

Article

A Method of Human Activity Recognition in Transitional Period

Lei Chen ¹, Shurui Fan ^{1,*}, Vikram Kumar ^{2,*} and Yating Jia ¹

¹ Tianjin Key Laboratory of Electronic Materials Devices, School of Electronics and Information Engineering, Hebei University of Technology, Tianjin 300401, China; 201931903018@stu.hebut.edu.cn (L.C.); ya1552715935@126.com (Y.J.)

² Indian Institute of Information Technology, Una 177220, India

* Correspondence: fansr@hebut.edu.cn (S.F.); vikram@iiit.ac.in (V.K.);
Tel.: +86-139-0201-3210 (S.F.); +91-98163-17024 (V.K.)

Received: 10 July 2020; Accepted: 26 August 2020; Published: 28 August 2020



Abstract: Human activity recognition (HAR) has been increasingly used in medical care, behavior analysis, and entertainment industry to improve the experience of users. Most of the existing works use fixed models to identify various activities. However, they do not adapt well to the dynamic nature of human activities. We investigated the activity recognition with postural transition awareness. The inertial sensor data was processed by filters and we used both time domain and frequency domain of the signals to extract the feature set. For the corresponding posture classification, three feature selection algorithms were considered to select 585 features to obtain the optimal feature subset for the posture classification. And We adopted three classifiers (support vector machine, decision tree, and random forest) for comparative analysis. After experiments, the support vector machine gave better classification results than other two methods. By using the support vector machine, we could achieve up to 98% accuracy in the Multi-class classification. Finally, the results were verified by probability estimation.

Keywords: activity recognition; posture transitions; inertial sensor; feature selection; support vector machine

1. Introduction

The human activity and posture transformation recognition is useful to provide users with valuable situational awareness, thus become one of the hotspots in many fields such as medical care, human-computer interaction, film and television production, and motion analysis [1]. The two dominant approaches for human activity classification used in literature are Vision-based systems and Wearable Sensor-based systems. Vision-based systems are widely used to detection of human parts and identification of daily activities [2]. These systems process the collected visual data for activity classification.

Wearable Sensor based systems consist of multiple inertial sensors connected to a human sensor network. After receiving and executing system commands, the raw human body data would be given feedback [3,4]. Inertial measurement (accelerometers and gyroscopes) units are used to measure the triaxle angular velocity and the triaxle acceleration signals generated during human body movement [5]. Sensors available in smartphones, such as temperature sensors and pressure sensors, are useful to know the surroundings [6]. The data collected from the sensors attached to the user and sensors installed in the surroundings are proceed to provide situational awareness to the user [7]. One of the problems of using accelerometer to detect the motion of an object is that it often affected by the gravitational field in the measurement, and its value ($g = 9.81 \text{ m/s}^2$) is relatively high. However, many studies have

found that gravity factors can be separated from body motion by filtering. When using three-axis accelerometer, the induced gravity vector can also help determine the direction of the object relative to the gravity axis [8]. The gyroscope measures the direction indirectly; that is, it first estimates the angular velocity, and integrates the angular velocity to obtain the direction. However, a reference initial angular position is needed to obtain the direction from the gyroscope [9]. Gyroscopes are also prone to noise, resulting in different offsets which can be eliminated by filtering.

At present, many scholars have studied the problem of human behavior recognition based on video data [10]. In [11], the authors proposed depth video-based HAR system to utilize skeleton joints features indoors. They used processed depth maps to track human silhouettes and produce body joints information in the of skeleton, then the hidden Markov model was trained by features calculated from the joint information. The trained model was adopted to recognize various human activities with a mean rate of 84.33% for nine daily routine activities of the elderly. Basbiker M, etc. [12] developed an intelligent human recognition system. In multiple stages of the system, a series of digital image processing technologies were used to extract the human activity feature data from the frame sequence, and a robust neural networks was established to classify the activity models by using a multi-layer feedforward perceptron network. However, the vision-based HAR is limited by spatial location, and video data is relatively complex. It is easier to cause privacy leakage. In contrast, data based on inertial measurement unit can avoid these problems very well, thus it is becoming a new trend of HAR.

The human activity recognition system has three types of feature extraction methods: temporal features, frequency features, and a combination of the two [13]. The authors of [14] put forward an algorithm named S-ELM-KRSL, which is more suitable for processing large-scale data with noises or outliers to identify the motion sequence of body. After experiment, the scheme could detect symptoms of mild cognitive impairment and dementia with satisfactory accuracy. In [15], Zhu, etc. proposed a semi-supervised deep learning approach using temporal Ensembling of deep long short-term memory to extract high-level features for human activity recognition. They investigated temporal Ensembling with some randomness to enhance the generalization of the neural networks. Besides the use of ensemble approach based on both labeled and unlabeled data, they also combined the supervised and unsupervised losses and demonstrated the effectiveness of the semi-supervised learning scheme in experimental results. The authors of [16] brought up a novel ensemble extreme learning machine (ELM) algorithm, in which Gaussian random projection is employed to initialize the input weights of base ELMs and more diversities had been generated to boost the performance of ensemble learning. The algorithm demonstrated recognition accuracies of 97.35% and 98.88% on two datasets. However, the training time of the algorithm is slightly longer. In [17], a feature selection algorithm based on fast correlation filtering was developed to achieve data preprocessing and demonstrated that the classification accuracy can reach up to 100%. However, the classification model only used the AIRS2 algorithm which may not be suitable for other classifier. Feature selection is based on well-defined evaluation criteria to select the original feature set, which eliminates small correlations and unnecessary features. The selected features don't change the original representation of the feature set, and feature selection helps online classification to be more flexible [18].

Most human behavior recognition systems developed in the past ignored posture transitions because the incidence of posture transitions is lower and the duration is shorter than other basic physical activities [19]. However, the above assumptions depend on different applications and are not applicable when multiple activities must be performed in a short period of time. On the other hand, in many practical scenarios, such as fitness or disability monitoring systems, determining posture transitions is critical because in these cases the user performs multiple tasks in a short period of time [20]. In fact, in the case of human behavior recognition system and transient posture perception, the classification will change slightly, and the absence of specified posture transformation may lead to poor system performance [21].

A posture transition is a finite duration event determined by its start and end times. In general, the time required for posture transitions between different individuals is different. The posture transition is limited by the other two activities and represents the transition period between the two activities [22]. Basic activities like standing and walking can be extended for a longer period of time than posture transitions. The data collection of the two types of activities is also different. The posture transformation needs to be repeated to obtain a separate sample. Since the basic activities are continuous, multiple window samples can be obtained from a single test according to the limitation of its time range [23].

The other works related to this paper are referred in [24,25]. We have researched a large number of features on HAR assisted by an inertial measurement unit in the past. The various activity features are classified hierarchical, and six basic activities can be identified with an average accuracy of 96.4%. However, the transition period of activities was out of account.

This paper focuses on Human Activity Recognition with postural transition awareness. In this paper, the motion of the human body was sensed by an accelerometer and a gyroscope of the inertial measurement unit. The magnitude and direction of the acceleration can be measured by vertically arranging the sensors in three-dimensional space. It can also be built on a single chip, and it is now common to use three-axis accelerometers in some commercial electronic devices [26]. First, we analyzed the six-axis signal data acquired by the inertial measurement unit, and then preprocessed to obtain a variety of signals that can represent the action. The various signals obtained from the preprocessing were extracted in the time domain and the frequency domain using various standard and original measurement methods to characterize each active sample. Thereafter, we perform feature selection according to the specific classification condition by using various feature selection algorithms. A variety of machine learning methods are used to classify and selected the one with the highest classification accuracy. Finally, we use support vector machine to classify the posture. Different kernel functions and specific parameters are used to optimize the model.

Figure 1 shows the framework followed in this paper for Activity Recognition. The framework consists of four modules: Data preprocessing, Feature Extraction and Selection, Classifier Selection, and Classifier Evaluation. The details of each module are given in next sections. In Section 2, we described the Data preprocessing, Feature extraction, and Data selection. Section 3 is focused on the Classifier Selection. In Section 4, we discussed Classifier Selection and Results. We concluded the paper in Section 5.

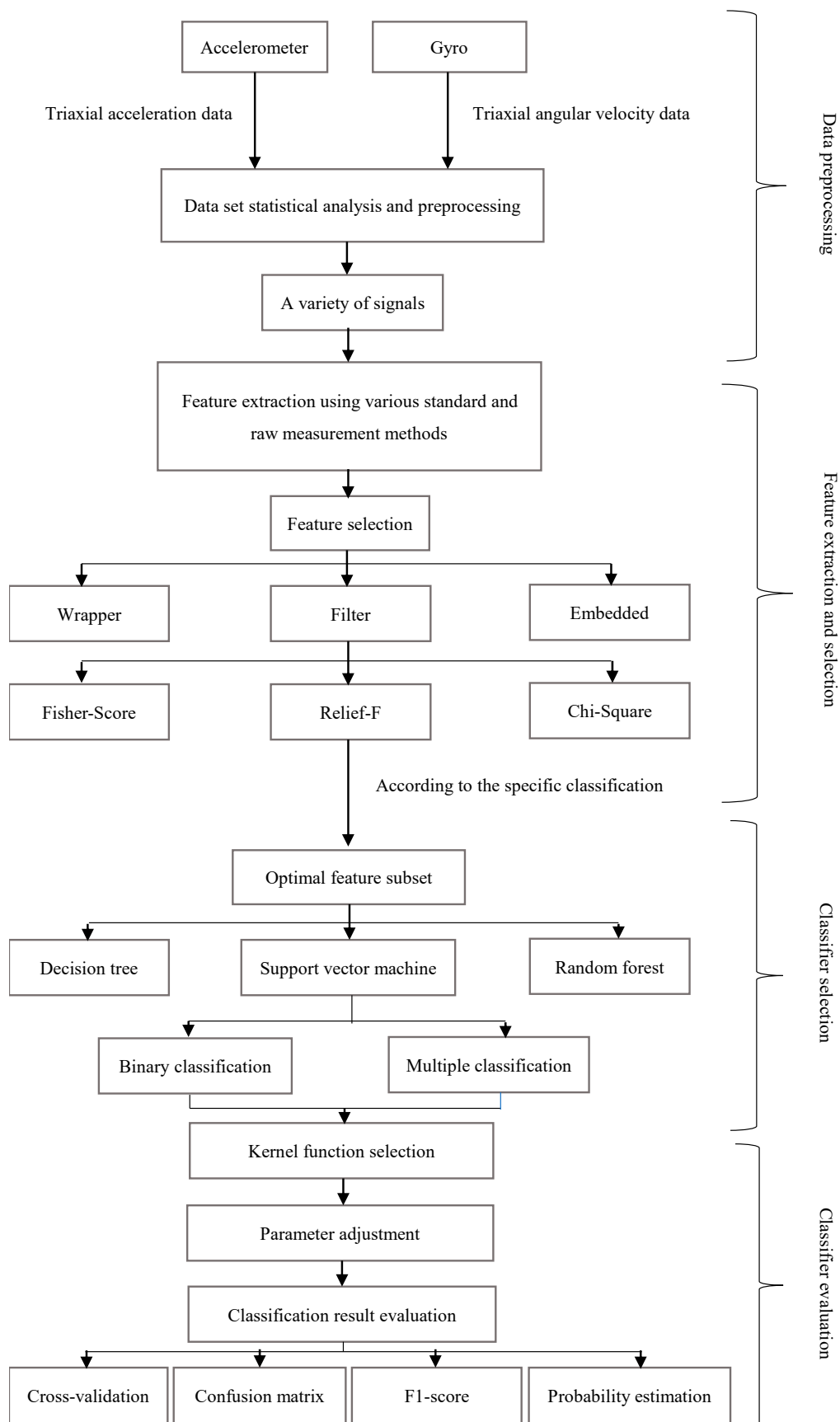


Figure 1. System Framework.

2. Data Preprocessing and Feature Selection

2.1. Data Preprocessing

The role of this module is to process the activity data received from the sensors and extract the variety of signals useful for activity recognition.

In this paper, we used the second generation human behavior recognition database available in the University of California Irvine (UCI) public platform [27]. The data set includes 6 basic activities: 3 static poses (standing, sitting, lying) and 3 dynamic poses (walking, downstairs, upstairs) for 30 different volunteers (everyone, aged between 19 and 48, who was instructed to follow the activity protocol when wearing an SGSII Smartphone at the waist as shown in Table 1), each volunteer was asked to do it twice. In addition, all possible pose transitions that occur between the existing three static poses are also available, including: standing-sitting (St-Si), sitting-standing (Si-St), sitting-lying (Si-Li), lying-sitting (Li-Si), standing-lying (St-Li), and lying-standing (Li-St). The frequency of the IMU was 100 Hz.

Table 1. Human activity recognition experiment protocol.

Serial Number	Static Poses	Time (s)	Serial Number	Dynamic Poses	Time (s)
0	Start (standing)	0	8	Walk (1)	15
1	Stand (1)	15	9	Walk (2)	15
2	Sit (1)	15	10	Downstairs (1)	12
3	Stand (2)	15	11	Upstairs (1)	12
4	Lay down (1)	15	12	Downstairs (2)	12
5	Sit (2)	15	13	Upstairs (2)	12
6	Lay down (2)	15	14	Downstairs (3)	12
7	Stand (3)	15	15	Upstairs (3)	12
			16	Stop	0

Table 1 shows all the activity tasks in order, and the corresponding time. In the process of experiment, every posture transformation performed twice by each volunteer. 60 labels were generated for each posture transformation which is accounting for 9% of all recorded experimental data. The duration of each posture transformation is different, and even reverse transitions (for example, Stand-Sit and Sit-Stand). The average duration of posture transition is 3.7 s, while the basic activity is about 20.1 s. The signals collected from one volunteer were extracted and the data of 12 movements (6 basic movements and 6 posture transformation) were statistically analyzed as shown in Figure 2.

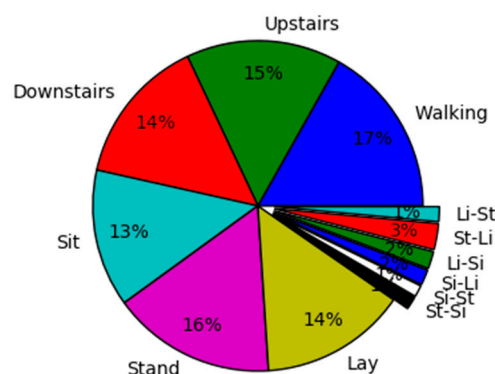


Figure 2. The statistics of posture data.

We process the original sensor signals obtained from the accelerometer ($ar(t)$) and the gyroscope ($wr(t)$) in three steps. First, we used a third-order median filter and a third-order low filter with a cutoff frequency of 20 Hz. Second, Battworth filter is applied for (transfer function is $H1(\omega)$) noise reduction, high-pass filter with a cutoff frequency of 0.3 Hz (transfer function is $H2(\omega)$) to eliminate

the influence of DC bias in the gyroscope. Third, the acceleration signal is divided into gravity $g(t)$ and object motion acceleration $a(t)$.

The sensor data is plotted as Figures 3 and 4. The red line is the acceleration signal in the X-axis, the green line is the acceleration signal in Y-axis, and the blue line is the acceleration signal in Z-axis. It is evident from Figures 3 and 4 that the sensor data in the attitude transition phase changes significantly. The units used for the accelerations are g's, while the gyroscope units are rad/seg. The horizontal axis describes the sampling points which is corresponding to the time. All the preprocessed signals are summarized in Table 2.

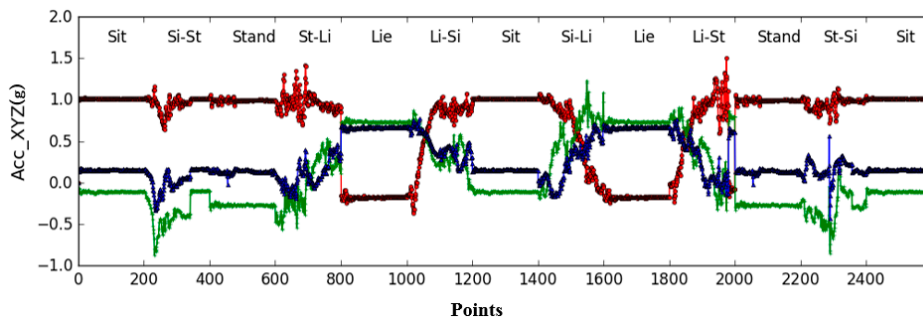


Figure 3. Acceleration X, Y, Z axis data.

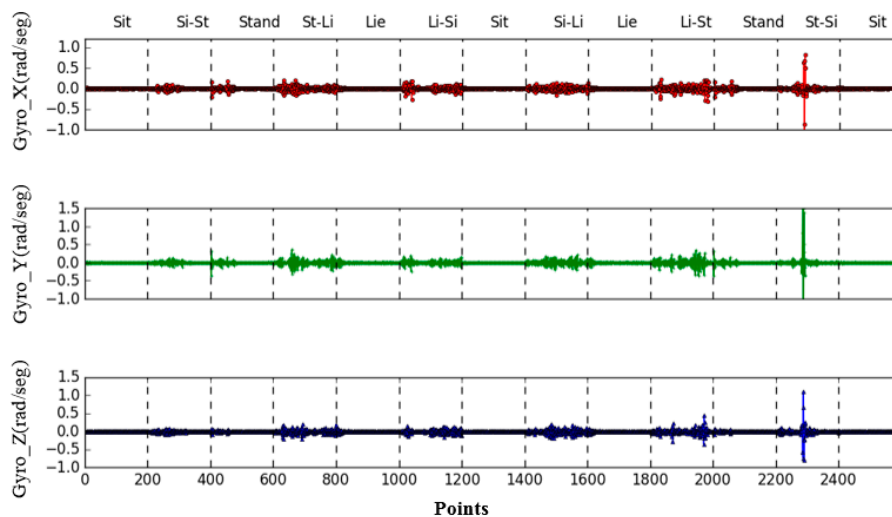


Figure 4. Angular velocity X, Y, Z axis data.

Table 2. Sensor inertial signal preprocessing.

Name	Quantity	Formula
Acceleration signal	tAcc (X,Y,Z)	$a_{\tau}(t) = H_1(a_r(t))$
Body acceleration signal	tAccBody (X,Y,Z)	$a(t) = H_2(a_{\tau}(t))$
Gravity signal	tGravity (X,Y,Z)	$g(t) = a_{\tau}(t) - a(t)$
Angular velocity signal	tGyro (X,Y,Z)	$\omega(t) = H_2(H_1\omega_r(t))$
Acceleration differential signal	tAccJerk (X,Y,Z)	$diff(a_{\tau}(t))$
Angular velocity differential signal	tGyroJerk (X,Y,Z)	$diff(\omega(t))$
Acceleration amplitude signal	tAccMag	$\ a_{\tau}(t)\ $
Angular velocity amplitude signal	tGyroMag	$\ \omega(t)\ $
Gravity amplitude signal	tGravityMag	$\ g(t)\ $
Acceleration and gravity angle signal	tAccAng	$\angle(a_{\tau}(t), g(t))$
Angular velocity and gravity angle signal	tGyroAng	$\angle(\omega(t), g(t))$
Acceleration frequency domain signal	fAcc (X,Y,Z)	$fft(a_{\tau}(t))$
Angular velocity frequency domain signal	fGyro (X,Y,Z)	$fft(\omega(t))$
Acceleration differential frequency domain signal	fAccJerk (X,Y,Z)	$fft(diff(a_{\tau}(t)))$

2.2. Feature Extraction

We used both the time and the frequency domain to extract the features. Table 3 shows the various measures and formulas used for generating feature sets on a fixed width window of length N, and there is 50% overlap between the two windows. The length of the window used in experiment is 2.56 s, since a person typically takes 1.5 steps per second on average, each window requires at least one full walking cycle.

Table 3. Feature Vector.

Function	Function Description	Formula
Mean (v)	Sample mean	$\bar{v} = \frac{1}{N} \sum_{i=1}^N v_i$
Var (v)	Sample variance	$\frac{1}{N-1} \sum_{i=1}^N (v_i - \bar{v})^2$
RMS (v)	Root mean square	$RMS = \left(\frac{1}{N} \sum_{i=1}^N v_i^2 \right)^{1/2}$
Energy (v)	Average of the sum of squares	$P(v) = \frac{1}{N} \sum_{i=1}^N (v_i)^2$
Entropy (v)	Information entropy	$E = - \sum_{i=1}^N v_i \log v_i$
Distance (v)	Euclidean distance	$L2 = \frac{1}{N-1} \sqrt{\sum_{i=2}^N (v_{i-1} - v_i)^2}$
MaxfreqInd (v)	Maximum frequency component	$\text{argmax}(v_i)$
MeanFreq (v)	Frequency signal weighted average	$\frac{\sum_{i=1}^N (iv_i)}{\sum_{j=1}^N v_j}$
EnergyBand (v,a,b)	Spectral energy in the [a,b] band	$\frac{1}{a-b+1} \sum_{i=a}^b v_i^2$

In our past work, we extracted a total of 585 features to describe each active window [25]. From the various features tabulated in Table 3, some new features are taken into account. These features are extracted from each axis of the acceleration signal and the angular velocity signal. The statistical features in Table 3 are also applicable to the x-axis, y-axis, z-axis, Mag, differential, and tilt angle of acceleration and angular velocity. Table 3 shows the feature representation form calculated by generating the metrics of the data set and the window signal of length 128. Taking the Mean (v) as an example to perform feature calculation on different processed signals and corresponding feature descriptions. Table 4 shows the characterization of the average value.

Table 4. Signal processing methods for feature average.

Characterization	Explanation
tAcc-X-Mean	The x-axis body acceleration signal after noise removal is averaged according to the window length
tAcc-Y-Mean	The y-axis body acceleration signal after noise removal is averaged according to the window length
tAcc-Z-Mean	The z-axis body acceleration signal after noise removal is averaged according to the window length
tGyro-X-Mean	The x-axis angular velocity signal after noise removal is averaged according to the window length
tGyro-Y-Mean	The y-axis angular velocity signal after noise removal is averaged according to the window length
tGyro-Z-Mean	The z-axis angular velocity signal after noise removal is averaged according to the window length
tGravityAcc-X-Mean	The gravity component of the x-axis acceleration signal is averaged according to the length of the window
tGravityAcc-Y-Mean	The gravity component of the y-axis acceleration signal is averaged according to the length of the window
tGravityAcc-Z-Mean	The gravity component of the z-axis acceleration signal is averaged according to the length of the window

Table 4. Cont.

Characterization	Explanation
tAccJerk-X-Mean	The derivative of the <i>x</i> -axis body acceleration signal is averaged according to the length of the window
tAccJerk-Y-Mean	The derivative of the <i>y</i> -axis body acceleration signal is averaged according to the length of the window
tAccJerk-Z-Mean	The derivative of the <i>z</i> -axis body acceleration signal is averaged according to the length of the window
tGyroJerk-X-Mean	The derivative of the gravity component of the <i>x</i> -axis acceleration signal is averaged according to the length of the window
tGyroJerk-Y-Mean	The derivative of the gravity component of the <i>y</i> -axis acceleration signal is averaged according to the length of the window
tGyroJerk-Z-Mean	The derivative of the gravity component of the <i>z</i> -axis acceleration signal is averaged according to the length of the window
tAccMag-Mean	The amplitude of the triaxial body acceleration signal is averaged according to the length of the window
tGyroMag-Mean	The amplitude of the three-axis angular velocity signal is averaged according to the window length
tGravityAccMag-Mean	The amplitude of the gravity component of the three-axis acceleration signal is averaged according to the length of the window
tAccAng-Mean	The angle between the acceleration signal and the direction of gravity is averaged according to the length of the window
tGyroAng-Mean	The angle between the angular velocity signal and the direction of gravity is averaged according to the length of the window

2.3. Feature Selection

The objective of this step is to select the significant features from the feature set obtained in the feature extraction module to the training model [28,29]. The feature selection methods adopted by most researchers include Filter, Embedded, Wrapper. In this step, we used the filtering methods in the feature selection algorithm. The basic principle of feature selection algorithm is shown in Figure 5.

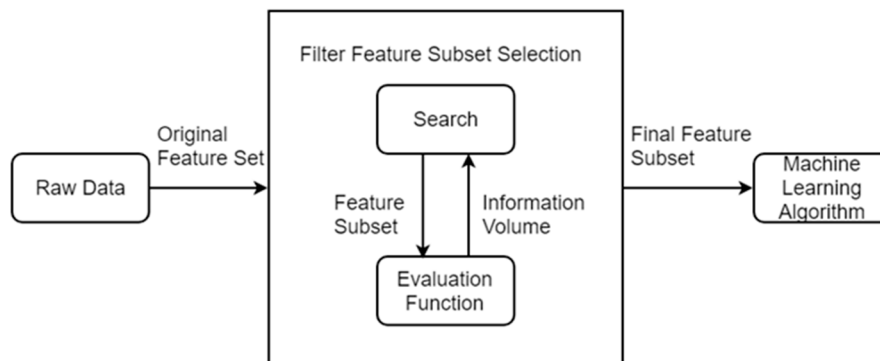


Figure 5. Filtered feature selection algorithm.

The algorithm uses divergence or correlation indicators to score each feature, and selects features with scores greater than a threshold or selects the top K features with the largest scores. Specifically, calculate the divergence of each feature, remove the features whose divergence is less than the threshold/select the top k features with the largest score; calculate the correlation between each feature and the label, and remove the features/selection with a correlation less than the threshold the top k features with the largest scores.

The advantages of the filtered feature selection algorithm are mainly versatility, low complexity, and fast running speed [30]. In this paper, three filtering feature selection algorithms, Relief-F, Fisher-Score, and Chi-Square, were applied to select the features.

The purpose of selected feature set is to classify the posture transformation between six basic movements (walking, going upstairs, downstairs, sitting, standing, and lying) and to achieve this, we selected 585 features. First, feature selection is made for the two categories: one is six basic actions, and another is six posture transformations. The results are shown in Figure 6. Secondly, the multiple classifications are characterized. The six basic movements are six categories, and another is all posture transformations. The results are shown in Figure 7.

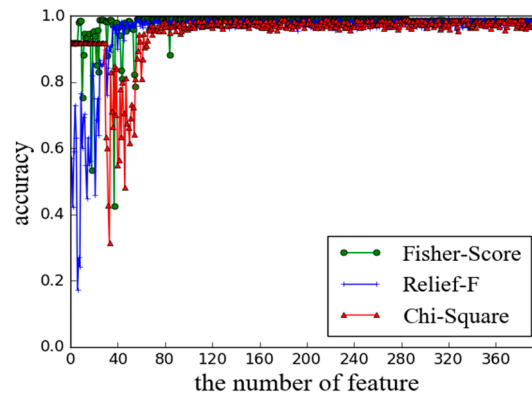


Figure 6. Three feature selection algorithm results in two categories.

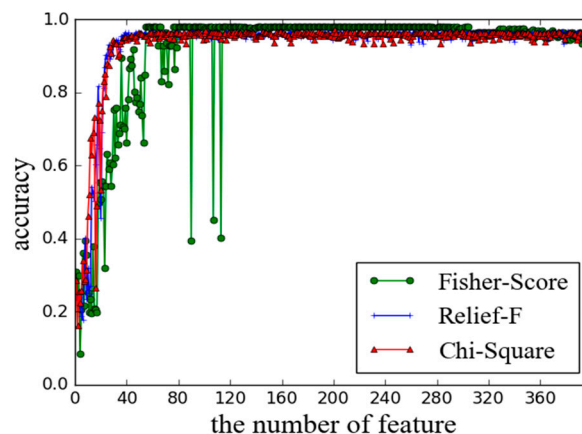


Figure 7. Three feature selection algorithm results in multiple classification.

In Figures 6 and 7, the abscissa refers to the number of features selected by the three feature selection algorithms, and the ordinate refers to the classification accuracy. It can be seen from Figures 6 and 7 that the classification accuracy increases gradually with increase in the number of selected features and approaches to 1. The ordering of the abscissa features in the three feature selection algorithms is sorted according to the scores of the features in the three algorithm principles.

In order to further select features of smaller dimensions to classify human poses with higher accuracy, we first input the first feature selected by each algorithm, that is, the three features into the classifier for training, obtain a classification model, and test it. If the test accuracy does not reach the ideal value, the first two features selected by each feature selection algorithm are selected for classification training, and so on, the feature combination with the highest classification accuracy is selected.

Finally, the features with highest score got from three feature selection methods were selected in the two categories: the maximum value in the fAcc (X) sequence, the frequency signal kurtosis in the fAcc (Y) sequence, and the sample range of the fAcc (X) sequence. In order to ensure classification accuracy in multiple classifications, 30 features (The top ten features selected by each feature selection method) were selected as shown in Table 5.

Table 5. Selected features.

SF. no.	Feature Description	Symbol Used
1, 2	sequence mean	tAcc (Y), tGravity (Y)
3, 4	sequence median	tAcc (Y), tGravity (Y)
5, 6	maximum value in the sequence	fAcc (Y), fGyro (X)
7	standard deviation	tAccJerk (X)
8	frequency signal skew	fAcc (Y)
9	range	fGyro (X)
10, 11, 12, 13	quartile of the sequence	tAccJerk (X), tGyroJerk (Z), tAccMag, tGravityMag
14, 15, 16, 17	10th percentile	tGyroJerk (Z), tAccMag, tGravityMag, tGyroAng
18, 19, 20	25th percentile	tAccMag, tGravityMag, fGyro (X)
21, 22	50th percentile	tAcc (Y), Gravity (Y)
23, 24, 25, 26	75th percentile	tGravity (Y), tGyroJerk (Z), tAccMag, tGravityMag
27, 28, 29, 30	90th percentile	tGyroJerk (Z), tAccMag, tGravityMag, tGyroAng

3. Classifier Selection

We used Support Vector Machine (SVM), which is a supervised machine learning algorithm developed in the last century and often used in statistical classification problems [31]. It was more often applied to the two-classification problem. The basic model is a linear classifier, which is transformed into a convex quadratic programming problem by maximizing the interval [32]. SVM is effective in high-dimensional space and suitable for situations where the dimensions are larger than the samples. Different kernel functions can be formulated for different scenarios. Linear separable samples can be classified by linear function. In diverse dimensions, the classifier shows different forms, such as a straight line for two-dimensions as shown in Figure 8, a plane for three-dimension and hyperplane for high-dimensional space.

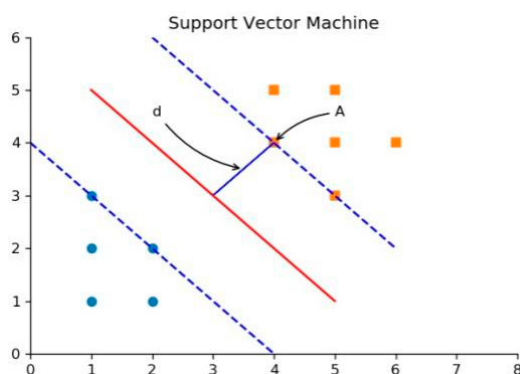


Figure 8. Support vector machine in two-dimensional space.

The decision tree is a tree that is constructed according to different strategies. By training the input data, the decision tree can be constructed, which can classify the unknown data efficiently, that is, predict the future based on the known [33]. It is a tree structure algorithm composed of root node, internal node, and leaf node. The core idea of the decision tree algorithm is to select attributes based on information gain and select the attribute with the largest information gain as the root [34]. The root is the top classification condition, each node of the tree acts as a test point on the property. The leaf node represents each category number, and the branch is on behalf of the output of each criteria. A binary tree has two branches on each node, while a node in a multi-tree has more than two branches.

The random forest algorithm is mainly based on the model aggregation idea, and has high precision in the classification and regression of high dimensional uncertainties [35]. The key idea under the random forest classifier is to grow a large number of unbiased decision trees from the guided samples, where each tree is voted for an activity class, and the random forest finally selects the most voted classification in the forest [36]. The random forest starts by selecting guide samples from the

original training data. Then learning each guide sample through the decision tree. Only a small number of variables are available for binary partitioning on each node.

In the previous section, three filtering feature selection algorithms were used to select three features for the two-category case, and 30 features were selected for the multi-classification case. Next, for the different classification cases, three features and 30 features were respectively applied to the three classifiers, and the test set classification accuracy is shown in Tables 6 and 7. According to the analysis of the classification results, there is no significant difference between the classification accuracy of the three sets of testers. We found that the results of the SVM are better than the other two. Precision, recall and F1-score is the evaluation index of the classification results. Avg/total calculates the mean value of entirety, which represents the overall situation of evaluation index. We used the features selected by Fisher-Score, Relief-F and Chi-Square to train the SVM, and the training set accuracy is shown in Table 8.

Table 6. Three classifier experimental results of two categories.

Test Set Classification Accuracy			Precision	Recall	F1-Score
SVM	1.0	Class 1	1.00	1.00	1.00
		Class 2	1.00	1.00	1.00
		Avg/total	1.00	1.00	1.00
Decision tree	0.9767	Class 1	0.97	1.00	0.99
		Class 2	1.00	0.75	0.86
		Avg/total	0.98	0.98	0.98
Random forest	0.9827	Class 1	0.97	1.00	0.99
		Class 2	1.00	0.75	0.86
		Avg/total	0.98	0.98	0.98

Table 7. Three classifier experimental results of multiple classification

Test set Classification Accuracy			Precision	Recall	F1-Score
SVM	0.9827	Avg/total	0.99	0.98	0.98
Decision tree	0.9792	Avg/total	0.98	0.98	0.98
Random forest	0.9801	Avg/total	0.98	0.98	0.98

Table 8. Training set results of SVM.

	Fisher-Score	Relief-F	Chi-Square
Accuracy	0.954	0.988	0.976

4. Classification Results Analysis and Improvement

4.1. Classifier Parameter Selection

In this Module, we used the support vector machine as a common classifier to classify the pose. The role of the kernel function is to map the input space to a high-dimensional space with certain rules, and construct an optimal separation hyperplane in it, and finally achieve the effect of separating nonlinear data [37]. We mainly used linear and Radial Basis Function.

If we learn and test the classifier model on the same subset of data, it will lead over-fitting phenomenon which can be avoided by cross-validation.

The data of 30 volunteers in the original data set were divided: the data of the first 15 people were used as the feature selection set, the data from th 16th to 26th person were used as the training set of the classifier, and the others were used as the test set of the classifier.

4.1.1. Classifier Linear Kernel Parameter Selection

A commonly used parameter in a linear kernel is the penalty factor C . When the value of C is large, the misclassification is less, the fitting to the sample is better, but it is easy to cause overfitting [38]. Although the possibility of misclassification becomes larger and the fit to the sample is degraded, the prediction effect may be more desirable due to the influence of noise between the samples [39].

First, based on the three features selected in the previous section, the linear kernel support vector machine is used to solve the two-class problem in behavior recognition. Figure 9 shows the selection process for parameter C in the two classifications. Next, based on the 30 features selected in the previous section, we used the linear kernel support vector machine to solve the seven classification problem in behavior recognition. Figure 10 shows the selection process of parameter C in the multi-class.

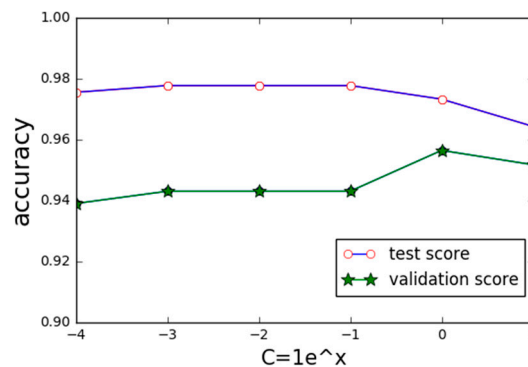


Figure 9. Selection of penalty factor C in the two categories.

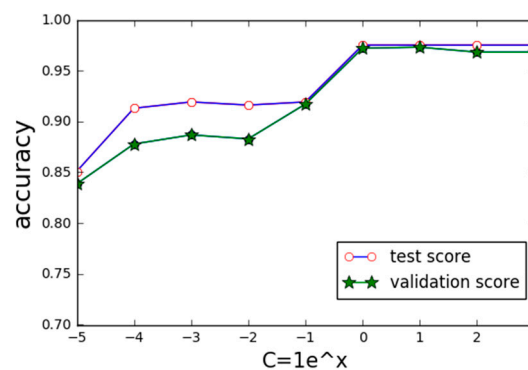


Figure 10. Selection of penalty factor C in multiple classification.

In Figures 9 and 10, the upper line represents the test set classification accuracy, and the lower line represents the cross-validation average. The abscissa shows the change of the penalty factor C , and the ordinate indicates the classification accuracy. It can be seen that with the increase of the penalty factor C , the classification accuracy and cross-validation average of the test set increase, but when the value of C is too large, the classification accuracy decreases slightly. In the process of processing the data, the larger the value of C , the more the error cannot be tolerated, and the time required for data processing will be longer. However, if the value of C is too small, we cannot guarantee that the parameter can be applied to other data sets. However, It still has a better effect. Therefore, considering the comprehensive consideration, we used the penalty factor value equals to 1. The 27th–29th people in the database were used for cross-validation to calculate the average precision value, mean value and standard deviation. We noticed that the classification accuracy of the test set is 0.973, the average cross-validation is 0.956, and the standard deviation of cross-validation is 0.042, which can achieve the desired effects. The factor C has a value of 1, and the classification accuracy of the test set is 0.975, the cross-validation average is 0.972, and the cross-validation standard deviation is 0.033, which can achieve the desired effect.

4.1.2. Classifier RBF Kernel Parameter Selection

The radial basis function (RBF) is a localized kernel function whose role is to map samples to high dimensional space. There are two main parameters in the classifier of RBF: the penalty factors C and σ [40]. The parameter σ reflects the clustering of the points after the mapping. The smaller the parameter σ , the distance between the mapped points tends to be equal, and the classification of the points will be finer, which will easily lead to overfitting. The larger the parameter σ , the coarser the classification will be, making it impossible to distinguish the data.

In the process of selecting the penalty factor C and the parameter σ , when the value of C is too large, over-fitting is easy to occur. When the value of σ is too small, the more support vectors are, the finer the classification is, and over-fitting easily occurs. And the increasing of the number of support vectors affects the speed of training and prediction [41]. The cross-validation is also used to determine whether the classification result has been over-fitted.

First, based on the three features selected in the previous section, the classifier of the radial basis kernel was used to solve the two-class problem in behavior recognition. Figure 11 shows the selection process of parameters C and σ in dichotomies. We used radial basis kernel support vector machine to solve the seven classification problem in behavior recognition based on the 30 features selected in the previous section. Figure 12 shows the selection process for parameters C and σ in the multi-category.

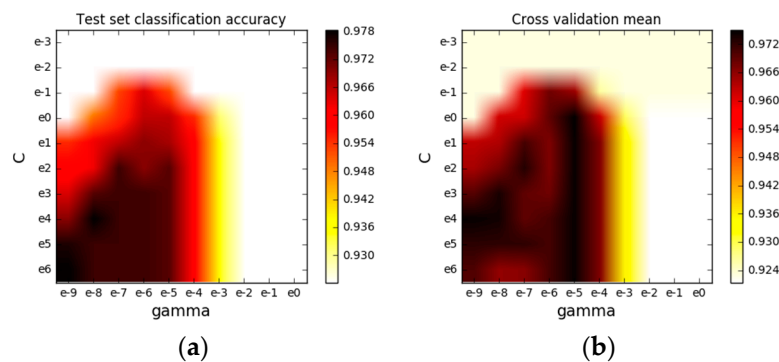


Figure 11. Selection of parameters C and σ in two categories: (a) Test set classification accuracy; (b) Cross validation mean.

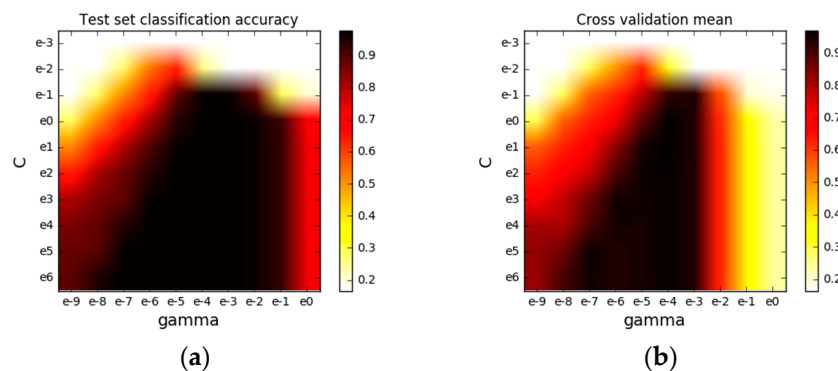


Figure 12. Selection of parameters C and σ in multiple classification: (a) Test set classification accuracy; (b) Cross validation mean.

There are two subgraphs in Figures 11 and 12. The abscissa shows the change of the parameter σ and the ordinate shows the change of the parameter C . While Figures 11 and 12a shows the classification accuracy of the test set, and Figures 11 and 12b represents the cross validation average. The darker the color, the larger the value. When the penalty factor C is too small and the parameter value σ is too large, the classification accuracy may not reach the ideal value. However, excessive pursuit of classification accuracy may cause computational complexity. Considering comprehensively, when the penalty factor

C is selected as 100 and the parameter is selected as 0.00001 in the second classification, the classification accuracy of the test set is 0.973, the cross-validation average is 0.975, and the cross-validation standard deviation is 0.011, which can achieve the desired effect, the penalty factor C in the seven classification. When the parameter is selected and the parameter is 0.001, the classification accuracy of the test set is 0.978, the average cross-validation is 0.938, and the cross-validation standard deviation is 0.057, which can achieve the desired effect.

4.2. Probability Estimation

Commonly used SVM can only generate categories without probability. The probability estimation can be used to transform the classification result of the support vector machine, that is, the probability that a sample belongs to each category [42].

The probabilistic calibration used in this study is isotonic regression, which is a nonparametric method. The core idea is to fit the deviation between the current classifier output and the real results. Isotonic regression is suitable for cases with large sample sizes, and over-fitting is prone to occur when the sample size is small. The Brier score can be used to evaluate the results of the probabilistic calibration. The Brier score is a loss, so the smaller score is better [43]. In all categories in which N predictions are aggregated, the Brier score measures the mean square error between the predicted probability and the actual probability assigned to the category. Therefore, for a set of predictions means the lower the Brier score, the better the prediction calibration effects.

In this paper, we used data of five volunteers on which we used the support vector machine to learn and classify, and then uses isotonic regression to probabilistically estimate the data compiled by the volunteers. Due to individual differences, they completed each activity in different time actually. In order to maintain the integrity of a whole set of actions, result of one volunteer was presented only in Figure 13. In Figure 13, the abscissa is the test set data corresponding to different postures randomly selected from the volunteer data, and the ordinate is the predicted probability value obtained by estimating the probability of the data. The seven different colored lines represent the probability that the data is predicted into seven categories.

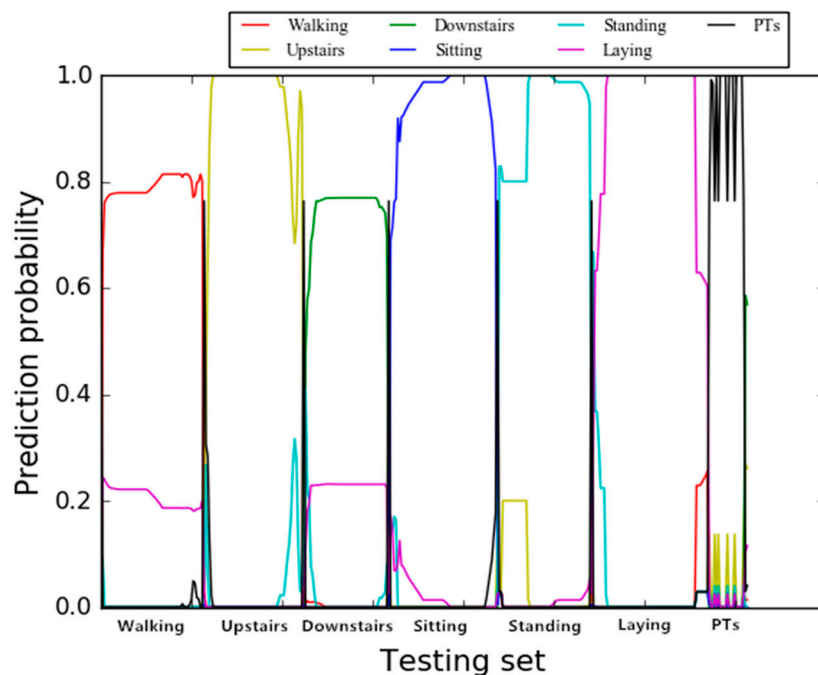


Figure 13. Seven classification probability estimation result.

The Brier score is then used to evaluate the results of the probability estimates. The average results of five volunteers are shown in Table 9. The column labels in the table represent which actions the selected data comes from, the row labels represent the seven categories, and the values in the table are the obtained Brier scores. The Brier score on the diagonal in the table is relatively small, so the result of the probability estimation achieves the desired effects. Comparing with the experiments adopted SVM in the literature [44], SVM with kernel parameter selection adjustment has a significantly higher effectiveness and accuracy in identifying “walking”, “upstairs”, and “downstairs”.

Table 9. Brier score evaluation result of probability estimation.

Brier-Score	Walking	Upstairs	Downstairs	Sitting	Standing	Laying	Posture Transitions
Walking	0.0105	0.2801	0.2432	0.2746	0.2332	0.2702	0.1980
Upstairs	0.2741	0.0107	0.2648	0.3023	0.2332	0.3344	0.2365
Downstairs	0.2532	0.2890	0.0204	0.2833	0.2374	0.2562	0.2087
Sitting	0.2820	0.3172	0.2741	0.0116	0.2714	0.3172	0.2293
Standing	0.2046	0.2231	0.1968	0.2341	0.0082	0.2701	0.1753
Laying	0.2820	0.3321	0.2840	0.3187	0.2622	0.0181	0.2543
Posture transitions	0.1861	0.2142	0.1602	0.2194	0.1793	0.2433	0.0356

5. Conclusions

In recent years, research on behavioral recognition methods for transitional attitude perception has become more and more widely used in many fields such as medical care. Based on the evaluated human behavior recognition data set it is found that the three-axis acceleration values of different static actions are significantly different, the three-axis angular velocity values are basically the same, and the posture conversion data between static actions changes significantly. It is undeniable that the data of the static posture is not always stable, as it cannot be guaranteed that the volunteer was completely still while sitting (or standing or lying) during the experiment.

We used Fisher-Score, Relief-F, and Chi-Square to select 585 features to obtain relatively good features set for classification. The features with higher scores were calculated using methods such as maximum value, minimum value, variance, skewness, kurtosis, and information entropy. The investigation shows that support vector machine gives better results than decision tree and random forest. In the second classification, the classification accuracy of the linear kernel ($C = 1$) is 97%, and the classification accuracy of the RBF kernel ($C = 1, \sigma = 0.001$) in the multi-class is 98%. Probability estimation overcomes some of the shortcomings of SVM and can directly output the probability that the data belongs to each category, thus making the results more intuitive.

Author Contributions: Funding acquisition, S.F.; resources, S.F.; project administration, S.F.; methodology, L.C.; writing—original draft preparation, L.C.; Visualization V.K.; Validation Y.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by (Key Research and Development Plan Project of Hebei province, China) grant number (19210404D); (Key research and development project from Hebei Province, China) grant number (20351802D); (Key Research Project of Science and Technology from Ministry of Education of Hebei Province, China) grant number (ZD2019010); (Graduate Innovation Funding Project of Hebei Province) grant number (CXZZSS2020031).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nweke, H.F.; Teh, Y.W.; Mujtaba, G.; Al-Garadi, M.A. Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions. *Inf. Fusion* **2019**, *46*, 147–170. [[CrossRef](#)]
2. Hōrak, H. Computer Vision-Based Unobtrusive Physical Activity Monitoring in School by Room-Level Physical Activity Estimation: A Method Proposition. *Information* **2019**, *10*, 269. [[CrossRef](#)]

3. Xu, K.; Lu, Y.; Takei, K. Multifunctional Skin-Inspired Flexible Sensor Systems for Wearable Electronics. *Adv. Mater. Technol.* **2019**, *4*, 4. [[CrossRef](#)]
4. Gao, W.; Ota, H.; Kiriya, D.; Takei, K.; Javey, A. Flexible Electronics toward Wearable Sensing. *Acc. Chem. Res.* **2019**, *52*, 523–533. [[CrossRef](#)] [[PubMed](#)]
5. Xu, H.; Pan, Y.; Li, J.; Nie, L.; Xu, X. Activity Recognition Method for Home-Based Elderly Care Service Based on Random Forest and Activity Similarity. *IEEE Access* **2019**, *7*, 16217–16225. [[CrossRef](#)]
6. Sony, S.; LaVenture, S.; Sadhu, A. A literature review of next-generation smart sensing technology in structural health monitoring. *Struct. Control Health Monit.* **2019**, *26*, e2321. [[CrossRef](#)]
7. Li, J.H.; Tian, L.; Wang, H.; An, Y.; Wang, K.; Yu, L. Segmentation and Recognition of Basic and Transitional Activities for Continuous Physical Human Activity. *IEEE Access* **2019**, *7*, 42565–42576. [[CrossRef](#)]
8. Liu, Y.; Wang, X.; Zhai, Z.; Chen, R.; Zhang, B.; Jiang, Y. Timely daily activity recognition from headmost sensor events. *ISA Trans.* **2019**, *94*, 379–390. [[CrossRef](#)]
9. Ma, C.; Li, W.; Cao, J.; Du, J.; Li, Q.; Gravina, R. Adaptive sliding window based activity recognition for assisted livings. *Inf. Fusion* **2020**, *53*, 55–65. [[CrossRef](#)]
10. Zhang, S.; Wei, Z.; Nie, J.; Huang, L.; Wang, S.; Li, Z. A Review on Human Activity Recognition Using Vision-Based Method. *J. Healthc. Eng.* **2017**, *2017*, 343. [[CrossRef](#)]
11. Kim, K.; Jalal, A.; Mahmood, M. Vision-Based Human Activity Recognition System Using Depth Silhouettes: A Smart Home System for Monitoring the Residents. *J. Electr. Eng. Technol.* **2019**, *14*, 2567–2573. [[CrossRef](#)]
12. Babiker, M.; Khalifa, O.O.; Htike, K.K.; Hassan, A.; Zaharadeen, M. Automated daily human activity recognition for video surveillance using neural network. In Proceedings of the 2017 IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA), Putrajaya, Malaysia, 28–30 November 2017; pp. 1–5.
13. De Leonardis, G.; Rosati, S.; Balestra, G.; Agostini, V.; Panero, E.; Gastaldi, L.; Knaflitz, M. Human Activity Recognition by Wearable Sensors: Comparison of different classifiers for real-time applications. In Proceedings of the 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Rome, Italy, 11–13 June 2018.
14. Chen, M.J.; Li, Y.; Luo, X.; Wang, W.P.; Wang, L.; Zhao, W.B. A Novel Human Activity Recognition Scheme for Smart Health Using Multilayer Extreme Learning Machine. *Cyber Enabled Intell.* **2019**, *6*, 239–258. [[CrossRef](#)]
15. Zhu, Q.; Chen, Z.; Soh, Y.C.; Yeng, C.S. A Novel Semisupervised Deep Learning Method for Human Activity Recognition. *IEEE Trans. Ind. Inform.* **2019**, *15*, 3821–3830. [[CrossRef](#)]
16. Chen, Z.; Jiang, C.; Xie, L. A Novel Ensemble ELM for Human Activity Recognition Using Smartphone Sensors. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2691–2699. [[CrossRef](#)]
17. Ridok, A.; Mahmudy, W.F.; Rifai, M. An improved artificial immune recognition system with fast correlation based filter (FCBF) for feature selection. In Proceedings of the 2017 Fourth International Conference on Image Information Processing (ICIIP), Shimla, India, 21–23 December 2017; pp. 1–6.
18. Truong, P.H.; You, S.; Ji, S.-H.; Jeong, G.-M. Wearable System for Daily Activity Recognition Using Inertial and Pressure Sensors of a Smart Band and Smart Shoes. *Int. J. Comput. Commun. Control* **2019**, *14*, 726–742. [[CrossRef](#)]
19. Jiang, S.; Lv, B.; Guo, W.C.; Zhang, C.; Wang, H.T.; Sheng, X.J.; Shull, P.B. Feasibility of Wrist-Worn, Real-Time Hand, and Surface Gesture Recognition via sEMG and IMU Sensing. *IEEE Trans. Ind. Inform.* **2018**, *14*, 3376–3385. [[CrossRef](#)]
20. Lu, J.; Tong, K. Robust Single Accelerometer-Based Activity Recognition Using Modified Recurrence Plot. *IEEE Sens. J.* **2019**, *19*, 6317–6324. [[CrossRef](#)]
21. Hsu, Y.; Yang, S.; Chang, H.; Lai, H. Human Daily and Sport Activity Recognition Using a Wearable Inertial Sensor Network. *IEEE Access* **2018**, *6*, 31715–31728. [[CrossRef](#)]
22. Aminikhanghahi, S.; Cook, D.J. Enhancing activity recognition using CPD-based activity segmentation. *Pervasive Mob. Comput.* **2019**, *53*, 75–89. [[CrossRef](#)]
23. Gani, M.O.; Fayezeen, T.; Povinelli, R.J.; Smith, R.O.; Arif, M.; Kattan, A.J.; Ahamed, S.I. A light weight smartphone based human activity recognition system with high accuracy. *J. Netw. Comput. Appl.* **2019**, *141*, 59–72. [[CrossRef](#)]
24. Fan, S.R.; Jia, Y.T.; Liu, J.H. Feature selection based on three-axis acceleration sensor for human body attitude recognition. *J. Appl. Sci.* **2019**, *37*, 427–436.

25. Fan, S.R.; Jia, Y.T.; Jia, C.Y. A Feature Selection and Classification Method for Activity Recognition Based on an Inertial Sensing Unit. *Information* **2019**, *10*, 290. [[CrossRef](#)]
26. Liu, B.; Cai, H.; Ju, Z.; Liu, H. Multi-stage adaptive regression for online activity recognition. *Pattern Recognit.* **2020**, *98*, 107053. [[CrossRef](#)]
27. Jorge-Lr, O.; Luca, O.; Albert, S.; Xavier, P.; Davide, A. Transition-aware human activity recognition using smartphones. *Neurocomputing* **2016**, *171*, 754–767.
28. Haryanto, A.W.; Mawardi, E.K.; Muljono. Influence of Word Normalization and Chi-Squared Feature Selection on Support Vector Machine (SVM) Text Classification. In Proceedings of the 2018 International Seminar on Application for Technology of Information and Communication, Semarang, Indonesia, 21–22 September 2018; pp. 229–233. [[CrossRef](#)]
29. Wang, A.; Chen, G.; Wu, X.; Liu, L.; An, N.; Chang, C.-Y. Towards Human Activity Recognition: A Hierarchical Feature Selection Framework. *Sensors* **2018**, *18*, 3629. [[CrossRef](#)] [[PubMed](#)]
30. Dai, H. Research on SVM improved algorithm for large data classification. In Proceedings of the 2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA), Shanghai, China, 9–12 March 2018; pp. 181–185.
31. Huo, Z.; Zhang, Y.; Shu, L.; Gallimore, M. A New Bearing Fault Diagnosis Method Based on Fine-to-Coarse Multiscale Permutation Entropy, Laplacian Score and SVM. *IEEE Access* **2019**, *7*, 17050–17066. [[CrossRef](#)]
32. Zhou, B.; Wang, H.; Hu, F.; Feng, N.S.; Xi, H.L.; Zhang, Z.H.; Tang, H. Accurate recognition of lower limb ambulation mode based on surface electromyography and motion data using machine learning. *Comput. Methods Programs Biomed.* **2020**, *193*, 105486. [[CrossRef](#)]
33. Sagi, O.; Rokach, L. Explainable decision forest: Transforming a decision forest into an interpretable tree. *Inf. Fusion* **2020**, *61*, 124–138. [[CrossRef](#)]
34. Clutterbuck, G.L.; Auld, M.L.; Johnston, L.M. High-level motor skills assessment for ambulant children with cerebral palsy: A systematic review and decision tree. *Dev. Med. Child Neurol.* **2020**, *62*, 693–699. [[CrossRef](#)]
35. Ishwaran, H.; Lu, M. Standard errors and confidence intervals for variable importance in random forest regression, classification, and survival. *Stat. Med.* **2019**, *38*, 558–582. [[CrossRef](#)]
36. Probst, P.; Wright, M.N.; Boulesteix, A.-L. Hyperparameters and tuning strategies for random forest. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2019**, *9*, e1301. [[CrossRef](#)]
37. Nanda, M.A.; Seminar, K.B.; Nandika, D.; Maddu, A. A Comparison Study of Kernel Functions in the Support Vector Machine and Its Application for Termite Detection. *Information* **2018**, *9*, 5. [[CrossRef](#)]
38. Chun-Yi, T.; Wen-Hsiu, C. Multiclass object classification using covariance descriptors with kernel SVM. *J. Comput.* **2018**, *29*, 244–249.
39. Han, Y.; Li, J.; Xing, H.W.; Yang, A.-M.; Pan, Y.-H. Demonstration of SVM Classification Based on Improved Gauss Kernel Function. *Adv. Intell. Syst. Comput.* **2018**, *613*, 189–195. [[CrossRef](#)]
40. Shi, L.J.; Sun, B.B.; Ibrahim, D.S. An active learning reliability method with multiple kernel functions based on radial basis function. *Struct. Multidiscip. Optim.* **2019**, *60*, 211–229. [[CrossRef](#)]
41. Zhao, T.; Pei, J.H.; Chen, H. Multi-layer radial basis function neural network based on multi-scale kernel learning. *Appl. Soft Comput.* **2019**, *82*, 105541. [[CrossRef](#)]
42. Konopko, K.; Janczak, D. Classification method based on multidimensional probability density function estimation dedicated to embedded systems. *IFAC Pap.* **2018**, *51*, 318–323. [[CrossRef](#)]
43. Goldenholz, D.M.; Goldenholz, S.; Romero, J.; Moss, R.; Sun, H.Q.; Westover, B. Development and Validation of Forecasting Next Reported Seizure Using e-Diaries. *Ann. Neurol.* **2020**, *88*, 588–595. [[CrossRef](#)]
44. Shi, J.; Zuo, D.; Zhang, Z. Transition Activity Recognition System Based on Standard Deviation Trend Analysis. *Sensors* **2020**, *20*, 3117. [[CrossRef](#)]

