

## Research Article

# An Efficient Feature Selection Method for Video-Based Activity Recognition Systems

Muhammad Hameed Siddiqi  and Amjad Alsirhani 

*Department of Computer Science, Jouf University, Sakaka, Saudi Arabia*

Correspondence should be addressed to Muhammad Hameed Siddiqi; [mhsiddiqi@ju.edu.sa](mailto:mhsiddiqi@ju.edu.sa)

Received 4 December 2021; Revised 26 January 2022; Accepted 1 February 2022; Published 22 February 2022

Academic Editor: Dost Muhammad Khan

Copyright © 2022 Muhammad Hameed Siddiqi and Amjad Alsirhani. 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.

Human activity recognition (HAR) is the examination of gestures and actions of humans from various resources such as depth or RGB cameras. In this work, we have designed a dynamic and robust feature selection algorithm for a HAR system, through which the system accurately recognizes various kinds of activities. In the proposed approach, we employed mutual information algorithm, which selects the prominent features from the extracted features. The proposed algorithm is the expansion of two methods like max-relevance and min-redundancy, respectively. This method has the capability to gather the assets of various extraction algorithms. But the procedure of selection may be unfair due to the dissimilarity between the classification power and redundancy of the features. To resolve this type of unfair selection, we stabilize both parts through the proposed algorithm that has autonomous upper limit of the mutual information function. Likewise, for the feature extraction and recognition, we used the symlet wavelet transform and hidden Markov model, respectively, for action classification. The proposed algorithm has been justified on depth-based database which has thirteen kinds of activities under comprehensive set of experiments. We showed that the proposed feature selection method achieved best classification accuracy against existing works.

## 1. Introduction

Human activity recognition (HAR) refers to a technique that a computer machine utilizes to automatically classify what action(s) is/are being executed, given a series of activity frames. Video-based activity classification has achieved massive attention in healthcare telemedicine since last two decades [1–6], somewhat because of its applications in various areas like video surveillance, activity analysis in healthcare, and human-computer interaction (HCI) [7]. Activity recognition may utilize in healthcare and telemedicine in order to improve the quality of lives of elderly patients, through which the clinicians and medical experts remotely monitor the daily routines of the corresponding patients and may recommend suitable recommendations for them.

Commonly, the HAR system is the classification of activities performed by various subjects against RGB or depth cameras. In HAR, activities recorded by color cameras

are disturbed by lighting effects, occlusion, and dynamic backgrounds in real environments [8]. Activity recognition through wearable sensors like accelerometer or gyroscope has significant performances in healthcare and telemedicine [9]. However, the activity collection using accelerometers and gyroscopes reduces the comfort of human body and dismisses the naturalness of HCI [10]. Therefore, the focus of this study is to employ Kinect depth camera for data collection.

Therefore, HAR is one of the genuine modules in personalized healthcare and telemedicine systems, for the aged and disabled people [11]. In order to monitor the daily routine of aged people, various types of cameras might be employed in smart homes or hospitals to get activities from videos. Based on the survey of WHO, the number of the aged people is quickly growing in the whole world, and their living homes need much resources such as human and healthcare expenses. Hence, the intensive care services are required to tackle the wide use of resources in order to

improve the living styles of the aged people [12]. Lots of studies have been recommended for the intensive care services that might reduce the death rates for the aged people. For instance, in Europe, it is assessed that the survival ratio of aged people is growing while getting the intensive care services instead of getting care at their homes [13]. Therefore, a robust and efficient depth-based approach is required for HAR systems that monitors day and night the actions of aged people and delivers them an intellectual living home which makes their lives easy at home.

A recent method based on spatial temporal transformer network was proposed in [14], which is the representations of additions among joints against the operator of the spatial temporal transformer. In their model, they recognized the intra-state interactions between various parts of the body by employing spatial self-care (SSC). On the other hand, a temporal self-care module (TSC) technique was used to consider the inter-state associations. However, it is very difficult for such system to embed the skeleton data in the structure of graph, which means that this approach is not considering the graph convolution of the skeleton sequences in processing the specified motion features [15]. Similarly, an integrated method was designed in [16] for a HAR system, which was the combination of recurrent neural network and a heuristic feature selection method (like) meta-heuristic optimization. However, recurrent neural network is very difficult to train that cannot process long sequence of frames. Moreover, this approach has a common limitation like violation of the problem restrictions, which caused research community to resolve the optimization concerns with distinct variables [17]. An improved version of the spatio-temporal-based method was proposed in [18] for the activity recognition systems. In their systems, string detector was utilized in order to perceive the spatial position of every frame. However, the result of this approach may degrade due to considering the short length time sequences that make the statistical performance of correlation doubtful, and this is one the major limitations for this approach [19]. On the other hand, a novel approach was proposed in [20], which was based on motion information to categorize various kinds of activities. They employed Laplacian pyramid depth motion images to produce the multi-scale illustration of activities. However, the biggest limitation for this approach is its absence to preserve the topological structure of the frame, which is mismatched from the original frames because of the contraction procedure [21].

In this work, we have designed a dynamic and robust feature selection algorithm for a HAR system, through which the system accurately recognizes various kinds of activities. In the proposed approach, we employed mutual information algorithm, which selects the prominent features from the extracted features. The proposed algorithm is the expansion of two methods like max-relevance and min-redundancy, respectively. This method has the capability to gather the assets of various extraction algorithms. But the procedure of selection may be unfair due to the dissimilarity between the classification power and redundancy of the features. To resolve this type of unfair selection, we stabilize both parts through the proposed algorithm that has

autonomous upper limit of the mutual information function. Similarly, we thoroughly estimated the  $\beta$  value by taking the enlargement and diminishing issue, which means that the proposed algorithm utilized the filtering methodology in order to take advantage of low computational cost. Therefore, the proposed feature selection algorithm overcomes the limitations of the existing designed approach; hence, it delivers much better performance compared to the existing works, which is suitable for inclusive healthcare domains. We justified our designed approach on realistic depth-based activity database that was recorded by Kinect depth camera in a static and controlled situation. Based on comprehensive experiments, the proposed feature selection algorithm achieved significant accuracy compared with existing works.

The entire paper is organized as follows. In Section 2, we briefly prescribe the latest related works with their limitations. In Section 3, we comprehensively explain the concept of the proposed feature selection method. In Section 4, we provide the procedures for the proposed HAR system. In Section 5, we present a series of multiple experiments followed by some discussion. Finally, in Section 6, we summarize the paper followed by some future directions.

## 2. Literature Review

Human activity states the movement of one or many parts of the human body, which might be static or composed of numerous primitive actions accomplished in some successive order. Lots of work have been done for feature extraction; however, most of them have their own limitations. Also, there might be some redundant feature which may decrease the accuracy of a typical HAR system; therefore, feature selection performs a fascinating role in HAR systems and helps them to achieve significant accuracy.

Thus, the author in [22] developed a random projection-based approach for HAR system against depth information. This approach has the ability to classify the human activities along view direction. Furthermore, this approach was not affected from any kind of environmental factors such as illumination. However, the random projection-based conversion matrix is produced without taking the fundamental structure of the original data that commonly leads to comparatively high misrepresentation [23]. Similarly, a new model was developed in [24] for the categorization of human activities. In their approach, they suggested new features from the angles derived from the human body parts in order to extract best features, which relies on weight descriptor considered from activity frames against various poses. Then, these extracted features are utilized as the input for the classifier. However, this approach fails to show the significant performance in depth activities such as clapping and boxing that cannot be differentiated in RGB cameras [25].

A state-of-the-art hierarchical approach was developed in [26] for depth-based sequences that might combine simultaneously the spatial and temporal information at various temporal scales. However, wrapper communicates with classifier that might have a risk of overfitting in the model [27]. Likewise, a wrapper-based feature selection method

was developed in [28] that claimed highest recognition rate; however, it did not consider the global optimum and mainly relied on local optimums [29]. Furthermore, the authors of [30, 31], respectively, used wavelet transform along with hidden conditional random fields for HAR systems. They utilized full covariance matrix in order to tackle the existing limitations. However, this approach was time consuming because of full covariance matrix. On the other hand, a latest study compared various kinds of classifiers against HAR in order to show their respective performances. They claimed that radial basis kernels and support vector classifier achieved the highest accuracies [32]. Practically, this method is simple and strong method but very difficult to train [33]; also, SVM neglects the temporal information among the activity frames, and hence every frame is projected to be statistically autonomous from the remaining [34].

Another recent approach was proposed in [35], where the human silhouettes were extracted from various activity frames, and then from them a specific sequence is generated using motion detection and tracking techniques. They utilized histogram of the oriented gradient in order to recognize the human activities. However, histogram of oriented gradient is very sensitive to the frame variation; hence, this method is not a better option for activity recognition [36]. Similarly, a two-stage approach was suggested in [37] for activity recognition. In the first stage, deep neural network was used due to which a gesture detector is executed to find the correct location of the pixels of significant key points of the body, while, in another stage, the architecture of a neural search algorithm is utilized to determine an optimum network structure which further utilized in the model for the evolution of spatio-temporal against the corresponding gestures. However, deep neural network has a common disadvantage, and it is very much expensive to train because of the complex data models. Moreover, the evolution of the spatiotemporal approach may degrade the results of the system due to considering the short length time sequences that make the statistical performance of correlation doubtful, and this is one of the major limitations for this approach [19].

A novel convolutional neural network-based HAR system was designed in [38], which employed two types of input descriptors such as depth motion image (which gathers successive depth frames) and the moving joint descriptor that indicates the movement of joints against time. However, depth motion images are less accurate because of real-time latency. Moreover, the major limitation of moving joint descriptor is obstinacy and lesser accuracy [39]. Similarly, a novel approach was designed in [40] that extracts high level movement features by scale invariant feature transform and deep features to classify the activities [40]. However, there were some errors in the corresponding feature points attained by this method, and because of this, the accuracy was gradually decreased [41].

Therefore, we have designed a dynamic and robust feature selection algorithm for a HAR system, through which the system accurately recognizes various kind of activities. In the proposed approach, we employed mutual

information algorithm, which selects the prominent features from the extracted features. The proposed algorithm is the expansion of two methods like max-relevance and min-redundancy, respectively. This method has the capability to gather the assets of various extraction algorithms.

### 3. Proposed Feature Selection Approach

The overall procedure of the proposed technique is presented in Figure 1.

The proposed technique is based on normalized mutual information, which is inherited from the maximum-relevance and minimum-redundancy approach. The proposed technique has one of the limitations to diminish the domination of redundancy and significance. Therefore, we expand this method by designing a scheme in order to integrate the functionality to propose mutual information utilized in this technique. For this purpose, we consider the random variable for the upper limit of mutual information by defining a bound. Assume that  $(i, j)$  is the pair of random variables that is considered as an input with their peripheral and combined distributions, which is calculated as follows:

$$M(I, J) \leq \min(E(I, J)), \quad (1)$$

where  $M$  represents the collected mutual information of  $I$  and  $J$ , while  $E$  indicates the function of entropy which is given below:

$$E(I) \leq \log_2 \left( \sum_{i \in \Omega I} p(i) \frac{1}{p(i)} \right), \quad (2)$$

$$E(I) \leq \log_2(|\Omega I|).$$

Substitute (1) and (2):

$$M(I, J) \leq \min(\log_2(|\Omega I|), \log_2(|\Omega J|)). \quad (3)$$

The proposed approach attains the predicted error of quantization  $\zeta$  through computing every feature under the similar number of levels ( $N$ ), which is described by Algorithm 1.

It may be perceived that the number of the quantization levels slowly increases till the error of quantization reaches lesser than the predefined threshold ( $\zeta$ ). Based on multiple experiments, the threshold value is set to 0.08 in our assessment. During evaluation, a value lower than 0.08 leans towards the increment in computational cost due to which we might not improve the performance. Hence, we set the minimum value of threshold to 0.08, which helps the algorithm to keep improvement without extra computational cost. For each feature  $I$  where  $|\Omega I| = Z$ ,

$$M(I, J) \leq \log_2(Z), \quad (4)$$

where  $\log_2(Z)$  is the upper limit of the independent feature which does not depend upon  $I$  and  $J$ . Likewise,  $\log_2(Z)$  is employed to diminish the concern of unacceptable normalizing weights. Hence, we estimate the standardize head-to-head mutual information,  $ZM(I, J)$ , as

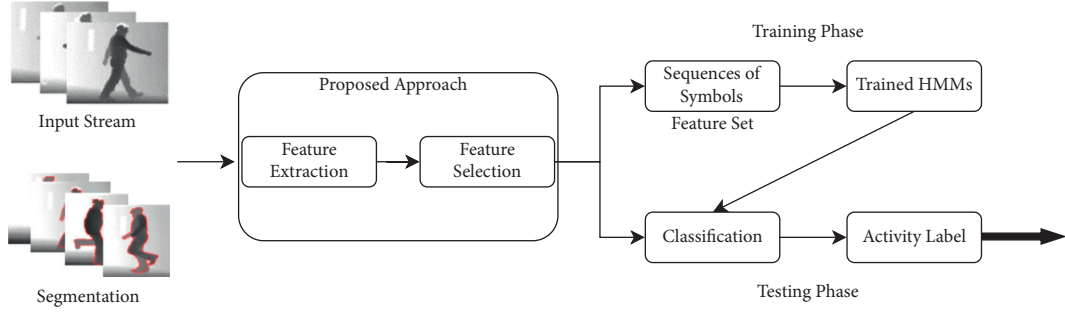


FIGURE 1: Flow diagram for the proposed approach.

```

Input:  $Y, I(1...K)$ , and  $\zeta$ 
Output:  $Z$ , and  $J(1...K)$ 
begin
  Let  $Z = 2$ 
  while 1 do
     $Max^{Er} = -1e + 8$ 
    for  $x = 1$  to  $K$  do
       $lower = \min(I(x))$ 
       $upper = \max(I(x))$ 
       $Step = (upper - lower) / Z$ 
       $Fac = [lower:Step:upper]$ 
       $BC = [lower - Step, lower:Step:upper]$ 
       $Quantiz(I(x), Q^{Er}) = Quantiz(I(x), Divider, BC)$ 
      if  $Q^{Er} > Max^{Er}$  then
         $Max^{Er} = Q^{Er}$ 
      end
    end
    if  $Max^{Er} < \zeta$  then
      Break
    end
     $Z = K + 1$ 
  end
end

```

ALGORITHM 1: Procedures of quantization.

$$ZM(I, J) = \frac{M(I, J)}{\log_2(Z)}, \quad (5)$$

where  $ZM(I, J)$  in (5) lies within the range of  $[0, 1]$ . Therefore, the class feature of mutual information (FCMI) is divided by  $\log_2|\Omega_{C_L}|$  in order to attain a balance among the redundancy and the relevance, which is described as

$$ZM(C_L, I) = \frac{M(C_L, I)}{\log_2|\Omega_{C_L}|}. \quad (6)$$

The potential feature is calculated through combining (5) and (6), as given below:

$$f_p(I_x) = ZM(C_L, I) \frac{1}{|S_{x-1}|} \sum_{x \in S_{x-1}} ZM(I_S, I_x), \quad (7)$$

where  $S = I_1, I_2, \dots, I_x$  is the feature set. In this work, we also integrated the normalized FCMI along with  $ZM(I, J)$  in order to evaluate the consequences of the inequality between the relevance and redundancy. Therefore, the best features are calculated as

$$f_p(I_x) = ZM(C_L, I) \frac{1}{|S_{x-1}|} \sum_{x \in S_{x-1}} \frac{M(I_S, I_x)}{\min(G(I_S), H(I_x))}. \quad (8)$$

The entire selection procedure depends upon the greedy forwarding approach, which is described in Algorithm 2.

As described before, we utilized symlet wavelet transform and hidden Markov model (HMM) for feature extraction and recognition, respectively, from [30].

```

Input:  $Z^s, K, I^{xy}, C^y$ ,
Output:  $S^k$ 
begin
   $S = \phi$ 
  for  $k = 1$  to  $K$  do
     $\sigma = \text{St. Dev. of } I^k$ 
     $\mu^k = \text{Average value of } I^k$ 
     $I^k = I^k - \mu^k$ 
     $I^n = I^k / \sigma^k$ 
  end
   $\bar{I} = \text{Quantiz}(I)$ 
  for  $i = 1$  to  $K$  do
    for  $j = 1$  to  $Z$  do
      Calculate  $f^p(\bar{I}_j)$ 
    end
     $s = \text{argmax } i \notin s_{fp}(\bar{I}_i)$ 
     $S = S \cup s$ 
  end
end

```

ALGORITHM 2: Mutual information-based feature selection along with greedy forwarding algorithm.

## 4. System Justification

The developed feature selection method was verified and validated against depth-based database. The entire experiments were performed in MATLAB in controlled environment (under laboratory settings) with fixed camera settings. Every subject performed the corresponding activities based on the training provided by the instructor. The whole experiments are performed in the following settings.

**4.1. Depth-Based Dataset.** In this depth-based database, there were totally 670 video sequences that were collected using Kinect depth camera. The entire database was created with the help of 70 subjects (male and female) who were the students of the university and their age range was 25 to 60. They performed thirteen various activities such as bending, jacking, place-jumping, running, side movement, skipping, walking, one hand waving, two hand waving, jumping, clapping, boxing, and sitting and standing. The subjects performed the activities under the instructions provided by the corresponding instructor. In this database, some activities are recorded in hospitals from the stroke patients. Entire frames were in different sizes, so we have normalized them by reducing their sizes to  $100 \times 100$ .

### 4.2. Procedure of Experiments

- (i) In the first experiment, the proposed feature selection method was tested against the depth-based database to check its significance. Based on multiple experiments, we selected and employed the 5-fold cross validation scheme for this experiment, which means that out of five activities, four activities were utilized for training and the remaining one is used for testing. This procedure will be repeated five

times due to which every activity will be used once for training and testing.

- (ii) In the second experiment, we employed existing different types of feature selection methods in the HAR system; however, we did not utilize the proposed feature selection method. The reason for doing this type of comprehensive experiments is to show the importance of the proposed feature selection method.
- (iii) Finally, in the third experiment, we compared the significance of the proposed HAR system under the presence of the proposed feature selection method against latest activity recognition approaches.

## 5. Results and Discussion

**5.1. First Experiment.** As described before, in this experiment, we represented the significance of the proposed feature selection method on depth-based database. The whole results are shown in Table 1.

The proposed HAR system under the presence of the proposed feature selection method showed best recognition rate on depth-based database as illustrated in Table 1. The reason is that in the proposed feature selection technique, the potential of the features is calculated through information measurement. Likewise, the greedy forward selection method was utilized, in which every feature is added to the feature set depending upon its value which approves that the proposed technique is more vigorous than existing works regarding recognition rate. In conclusion, the proposed feature selection algorithm overcomes the limitations of the existing designed approaches; hence, it delivers much better performance compared to the existing works.

**5.2. Second Experiment.** As described before, in this experiment, we represented the importance of the proposed

TABLE 1: Accuracy of the designed approach using depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>99</b>	0	0	0	1	0	0	0	0	0	0	0	0
JK	0	<b>100</b>	0	0	0	0	0	0	0	0	0	0	0
PJU	0	0	<b>98</b>	0	0	1	0	0	0	1	0	0	0
RUN	0	0	0	<b>99</b>	0	0	0	0	1	0	0	0	0
SMO	0	0	1	0	<b>97</b>	0	1	0	0	0	1	0	0
SKI	0	0	0	0	0	<b>99</b>	0	0	0	1	0	0	0
WAK	0	0	1	0	2	0	<b>96</b>	0	1	0	0	1	0
1HW	0	0	0	0	0	0	0	<b>100</b>	0	0	0	0	0
2HW	0	1	0	0	0	0	0	0	<b>99</b>	0	0	0	0
JU	1	0	0	0	1	0	1	0	0	<b>97</b>	0	0	0
CL	0	0	1	1	0	0	0	0	0	0	<b>98</b>	0	0
BOX	0	0	0	0	1	1	0	0	1	0	0	<b>97</b>	0
SS	0	0	0	0	0	0	2	0	0	0	0	0	<b>98</b>
Average							<b>98.2%</b>						

BD represents bending, JK represents jacking, PJU represents place-jumping, RUN represents running, SDM represents side movement, SKI represents skipping, WAK represents walking, 1HW represents one hand waving, 2HW represents two hand waving, JU represents jumping, CL represents clapping, BOX represents boxing, and SS represents sitting and standing.

TABLE 2: Accuracy of the designed approach with the wrapper method on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>80</b>	0	2	1	0	2	0	2	0	1	0	2	0
JK	2	<b>77</b>	1	2	2	2	4	1	2	2	2	1	2
PJU	1	2	<b>81</b>	2	1	1	2	1	2	2	1	2	2
RUN	2	2	1	<b>78</b>	4	2	1	2	1	1	1	4	1
SMO	1	1	3	2	<b>76</b>	3	2	2	4	2	1	1	2
SKI	0	2	2	1	2	<b>83</b>	1	2	1	2	1	2	1
WAK	2	1	0	2	1	2	<b>85</b>	1	2	0	2	1	1
1HW	2	2	1	2	1	2	3	<b>79</b>	1	2	1	2	2
2HW	1	2	2	1	2	1	1	2	<b>82</b>	1	2	1	2
JU	2	1	1	2	1	2	2	1	2	<b>83</b>	1	2	0
CL	1	2	2	1	2	1	1	0	2	2	<b>85</b>	0	1
BOX	2	2	3	1	2	2	1	2	4	1	2	<b>77</b>	1
SS	1	2	2	4	1	2	3	1	2	2	3	2	<b>75</b>
Average							<b>80.1%</b>						

TABLE 3: Accuracy of the designed approach with pointwise mutual information on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>79</b>	1	2	2	1	2	1	2	2	1	2	3	2
JK	2	<b>82</b>	1	2	1	2	2	1	2	2	2	1	0
PJU	1	4	<b>75</b>	1	3	1	2	2	4	2	1	2	2
RUN	1	2	2	<b>76</b>	2	2	1	4	2	1	2	3	2
SMO	2	1	2	2	<b>80</b>	1	2	2	1	2	2	2	1
SKI	1	2	1	4	2	<b>78</b>	1	2	3	2	1	2	1
WAK	0	2	2	1	2	2	<b>84</b>	1	2	1	2	1	0
1HW	1	1	2	2	1	0	2	<b>85</b>	1	2	1	2	0
2HW	1	2	1	0	2	1	1	2	<b>87</b>	0	1	1	1
JU	2	1	1	4	2	2	2	1	4	<b>74</b>	2	3	2
CL	1	2	2	2	1	2	2	4	1	2	<b>77</b>	2	2
BOX	2	2	1	2	2	2	1	2	2	1	6	<b>75</b>	2
SS	2	1	6	2	1	2	2	1	2	4	1	2	<b>74</b>
Average							<b>78.9%</b>						

TABLE 4: Accuracy of the designed approach with the relief-based method on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>87</b>	0	2	1	1	2	1	2	1	0	2	2	0
JK	0	<b>89</b>	1	2	0	2	1	0	2	1	0	2	0
PJU	2	2	<b>82</b>	1	2	1	2	1	2	2	1	1	1
RUN	1	2	1	<b>85</b>	1	2	1	2	1	1	0	1	2
SMO	2	1	2	2	<b>79</b>	1	4	2	1	2	2	0	2
SKI	2	1	1	2	2	<b>84</b>	1	0	1	2	1	2	1
WAK	1	2	0	1	2	2	<b>86</b>	1	2	0	2	1	0
1HW	0	1	2	0	1	0	2	<b>90</b>	0	2	0	0	2
2HW	1	2	2	1	1	1	2	2	<b>81</b>	2	2	1	2
JU	2	0	1	2	2	1	2	1	1	<b>83</b>	2	2	1
CL	2	2	2	1	2	1	1	2	2	2	<b>79</b>	2	2
BOX	0	1	2	1	2	2	1	2	1	2	2	<b>81</b>	3
SS	2	2	0	2	1	1	2	1	4	1	0	2	<b>82</b>
Average							<b>77.3%</b>						

TABLE 5: Accuracy of the designed approach with greedy forward selection on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>75</b>	2	1	2	4	2	1	6	2	1	2	2	0
JK	2	<b>77</b>	1	2	2	2	4	1	2	2	1	2	2
PJU	1	2	<b>80</b>	1	0	2	2	1	1	2	2	4	2
RUN	2	1	2	<b>78</b>	2	1	1	4	2	1	3	2	1
SMO	0	2	1	2	<b>85</b>	2	2	0	1	2	0	2	1
SKI	4	1	1	2	4	<b>75</b>	1	2	1	2	1	3	3
WAK	1	2	6	1	3	2	<b>73</b>	2	2	1	4	2	1
1HW	2	4	2	2	1	2	3	<b>77</b>	1	2	1	1	2
2HW	1	2	1	2	2	1	2	2	<b>83</b>	1	2	1	0
JU	2	1	1	3	1	2	0	1	2	<b>82</b>	2	2	1
CL	0	2	2	1	2	1	2	2	1	2	<b>84</b>	0	1
BOX	2	2	1	1	6	2	1	4	2	1	0	<b>76</b>	2
SS	1	1	2	4	1	3	2	2	1	2	1	1	<b>79</b>
Average							<b>78.8%</b>						

TABLE 6: Accuracy of the designed approach with greedy backward selection on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>81</b>	2	1	2	3	2	1	2	2	1	0	1	2
JK	1	<b>88</b>	2	1	0	2	1	2	1	0	1	1	0
PJU	0	1	<b>90</b>	2	1	0	2	1	0	1	1	0	1
RUN	2	2	0	<b>84</b>	2	1	1	0	2	1	2	2	1
SMO	1	2	2	1	<b>83</b>	2	1	2	1	2	1	0	2
SKI	2	0	1	2	2	<b>85</b>	1	2	1	0	1	2	1
WAK	0	2	0	1	0	2	<b>89</b>	1	0	2	1	0	2
1HW	1	0	2	0	1	0	0	<b>92</b>	2	1	0	1	0
2HW	1	2	1	2	0	1	2	2	<b>84</b>	1	2	1	1
JU	2	2	1	1	2	2	1	1	4	<b>80</b>	1	2	1
CL	2	1	2	2	1	1	2	2	1	2	<b>81</b>	1	2
BOX	0	2	1	1	2	0	1	2	0	1	2	<b>87</b>	1
SS	1	0	2	0	1	2	0	0	2	0	1	0	<b>91</b>
Average							<b>85.8%</b>						

feature selection method in HAR system. Therefore, we utilized various types of existing feature selection methods, but we did not use the proposed technique. We employed the wrapper method, pointwise mutual information, relief-

based method, greedy forward selection, greedy backward selection, particle swarm optimization, targeted projection pursuit, scatter search, variable neighborhood search, probabilistic distance, entropy, consistency-based feature

TABLE 7: Accuracy of the designed approach with particle swarm optimization on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>82</b>	2	1	1	2	2	2	1	2	2	1	0	2
JK	0	<b>84</b>	1	2	1	2	1	2	2	1	1	2	1
PJU	1	2	<b>90</b>	0	2	0	1	1	0	0	2	0	1
RUN	0	2	0	<b>91</b>	0	1	1	0	2	1	0	2	0
SMO	2	1	2	1	<b>81</b>	2	2	4	1	0	2	0	2
SKI	1	2	1	2	0	<b>86</b>	1	0	1	2	1	2	1
WAK	1	0	2	0	2	1	<b>88</b>	2	0	2	0	1	1
1HW	0	2	1	1	1	2	0	<b>87</b>	1	1	2	0	2
2HW	1	2	2	1	2	1	4	2	<b>79</b>	2	1	1	2
JU	2	0	1	2	2	1	2	1	6	<b>78</b>	2	2	1
CL	1	2	2	1	0	1	2	2	1	2	<b>83</b>	1	2
BOX	1	2	0	1	2	2	1	1	2	0	2	<b>85</b>	1
SS	2	0	1	2	1	0	2	0	1	2	0	2	<b>87</b>
Average							<b>84.7%</b>						

TABLE 8: Accuracy of the designed approach with targeted projection pursuit on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>80</b>	1	2	1	2	1	2	1	2	2	3	1	2
JK	2	<b>77</b>	1	2	1	2	4	2	3	2	1	2	1
PJU	2	1	<b>79</b>	2	4	1	2	1	2	1	2	1	2
RUN	1	2	1	<b>82</b>	2	2	1	2	1	1	2	2	1
SMO	2	1	2	2	<b>76</b>	4	2	1	2	2	4	1	1
SKI	2	2	1	6	2	<b>74</b>	1	2	1	4	1	2	2
WAK	1	2	2	2	1	2	<b>78</b>	4	2	1	2	2	1
1HW	0	1	2	1	1	4	2	<b>83</b>	2	0	1	1	2
2HW	1	2	0	1	2	1	1	2	<b>86</b>	2	0	1	1
JU	1	2	1	0	1	0	2	1	2	<b>88</b>	1	0	1
CL	2	0	2	2	1	1	1	2	1	2	<b>82</b>	2	2
BOX	1	2	2	1	2	2	1	2	2	1	2	<b>78</b>	4
SS	2	2	1	2	1	2	6	1	1	2	4	1	<b>75</b>
Average							<b>79.8%</b>						

TABLE 9: Accuracy of the designed approach with scatter search pursuit on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>77</b>	1	2	2	1	2	4	2	2	1	3	2	1
JK	2	<b>80</b>	1	1	2	2	1	2	4	0	2	1	2
PJU	1	2	<b>79</b>	2	1	4	2	1	2	2	1	2	1
RUN	2	2	1	<b>73</b>	2	2	1	2	2	1	6	1	5
SMO	2	1	2	4	<b>74</b>	1	2	6	1	2	1	2	2
SKI	1	2	1	2	2	<b>78</b>	3	2	1	2	1	4	1
WAK	0	1	2	0	1	0	<b>91</b>	0	2	0	2	0	1
1HW	2	1	0	1	1	2	2	<b>84</b>	1	2	0	2	2
2HW	1	2	2	0	2	1	2	4	<b>82</b>	1	2	1	0
JU	2	1	1	2	1	2	2	1	4	<b>79</b>	1	2	2
CL	1	2	2	1	2	1	1	2	2	1	<b>84</b>	0	1
BOX	2	2	1	1	2	2	1	1	1	2	3	<b>80</b>	2
SS	0	1	2	2	1	2	0	2	0	1	1	2	<b>86</b>
Average							<b>80.5%</b>						

selection, correlation-based feature selection, recursive feature elimination, and genetic algorithm. The entire results for this set of experiments on depth-based database are shown in Tables 2–16, respectively.

As illustrated from these sets of experiments, the proposed feature selection technique has a significant part in the high classification of the HAR system. When the proposed method has been removed for the system, then abruptly the



TABLE 10: Accuracy of the designed approach with variable neighborhood search on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>88</b>	2	0	1	1	2	1	0	2	0	2	0	1
JK	0	<b>89</b>	1	2	1	2	0	1	0	2	0	2	0
PJU	1	0	<b>92</b>	1	0	0	2	1	2	0	1	0	0
RUN	0	2	1	<b>91</b>	0	1	1	0	0	1	0	1	2
SMO	2	0	0	1	<b>93</b>	0	0	1	2	0	1	0	0
SKI	0	2	1	0	2	<b>90</b>	1	0	1	2	1	0	1
WAK	1	0	2	2	0	2	<b>87</b>	2	1	0	2	1	0
1HW	2	1	1	2	1	0	2	<b>85</b>	0	2	0	2	2
2HW	1	0	2	1	0	1	0	1	<b>93</b>	0	1	0	0
JU	0	2	1	0	2	0	1	1	0	<b>91</b>	0	1	1
CL	1	0	0	2	0	1	0	0	1	0	<b>94</b>	1	0
BOX	1	2	0	1	2	0	1	2	0	1	0	<b>89</b>	1
SS	2	0	2	1	1	2	0	1	2	1	2	1	<b>85</b>
Average							<b>89.8%</b>						

TABLE 11: Accuracy of the designed approach with probabilistic distance on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>75</b>	2	2	1	4	2	1	2	2	1	2	4	2
JK	2	<b>80</b>	1	2	2	2	3	0	1	2	1	2	2
PJU	1	2	<b>77</b>	2	1	4	2	3	2	1	2	2	1
RUN	1	2	1	<b>82</b>	2	1	1	2	2	1	2	1	2
SMO	2	1	2	2	<b>80</b>	2	1	2	1	2	1	2	2
SKI	1	2	1	4	2	<b>79</b>	1	2	2	2	1	2	1
WAK	2	1	2	0	1	2	<b>83</b>	1	2	1	2	1	2
1HW	0	2	0	2	1	1	2	<b>87</b>	1	2	0	0	2
2HW	1	2	3	1	2	1	4	2	<b>77</b>	2	2	1	2
JU	2	0	1	2	1	2	0	1	2	<b>84</b>	2	2	1
CL	1	1	2	1	2	0	2	2	1	1	<b>86</b>	1	0
BOX	1	2	2	1	2	5	1	2	2	4	1	<b>75</b>	2
SS	2	2	1	2	1	2	2	0	2	1	2	2	<b>81</b>
Average							<b>80.4%</b>						

TABLE 12: Accuracy of the designed approach with entropy on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>84</b>	1	2	2	0	2	1	2	1	2	2	1	0
JK	0	<b>88</b>	1	0	1	2	2	1	2	0	1	0	2
PJU	1	2	<b>90</b>	1	2	0	0	1	0	1	0	2	0
RUN	1	2	2	<b>83</b>	1	2	1	2	2	1	2	0	1
SMO	2	1	1	2	<b>82</b>	1	2	0	2	2	1	2	2
SKI	2	2	1	2	2	<b>80</b>	1	2	1	2	1	3	1
WAK	1	2	2	1	1	2	<b>82</b>	2	2	1	2	0	2
1HW	1	0	2	2	1	1	2	<b>86</b>	1	0	1	2	1
2HW	1	2	0	1	2	1	0	1	<b>89</b>	2	0	1	0
JU	0	2	1	0	1	0	2	1	0	<b>91</b>	1	0	1
CL	2	1	2	1	2	1	2	2	1	2	<b>81</b>	2	1
BOX	1	2	2	1	1	2	1	0	2	1	2	<b>85</b>	0
SS	1	2	1	2	0	1	2	1	0	2	1	1	<b>86</b>
Average							<b>85.1%</b>						

accuracies are decreased. The results shown in Tables 2–16 confirm the issue of highest resemblance between the features of various activities. These results approved our

examination and offered strong evidence, letting us to accomplish that the proposed algorithm selects the best set of features that improved the recognition rate of HAR systems.

TABLE 13: Accuracy of the designed approach with consistency-based feature selection on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>91</b>	1	0	2	0	2	0	1	0	2	0	1	0
JK	2	<b>88</b>	1	0	1	2	1	0	2	0	1	0	2
PJU	0	1	<b>89</b>	2	1	0	2	1	1	2	0	1	0
RUN	1	2	0	<b>90</b>	0	1	1	0	1	1	2	0	1
SMO	2	0	2	1	<b>84</b>	2	1	2	2	1	0	1	2
SKI	1	2	1	0	2	<b>88</b>	1	0	1	0	1	2	1
WAK	2	1	2	2	1	2	<b>83</b>	1	0	2	1	1	2
1HW	1	0	1	2	1	0	0	<b>91</b>	2	0	2	0	0
2HW	1	2	0	1	0	1	2	0	<b>89</b>	2	0	1	1
JU	2	1	1	0	2	2	0	1	2	<b>86</b>	1	0	2
CL	0	2	0	2	1	0	2	2	1	2	<b>85</b>	2	1
BOX	1	2	1	1	0	2	1	0	0	1	1	<b>90</b>	0
SS	2	0	2	1	1	0	2	1	2	0	0	2	<b>87</b>
Average							<b>87.8%</b>						

TABLE 14: Accuracy of the designed approach with correlation-based feature selection on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>80</b>	2	2	1	1	2	2	1	2	2	1	2	2
JK	2	<b>85</b>	1	2	0	2	1	2	1	0	2	1	1
PJU	1	0	<b>89</b>	1	2	1	2	1	0	2	0	1	0
RUN	1	2	0	<b>87</b>	1	2	1	0	2	1	2	0	1
SMO	0	1	2	0	<b>91</b>	1	0	2	0	0	1	2	0
SKI	2	1	1	2	2	<b>82</b>	1	2	1	2	1	2	1
WAK	1	2	2	1	1	2	<b>83</b>	1	2	1	2	1	1
1HW	2	0	2	1	1	2	1	<b>85</b>	2	2	0	1	1
2HW	1	2	0	2	0	1	0	2	<b>88</b>	0	1	1	2
JU	0	1	1	0	2	0	2	1	2	<b>89</b>	2	0	0
CL	1	0	0	2	0	1	0	0	0	2	<b>92</b>	1	1
BOX	1	2	0	1	1	0	1	2	0	0	1	<b>91</b>	0
SS	2	1	2	2	1	2	2	1	2	1	2	1	<b>81</b>
Average							<b>86.4%</b>						

TABLE 15: Accuracy of the designed approach with recursive feature elimination on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>90</b>	2	0	1	0	2	0	1	0	2	0	1	1
JK	1	<b>89</b>	1	0	1	2	1	0	2	0	2	1	0
PJU	0	1	<b>91</b>	2	1	0	0	1	0	2	0	0	2
RUN	1	2	0	<b>88</b>	0	1	1	0	2	1	2	1	1
SMO	2	0	1	1	<b>86</b>	0	2	2	1	2	0	2	1
SKI	1	0	1	0	2	<b>92</b>	1	0	1	0	1	0	1
WAK	0	2	0	2	1	2	<b>86</b>	1	0	2	0	2	2
1HW	2	0	1	0	1	1	1	<b>89</b>	2	0	2	1	0
2HW	1	2	0	1	0	1	0	1	<b>91</b>	0	0	1	2
JU	1	0	1	0	2	0	2	1	0	<b>90</b>	2	0	1
CL	2	1	0	2	0	1	0	0	2	1	<b>87</b>	2	2
BOX	2	2	1	1	2	1	1	2	1	0	1	<b>85</b>	1
SS	1	0	2	2	1	0	1	1	0	2	0	1	<b>89</b>
Average							<b>88.7%</b>						

5.3. *Third Experiment.* Lastly, in this experiment, the weighted average classification accuracy of the proposed work has been compared with existing latest studies. Most of the existing systems are implemented using the same settings

as explained in their corresponding articles. The overall results for this experiment are presented in Table 17.

It is indicated in Table 17 that the proposed algorithm accomplished highest accuracy compared with latest existing

TABLE 16: Accuracy of the designed approach with genetic algorithm on depth database.

Activities	BD	JK	PJU	RUN	SMO	SKI	WAK	1HW	2HW	JU	CL	BOX	SS
BD	<b>92</b>	2	0	1	0	2	0	0	1	0	0	2	0
JK	1	<b>89</b>	1	0	2	2	1	1	0	0	2	0	1
PJU	2	0	<b>90</b>	2	0	1	0	1	2	0	0	1	1
RUN	1	2	0	<b>91</b>	1	0	1	0	1	1	0	0	2
SMO	2	1	0	0	<b>93</b>	0	0	1	0	1	0	0	2
SKI	0	0	1	2	0	<b>92</b>	1	0	1	0	1	2	0
WAK	1	2	0	0	2	0	<b>89</b>	2	0	2	0	1	1
1HW	2	0	1	2	1	1	2	<b>87</b>	1	0	2	0	1
2HW	1	1	2	0	1	1	0	2	<b>88</b>	1	0	1	2
JU	2	0	1	1	0	0	1	1	0	<b>91</b>	2	0	1
CL	2	1	1	0	2	1	0	1	2	0	<b>86</b>	2	2
BOX	1	2	0	1	1	2	1	2	1	2	0	<b>87</b>	0
SS	1	0	2	0	1	0	0	1	1	0	2	2	<b>90</b>
Average							<b>89.6%</b>						

TABLE 17: Comparison of the proposed algorithm with latest HAR approaches.

Existing works	[42](%)	[43] (%)	[44] (%)	[45] (%)	[46] (%)	[47] (%)	[48] (%)	[49] (%)	Proposed approach (%)
Accuracies	79.4	86.1	79.4	83.8	90.0	92.5	75.9	88.4	<b>98.2</b>

approaches. This is because the proposed algorithm selects the informative features from the activity frames in the presence of various environmental factors that may cause misclassification occlusion.

## 6. Conclusions

The activity recognition is an important and significant concern in the field of healthcare and telemedicine. The common HAR system involves three basic modules: segmentation, feature extraction, and recognition. In this research, we have studied some latest mutual information feature selection techniques. We described the shortcomings of every technique, and hence we designed an algorithm that is inherited from the standardized mutual information-based feature selection along with two enhancements: the regularization of the mutual information and the independent feature regularizing weights. We used a mutual information algorithm to choose the most prominent features from the extracted features in the suggested method. The suggested approach is a combination of two methods: max-relevance and min-redundancy. This approach can collect the assets from a variety of extraction algorithms. However, due to the disparity in classification power and feature redundancy, the selection approach may be unjust. As a noble solution to this problem of biased selection, we stabilize both sections using the suggested approach which has an autonomous upper limit of the mutual information function. Similarly, we comprehensively calculated the  $\beta$  value by considering the enlargement and diminishing issue, implying that the proposed algorithm used filtering methodology to benefit from low computing cost. As a result, the suggested feature selection algorithm overcomes the limits of the existing planned techniques, so it provides significantly better performance than previous studies, making it ideal for inclusive healthcare domains. On a realistic depth-

based activity database captured by the Kinect depth camera in a static and controlled setting, we justified our designed approach. Similarly, for the feature extraction and recognition, we employed the symlet wavelet transform and hidden Markov model, respectively, for action classification. Based on extensive testing, the suggested feature selection algorithm outperformed previous work by a wide margin.

In real domains, there are some parameters like dynamic background, lighting effects, blurring, and spontaneous factors, which may reduce the accuracy of the proposed feature selection algorithm. Therefore, in the future, the validation of the proposed algorithm needs some investigation to keep the same accuracy in real domain under the presence of the aforementioned environmental factors. Moreover, we will employ the proposed technique in healthcare domains in order to resolve the privacy concern.

## Data Availability

The data used for this study and simulation will be provided on demand.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was funded by the Deanship of Scientific Research at Jouf University under grant no. DSR-2021-02-0343.

## References

- [1] L. Wang, Y. Xiong, Z. Wang et al., "Temporal segment networks for action recognition in videos," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 11, pp. 2740-2755, 2018.

- [2] J. Carreira and A. Zisserman, "Quo vadis, action recognition? a new model and the kinetics dataset," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6299–6308, Honolulu, HI, USA, July 2017.
- [3] D. Tran, H. Wang, L. Torresani, J. Ray, Y. LeCun, and M. Paluri, "A closer look at spatiotemporal convolutions for action recognition," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 6450–6459, Salt Lake City, UT, USA, June 2018.
- [4] L. Wang, W. Li, W. Li, and L. Van Gool, "Appearance-and-relation networks for video classification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1430–1439, Salt Lake City, UT, USA, June 2018.
- [5] S. Xie, C. Sun, J. Huang, Z. Tu, and K. Murphy, "Rethinking spatiotemporal feature learning: speed-accuracy trade-offs in video classification," in *Proceedings of the European conference on computer vision (ECCV)*, pp. 305–321.
- [6] X. Wang, R. Girshick, A. Gupta, and K. He, "Non-local neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7794–7803, 2018.
- [7] Y.-D. Zheng, Z. Liu, T. Lu, and L. Wang, "Dynamic sampling networks for efficient action recognition in videos," *IEEE Transactions on Image Processing*, vol. 29, pp. 7970–7983, 2020.
- [8] V. Osmani, S. Balasubramaniam, and D. Botvich, "Human activity recognition in pervasive health-care: supporting efficient remote collaboration," *Journal of Network and Computer Applications*, vol. 31, no. 4, pp. 628–655, 2008.
- [9] J. E. Dunn, M. A. Rudberg, S. E. Furner, and C. K. Cassel, "Mortality, disability, and falls in older persons: the role of underlying disease and disability," *American Journal of Public Health*, vol. 82, no. 3, pp. 395–400, 1992.
- [10] U. E. Reinhardt, "Does the aging of the population really drive the demand for health care?" *Health Affairs*, vol. 22, no. 6, pp. 27–39, 2003.
- [11] C. Bin, "Shi Yan. Research and development of somatosensory virtual mouse based on Kinect," *Journal of Software*, vol. 37, no. 2, pp. 46–49, 2016.
- [12] J. An, X. Cheng, Q. Wang, H. Chen, J. Li, and S. Li, "Human action recognition based on Kinect," in *Journal of Physics: Conference Series*, vol. 1693, no. 1, IOP Publishing, Article ID 012190, 2020.
- [13] Y. J. Mao, *Research on Continuous Action Recognition of Human Body Based on Kinect*, University of Electronic Science and Technology of China, Chengdu, China, 2017.
- [14] C. Plizzari, M. Cannici, and M. Matteucci, "Skeleton-based action recognition via spatial and temporal transformer networks," *Computer Vision and Image Understanding*, vol. 208, Article ID 103219, 2021.
- [15] Z. Huang, X. Shen, X. Tian, H. Li, J. Huang, and X. S. Hua, "Spatio-temporal inception graph convolutional networks for skeleton-based action recognition," in *Proceedings of the 28th ACM International Conference on Multimedia*, pp. 2122–2130, 2020, October.
- [16] P. Khatiwada, M. Subedi, A. Chatterjee, and M. W. Gerdes, "Automated human activity recognition by colliding bodies optimization-based optimal feature selection with recurrent neural network," 2020, <https://arxiv.org/abs/2010.03324>.
- [17] M. Sheikhi Azqandi, "A novel hybrid genetic modified colliding bodies optimization for designing of composite laminates," *Mechanics of Advanced Composite Structures*, vol. 8, no. 1, pp. 203–212, 2021.
- [18] L. Yuan, Y. Zhou, S. Chang et al., "Toward accurate person-level action recognition in videos of crowded scenes," in *Proceedings of the 28th ACM International Conference on Multimedia*, pp. 4694–4698, Seattle WA USA, October 2020.
- [19] L. N. Ferreira, D. A. Vega-Oliveros, M. Cotacallapa et al., "Spatiotemporal data analysis with chronological networks," *Nature Communications*, vol. 11, no. 1, pp. 1–11, 2020.
- [20] C. Li, Q. Huang, X. Li, and Q. Wu, "Human action recognition based on multi-scale feature maps from depth video sequences," 2021, <https://arxiv.org/abs/2101.07618>.
- [21] D. Shao, L. A. Mateos, and W. G. Kropatsch, *Irregular Laplacian Graph Pyramid*, Computer Vision Winter Workshop 2010, Nove Hradý, 2010.
- [22] G. Zhao, "Feature recognition of human motion behavior based on depth sequence analysis," *Complexity*, vol. 2021, 2021.
- [23] H. Xie, J. Li, and H. Xue, "A survey of dimensionality reduction techniques based on random projection," 2017, <https://arxiv.org/abs/1706.04371>.
- [24] G. Wu, Z. Liu, and X. Chen, "Weighted classification of machine learning to recognize human activities," *Complexity*, vol. 2021, 2021.
- [25] M. H. Siddiqi, N. Almashfi, A. Ali et al., "A unified approach for patient Activity recognition in healthcare using depth camera," *IEEE Access*, vol. 9, pp. 92300–92317, 2021.
- [26] H. Wu, X. Ma, and Y. Li, "Hierarchical dynamic depth projected difference images-based action recognition in videos with convolutional neural networks," *International Journal of Advanced Robotic Systems*, vol. 16, no. 1, Article ID 1729881418825093, 2019.
- [27] Y. B. Wah, N. Ibrahim, H. A. Hamid, S. Abdul-Rahman, and S. Fong, "Feature selection methods: case of filter and wrapper approaches for maximising classification accuracy," *Pertanika Journal of Science & Technology*, vol. 26, no. 1, 2018.
- [28] J. Chong, P. Tjurin, M. Niemelä, T. Jämsä, and V. Farrahi, "Machine-learning models for activity class prediction: a comparative study of feature selection and classification algorithms," *Gait & Posture*, vol. 89, pp. 45–53, 2021.
- [29] M. H. Siddiqi, M. Alruwaili, and A. Ali, "A novel feature selection method for video-based human activity recognition systems," *IEEE Access*, vol. 7, pp. 119593–119602, 2019.
- [30] M. Siddiqi, R. Ali, M. Rana, E.-K. Hong, E. Kim, and S. Lee, "Video-based human activity recognition using multilevel wavelet decomposition and stepwise linear discriminant analysis," *Sensors*, vol. 14, no. 4, pp. 6370–6392, 2014.
- [31] M. H. Siddiqi, M. Alruwaili, A. Ali, S. Alanazi, and F. Zeshan, "Human activity recognition using Gaussian mixture hidden conditional random fields," *Computational Intelligence and Neuroscience*, vol. 2019, p. 8590560, 2019.
- [32] I. A. Bustoni, I. Hidayatulloh, A. M. Ningtyas, A. Purwaningsih, and S. N. Azhari, "Classification methods performance on human activity recognition," *Journal of Physics: conference Series*, vol. 1456, no. 1, Article ID 012027, 2020.
- [33] A.-N. Sharkawy, "Principle of neural network and its main types: review," *Journal of Advances in Applied & Computational Mathematics*, vol. 7, no. 1, pp. 8–19, 2020.
- [34] M. H. Siddiqi, R. Ali, A. M. Khan, Y. T. Young-Tack Park, and S. Sungyoung Lee, "Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields," *IEEE Transactions on Image Processing*, vol. 24, no. 4, pp. 1386–1398, 2015.
- [35] O. Elharrouss, N. Almaadeed, S. Al-Maadeed, A. Bouridane, and A. Beghdadi, "A combined multiple action recognition and summarization for surveillance video sequences," *Applied Intelligence*, vol. 51, no. 2, pp. 690–712, 2021.

- [36] M.-K. Cheon, W.-J. Lee, C.-H. Hyun, and M. Park, "Rotation invariant histogram of oriented gradients," *International Journal of Fuzzy Logic and Intelligent Systems*, vol. 11, no. 4, pp. 293–298, 2011.
- [37] H. H. Pham, H. Salmane, L. Khoudour, A. Crouzil, S. A. Velastin, and P. Zegers, "A unified deep framework for joint 3d pose estimation and action recognition from a single rgb camera," *Sensors*, vol. 20, no. 7, p. 1825, 2020.
- [38] A. Kamel, B. Sheng, P. Yang, P. Li, R. Shen, and D. D. Feng, "Deep convolutional neural networks for human action recognition using depth maps and postures," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 9, pp. 1806–1819, 2018.
- [39] A. I. Faisal, S. Majumder, T. Mondal, D. Cowan, S. Naseh, and M. J. Deen, "Monitoring methods of human body joints: state-of-the-art and research challenges," *Sensors*, vol. 19, no. 11, p. 2629, 2019.
- [40] S. C. B. Dash, S. R. Mishra, K. Srujan Raju, and L. V. Narasimha Prasad, "Human action recognition using a hybrid deep learning heuristic," *Soft Computing*, vol. 25, no. 20, pp. 13079–13092, 2021.
- [41] T. Zhang, R. Zhao, and Z. Chen, "Application of migration image registration algorithm based on improved SURF in remote sensing image mosaic," *IEEE Access*, vol. 8, pp. 163637–163645, 2020.
- [42] T. T. Zin, Y. Htet, Y. Akagi et al., "Real-time action recognition system for elderly people using stereo depth camera," *Sensors*, vol. 21, no. 17, p. 5895, 2021.
- [43] J. Trelinski and B. Kwolek, "CNN-based and DTW features for human activity recognition on depth maps," *Neural Computing & Applications*, pp. 1–13, 2021.
- [44] N. Tasnim, M. K. Islam, and J.-H. Baek, "Deep learning based human activity recognition using spatio-temporal image formation of skeleton joints," *Applied Sciences*, vol. 11, no. 6, p. 2675, 2021.
- [45] K. Sidor and M. Wysocki, "Recognition of human activities using depth maps and the viewpoint feature histogram descriptor," *Sensors*, vol. 20, no. 10, p. 2940, 2020.
- [46] Z. Ramezanpanah, M. Mallem, and F. Davesne, "Human action recognition using laban movement analysis and dynamic time warping," *Procedia Computer Science*, vol. 176, pp. 390–399, 2020.
- [47] H. Mohammadzade, S. Hosseini, M. R. Rezaei-Dastjerdehei, and M. Tabejamaat, "Dynamic time warping-based features with class-specific joint importance maps for action recognition using Kinect depth sensor," *IEEE Sensors Journal*, vol. 21, no. 7, pp. 9300–9313, 2021.
- [48] H. Wang, Z. Song, W. Li, and P. Wang, "A hybrid network for large-scale action recognition from rgb and depth modalities," *Sensors*, vol. 20, no. 11, p. 3305, 2020.
- [49] M. F. Bulbul, S. Tabussum, H. Ali, W. Zheng, M. Y. Lee, and A. Ullah, "Exploring 3D human action recognition using STACOG on multi-view depth motion maps sequences," *Sensors*, vol. 21, no. 11, p. 3642, 2021.