the average recognition rate of sports dance movements can reach 80.21%, which verifies its practicability.

1. Introduction

With the rapid development of information science and technology and the rapid improvement of the preparation level of hardware facilities, computer technology continues to mature and has brought a huge impact on people's daily life and work life. As one of the important components in the field of computer technology, action recognition technology has great significance and value in the current research field. Human action recognition research can be divided into simple and complex levels according to content and type. Human action recognition at the simple level often contains a small number of people, while the human body recognition at the complex level not only contains a large number of people but also includes different types of visual recognition, mainly including movement, action, and behavior. The current research field of action technology is still at the level of action. That is to use machine technology to learn and train information data, then extract action features

in the dataset, input the extracted features into a supervised or unsupervised classification model, and then count the classification results of the features. Different from simple sports movements, sports dance movements are more complex, and in the process of sports, the costume shape and background often change, and there is no uniform specification. This type of action recognition often requires high details, so there are very few professional action studies such as footprint extraction and sports dance movements.

The artificial intelligence distributed edge computing technology has become one of the representatives of the new technology field under the progress of the times and is becoming more and more mature. From conceptual consensus to industrial practice, from pilot exploration to commercial application, empowering all walks of life, it has become an important starting point for promoting the digital transformation of the industry. At present, edge computing has been widely used in many important fields. For example, the application of edge computing technology can be seen

in value industries such as operators, power energy, industrial manufacturing, intelligent transportation, smart cities, and digital entertainment. In the foreseeable future, there will be more and more application scenarios of edge computing, and the value will be brought into play. The introduction of edge computing into footprint extraction and sports dance action recognition can collect, process, and analyze dance skill information and dancer's action data information, realize personalized and precise training, and improve the level and quality of sports dance training movements.

At present, there are few studies that combine information science technology with footprint extraction and sports dance movement recognition. This paper proposes a novel research direction of footprint extraction and sports dance movement recognition based on artificial intelligence distributed edge computing technology. This technology can effectively make accurate judgments on footsteps and dance movements, provide a perfect and improved development suggestion for the intelligent development of the dance industry, and provide new ideas for in-depth research on movement preservation technology.

2. Related Work

In recent years, many scholars have carried out research on edge computing technology as well as footprint extraction and sports dance movement recognition methods. TranTX envisions a real-time, context-aware collaboration framework. It is located at the edge of the RAN, consists of MEC servers and mobile devices, and integrates heterogeneous resources at the edge. Tran et al. discuss the key technical challenges and open research issues that need to be addressed in order to effectively integrate MEC into the 5G ecosystem [1]. Ke et al. propose a cloud-based mobile edge computing (MEC) offloading framework in in-vehicle networks and studies the effectiveness of computational transport strategies for vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) communication modes. Considering the time consumption of computing task execution and the mobility of vehicles, they proposed an effective prediction combination mode degradation scheme; the proposed scheme greatly reduces the computational cost and improves the task transmission efficiency [2]. Hirokatsu et al. propose an emerging fine-grained pedestrian action recognition problem that induces advanced preload safety to estimate pedestrian intent in advance. Fine-grained action recognition can induce pedestrian intent estimation for useful advanced driver assistance systems (ADAS). Several types of configurations were evaluated to explore an efficient approach for fine-grained pedestrian action recognition without the need for large-scale databases [3]. Yanhua et al. proposed a discriminative multi-instance multitask learning (MIMTL) framework to discover the intrinsic relationship between joint configurations and action classes. He extensively evaluates MIMTL using three benchmark 3D action recognition datasets, and experimental results show that the proposed MIMTL framework performs well compared to several state-of-the-art methods [4]. Li et al. pro-

pose an efficient and simple method to encode the spatiotemporal information of skeleton sequences into color texture images, called joint distance map (JDM). They employ convolutional neural networks to leverage discriminative features from JDM for human action and interaction recognition. State-of-the-art results on large RGB+D datasets and small UTD-MHAD datasets have validated the effectiveness of the proposed method in both single-view and cross-view settings [5]. Ke et al. introduced SkeletonNet to extract body part-based features from each frame of a skeleton sequence. The proposed features are translation, rotation, and scale invariant compared to the original coordinates of the skeleton joints. They convert features into images and feed them to the proposed deep learning network [6]. To sum up, after several years of exploration, action recognition and edge computing technology have been deeply studied by many scholars, but there are not many studies on the combination of the two aspects. Therefore, in order to promote the in-depth development of the sports dance industry, the research on footprint extraction and sports dance action recognition methods based on artificial intelligence distributed edge computing is urgent.

3. Footprint Extraction and Sports Dance Action Recognition Method Based on Edge Computing

3.1. Overview of Edge Computing Theory. With the interconnection and integration of the Internet and various industries, the number of network devices connected and the data of network communication users have exploded, and users' consumption demands for various mobile application services are increasing. The inherent limitations of the existing cloud computing technology are gradually exposed in the process of the development of the times, and it can no longer carry such a huge data system alone. Therefore, in order to meet the development of the times and fill the technical gap and the growing needs of the people, the distributed edge computing technology based on artificial intelligence technology was born. There is no fixed standard for the definition of edge computing. Many European and American countries believe that the essence of edge computing is a data platform that integrates multiple core technologies (computer control technology, data storage technology, communication technology, etc.) and can realize multiple services. It has a more refined structure, as shown in Figure 1. This technology is located at the edge layer of the network where people and devices are located and can solve the problems of network delay and serious node traffic consumption in cloud computing [7].

3.1.1. Features of Edge Computing. Based on the understanding of the definition of edge computing and its unique excellent performance, the characteristics of edge computing can be summarized into 5 points, as shown in Table 1.

- (1) Geographically intensive: edge computing technology is distributed on multiple computing service platforms, making services closer to users

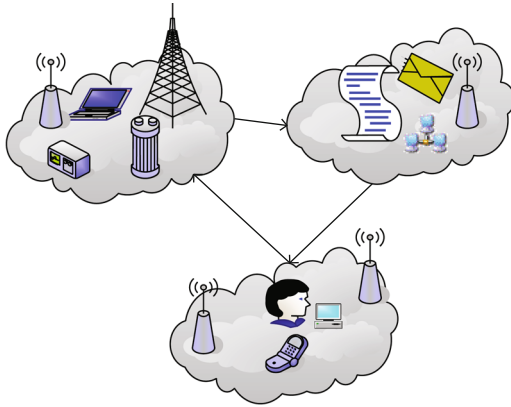


FIGURE 1: Edge computing network structure.

- (2) Mobile support: it allows for mobile support for both the host and the device, for the user
- (3) Perception: since edge computing technology is located at the network edge layer, close to the mobile user's device terminal, it allows the mobile user to select the edge server closest to their physical location to inquire about services
- (4) Real-time data processing: real-time performance is determined by the definition of edge computing. It integrates the three capabilities of data collection, processing and execution to avoid the delay caused by data upload and download, and to improve the processing capability and response speed of local IoT devices
- (5) Diversification of application development: in the future, more than half of the data will be processed at the source of its generation, and there will be various application scenarios, such as industrial manufacturing, smart cars, smart cities, smart homes, etc. Users can customize IoT applications according to their own business needs

3.1.2. *Edge Computing Platform.* There are multiple computing platforms in edge computing, including ParaDrop, Cloudlet, PCloud, Firework, and Haiyun computing system. Firstly, each computing platform is briefly introduced, and then, its application fields and characteristics are compared [8].

(1) *Airdrop Platform (ParaDrop).* The airdrop platform is composed of several gateway devices, which rely on airdrop technology to combine the devices, thus deriving a new edge computing platform. This platform differs from traditional edge computing platforms in that it relies on container technology to maintain the independence of the application runtime environment. All operations such as application setting, distribution, operation, and control are carried out on the background service control system in the cloud [9]. The components of the entire system mainly include three structures, as shown in Figure 2, which are the platform backend, the platform gateway, and the developer, and the three struc-

tures are responsible for different tasks. The back-end is mainly responsible for the allocation and management of information and data resources of the entire platform, the maintenance and control of each network management node, and the management and control of users and services. As a hardware engine, the gateway is mainly responsible for performing work, such as providing resource environments such as CPU and memory to ensure the normal operation of the system platform.

(2) *Micro Cloud Platform (Cloudlet).* Since the concept of the microcloud platform was first proposed in 2009, it has attracted widespread attention from the masses and scholars. The microcloud platform is realized by relying on a host with powerful data processing capabilities and rich resources. This host is located at the edge layer of the network and allows all nearby mobile terminal devices to access and connect with each other, thereby realizing applications and services, as shown in Figure 3. It differs from traditional terminal equipment which only contains two layers in that it has three layers, which can be attached to lower-level servers, even on personal computers [10]. This microcomputing platform at the edge of the network can not only provide a storage space for the data transmitted from the cloud but also automate the management of these data information. The rich resources of the platform itself can provide services for multiple users at the same time, which shortens the distance between the computing platform and users and alleviates the problems of network delay and bandwidth.

(3) *Particle Cloud Platform (PCloud).* The particle cloud platform is a computing platform that integrates the data resources of the network edge layer and the cloud to achieve the purpose of complementary resources. The inherent limitations of cloud computing technology make it impossible to provide users with low-latency network services, but the particle cloud platform can solve this. At the same time, it can combine its surrounding terminal system resources with the rich resources of cloud computing itself to provide more diverse services [11].

(4) *Firework Platform (Firework).* The firework platform is a programming model proposed on the basis of edge computing technology, and its frame structure is shown in Figure 4. The firework computing platform has two layers of information data analysis and processing scheduling, which can automatically divide the analysis and processing activities of a type of information data into several subactivities of information data analysis and processing. Among them, the first layer of scheduling is the same subactivity scheduling, and the gateway of the platform will gather all nearby nodes processing the same subactivity with the fastest response speed for cooperative processing according to the characteristics of the surrounding environment. The nodes in the surrounding idle state will also be called to share tasks. The second layer of scheduling is computing layer scheduling, which is similar to the first layer of scheduling. The gateway nodes of the second layer also cooperate with each other when performing computing tasks [12].

TABLE 1: Edge computing features.

Sequence	Features	Effect
1	Geographically dense	Deploy many computing platforms to bring services closer to users
2	Mobile support	Separation of host identity and location identity for mobile support
3	Perceptual	Allow mobile users to access from the edge server closest to their physical location Ask service
4	Proximity	Reduce latency in accessing services and improve experience
5	Heterogeneity	Diversity of communication technologies for edge service delivery

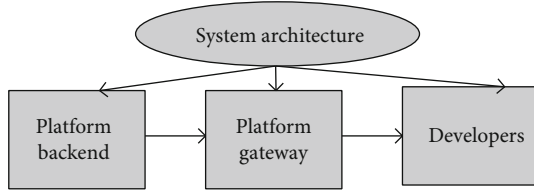


FIGURE 2: ParaDrop platform structure.

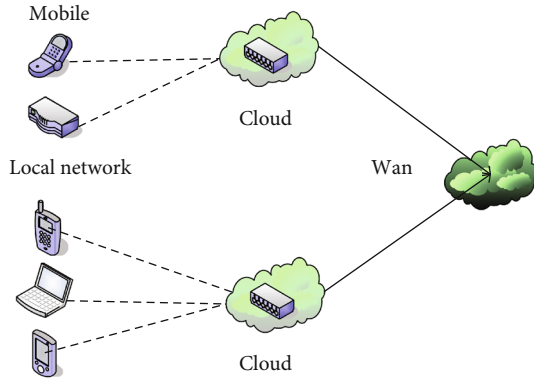


FIGURE 3: Cloudlet architecture.

(5) *Haiyun Computing System*. The sea cloud computing system can be divided into two parts: sea computing and cloud computing. But in detail, it is composed of four parts, namely, computing model, storage system, data center, and elastic processor. The computing model follows the REST architecture of the web computing technology; the access from the sea end to the cloud is realized by this interface, and the client mobile device application is also run through this interface; the storage system is used to process a large amount of data, the data center supports the stable operation of the thread, and the elastic processor can make the entire system process trillions of data information per second, which is very efficient [13].

Table 2 shows the performance and differences between the airdrop platform, micro cloud platform, particle cloud platform, the firework platform, and the Haiyun computing system from the application field, service mobility, virtualization technology, and system characteristics to the main edge computing systems.

3.2. *Classification of Action Recognition Technology*. When motion recognition technology was first proposed, it only

used intelligent machines such as computers to identify whether the target was moving. With the in-depth development of technology, motion recognition technology can not only distinguish whether an object is in a state of motion but also distinguish the type of motion; this includes actions such as jumping, running, or walking. At present, in the field of recognition, most of the recognition tasks are carried out by relying on video images, that is, the action of the target object in the video image is defined as a complete sequence. The sequence contains several relatively simple movements, which are composed of postures transformed by the body. The fluency and coherence of each movement change are also formed by the excessive posture [14].

The similarity between dance and sports lies in the composition of movements. Dance movements are also composed of many simple movements, but these simple movements are different from the simplicity of sports movements, and their postures are more complex. These complex poses form a complete action sequence. When performing action recognition, the sequence needs to be divided into short but complete subsequences. But the challenge of this work is that it is difficult to quickly find the split points of the sequence. Dance movements are more coherent and smooth than sports movements, and only one key frame will contain multiple movement information. Therefore, it is difficult to achieve segmentation only by extracting key frames in the video sequence. Therefore, dance movement recognition in reality usually relies on manual processing. The main part in the dance video sequence is manually segmented, cropped into multiple small videos, and then identified in sequence [15].

The formation of action edges is caused by inconsistent speeds and incoherent action curvatures, and the edge computing recognition technology in the field of artificial intelligence is based on this principle. It is aimed at detecting the boundaries of actions on the edge side of action sequences to identify dance movements [16].

3.2.1. Action Segmentation Model

(1) *PCA Segmentation Model*. The basic principle of the unsupervised PCA method is to use the correlation (for example, when the left arm swings forward during walking, the right arm will swing backward) of the motion of each joint of the human body and use the PCA method to reduce the dimension of the data to extract the main representation information for human movements. Two different action types will have different principal components, so the

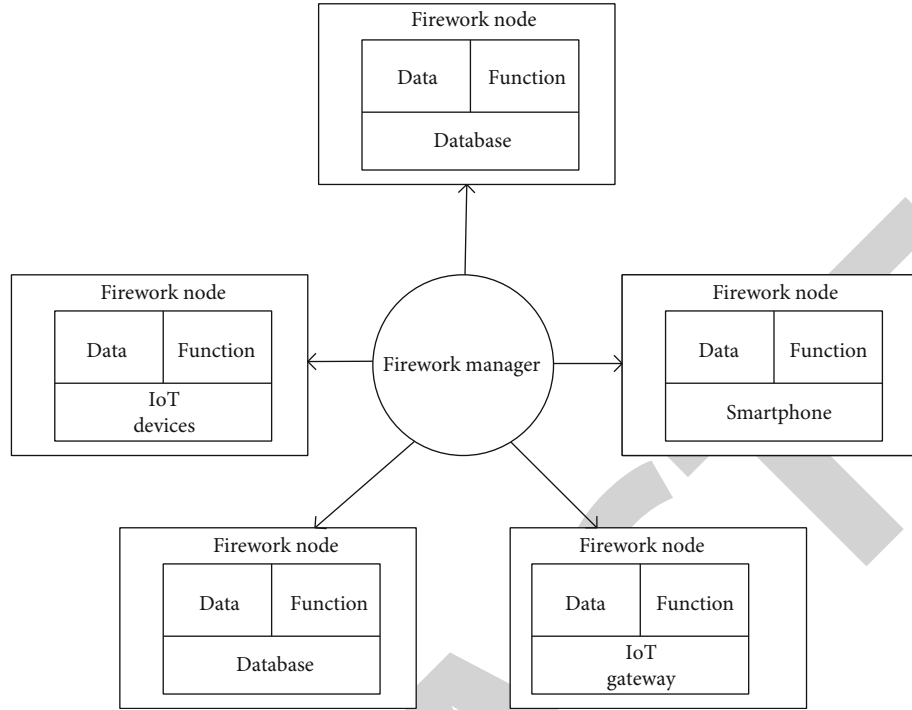


FIGURE 4: Firework framework.

TABLE 2: Edge computing platform comparison.

Sequence	System	Application field	Mobile service	Virtualization technology	Features
1	ParaDrop	Data-intensive applications	Mobile service issues are not considered	Container technology is suitable for IoT applications	Suitable for IoT applications
2	Cloudlet	Computationally intensive applications	Provide deployment scenarios for mobile services	Virtualization of physical machines	Cloud services at the edge
3	PCloud	Supports multiple types of applications	Combining user equipment with the cloud	Virtualization of physical machines	Unified management of cloud resources and edge resources
4	Firework	Data sharing and processing	Mobility is not considered	Different devices use different virtual technologies	Implement the concept of computational flow
5	SeaCloud computing system	For offshore applications	Mobility is not considered	Different scenarios use different virtual technologies	Focus on improving performance

behavior can be distinguished [17]. In the segmentation process of the PCA video frame sequence, the motion data in an interval is firstly extracted by PCA, and then, the window length of this interval is lengthened. If it is found that the extracted components are quite different from the components of the previous interval when extending to a certain data frame, then it is concluded that the movement has switched behaviors here, and the model is shown in Figure 5.

Assuming that there is M to represent the video frame sequence, the final sequence segmentation result is specified as M_1, \dots, M_s , and the number and boundary of actions are determined by S . Each frame is represented by using all joint rotations (relative to the parent node in the body's hierar-

chy) at a specified time, where rotations refer to quaternions [18]. Ignoring absolute body position and body orientation information, this method will be independent of the specific body position and orientation in world coordinates.

Each frame $x_i (i = 1, 2, \dots, n)$ has num joints at the body level, and each joint is assigned a quaternion, so the frame is represented as a point in the $4 * \text{num}$ -dimensional space, denoted as $R^{4*\text{num}}$. The motion sequence corresponding to the trajectory motion center of $R^{4*\text{num}}$ is defined as

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (1)$$

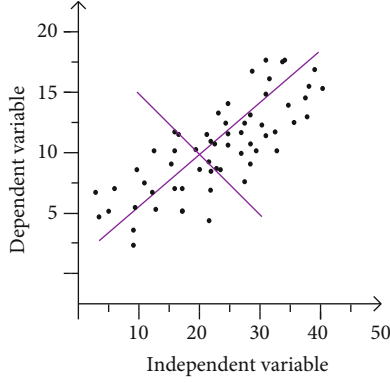


FIGURE 5: PCA segmentation model.

Frames mainly appear on some two-dimensional hyperplanes containing x_i , so the $4 * \text{num}$ -dimensional data is highly correlated, that is, the data representation of all relevant nodes can be used for key data extraction. Assuming the dimension is r , the frame can be approximated as

$$x_i' = \bar{x} + a_{i1}v_1 + a_{i2}v_2 + \dots + a_{ir}v_r. \quad (2)$$

v_1, v_2, \dots, v_r is the unit orthogonal vector generated in the linear subspace, which corresponds to the two-dimensional hyperplane at the same time, and the definite feature frame in the video sequence is denoted as $a_{i1}, a_{i2}, \dots, a_{ir}$. The more inconsistent the recognition result is with the original action, the more dimensions are required [19].

The frame x_i is projected on the included \bar{x} hyperplane, the projection result is denoted as x_i' , and the projection error is expressed as

$$e = \sum_{i=1}^n \|x_i - x_i'\|^2. \quad (3)$$

$\|x\|$ is the standard Euclidean norm in $R^{4 * \text{num}}$. Finding the r -dimensional hyperplane with the smallest e .

First calculate the size of the singular value, then divide the value, and then calculate the PCA, which is to summarize the core of the action from the video sequence frame. The core sequence frame is then constructed into a matrix D of size $n \times 56$, where $n \gg 56$, n is the number of frames.

After the singular values are divided, a matrix U, V, Σ will be formed, as shown in the formula [20]:

$$D = UV\Sigma^T. \quad (4)$$

In formula (4), the columns of U and V are orthogonal unit vectors, and matrix Σ is a diagonal matrix of size $4 * \text{num} \times 4 * \text{num}$. The nonnegative decreasing singular values on the diagonal of the diagonal matrix Σ are denoted as σ_i . The r th column in matrix V represents the basis in the hyperplane. In the video projection, all the video sequence frames belonging to the optimal hyperplane discard the singular values but do not include the singular values in the

maximum r , and the projection error of the video sequence is expressed as

$$e = \sum_{i=1}^n \|x_i - x_i'\|^2, \quad (5)$$

$$e = \sum_{j=r+1}^{56} \sigma_j^2, \quad (6)$$

where E_r represents the ratio of information preserved on the optimal r -dimensional hyperplane by projecting frames:

$$E_r = \frac{\sum_{j=1}^r \sigma_j^2}{\sum_{j=1}^{56} \sigma_j^2} \quad (7)$$

Different types of actions will use different r , and using a fixed r will increase the error [21].

Even for a relatively simple pose transformation or action, the error will increase in the form of an approximately constant slope. If it is of the same type as the action, the same error occurs when it adds a new action sequence frame. If a new action sequence frame is added and the action is of a different type, the error will increase dramatically. We usually discern action correlations by looking at discrete derivatives:

$$d_i = e_i - e_{i-1}, \quad (8)$$

where l must be a large value to avoid noise in the data. The derivative d_i rises sharply and exceeds a constant value at uncorrelated actions.

(2) *HMM Segmentation Model*. Data flow is assumed to be a sequential or temporal structure that consists of consecutive blocks of data, with data points in each block originating from the same underlying distribution [22]. The task of segmentation is performed in an unsupervised manner, that is, without any a priori given labels or segmentation boundaries, and its model is shown in Figure 6.

Assuming that there is a data sequence $\{\vec{x}_t\}_{t=1}^T$, estimate the probability density function of a sliding window of length W and obtain the corresponding sequence $\{p_t(x)\}_{t \in S}$, $S = \{W, \dots, T\}$ of pdfs, which can be estimated using a multivariate Gaussian kernel standard density [23]:

$$pt(x) = \frac{1}{w} \sum_{w=0}^{W-1} \frac{1}{(2\pi\sigma^2)^{d/2}} \exp\left(-\frac{(x - \vec{x}_{t-w})^2}{2\sigma^2}\right), \quad (9)$$

where \vec{x}_t represents a specific vector-valued point and x represents a vector-valued variable.

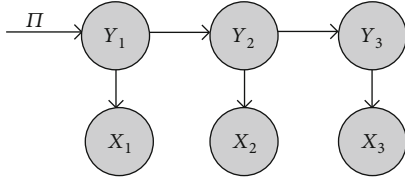


FIGURE 6: HMM segmentation model.

Then, establish an HMM, each pdfs is represented by a state $s \in S$, where S is the set of states in HMM, and define a continuous observation probability distribution for each state s to observe a pdf in state s :

$$P(p_t(x)|s) = \frac{1}{\sqrt{2\pi}\delta} \exp\left(-\frac{d(p_s(x), p_t(x))}{2\delta^2}\right). \quad (10)$$

The definition of each parameter is shown in Table 3.

The goal is to find an pdfs representative of a given sequence based on a small part of the sequence representing pdfs, called the prototype, which exhibits only a small number of prototype changes. So defining A as transforming into the same state is K times more likely than transforming into any other state:

$$p_{ij} = \frac{k}{k+N-1}, i=j, \quad (11)$$

or

$$p_{ij} = \frac{k}{k+N-1}, i \neq j. \quad (12)$$

Using the Viterbi algorithm on HMM to compute the optimal, that is, the most probable prototype pdfs state sequence obtained from pdfs of the given sequence, which represents the obtained segment.

3.2.2. Background Extraction Method. Many dance video datasets are directly recorded with high resolution and also contain a lot of noise during conversion, which will affect the extraction of edge features in dance video images. Therefore, preprocessing operations, including background subtraction operations, need to be performed on the dance dataset first. This method aims to build a background model and learn the parameters of the model at the same time. After learning the parameters, the image of the current frame is compared with the background model [24]. Any area with a large difference is considered a foreground object. In recent years, researchers have proposed many methods for background subtraction, among which Gaussian mixture model is widely used because of its ability to model complex background well and its adaptability.

The Gaussian mixture model method mainly regards the video image sequence as a combination of multiple single Gaussian models, which maintains a multidistributed density function for each pixel in the image. Therefore, the method based on Gaussian mixture model can handle the multimodel background distribution problem well [25, 26].

TABLE 3: Formula parameter meaning.

Sequence	Parameter	Implication
1	$\{\pi_s\}_{s \in S}$	The initial state of HMM
2	$\pi_s = \frac{1}{N}$	Number of states
3	$N = S $	Number of states
4	$A = (p_{ij})_{i,j \in S}$	The transition matrix of HMM

Gaussian mixture model methods usually describe a video image sequence as a pixel probability distribution function. The specific process of the Gaussian mixture model is as follows.

The first is to build a model. Assuming that the value of a certain pixel point at time t is X_t , the probability of occurrence of X_t can be obtained from

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} \cdot \eta(X_t, \mu_{i,t}, \sigma_{i,t}). \quad (13)$$

In Formula (13), $\omega_{i,t}$ is the weight of the i th Gaussian distribution at time t , $\eta(X_t, \mu_{i,t}, \sigma_{i,t})$ is the corresponding probability density function, $\mu_{i,t}$ is the corresponding mean, and similarly $\sigma_{i,t}$ is the variance. At the same time, the specific representation of $\eta(X_t, \mu_{i,t}, \sigma_{i,t})$ is expressed as

$$\eta(X_t, \mu_{i,t}, \sigma_{i,t}) = \frac{1}{\sqrt{2\pi}|\sigma_{i,t}|} e^{-(1/2)(X_t - \mu_{i,t})^T \sigma_{i,t}^{-1} (X_t - \mu_{i,t})}. \quad (14)$$

Firstly, each pixel value of the first frame image of the video is assigned to the mean of K Gaussian distributions; secondly, a larger value is assigned to their variance, and their weights are assigned the same value.

The second is to update the model. Assuming that the value of a pixel in a newly input frame of image is X_t , formula (15) is used to determine whether the pixel matches the established K Gaussian distributions.

$$|X_t - \mu_{i,t-1}| \leq 2.5\sigma_{i,t-1}. \quad (15)$$

If there is one of the K Gaussian distributions that satisfies the condition of formula (15), it is determined that the pixel is matched with it, and the mean, variance, and weight of the Gaussian distribution are updated. The update process is shown in the formula:

$$\omega_{i,t} = (1 - \alpha)\omega_{i,t-1} + \alpha, \quad (16)$$

$$\mu_{i,t} = (1 - \beta)\mu_{i,t-1} + \beta X_{i,t}, \quad (17)$$

$$\sigma_{i,t} = (1 - \beta)\sigma_{i,t-1} + \beta(X_{i,t} - \mu_{i,t})^T (X_{i,t} - \mu_{i,t}), \quad (18)$$

TABLE 4: Footprint deature extraction.

Characteristic parameters	Representation	Calculation method	Unit
Heel centroid to first metatarsal centroid	x_1	Distance formula	Pixel
Heel centroid to thumb centroid	x_2	Distance formula	Pixel
1st plantar centroid to thumb centroid	x_3	Distance formula	Pixel
Follow the center of mass to the origin of the coordinates	x_4	Distance formula	Pixel
The centroid of the first metatarsal to the coordinate origin	x_5	Distance formula	Pixel

TABLE 5: Sample data extraction results.

Sample sequence	x_1	x_2	x_3	x_4	x_5
1	1333.2	1695.8	327.1	2480.4	288.9
2	1335.6	1640.1	314.8	2473.6	276.1
3	1334.4	1662.3	332.4	2277.7	269.8
4	1332.9	1647.6	317.5	2358.1	283.2
5	1341.6	1636.7	325.7	2465.2	273.5
6	1345.2	1635.9	329.6	2396.4	277.6

$$\beta = \alpha \eta(X_t, \mu_{i,t}, \sigma_{i,t}). \quad (19)$$

In the above formula, α is the learning rate of the Gaussian mixture model and $0 \leq \alpha \leq 1$; α determines the update speed of the background model. β is the parameter update factor, which determines the speed of parameter update.

If none of the K Gaussian distributions meet the conditions of formula (15), it is considered that the pixel does not match the K Gaussian distributions. Then, define a new Gaussian distribution with X_t as the mean and delete the distribution with the smallest priority $\lambda_{i,t} = \omega_{i,t}/\sigma_{i,t}$, so the final number of Gaussian distributions is still K .

Then, there is foreground detection. When the training of the background model is completed, the K Gaussian distributions are arranged according to the size of $\lambda_{i,t}$, and the first B distributions with high priority are taken, and then, the background is generated using the formula:

$$B = \operatorname{argmin} \left(\sum_{k=1}^b \omega_k > T \right). \quad (20)$$

In formula (20), T is the threshold. When the Gaussian distribution with the nearest new pixel value belongs to any one of the above B Gaussian distributions, it is considered as a background pixel; otherwise, it is a foreground pixel.

4. Footprint Extraction and Sports Dance Movement Recognition Experiment

The experiment in this paper is divided into two parts: one is the footprint extraction test, and the other is the sports dance movement recognition test. The identification method based on artificial intelligence distributed edge computing technology is compared with the traditional identification

method to verify the effectiveness and feasibility of the proposed algorithm [27, 28].

4.1. Footprint Extraction Test. The footprint extraction test takes the tester's footprint characteristics as the research object, and the sample size is 20. The footprint feature data (morphological features, length features) of 20 testers was collected as a training set to build a footprint discrimination model, as shown in Table 4.

Six out of 20 testers were randomly selected to recollect stereoscopic footprints and use them as test samples to test the accuracy of the algorithmic footprint discrimination model. The extracted sample data are shown in Table 5.

4.1.1. Morphological Feature Extraction Accuracy. Figure 7(a) shows the accuracy of morphological feature extraction of edge computing technology.

Figure 7(b) shows the accuracy of morphological feature extraction by traditional recognition technology.

It can be seen from Figure 7 that the overall mean of the extraction accuracy of the footprint morphological features of the sample data by the recognition method based on edge computing technology is 98.98 [29, 30]. The overall mean of the extraction accuracy of the footprint morphological features of the sample data by the identification method based on traditional technology is 86.84 [31, 32].

4.1.2. Length Feature Extraction Accuracy. Figure 8(a) shows the accuracy of length feature extraction of edge computing technology.

Figure 8(b) shows the accuracy of feature extraction of traditional recognition technology.

It can be seen from Figure 8 that the overall mean of the extraction accuracy of the footprint length feature of the sample data by the identification method based on edge computing technology is 99.12. The overall mean of the extraction accuracy of the footprint length feature of the sample data by the identification method based on traditional technology is 87.63.

4.2. Sports Dance Movement Recognition Test. Since the research on the combination of action recognition technology and dance has just started, the available dance datasets are still relatively small. Therefore, only two dance datasets are used in this experiment to verify the recognition rate of the recognition method and the training accuracy of the sample set, namely the DanceDB dataset and the FolkDance dataset. There are four groups of dance movements in the two dance datasets, and each group of dance movements

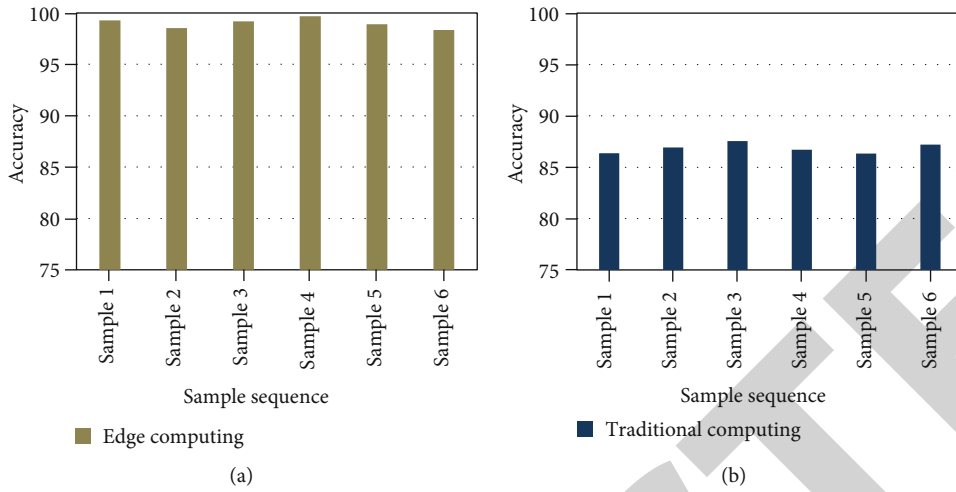


FIGURE 7: Morphological feature extraction accuracy.

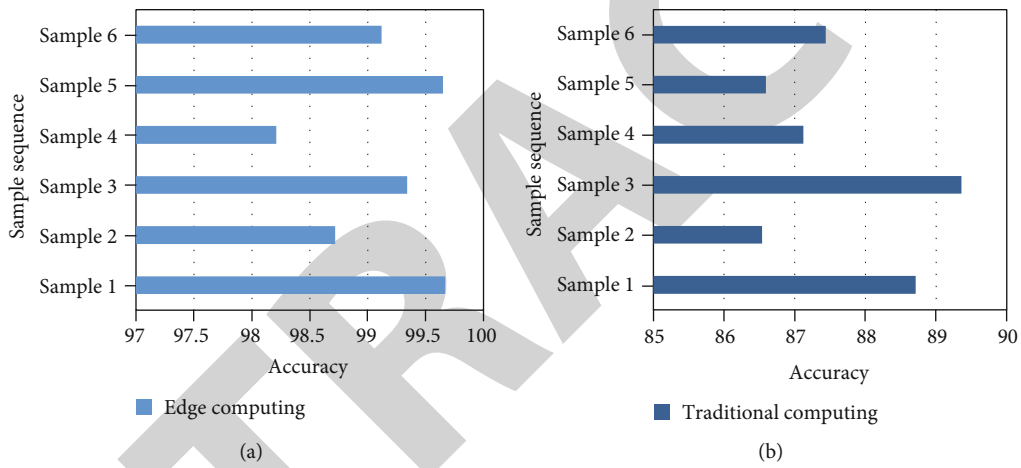


FIGURE 8: Length feature extraction accuracy.

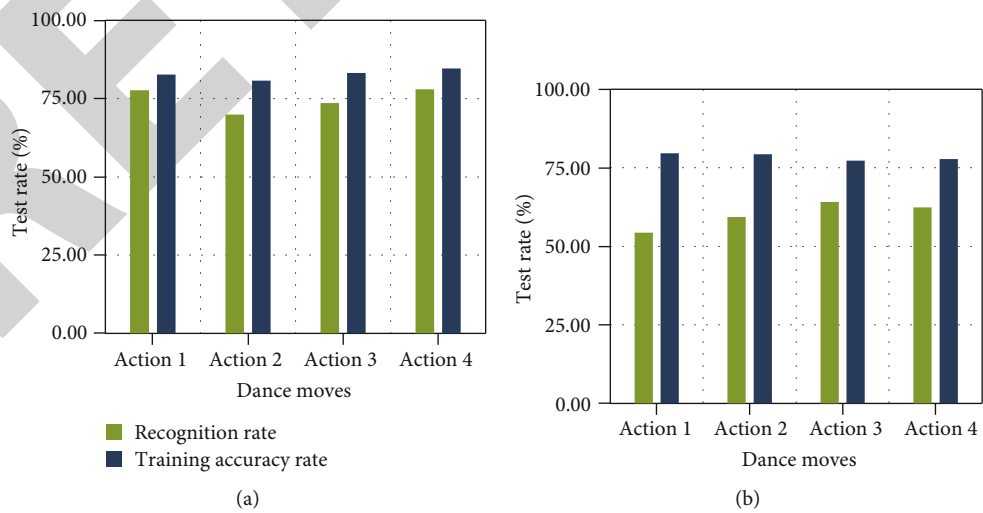


FIGURE 9: DanceDB dataset recognition test results.

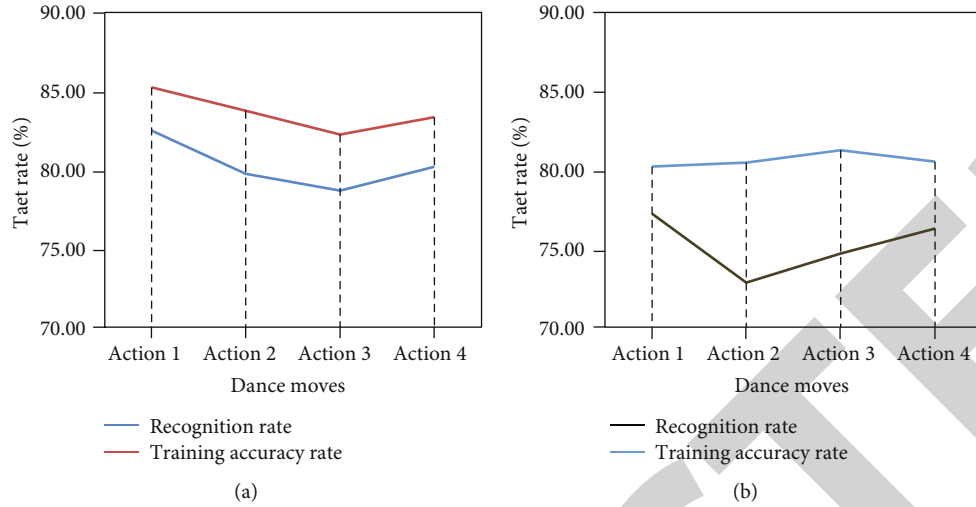


FIGURE 10: FolkDance dataset recognition test results.

contains several refined dance movements. The movement categories are relatively rich, and each group of dance movements is relatively complex and challenging.

4.2.1. DanceDB Dataset Identification Test. Figure 9(a) shows the recognition effect of edge computing technology.

Figure 9(b) shows the recognition effect of traditional recognition technology.

It can be seen from Figure 9 that in the DanceDB dataset identification test, the overall mean of the recognition rate of the dance movements in the sample set by the recognition method based on edge computing technology is 74.73%, and the overall mean of the training accuracy of the sample set is 82.80%. However, the overall mean of the recognition rate of the dance movements in the sample set by the identification method based on the traditional technology is 60.07%, and the overall mean of the training accuracy of the sample set is 78.51%.

4.2.2. FolkDance Dataset Recognition Test. Figure 10(a) shows the recognition effect of edge computing technology.

Figure 10(b) shows the recognition effect of traditional recognition technology.

It can be seen from Figure 10 that in the FolkDance dataset recognition test, the recognition method based on edge computing technology has an overall average recognition rate of 80.21% for dance movements in the sample set and an overall average of 83.56% for the training accuracy of the sample set. However, the overall mean of the recognition rate of the dance movements in the sample set by the identification method based on the traditional technology is 75.21%, and the overall mean of the training accuracy of the sample set is 80.54%.

5. Discussion

Through the comparative experimental data of edge computing technology identification method and traditional identification method, the following conclusions can be drawn:

- (1) In terms of the accuracy of footprint morphological feature extraction, the overall average test of the recognition method based on edge computing technology is 12.14 higher than that of the traditional recognition technology test
- (2) In terms of the accuracy of footprint length feature extraction, the overall mean of the edge computing-based identification method is 11.49 higher than that of the traditional identification technology
- (3) In the DanceDB dance movement dataset test, the overall average recognition rate of the recognition method based on edge computing technology is 13.66% higher than the overall average recognition rate of the traditional recognition technology. The overall mean of training accuracy based on edge computing technology identification method is 4.29% higher than the overall mean of traditional identification technology training accuracy
- (4) In the FolkDance dance movement dataset test, the overall average recognition rate of the recognition method based on edge computing technology is 5% higher than the overall average recognition rate of the traditional recognition technology. The overall mean of the training accuracy of the recognition method based on edge computing technology is 3.04% higher than the overall mean of the training accuracy of the traditional identification technology

The whole comparative experimental data shows that under the condition of keeping other experimental conditions the same, the footprint extraction and sports dance action recognition tests of different algorithm recognition technologies are performed. Whether it is in the extraction of morphological features and length features of footprints or in the recognition rate and training accuracy of sports dance movements, the distributed edge computing technology based on artificial intelligence is more superior. It shows that the recognition technology under edge computing based

on artificial intelligence technology can be effectively applied to the movement recognition of footstep extraction and sports dance.

6. Conclusion

The research of footprint extraction and action recognition has always been a very challenging topic in the academic field. With the rapid maturity of information technology, people are more and more closely integrated with intelligent technology. At the same time, the requirements for it are getting higher and higher. Whether it is in the spiritual life of the people or in the material life, the recording and preservation of footprint extraction and motion recognition are of great value and significance. Distributed edge computing technology based on artificial intelligence technology has powerful learning capabilities and data processing and analysis capabilities. It can effectively extract information features, remove invalid movements, reduce the amount of calculation, and improve the accuracy of footprint extraction and dance movement recognition efficiency. It is believed that with the maturity and improvement of the algorithm, the edge computing recognition technology in the field of artificial intelligence technology will be more and more high-quality and high-level development. Although this paper has carried out in-depth research on footprint extraction and action recognition using distributed edge computing technology under artificial intelligence technology, there are still many shortcomings. The depth and breadth of the research in this paper is not enough. In the process of this research, the selection and acquisition of experimental data are carried out under absolutely ideal conditions. The completeness and validity are not enough, some interference factors involved in the sample collection process are not considered, and the test evaluation is also restricted by many factors. The author's academic level research is also limited, and the research on edge computing technology is still in the preliminary stage. In the future work, the methods and means of feature extraction and action recognition will be studied from more perspectives based on the existing technology and level, and the algorithm will be continuously optimized.

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Conflicts of Interest

The author states that this article has no conflict of interest.

References

- [1] T. X. Tran, A. Hajisami, P. Pandey, and D. Pompili, "Collaborative mobile edge computing in 5G networks: new paradigms, scenarios, and challenges," *IEEE Communications Magazine*, vol. 55, no. 4, pp. 54–61, 2017.
- [2] Z. Ke, Y. Mao, S. Leng, Y. He, and Y. Zhang, "Mobile-edge computing for vehicular networks: a promising network paradigm with predictive off-loading," *IEEE Vehicular Technology Magazine*, vol. 12, no. 2, pp. 36–44, 2017.
- [3] K. Hirokatsu, S. Yutaka, A. Yoshimitsu, S. Oikawa, and Y. Matsui, "Temporal and fine-grained pedestrian action recognition on driving recorder database," *Sensors*, vol. 18, no. 2, pp. 627–627, 2018.
- [4] Y. Yanhua, C. Deng, S. Gao, W. Liu, D. Tao, and X. Gao, "Discriminative multi-instance multitask learning for 3D action recognition," *IEEE Transactions on Multimedia*, vol. 19, no. 3, pp. 519–529, 2017.
- [5] C. Li, Y. Hou, P. Wang, and W. Li, "Joint distance maps based action recognition with convolutional neural networks," *IEEE Signal Processing Letters*, vol. 24, no. 5, pp. 624–628, 2017.
- [6] Q. Ke, S. An, M. Bennamoun, F. Sohel, and F. Boussaid, "SkeletonNet: mining deep part features for 3-D action recognition," *IEEE Signal Processing Letters*, vol. 24, no. 6, pp. 731–735, 2017.
- [7] X. Wang, L. Gao, J. Song, and H. Shen, "Beyond frame-level CNN: saliency-aware 3-D CNN with LSTM for video action recognition," *IEEE Signal Processing Letters*, vol. 24, no. 4, pp. 510–514, 2017.
- [8] T. Kerola, N. Inoue, and K. Shinoda, "Cross-view human action recognition from depth maps using spectral graph sequences," *Computer Vision & Image Understanding*, vol. 154, pp. 108–126, 2017.
- [9] X. Lu, H. Yao, S. Zhao, X. Sun, and S. Zhang, "Action recognition with multi-scale trajectory-pooled 3D convolutional descriptors," *Multimedia Tools and Applications*, vol. 78, no. 3, pp. 1–17, 2019.
- [10] S. Singh, C. Arora, and C. V. Jawahar, "Trajectory aligned features for first person action recognition," *Pattern Recognition*, vol. 62, pp. 45–55, 2017.
- [11] S. Yu, Y. Cheng, S. Su, G. Cai, and S. Li, "Stratified pooling based deep convolutional neural networks for human action recognition," *Multimedia Tools and Applications*, vol. 76, no. 11, pp. 13367–13382, 2017.
- [12] T. Hao, D. Wu, Q. Wang, and J. S. Sun, "Multi-view representation learning for multi-view action recognition," *Journal of Visual Communication and Image Representation*, vol. 48, pp. 453–460, 2017.
- [13] Z. Gao, S. H. Li, G. T. Zhang, Y. J. Zhu, C. Wang, and H. Zhang, "Evaluation of regularized multi-task learning algorithms for single/multi-view human action recognition," *Multimedia Tools and Applications*, vol. 76, no. 19, pp. 20125–20148, 2017.
- [14] Z. Tu, W. Xie, J. Dauwels, B. Li, and J. Yuan, "Semantic cues enhanced multi-modality multi-stream CNN for action recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 99, pp. 1423–1437, 2018.
- [15] M. Asadi-Aghbolaghi and S. Kasaei, "Supervised spatio-temporal kernel descriptor for human action recognition from RGB-depth videos," *Multimedia Tools and Applications*, vol. 77, no. 3, pp. 1–21, 2017.
- [16] Y. Fu, T. Zhang, and W. Wang, "Sparse coding-based space-time video representation for action recognition," *Multimedia Tools and Applications*, vol. 76, no. 10, pp. 12645–12658, 2017.
- [17] S. Hadfield, K. Lebeda, and R. Bowden, "Hollywood 3D: what are the best 3D features for action recognition?," *International Journal of Computer Vision*, vol. 121, no. 1, pp. 95–110, 2017.
- [18] S. Yu, Y. Cheng, L. Xie, Z. Luo, M. Huang, and S. Li, "A novel recurrent hybrid network for feature fusion in action

- recognition,” *Journal of visual communication & image representation*, vol. 49, pp. 192–203, 2017.
- [19] M. M. Moussa, E. E. Hemayed, H. Nemr, and M. B. Fayek, “Human action recognition utilizing variations in skeleton dimensions,” *Arabian Journal for Science & Engineering*, vol. 43, no. 5, pp. 1–14, 2017.
- [20] Y. Zheng, H. Yao, X. Sun, X. Jiang, and F. Porikli, “Breaking video into pieces for action recognition,” *Multimedia Tools & Applications*, vol. 76, no. 21, pp. 22195–22212, 2017.
- [21] J. Joo and H. Choi, “Improving parallelism for video action recognition model using one-dimensional convolutional neural network,” *KIISE Transactions on Computing Practices*, vol. 27, no. 4, pp. 216–220, 2021.
- [22] L. Zhi, W. M. Wang, L. Guo, J. He, and G. Q. Huang, “Toward open manufacturing: a cross-enterprises knowledge and services exchange framework based on blockchain and edge computing,” *Industrial Management & Data Systems*, vol. 118, no. 9, pp. 303–320, 2018.
- [23] L. He, K. Ota, and M. Dong, “Learning IoT in edge: deep learning for the Internet of Things with edge computing,” *IEEE Network*, vol. 32, no. 1, pp. 96–101, 2018.
- [24] H. Nagano, “Big data, information and communication technology, artificial intelligence, internet of things: how important are they for gastroenterological surgery?,” *Annals of Gastroenterological Surgery*, vol. 2, no. 3, p. 166, 2018.
- [25] I. Giuffrida, F. Lederer, and N. Vermeys, “A legal perspective on the trials and tribulations of AI: how artificial intelligence, the internet of things, smart contracts, and other technologies will affect the law,” *Case Western Reserve Law Review*, vol. 68, no. 3, p. 14, 2018.
- [26] H. Zhu, H. Wei, B. Li, X. Yuan, and N. Kehtarnavaz, “Real-time moving object detection in high-resolution video sensing,” *Sensors*, vol. 20, no. 12, p. 3591, 2020.
- [27] X. Xiang, Q. Li, S. Khan, and O. I. Khalaf, “Urban water resource management for sustainable environment planning using artificial intelligence techniques,” *Environmental Impact Assessment Review*, vol. 86, article 106515, 2021.
- [28] K. A. Ogudo, D. Muwawa Jean Nestor, O. Ibrahim Khalaf, and H. Daei Kasmaei, “A device performance and data analytics concept for smartphones’ IoT services and machine-type communication in cellular networks,” *Symmetry*, vol. 11, no. 4, pp. 593–609, 2019.
- [29] Z. Lv, D. Chen, R. Lou, and Q. Wang, “Intelligent edge computing based on machine learning for smart city,” *Future Generation Computer Systems*, vol. 115, pp. 90–99, 2021.
- [30] S. Wan, L. Qi, X. Xu, C. Tong, and Z. Gu, “Deep learning models for real-time human activity recognition with smartphones,” *Mobile Networks and Applications*, vol. 25, no. 2, pp. 743–755, 2020.
- [31] G. Suryanarayana, K. Chandran, O. I. Khalaf, Y. Alotaibi, A. Alsufyani, and S. A. Alghamdi, “Accurate magnetic resonance image super-resolution using deep networks and Gaussian filtering in the stationary wavelet domain,” *In IEEE Access*, vol. 9, pp. 71406–71417, 2021.
- [32] L. Jiang, “Footprint extraction and sports training action recognition based on wireless network communication,” *Wireless Communications and Mobile Computing*, vol. 2022, 2022.