Research Article

Weighted Classification of Machine Learning to Recognize Human Activities

Guorong Wu, Zichen Liu, and Xuhui Chen

School of Art and Design, Nanchang University, Nanchang, China

Correspondence should be addressed to Xuhui Chen; xuhui_chen202102@126.com

Received 3 February 2021; Revised 24 March 2021; Accepted 1 April 2021; Published 13 April 2021

Academic Editor: M. Irfan Uddin

Copyright © 2021 Guorong Wu 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.

This paper presents a new method to recognize human activities based on weighted classification for the features extracted by human body. Towards this end, new features depend on weight taken from image or video used in proposed descriptor. Human pose plays an important role in extracted features; then these features are used as the weight input with classifier. We use machine learning during two steps of training and testing images of standard dataset that can be used during benchmarking the system. Unlike previous methods that need size or length of shapes mainly to represent the cues when machine learning is used to recognize human activities, accurate experimental results coming from appropriate segments of the human body proved the worthiness of proposed method. Twelve activities are used in challenging of availability comparison with dataset to demonstrate our method. The results show that we achieved 87.3% in training set, while in testing set, we achieved 94% in terms of precision.

1. Introduction

In the era of Information Technology (IT), recognition of human activities plays a paramount role in many applications and study of computers. Recently, machine learning has aimed to take the features of the human body, analyze them, and then extract the informative data that achieve the reality of activities. Each human activity has certain features that indicate its type. Therefore, automatic learning depends mainly on the weight of these features in order to classify these features and thus obtain right detection of the activity.

Human Activities Recognition (HAR) using video and images has been widely carried out especially due to flexibility [1]. Anyway HAR in still image got less attention because of missing the prior information during processing and nature of the image. Hence, the image has fixed and rigid information difficult to deal with [2, 3]. The information needed to process any HAR system is extracted from both image and video, and the preprocessing needs to be done before implementation to get prior knowledge considering features extracted in advance. However, ground truth related to HAR considers both video and image and the challenging one is the image; and as known video consists mainly of a

sequence of images, predata and postdata are provided for extraction. The main goal of HAR system that uses image or video is to develop an automatic rendering system of gathered data to precisely determine the behavior of the man. HAR is useful in a wide range of applications, especially when appropriate implementation is suitable for human computer interfaces [4]. HAR scheme is mostly used in applications of using surveillance security system in terms of automation.

Every HAR system needs a classifier to classify the features for activity detection, and SVM is one of the best classifiers in this regard due to its easiness to use and reliability that can improve it to be suitable for the proposed method.

Systems that deal with pattern recognition must consist of inevitable stages like preparation stage (noise reduction and enhancement), segmentation (or partitioning), feature extraction, and classification [5]. In image, noise removal is necessary, and many techniques suggested in literature such as fuzzy system are related to image and are worth addressing [6]. Segmentation is also useful for segmenting images in video or the images themselves, machine learning depends on the features extracted from the human body,

and, in order to extract the data, we must first segment the human image because the segmentation gives the system a clearer view that the machine learning algorithm understands. [7]. Machine learning should extract important features that are fed to the classifier, which would be data estimated for machine during learning; then another inheriting subinformation within will be derived, [8] actually selecting best features considering the main issue to control HAR system and make the system useful or not [9, 10].

Recently, many classifiers have been suggested and developed accurately for real-time applications, achieving good results [11]; the majority of these classifiers were developed mathematically and the famous one is Support Vector Machine (SVM) which will also be developed in depth in this paper. In addition to several classifiers such as K-Nearest Neighbor (KNN) and Linear Discriminant Analysis (LDA) which perform good classification of human activities [12, 13], real-time applications need high response time to be suitable; few classifiers used large features for their processes to achieve accurate results. Generally, the system that deals with pattern recognition or classification consists of three main steps: preprocessing, feature extraction, and finally classification by suitable classifier [10, 14].

The purpose of classification is to recognize the human activities behavior; in this regard understanding movement of human is necessary for understanding and executing all activities. HAR system should take the real analysis of each movement with its response. Sometimes the shape of human body takes shapes of feature not easy to understand due to similarity of two activities like walking and running, where human body takes most likely same position and shape. Figure 1 depicts the visualized the activities by each action like running, walking, and other activities, so one part of human body is easy to detect, while other parts are difficult to detect for some activities. Some motion activities are difficult to detect by a special domain like smiling, crying, or sadness; thus a frequency domain (DT or CWT filters) is needed for feature extraction [15, 16].

Video and static image are the two main domains focused on in this paper; the challenging one is still image; therefore, the proposed method will contribute to both video and image. Due to the fact that the recognition of activity from single or still image actually somewhat has ignored prior knowledge, the results will be harder. For this, existing methods were improved and contributed to this regard [17]. Extracting feature for classification of human activities can be both hardware and software. Hardware used wearable sensors to detect the activities [18, 19], while software used extracting features within programming tools. Figure 2 illustrates the two different domains mentioned above.

Recently, video surveillance camera has been widely used outdoors and indoors for security reasons in many fields: industry, medicine, education and so forth, and the camera-produced video needs to be processed outside the circuit to handle it [20]. Most of images need to reduce noise, which is considered as preprocessing stage to prepare for extracting useful data from it; many algorithms for noise reduction were suggested in literature [21] and used to prepare the image without contribution.

HAR is considered as an important topic; therefore many surveys have been introduced and published [9, 22]. Some survey papers are categorized as classifying the activities and clustering, while the others consider features extracted with their types; anyway most of them introduce detailed HAR system as a big issue, since HAR system is complex due to including many necessary stages. The most important studies in literature consider the applications of HAR system and where we can apply it in different life fields [23].

HAR system consists of three main stages: preparation stage, feature extraction stage, and classification stage. Deep learning is widely spread nowadays and has become necessary due to its use in many modern applications. HAR system used within deep learning algorithms such as any system has to learn the machine to have an effect on the human jobs and aims to use the source of computational machine perfectly. Prediction of future results is needed because the parameters of each system increased and became necessary to be controlled by deep learning such as traffic flow congestion [24, 25]. Automated behavior of machine has become worthy in this world, so machine learning is important here, which coincided with the economic development of some countries like China [26].

2. Literature Review

Many researchers proposed different techniques in literature. We present the most important studies in the field of machine learning. Complex features are extracted by using end-to-end machine learning and multimodal temporal fusion in smartphone [27]. HAR method based on wearable sensors was suggested in [28] and used stacked autoencoder to extract powerful features from human body. Mobile edge computing in video but used frame by frame as image, sensors of multifunctional simulated by smart phone to gather information from human body to recognize the activities in medical issue [29]. Monocular images extracted from video in 3D and 2D poses to verified from which activity was presented in [30]; the authors used two high parameters in still and video images. Random forest model was presented to enhance the deep learning, and 40 activities were recognized with good performance of HAR system [31].

3. HAR System in Machine Learning

Acquired data were collected from various types such as embedded and vision sensors; data will be analyzed in the preprocessing stage. Information coming from devices is classified into two categories; data comes from vision sensors like camera or wearable sensors; there is a difference in preprocessing between each category; data from sensors need to reduce noise and enhance the image, while segmentation was required for the data that come from camera [13]. Simple architecture illustrated in Figure 3 shows the main data entry to the system with related preprocessing action.

FIGURE 1: Different activities with KTM dataset.

FIGURE 2: Both domains for extracting features to classify the activities of human behavior.

FIGURE 3: General system of HAR including the most important stages.

3.1. Collecting Data. Information dealing with HAR system consists of two categories: the first is sensor based HAR system and the second is vision based HAR system. One can predefine a set of n activities or actions A ; then the following can be noted:

$$A = \{A_0, A_1, \dots, A_{n-1}\}, \quad (1)$$

and multiple groups of sensors are used to measure the number of attributes S in k interval time series of $I = [t_a, t_w]$, so these activities can be yielded:

$$S = \{S_0, S_1, \dots, S_{k-1}\}. \quad (2)$$

The purpose of the HAR system is to find partition of temporal of I such as $[I_0, I_1, \dots, I_{r-1}]$ and list of attributes S ;

classes collection will describe activities during each time partition I_k , and the collection of time can be defined as

$$U_{k=0}^{r-1} I_k = I. \quad (3)$$

Vision based HAR depends on sensing device technology like CCTV and camera with recording capabilities to record activities of the human [32]. In this approach, it does not need wearable sensors or smartphone but depends mainly on image quality taken from camera. Quality of image is determined by image resolution, illumination change, and environmental lighting. Data collected behave like sequence of images or audio signal or vision of computer. After collecting data, they will pass through processes such as feature extracting, modeling with segmentation for activities, and then activities classification and tracking.

3.2. Machine or Deep Learning Based HAR. Signal processing is a technique that has been used for a long time to analyze collected data from sensors [33]. This approach is implemented to perform preprocessing and extract engineering features from video or image; next, extracted features are trained in machine learning (ML) algorithm for classification to decide the activity [26]. An easy method for classification using logistic regression to distribute the probability under the formula behaves the same as classification model:

$$\begin{aligned} P(A = 1|x) &= \frac{\exp(mx + n)}{1 + \exp(mx + n)}, \\ P(A = 0|x) &= \frac{\exp(mx + n)}{1 + \exp(mx + n)}. \end{aligned} \quad (4)$$

Extracting feature engineering will be analyzed manually for selecting suitable set of features and then will be implemented to reduce feature space. Deep learning has become an interesting research topic in the past decades; then it will deal with human performance in different research areas including the HAR system. For traditional ML, normal extraction of features is performed in HAR system but with large datasets and real-time HAR system needs to improve and increase the efficiency. There are different terms and characteristics of machine learning and deep learning, which are presented in Table 1. The proposed HAR system relies on machine learning and its conditions and requirements are met [25].

Other classifiers are used in HAR system such as naïve Bayes but the ability to manipulate increasing of features is not good and the classifier itself is difficult to modify with new type of features. The decision tree classifier is not able to be adaptive with deep learning method due to the fact that deep learning is changing the process during running the system and adaptive processing is needed. The biggest issues with classifiers are the features extraction and what kind of processing these features require.

3.3. Learning HAR System. There has been an exponential growth in using the HAR system in the recent years; for getting more straightforward and efficient data process,

learning algorithm was improved and developed involving collecting with analyzing data, pattern recognition, and useful features that are fed to the system of ML algorithm [34].

Many algorithms are introduced in literature in the field of ML and they are generally classified into the following categories:

- (1) Classification algorithm: represented by naïve Bayes and decision tree algorithm
- (2) Rule extraction algorithm: deals with prior knowledge algorithm
- (3) Clustering algorithm: such as K-means algorithm and EM algorithm
- (4) Special support vector machine: defined as SVM algorithm with its improvement
- (5) Neural network with hidden layers
- (6) Genetic algorithm

An important stage in HAR system classification in this article is using an improved SVM classifier to classify the activities. One of the advantages of using such classifier (SVM) is being flexible and developable [35]. Support vector classification gets vector training $x_i \in \mathbb{R}^n$ $i = 1, \dots, l$ for two categories or classes and the vectors $y \in \mathbb{R}^l$ and $y_i \in \{1, -1\}$, so we can define the primitive problem:

$$\begin{aligned} \min \quad & \frac{1}{2} w^T w - \nu p + \frac{1}{l} \sum_{i=1}^l \epsilon_i \\ Y_i(w^T \phi(x_i) + b) & \geq p - \epsilon_i \\ \epsilon_i & \geq 0 \quad i = 1, \dots, l \quad p \geq 0. \end{aligned} \quad (5)$$

Actually, this vector has truth of $0 \leq \nu \leq 1$ and the training is on vector x_i and it is mapped in ϕ .

To improve the SVM classifier, first there is a need to play with factors of the classifier and find the new path so the algorithm with required results can be adapted. There are linear and nonlinear classifiers; the simple linear SVM is defined as (x_i, y_i) , where $i = 1, \dots, N$ and $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$ with attribute of i -th sample; if considering $y_i \in \{-1, 1\}$, which is class label, then the decision will be

$$w^T x + b = 0. \quad (6)$$

Since w is weight vector and b is the bias, when the training is linearly separated, coordinate (w, b) will be defined as

$$\begin{cases} w^T x + b \geq 1 & \text{if } y_i = 1, \\ w^T x + b < 1 & \text{if } y_i = -1. \end{cases} \quad (7)$$

And the function of linear classifier is defined as follows:

$$F(x) = \text{sign}(w^T x + b). \quad (8)$$

So, during training for a given dataset to decide the i -th plane margin of sample x_i when defining the plane (w, b) can be defined as

TABLE 1: The different characteristics of ML and deep learning as regards the HAR system.

Process	ML	Deep learning
Preprocessing	Preprocessing is required and necessary for the performance	Preprocessing is not required; also normalization method
Features collection	(i) Depends on manual extraction (ii) Avoids complicated activities (iii) Considers feature selection with reduction technique (iv) Fails to handle variability and similarity class	(i) Learns to extract the features directly (ii) Temporal and spatial dependence with scaling features automatically
Learning process	(i) Considers small training data (ii) Less memory and computational time are needed	(i) Works with large data (ii) Works well with high complexity and computation (iii) Needs special hardware to perform training

$$\gamma_i = y_i(w^T x + b). \quad (9)$$

Function margin given image from dataset during training to make decision boundary that consists of two lines that is gap called margin and width of this margin is controlled critically, where the width is defined as w in equation (9).

Minimizing the margin makes the classifier more accurate and the classifier has to be optimized as given information below.

Control w , b , and $\varepsilon_i \geq 0$.

$$\Phi(w) = w^T w + c \sum_{i=1}^N \varepsilon_i. \quad (10)$$

After minimizing, the function will be

$$f(x) = \sum a_i y_i x_i^T x + b. \quad (11)$$

3.4. Nonlinear HAR System. In the proposed method, there is a need to get the optimal hyperplane which is defined as maximizing the generalization. The problem is when the training gains the data separated not linearly which means when running the classifier linearly it cannot reach good result. Due to feature space which mapped as a high dimensional not allowed to perform the classifier well. For this reason, the proposed method provides efficient way to solve higher-dimensional space called kernel trick [36] or weighted features. When the linear system depends on dot product, develop classifier to be inner product to kernel functions for training data. In this regard, the proposed method is designed to find $a_1, \dots, a_n$ in the following equation:

$$Q(a) = \sum_{i=1}^N a_i - \frac{1}{2} \sum_{i,j=1}^N a_i a_j y_i y_j K(x_i, x_j). \quad (12)$$

Here is the maximization with respect to a factor, since

$$\sum_{i=1}^N a_i y_i = 0 \quad C \geq a_i \geq 0. \quad (13)$$

So the classifier can be applied by using the following equation:

$$f(x) = \sum a_i y_i K(x_i, x_j) + b. \quad (14)$$

For given images in dataset provided have some activities try to train the system using improved classifier to classify six activities. The system can run for more activities through training; then proceed with testing the system. Weighting features extracted are also used as the factor of classification $\Gamma_{\text{all}} = \{\Gamma_1, \Gamma_2, \dots, \Gamma_m\}$.

$$f(x) = \sum a_i y_i \Gamma_m K(x_i, x_j) + b. \quad (15)$$

Vector of weighted features can be variable and jumping among activities x . Weight in one activity differs from the others duo to the nature of the weight, with proposed method feature used as angles between human body and terminals like arms, legs, and so on.

4. Method and Discussion

In each HAR system for machine learning there exist three main stages, which are preprocessing, feature extraction, and finally classification which is our main concern. Images taken from different sources like camera or scanner are considered as input to the HAR system. Two interesting datasets are used such as INRIA and KTH getting from public domain of six activities which are walking, running, jogging, boxing, clapping, and waving. Images are normalizing the size as recognition rate estimation. Recognition rate can be calculated by the following formula:

$$\text{Recognition rate} = \frac{\text{Number of corrected activities}}{\text{Number of given activities}} \times 100\%. \quad (16)$$

Any HAR model consists of stages illustrated in Figure 4.

The system starts with image coming from standard dataset which contains more than 1580 activities, each with 9 actions. The proposed method analyzes the activities automatically via the designed algorithm to reduce effort of the human and get instant response.

Schematic of the proposed method is shown in Figure 4, which consists of preprocessing stage of segmentation of the image. This stage is sometimes called preparation stage to prepare the properties of given image to the next step; this

FIGURE 4: Structure of HAR system in the proposed method.

stage must be taken into consideration as an important stage due to the fact that processing here decides whether the features are extracted properly or not. Any mistake in this process can disrupt the system and may not give accurate results. Some images from datasets come with noise or plankton; to get rid of this noise, a powerful expression like the one in equation (16) needs to be used.

$$f(x, y) = \frac{1}{mn - d} \sum_{(s, t) \in S_{xy}} \text{Id}(s, t), \quad (17)$$

where S_{xy} represents neighbor pixels of image I 's size ($M \times N$) for corresponding coordinate (x, y) ; Id represents remaining gray level value such as $d/2$ detected size.

Some images got noise and others not, but the majority are clear images; we apply noise reduction for integrity of the system. Another issue in preparation step is segmentation; in this process, there are two conditional actions; the first is background subtraction which means we extract the desired foreground object, in our case the human body, by background subtraction from the image to get human shape; this process is yielded when applying the expression in equation (17) to a given image.

$$\text{St}(x, y) = \begin{cases} 1 & f(x, y) \text{ belong to} \\ & \text{foreground at time } t \\ 0 & \text{otherwise.} \end{cases} \quad (18)$$

Another segment type considers part of the image and relation of object (human body) with the image environment; this process is variable from one HAR system to another and most studies in literature contribute to this stage. Process in this stage provides action to next stage, which is feature extraction stage.

Feature extraction gains its importance from the fact that any classification depends on features extracted in addition to the fact that the classifier cannot reach any result without vectors of good features. Features are extracted from both

the object itself and position of each part. Figure 5 shows how human body will be segmented within given image.

Information extracted from segment image is stored as vector of $f(\mu) = \{\mu_1, \mu_2, \dots, \mu_n\}$, where μ is the data extracted from one segment. What applies to one activity automatically applies to the other activities.

Other features extracted using triangle scan for all images are called triangle features, and triangle scan is applied flexibly and is resizable in addition to rotational case as shown in Figure 6.

Triangle can scan all the images starting from top left to down right with 360-degree rotation for each step. Other features are angles of vertical elevation of the object with the ground horizon line, angles of the triangle, and vertical elevation stored also in vectors for later classification with the classifier.

With the proposed method, novel features are extracted from human body and then their weights are used in nonlinear classifier such as improved SVM. Features control the results of classifier; when getting certain results, arrangement of features starts to change according to their weights, and changing the weights occurs during training iteration for classifier. During segmentation of the image into subsegment parts, the system detects whether there is any part of human body. As human body parts change their positions during activities, we can detect the new positions accordingly, as shown in Figure 7.

5. Results and Discussion

Two interesting datasets are used to evaluate the proposed method, which are obtained from public domain. Datasets used are KTH and INRIA, all with six activities, which are walking, running, jumping, boxing, waving, and clapping. Eight scenarios are contained inside the activities. The image used in the proposed method was first normalized to 64×128 pixels, with front side considered for waving and clapping, while for the rest of activities side view is considered. These six activities are chosen due to their popularity in static image of datasets.

Nonlinear SVM classifier is used to get highly accurate recognition rate when dynamic features are extracted from the image considered. Some activities like walking, running, and jogging get high results as regards recognition rate due to the fact that the features extracted as angles are affected directly with these activities. When the rest of activities get less recognition rate because of the nature of such activities, features gain less in such activities. Applicability of the proposed method is limited with online recognition rate and time computation through running which is longer than the others. This method can get accurate result in still image without online processing by using proposed features in terms of richness. The proposed method obtains a high recognition rate of 87.3%; it is the ratio of the correct classification to the wrong one and is considered as a good result in HAR system. In still images, this result is considered to be high due to the lack of reliable previous information in

FIGURE 5: Segmentation of the human body within one activity.

FIGURE 6: How features are extracted from human object through an image.

FIGURE 7: Behavior of classifier with corresponding weights.

TABLE 2: Recognition rate of the proposed method by applying extracted features.

Running	0.86	0.02	0.03	0.01	0.04	0.03
Walking	0.13	0.87	0.0	0.09	0.03	0.02
Boxing	0.0	0.01	0.84	0.1	0.0	0.0
Jogging	0.03	0.09	0.12	0.87	0.0	0.0
Waving	0.0	0.01	0.02	0.03	0.86	0.11
Clapping	0.01	0.02	0.01	0.01	0.11	0.85
	Running	Walking	Boxing	Jogging	Waving	Clapping

FIGURE 8: Machine learning of the proposed method including training and testing modes.

still images, and it is not like in video, where previous information can be relied upon, such as the previous frame, which will of course change the result.

Data in Table 2 are called confusion matrix of supervised method. Results depicted inside the confusion matrix show that some activities get more interest such as walking, running, and jogging because of being adapted with suggested features (angles), so the system is able to recognize them perfectly.

The proposed model within features extraction focuses on arms and legs with respect to the direction of the vertical body. In contrast, other activities such as clapping and waving focus on triangle derived from the arms and legs.

Training system with collected datasets for six activities reaches acceptable results when using improved SVM classifier as compared with existing methods.

The proposed method allows evaluating the system's accuracy for six activities using two datasets evaluating stand with activities shared between these datasets. Training system got good response time but delay occurs when changing image to gray scale followed by noise reduction and segmentation. Training depends also on number of images inside the dataset; training will be with process samples of known activities identified by dataset; then when finishing the training, start with testing unknown images according to features vectors belonging to a particular class. Features are entered into the classifier by performing analysis in the following equation:

$$C = (x_i, y_i) | x_i \in \mathcal{R}, \quad (19)$$

where $y_i \in \{1, -1\}$ for all class $i = 1, 2, \dots, n$.

Each x_i in a vector of real n-dimensional and y_i indicates whether the feature belongs to = 1 or else = -1, which means that two classes A and B in this case are categorized as separated data. During classification of machine learning,

decision rule will created as $d(x): \mathcal{R}^d \rightarrow \{w_1, w_2, \dots, w_c\}$ to classify certain X belonging to one class. Observation of training set yields building rule in the training. Prediction stage, also named as testing stage, takes the given image and processes it to apply machine learning during training set. The system will compare the features extracted from the given image with the features extracted during training process to find the suitable match or closer counterpart to recognize the activity. Training and testing modes are presented in Figure 8.

A number of features suggested are considered as powerful features supported by HAR system. Most of study improve and enhance in this section because of this part still need to improve in additional to dataset used limit the feature used according to the nature of the object inside the image itself. Using accurate classifier also contributes to achieving good results in a way of adapting it with the proposed method.

6. Conclusion

We report a system of HAR machine learning that recognizes human activities such as walking, running, jogging, walking, clapping, and waving. The proposed system mainly consists of three main stages, which are preprocessing, feature extraction, and classification. Our contribution is to the last two stages for system integration. In terms of machine learning classifier considered as the heart of the HAR system, improvement in this context yields good results. New suggested features like angles derived from triangles and human body verticality with the horizon line play an important role in HAR system. The proposed HAR system achieved good results in terms of recognition rate by classifying the activities and accuracy, while real-time response was not considered in this study.

Data Availability

Two standard datasets were used in the proposed system, each with more than 890 available images, to recognize six activities. Two types of features were derived from the human body via angles used to run SVM classifier. Improving the mathematical issue of SVM classifier was helpful to improve the system. The image used in the HAR system uses normalization scale of 24×128 pixels.

Conflicts of Interest

The authors confirm that there are no conflicts of interest regarding the publication of this paper.

References

- [1] H. Mliki, F. Bouhlef, and H. Mohamed, "Human activity recognition from UAV-captured video sequences," *Pattern Recognition*, vol. 100, Article ID 107140, 2020.
- [2] M. Kinoshita, "Image processing device, image processing method, and program," *U.S. Patent Application*, vol. 16/713, p. 766, 2020.
- [3] G. Sulong and A. Mohammedali, "Recognition of human activities from still image using NOVEL CLASSIFIER," *Journal of Theoretical & Applied Information Technology*, vol. 71, p. 1, 2015.
- [4] M. Lupión, J. Medina-Quero, J. F. Sanjuan, and P. M. Ortigosa, "DOLARS, a distributed on-line activity recognition system by means of heterogeneous sensors in real-life deployments-A case study in the smart lab of the university of Almeria," *Sensors*, vol. 21, no. 2, p. 405, 2021.
- [5] P. Tu, J. Li, H. Wang, T. Cao, and K. Wang, "Non-linear chaotic features-based human activity recognition," *Electronics*, vol. 10, no. 2, p. 111, 2021.
- [6] J. Guan, R. Lai, A. Xiong, Z. Liu, and L. Gu, "Fixed pattern noise reduction for infrared images based on cascade residual attention CNN," *Neurocomputing*, vol. 377, pp. 301–313, 2020.
- [7] J. Yan, F. Angelini, and S. M. Naqvi, "Image segmentation based privacy-preserving human action recognition for Anomaly detection," in *Proceedings of ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, Barcelona, Spain, May 2020.
- [8] G. Sulong and A. Mohammedali, "Human activities recognition via features extraction from skeleton," *Journal of Theoretical & Applied Information Technology*, vol. 68, no. 3, 2014.
- [9] A. Mohite, P. Rege, and D. Chakravarty, "Human activity recognition using positioning sensor and deep learning technique," *Advances in Signal and Data Processing*, pp. 473–489, Springer, Singapore, 2021.
- [10] L. Wang and R. Liu, "Human activity recognition based on wearable sensor using hierarchical deep LSTM networks," *Circuits, Systems, and Signal Processing*, vol. 39, no. 2, pp. 837–856, 2020.
- [11] B. T. Atiyha, S. Aljabbar, A. Ali, and A. Jaber, "An improved cost estimation for unit commitment using back propagation algorithm," *Malaysian Journal of Fundamental and Applied Sciences*, vol. 15, no. 2, pp. 243–248, 2019.
- [12] A. Mohammedali, *Human Activities Recognition in Still Image*, LAP LAMBERT Academic Publishing, Sunnyvale, CA, USA, 2016.
- [13] L. M. Dang, "Sensor-based and vision-based human activity recognition: a comprehensive survey," *Pattern Recognition*, vol. 108, Article ID 107561, 2020.
- [14] F. Zhang, K. Niu, J. Xiong et al., "Towards a diffraction-based sensing approach on human activity recognition," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 1, pp. 1–25, 2019.
- [15] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: a survey," *Pattern Recognition Letters*, vol. 119, pp. 3–11, 2019.
- [16] J. Shukla, M. Barreda-Ángeles, and J. Oliver, "Feature extraction and selection for emotion recognition from electrodermal activity," *IEEE Transactions on Affective Computing*, vol. 99, 2019.
- [17] Z. Qin, Y. Zhang, S. Meng, Z. Qin, and K.-K. R. Choo, "Imaging and fusing time series for wearable sensor-based human activity recognition," *Information Fusion*, vol. 53, pp. 80–87, 2020.
- [18] K. Wang, J. He, and L. Zhang, "Attention-based convolutional neural network for weakly labeled human activities' recognition with wearable sensors," *IEEE Sensors Journal*, vol. 19, no. 17, pp. 7598–7604, 2019.
- [19] K. Li, "Applying multivariate segmentation methods to human activity recognition from wearable sensors' data," *JMIR mHealth and uHealth*, vol. 7, no. 2, p. e11201, 2019.
- [20] Q. Chen and M. Dong, "Detection and adaptive video processing of hyperopia scene in sports video," *Complexity*, vol. 2021, Article ID 6610760, 13 pages, 2021.
- [21] Y. Zhou, G. Gu, and Q. Ren, "Research on the relationship between business cycle and industrial fluctuations in north-east China based on complete ensemble empirical mode decomposition with adaptive noise," *Complexity*, vol. 2021, Article ID 8832201, 16 pages, 2021.
- [22] F. Demrozi, "Human activity recognition using inertial, physiological and environmental sensors: a comprehensive survey," 2004.08821, 2020.
- [23] M. M. Hassan, M. Z. Uddin, A. Mohamed, and A. Almogren, "A robust human activity recognition system using smartphone sensors and deep learning," *Future Generation Computer Systems*, vol. 81, pp. 307–313, 2018.
- [24] Y. Shi, "A survey of hybrid deep learning methods for traffic flow prediction," in *Proceedings of the 2019 3rd International Conference on Advances in Image Processing*, Chengdu, China, November 2019.
- [25] Y. Hou, Z. Deng, and H. Cui, "Short-term traffic flow prediction with weather conditions: based on deep learning algorithms and data fusion," *Complexity*, vol. 2021, 14 pages, 2021.
- [26] H. Sun and X. Zhong, "Impact of financial R&D resource Allocation efficiency based on VR technology and machine learning in complex systems on total factor productivity," *Complexity*, vol. 2020, 15 pages, 2020.
- [27] M. Gjoreski, V. Janko, G. Slapničar et al., "Classical and deep learning methods for recognizing human activities and modes of transportation with smartphone sensors," *Information Fusion*, vol. 62, pp. 47–62, 2020.
- [28] T. Ozcan and A. Basturk, "Human action recognition with deep learning and structural optimization using a hybrid heuristic algorithm," *Cluster Computing*, vol. 23, pp. 1–14, 2020.
- [29] 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.

- [30] D. Luvizon, D. Picard, and H. Tabia, "Multi-task deep learning for real-time 3D human pose estimation and action recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, inprint, 2020.
- [31] Y. Guo, J. Yu, and W. Wu, "Research on still image activity recognition based on deep learning," in *Proceedings of Eleventh International Conference on Graphics and Image Processing (ICGIP 2019)*, vol. 11373, International Society for Optics and Photonics, Hangzhou, China, October 2020.
- [32] K. Prastika, "Application of individual activity recognition in the room using CNN Alexnet method," *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, vol. 1007, no. 1, 2020.
- [33] A. Subasi, D. H. Dammas, R. D. Alghamdi et al., "Sensor based human activity recognition using adaboost ensemble classifier," *Procedia Computer Science*, vol. 140, pp. 104–111, 2018.
- [34] L. Huang, "Personnel scheduling problem under hierarchical management based on intelligent algorithm," *Complexity*, vol. 2021, 14 pages, 2021.
- [35] L. Wu and M. Li, "Applying a probabilistic network method to solve business-related few-shot classification problems," *Complexity*, vol. 2021, Article ID 6633906, 12 pages, 2021.
- [36] S. S. Tabrizi, S. Pashazadeh, and V. Javani, "Comparative study of table tennis forehand strokes classification using deep learning and SVM," *IEEE Sensors Journal*, vol. 20, no. 22, pp. 13552–13561, 2020.