Robust human activity recognition using single accelerometer via wavelet energy spectrum features and ensemble feature selection

Yiming Tian, Jie Zhang, Jie Wang, Yanli Geng & Xitai Wang

To cite this article: Yiming Tian, Jie Zhang, Jie Wang, Yanli Geng & Xitai Wang (2020) Robust human activity recognition using single accelerometer via wavelet energy spectrum features and ensemble feature selection, Systems Science & Control Engineering, 8:1, 83-96

To link to this article: https://doi.org/10.1080/21642583.2020.1723142

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 18 Feb 2020.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=tssc20
Robust human activity recognition using single accelerometer via wavelet energy spectrum features and ensemble feature selection

Yiming Tian\textsuperscript{a}, Jie Zhang\textsuperscript{b}, Jie Wang\textsuperscript{a}, Yanli Geng\textsuperscript{a} and Xitai Wang\textsuperscript{a}

\textsuperscript{a}School of Artificial Intelligence & Data Science, Hebei University of Technology, Tianjin, People’s Republic of China; \textsuperscript{b}School of Engineering, Merz Court, Newcastle University, Newcastle upon Tyne, UK

ABSTRACT
Wearable sensor-based human activity recognition has been widely used in many fields. Considering that a multi-sensor based recognition system is not suitable for practical applications and long-term activity monitoring, this paper proposes a single wearable accelerometer-based human activity recognition approach. In order to improve the reliability of the recognition system and remove redundant features that have no effect on recognition accuracy, wavelet energy spectrum features and a novel feature selection method are introduced. For each activity sample, wavelet energy spectrum features of the acceleration signal are extracted and the activity is represented by a feature set including wavelet energy spectrum features and features of other attributes. Then, considering the limitation of single filter feature selection method, this paper proposes an ensemble-based filter feature selection (EFFS) approach to optimize the feature set. Features that are robust to sensor placement and highly distinguishable for different activities are selected. In the experiment, the acceleration data around waist is collected and two classifiers: k-nearest neighbour (KNN) and support vector machine (SVM) are utilized to verify the effectiveness of the proposed features and EFFS method. Experiment results show that the wavelet energy spectrum features can increase the discrimination between different activities and significantly improve the activity recognition accuracy. Compared with other four popular feature selection methods, the proposed EFFS approach provides higher accuracy with fewer features.

1. Introduction
In recent years, human activity recognition (HAR) has become a popular research area in pattern recognition and machine learning. This technology benefits numerous applications, such as health-care monitoring, fitness, smart home, etc. For example, it can be used in monitoring daily activities of elderly people living alone (Chenbumroong, Shuang, & Yu, 2015) and providing technical support to improve the speed of swimmers (Mooney, Corley, Godfrey, Quinlan, & Ó Laighin, 2016). By controlling the indoor temperature according to the amount of human activity, effective energy control can be achieved in smart homes (Javed, Larijani, Ahmadinia, & Gibson, 2017). Human activity recognition also plays an important role in developing somatosensory animations and games (Tachibana, Noah, Bronner, Ono, & Onozuka, 2011).

In activity recognition, video-based activity recognition has experienced great breakthroughs. Distributed cameras were utilized for tracking and activity recognition (Song et al., 2010). Microsoft Kinect camera was utilized by Huynh-The et al. (Huynh-The et al., 2018) for developing an effective indoor activity analysis approach. This approach includes skeleton-based feature extraction and topic model-based learning, which are two major contributions. Ma et al. (Ma et al., 2014) also used Kinect camera and built a dataset that consists of six types of activities, curvature scale space (CSS) features were extracted from each frame and each activity was represented by a feature bag. At last, a parameter-optimized extreme learning machine was utilized to identify fall from other activities. Kwolek et al. (Kwolek & Kepski, 2015) proposed a fall detection method that combined Kinect and acceleration sensor. The acceleration data was used to monitor the trend of falling and the depth of the image was used to finalize and validate the fall warnings from accelerated data. But this method fails when the subject leaves Kinect’s detection range. Video-based activity recognition methods have some disadvantages: first, the performance of these methods will be greatly affected by external conditions, especially when there is high illumination in the external environment. Moreover, the cameras are not suitable for installation in some places, such
as bathrooms and bedrooms, which would invade personal privacy. Last but not least, the video-based system is complicated and cameras with high performance are expensive.

Sensor-based activity recognition which can get rid of the influence of light and disturbances shade has attracted more attentions from researchers (Cornacchia, Ozcan, Zheng, & Velipasalar, 2017). Previous works in this area can be roughly divided into two categories according to the composition of the sensor hardware system. In the first category, the general idea is to use the networked sensor layout to enhance the system's recognition capability (Lai et al., 2013). The final decision can be obtained by using fusion theory, such as data-level fusion (Gravina, Alinia, Ghasemzadeh, & Fortino, 2017), feature-level fusion (Liu, Gao, John, Staudenmayer, & Freedson, 2012), and decision-level fusion (Mo, Liu, Gao, & Freedson, 2013). For example, Zhu et al. (Zhu & Sheng, 2009) proposed a data-level fusion approach for activity recognition and the acceleration data from feet and waist were utilized. The experiments showed that this method achieved an accuracy of 88.1% when using SVM, which was 12.3% higher than that of only using the acceleration data from hip. A gait analysis system based on body sensor network (BSN) is proposed in (Wang, Qiu, Cao, & Jiang, 2013) to measure quantitative gait parameters. The proposed BSN contains two wireless inertia sensors fixed to the left and right ankles, acceleration and angular velocity are obtained from both sides. Lai et al. (Lai, Chang, Chao, & Huang, 2011) designed a multi-accelerometer based fall detection system. Six accelerometers were fixed on different parts of the body when the accelerometer detected acceleration value was significantly higher than the normal range, the system issued a fall warning. However, these threshold-based fall detection methods are susceptible to noise, wear location and individual differences. The application of BSN also covers sports training. For example, a badminton training system which consists of left wrist sensor, right wrist sensor, waist sensor and right ankle sensor is designed in (Wang, Guo, & Zhao, 2016). The proposed system fused the data of four sensors at the feature level and utilized the double-layer hidden markov model (HMM) classification algorithm to recognize 14 types of badminton batting activities. Despite the multi-sensor system have made outstanding achievement. Excessive sensors on the body will cause inconvenience to the person's daily life, especially in long-term monitoring. Moreover, using multiple sensors will significantly increase the cost of the equipment.

Comparatively, the second category which utilizes a single sensor for activity recognition has attracted more attention because of its convenience and practicality (Shoaib, Bosch, Incel, Scholten, & Havinga, 2015). Optimization work is always performed in the feature set so that the classifier can achieve good recognition performance and generalization in the single sensor-based recognition system. Khan et al. (Khan, Lee, Lee, & Kim, 2010) proposed a feature extraction method for real-time activity recognition on mobile phones. However, the system has poor performance when the phone is in the pocket of pants. A cascade classifier was proposed by Cheng (Cheng & Jhan, 2013) for fall detection by using a single triaxial accelerometer, different positions had been tested for the proposed cascade classifier. Experiments showed that the best result of 98.48% accuracy rate had been obtained when the single accelerometer was fixed to the chest. Wang et al. (Wang, Chen, Yang, Zhao, & Chang, 2016) explored the power of triaxial accelerometer built-in a smartphone in recognizing human physical activities. Experiments show that the accelerometer has superiority over the gyroscope in distinguishing some daily activities such as lying from sitting and standing. Acceleration data from lumbar position was used to recognize the four activities in (Li & Yao, 2016). However, this method is only effective when the upper body changes from vertical to level position. Margarito et al. (Margarito, Helaoui, Bianchi, Sartor, & Bonomi, 2016) researched the performance of template matching method on activity recognition by using single accelerometer placed on the wrist. Its performance is compared with the statistical learning classifier through experiments and it is found to be suitable as a supplement to the statistical learning recognition method. In (Gibson, Amira, Ramzan, Casaseca-De-La-Higuera, & Pervez, 2016), wavelet features of the accelerometer data from the subject's chest were extracted and a multi-classifier frame which fuses the decision of each classifier by voting method was utilized for fall detection. However, this method requires more complex equipment, which is not conducive to its promotion and popularization, in addition, a multi-classifier system may reduce the real-time of detection. Rodríguez-Martin et al. (Rodríguez-Martin et al., 2015) proposed a hierarchical classifier to recognize the posture transition of patients with Parkinson's disease by using a single accelerometer fixed to the waist. The first level classifier identifies whether a person has executed a posture transition and the second level classifier recognize which type of posture transition has occurred.

Though most of the aforementioned researches have attempted to improve the recognition performance and reliability of activity recognition system, most of them ignored that the accelerometer is worn on the human body and is usually removed when not in use and is worn again before use. Therefore, wearing it once again does not guarantee complete and exact alignment with the previous position, which usually brings a deviation from
the standard position. Some works (Wang, Wu, Chen, Ghoneim, & Hossain, 2016), (Chen, Zhu, Chai, & Zhang, 2017) have begun to study the robustness of recognition performance when positional changes occur. This problem is easy to ignore and is likely to cause a decline in recognition performance. Therefore, this paper focused on the reliability of single-accelerometer based human activity recognition and selecting optimum number of features robust to the position around waist.

At present, wavelet energy spectrum features based on wavelet analysis have been widely applied to fault diagnosis (Yan, Gao, & Chen, 2014) and biological signal-based disease diagnosis (Shriram, Sundhararajan, Shete, & Daimiwal, 2017), (Patidar, Pachori, & Acharya, 2015). As a kind of time–frequency feature, wavelet energy spectrum is more advantageous to deal with the mode mixing problem compared with some traditional time and frequency features. However, to our best knowledge, wavelet energy spectrum is rarely used in activity recognition problems. Therefore, in this paper, the acceleration wavelet energy spectrum features were utilized to construct feature set, enhancing the discrimination between different kinds of activities.

On the other hand, high-quality features are critical to improving the recognition accuracy of pattern recognition. There are uncorrelated or redundant features in the feature set that do not contribute to the recognition accuracy. Furthermore, these redundant features will lead to higher computational cost. Therefore, it is necessary to optimize the optimal feature subset to achieve good recognition performance and reduce computational cost. In addition, a good feature subset also needs to be robust to the position of the sensor’s offset. Compared with wrapper method, filter feature selection methods can be easily applied to high-dimensional datasets and have high computational efficiency (Vergara & Estévez, 2014). However, most of these methods can only evaluate features from an evaluation criterion and they can’t measure the features from the perspective of the feature set. Recently, some researches (Chandrashekar & Sahin, 2014), (Apolloni, Leguizamón, & Alba, 2016), (Abeel, Helleputte, & Dael, 2015). As a kind of time–frequency feature, wavelet energy spectrum is more advantageous to deal with the mode mixing problem compared with some traditional time and frequency features. However, to our best knowledge, wavelet energy spectrum is rarely used in activity recognition problems. Therefore, in this paper, the acceleration wavelet energy spectrum features were utilized to construct feature set, enhancing the discrimination between different kinds of activities.

On the other hand, high-quality features are critical to improving the recognition accuracy of pattern recognition. There are uncorrelated or redundant features in the feature set that do not contribute to the recognition accuracy. Furthermore, these redundant features will lead to higher computational cost. Therefore, it is necessary to optimize the optimal feature subset to achieve good recognition performance and reduce computational cost. In addition, a good feature subset also needs to be robust to the position of the sensor’s offset. Compared with wrapper method, filter feature selection methods can be easily applied to high-dimensional datasets and have high computational efficiency (Vergara & Estévez, 2014). However, most of these methods can only evaluate features from an evaluation criterion and they can’t measure the features from the perspective of the feature set. Recently, some researches (Chandrashekar & Sahin, 2014), (Apolloni, Leguizamón, & Alba, 2016), (Abeel, Helleputte, Van de Peer, Dupont, & Saey, 2010) have found that ensemble feature selection can effectively identify features with high discrimination in feature subsets and determine features with high correlation with output. Thus, an ensemble-based filter feature selection (EFFS) method is proposed in this paper to select the optimum number of features which are relevance and robust to the deviation position.

The structure of the paper is as follows: after the introduction, Section 2 introduces the wavelet energy spectrum features and the feature set that used in this paper. Section 3 presents the proposed ensemble-based filter feature selection (EFFS) method. Section 4 describes the experimental setup and data sets. The proposed features and the EFFS method are experimentally evaluated and analyzed in Section 5. Some concluding remarks are given in Section 6.

2. Feature extraction

2.1. Wavelet packet decomposition

Based on wavelet transform, wavelet packet decomposition (Rioul & Vetterli, 1991) can decompose the acceleration signal into low-frequency (H) and high-frequency (G) components. The low-frequency approximation signal and high-frequency detail signal can be obtained by using wavelet packet decomposition. The acceleration signals of different activities have different frequency distributions (Bouten, Koekkoek, Verduin, Kodde, & Janssen, 1997). For example, the frequency of activity ‘walk’ is almost in the range of 1-6 Hz, and the frequency of activity ‘jogging’ is almost in the range of 1.7-16 Hz. The wavelet packet decomposition can better reflect the different characteristics of the high-frequency and the low-frequency part of the acceleration signal under different activities. The three-layer wavelet packet decomposition process is shown in Figure 1.

The specific process of wavelet packet decomposition is as follows: given the orthogonal scaling function \( \phi(t) \) and the wavelet function \( \psi(t) \), the two scale relationship is:

\[
\begin{align*}
\phi(t) &= \sqrt{2} \sum_{k} h_{0k} \phi(2t - k) \\
\psi(t) &= \sqrt{2} \sum_{k} h_{1k} \phi(2t - k)
\end{align*}
\]  

(1)

Where \( h_{0k} \) and \( h_{1k} \) are filter coefficients in multiresolution analysis. To further generalize the two-scale equation, define the following recursive relations:

\[
\begin{align*}
w_{2n}(t) &= \sqrt{2} \sum_{k \in \mathbb{Z}} h_{0k} w_{2n-1}(2t - k) \\
w_{2n+1}(t) &= \sqrt{2} \sum_{k \in \mathbb{Z}} h_{1k} w_{2n-1}(2t - k)
\end{align*}
\]  

(2)

Figure 1. The sketch map of three-layer wavelet packet decomposition.
When \( n = 0, w_0(t) = \phi(t), w_1(t) = \varphi(t) \). The set of functions defined above \( \{w_n(t)\}_{n \in \mathbb{Z}} \) is the wavelet packet determined by \( w_0(t) = \phi(t) \). The wavelet packet \( \{w_n(t)\}_{n \in \mathbb{Z}} \) is a set of functions including the scaling function \( w_0(t) \) and the wavelet function \( w_1(t) \). Therefore, the wavelet packet coefficient recursion formula is:

\[
\begin{cases}
d^{i+1,2n}_k = \sum_l h_0(2l-k)d^{i,n}_l \\
d^{i+1,2n+1}_k = \sum_l h_1(2l-k)d^{i,n}_l
\end{cases}
\]  

The reconstruction formula of the wavelet packet is:

\[
d^{i,n}_j = \sum_k [h_0(l-2k)d^{i+1,2n}_{k} + h_1(l-2k)d^{i+1,2n+1}_{k}]
\]

\[
= \sum_k g_0(l-2k)d^{i+1,2n}_{k} + \sum_k g_1(l-2k)d^{i+1,2n+1}_{k}
\]  

(4)

Where \( d^{i+1,2n}_{k} \) and \( d^{i+1,2n+1}_{k} \) are the \( k \)th decomposition sequences obtained by decomposing the original signal through the \( j \)-layer wavelet packet; \( h_0, h_1, g_0, g_1 \) are low-pass and high-pass filter coefficients, respectively.

### 2.2. Wavelet energy spectrum features

The wavelet energy spectrum is the result of wavelet packet decomposition from the energy point of view. For different types of activities, the energy distribution of the acceleration signals in each frequency band will have a large diversity. Besides, the acceleration signal of human activity has the characteristics of non-linearity. So the wavelet energy spectrum can be used as the feature information and is suitable for distinguishing different activities. The relationship between the energy of the wavelet packet transform and the energy of the original signal is as follows:

\[
\int_{-\infty}^{+\infty} |f(x)|^2 dx = \sum |d(j,k)|^2
\]  

(5)

Therefore, the sum of squared signals in each frequency band after wavelet packet decomposition can be selected as the wavelet packet energy spectrum. The wavelet decomposition result is represented by \( d_{j,k}(k) \), and the energy of the signal in each frequency band is:

\[
E_{ij} = \sum_{k=1}^{N} |d_{ij}(k)|^2; j = 0, 1, \ldots, 2^l - 1
\]  

(6)

Where \( N \) represents the length of the original signal, all \( E_{i,j} \) constitutes the wavelet packet energy spectrum:

\[
E = [E_{i,0}, E_{i,1}, \ldots, E_{i,j}, \ldots, E_{i,2^l-1}]
\]  

(7)

After normalization, the feature vector is:

\[
E' = \frac{E}{\sum E}
\]  

(8)

Where

\[
\sum E = E_{i,0} + E_{i,1} + \ldots + E_{i,j} + \ldots + E_{i,2^l-1}.
\]

The number of layers of the wavelet packet decomposition is directly related to the number of sub-bands obtained and the dimension of the eigenvector. Through a large number of experimental tests, we found that when the three-layer wavelet packet of the acceleration signal is decomposed, the energy distribution of each sub-band can effectively distinguish the activities. So in this paper, the dney wavelet function is used to perform three-layer wavelet packet decomposition on the \( x, y, z \) triaxial acceleration signals. Since the sampling frequency of our input signal is 150 Hz, according to the sampling theorem, the frequency at the root node of the wavelet packet decomposition is set to 0-64 Hz. According to this, the frequency distribution of the first band of the third layer is 0-8 Hz, and the frequency of the subsequent band is sequentially incremented by 8 Hz. A total of 24 dimensional energy spectrum features are obtained. The distribution of the wavelet packet energy spectrum under six randomly selected activity samples is shown in Figure 2.

It can be seen from Figure 2 that the eigenvectors formed by the wavelet energy spectrum can clearly distinguish the six activities. Among them, the energy distribution of activity ‘walking’ is more gradual than the other activities and the activity ‘ascending stairs’ shows the most obvious fluctuation of the energy in the 24 frequency bands. The activity ‘jump’ has lower energy values in the first 8 bands than other activities and the activity ‘descending stairs’ is better distinguished from other activities in the last 6 bands. The overall trend of the energy distribution of the two running activities is similar, however, there is a significant difference in the energy amplitude in some frequency bands.

### 2.3. Feature sets

In addition to wavelet energy spectrum features mentioned above, features of other attributes which are proved effective for activity recognition in the previously published researches, such as mean, variance and AR-coefficients have also been extracted. All these features have been listed in Table 1. The proposed feature is marked in bold.
Table 1. Selected features.

| Feature No. | Feature | Feature ID |
|-------------|---------|------------|
| 1           | x, y, z acceleration mean (Chernbumroong et al., 2015) | 1–3 |
| 2           | x, y, z acceleration root mean square (Chernbumroong et al., 2015) | 4–6 |
| 3           | Correlation between axes (Chernbumroong et al., 2015) | 7–9 |
| 4           | x, y, z quartile difference (Chernbumroong et al., 2015) | 10–12 |
| 5           | x, y, z acceleration variance (Wang, Guo, et al., 2016) | 13–15 |
| 6           | Signal magnitude vector (SMV) (Cheng & Jhan, 2013) | 16 |
| 7           | Signal magnitude area (SMA) (Khan et al., 2010) | 17 |
| 8           | x, y, z acceleration kurtosis (Wang, Guo, et al., 2016) | 18–20 |
| 9           | x, y, z acceleration skewness (Wang, Guo, et al., 2016) | 21–23 |
| 10          | AR-coefficients (Wang, Chen, et al., 2016) | 24–35 |
| 11          | x, y, z wavelet energy spectrum* | 36–59 |
| Total       |         | 59 |

3. The proposed ensemble-based filter feature selection (EFFS) method

Ensemble feature selection is a new technique proposed to optimize the feature set (Solorio-Fernández, Carrasco-Ochoa, & Martínez-Trinidad, 2016). The superiority of ensemble feature selection has been demonstrated by many research fields, such as molecular signatures (Haury, Gestraud, & Vert, 2011), cancer diagnosis (Abeel et al., 2010) and load forecasting (Hua, Bao, Xiong, & Chiong, 2015). To our best knowledge, there are currently no reported researches applying ensemble feature selection to human activity recognition. Therefore, in this paper, EFPS method is proposed and utilized for feature selection in activity recognition. This section first introduces the four filter feature selection methods that we used in the EFPS method.

3.1. Information gain

Information gain (Agarwal & Mittal, 2013) as a classical filter feature selection method can determine the relevant attributes of each feature from a set of feature sets. This method ranks and selects the features based on information theory to reduce the feature dimensions of the machine learning algorithm. By measuring the entropy of the distribution, the features’ uncertainty are taken into consideration. Then, it ranks the features based on the relationship between feature and class. The inadequacy of the method is that it ignores the uniformity of the distribution of feature items between classes. The entropy of variable $X$ can be defined as:

$$H(X) = -\sum_i P(x_i)\log_2(P(x_i))$$

Let $P(x_i)$ denote the prior probability value of $X$. After obtaining the value of another variable $Y$, the entropy of $X$ is defined as:

$$H(X/Y) = -\sum_j P(y_j)\sum_i P(x_i|y_j)\log_2(P(x_i|y_j))$$

In the above equation, when $Y$ is known, the posterior probability of $X$ and $Y$ can be expressed as $P(x_i|y_j)$. The information gain can be defined as the reduction of $X$ entropy due to the additional information $Y$ brings to $X$, which is defined as:

$$IG(X|Y) = H(X) - H(X|Y)$$

Based on the above theoretical analysis, Equation (11) can be used to calculate the feature relevance ranking, which will be utilized to select the robust features with better distinguishing performance.
3.2. Gain ratio

Gain ratio (Ibrahim, Badr, & Shaheen, 2012) is an improvement of the information gain method. This method is a supplement to information gain to cope with the characteristics of features with large diversity. This kind of deviation may reduce the generalization ability of learning algorithm. The gain ratio improves information gain by determining the feature through using the number and size of branches and taking intrinsic information from the feature distribution entropy. The following formula can be used to calculate the gain ratio for a given feature x and y.

\[
\text{Gain Ratio}(y,x) = \frac{\text{Information Gain}(y,x)}{\text{Intrinsic Value}(x)}
\]

(12)

Where Intrinsic Value(x) = \displaystyle - \sum_{\bar{S}} |S| \times \log_2 \frac{|S|}{|\bar{S}|}, |S| is the number of possible values of feature x, and |\bar{S}| is the number of actual values of feature x.

3.3. Chi-squared statistic

The Chi-squared (\( \chi^2 \)) statistic (Nissim, Moskovitch, Rokach, & Elovici, 2012) measures the independence of two variables by calculating the Chi-squared score. The independence of the feature item relative to the one activity can be judged by Chi-squared. A low value of \( \chi^2 \) indicates that the features and class are independent whereas a large value indicates that there is a high degree of dependency between the features and the class. The Chi-squared statistic has the advantage of small computational complexity. However, this method has the negative correlation between the frequency of occurrence of features and the category, which may affect the feature selection. Chi-square statistic can be defined as:

\[
\chi^2(r,c_i) = \frac{N[P(r,c_i)P(\tilde{r},\tilde{c_i}) - P(r,\tilde{c_i})P(\tilde{r},c_i)]^2}{P(r)P(\tilde{r})P(c_i)P(\tilde{c_i})}
\]

(13)

Where N represents the entire dataset, r represents the presence of the feature and c_i represents the activity category. P(r, c_i) indicates the probability that r occurs in the c_i and P(\tilde{r}, c_i) indicates the probability that r does not appear in the c_i. Also, P(r, \tilde{c_i}) and P(\tilde{r}, \tilde{c_i}) respectively indicate the probability that a feature occurs and does not occur in an activity category that is not marked as c_i. P(r) represents the probability of a feature appearing in the dataset and P(\tilde{r}) is the probability that the feature does not appear in the dataset. P(c_i) is the probability that the dataset is marked as c_i.

3.4. ReliefF

ReliefF (Kononenko, 1994) is a multi-category feature selection method proposed by Kononenko in 1994 based on Kira’s work. ReliefF is mainly used to solve problems such as multi-classification, data loss and data noise (Reyes, Morell, & Ventura, 2015). ReliefF can handle incomplete data and reduce the effects of noise in the data. However, its random selection of sample strategies affects the features that have a significant effect on the separation of small samples. The main point of the ReliefF is to perform feature evaluation based on the distinguishing ability of features on close-range samples. The high-quality features should make the same kind of samples close to each other, and the different types of samples are far away from each other. The ReliefF updates the weight of the feature using equation (14):

\[
W[i] = W[i] - \sum_{j=1}^{k} \frac{\text{diff}(i, R_s, H_j)/(mk)}{mk}
\]

\[
+ \sum_{C \neq \text{class}(R_s)} \frac{p(C)}{1-p(\text{class}(R_s))} \sum_{j=1}^{k} \text{diff}(i, R_s, M_j(C))
\]

(14)

Where W[i] is the weight of feature i, R_s are the samples randomly selected in the training set, H_j is the jth sample of the k nearest neighbour samples that is the same kind as H_j, and M_j is the jth sample of the k nearest neighbour samples that are heterogeneous with R_s. P(c) is the probability of occurrence of a class C sample.

3.5. The framework of proposed EFFS method

The EFFS method is a data preprocessing stage before implementing the machine learning algorithm. Our proposed EFFS method is based on the feature ranking obtained by the above-mentioned filter method. It utilizes a single-party filter method to perform initial selection of the original feature space, this includes that the four feature selection method mentioned above are utilized to obtain the results of feature ranking according to the degree of feature importance. The ranking result of each filter method can be expressed as:

\[
\begin{align*}
AIG &= \{X(1), X(2), \ldots, X(k), \ldots, X(m)\} \\
AGR &= \{XG(1), XG(2), \ldots, XG(k), \ldots, XG(m)\} \\
ACS &= \{XC(1), XC(2), \ldots, XC(k), \ldots, XC(m)\} \\
A\text{ReliefF} &= \{XR(1), XR(2), \ldots, XR(k), \ldots, XR(m)\}
\end{align*}
\]

(15)

Where X(i,k), XG(i,k), XC(i,k) and XR(i,k) represents the ranking number of the kth feature in these four methods, respectively. AIG, AGR, ACS and AReliefF represents the ordering of features by the four feature selection methods, respectively.

Based on the feature ranking obtained by the above four filter-based feature selection methods, EFFS method
utilizes the ‘linear summation’ rule to combines the output of each filter method. The ranks of each feature in the four filter feature selection methods will be added and EFFS utilizes the sum of ranking obtained as an index to measure the importance of the features. Each feature is sorted according to the new index from the smallest to the largest, so as to obtain the ensemble ranking of features under the four methods, which can be expressed as:

\[ A = [X(1), X(2), \cdots, X(k), \cdots, X(m)] \] (16)

Where

\[ X(k) = \alpha X_I(k) + \beta X_G(k) + \gamma X_CS(k) + \epsilon X_R(k) \]

In this formula, \( \alpha, \beta, \gamma, \epsilon \) are the weights corresponding to the four methods. These weights in the formula are obtained from multiple experiments to obtain a better weight combination of the four feature selection methods. Different weight combinations including (2, 1, 2, 1), (5, 3, 2, 3) and (2, 5, 4, 3) were tested by experiments. The experimental results of these three weight combinations are detailed in Section 5.3. \( X(k) \) is the index used to sort features by the EFFS method, and the features are ranked according to the value of \( X(k) \) from small to large. Figure 3 shows the specific flow of the proposed EFFS method.

4. Experimental equipment and data

The experimental dataset has been acquired by using the TRIGNO EMG-acceleration signal sensing device with an acceleration range of ±6g and a sampling frequency of 150Hz. The data acquisition platform of the device is shown in Figure 4(a). Since the data collection node, as shown in Figure 4(b) has wireless transmission function, the signal can be transmitted to the data acquisition platform in real time. The whole acquisition process is performed under wireless conditions, as shown in Figure 4(c). Five healthy students took part in the experiment and they were asked to perform the activities shown in the second column of Table 2 in sequence. Each volunteer was asked to tie the data collection node to the specified standard position (as shown at the top of Figure 5) and the simulated deviation position (shown at the bottom of Figure 5). The samples acquisition under the specified standard position are utilized for training and the samples acquisition under the simulated deviation position are utilized for testing the robustness of the proposed method to the sensor position deviation. So we can improve the reliability of the recognition system to the sensor deviation from the feature selection aspect. The right side of Figure 5 shows the acceleration signals for a set of ‘walking’ activity collected by the sensor in both positions. After pre-processing of the collected data, we utilized the sliding window technique to split the activity

| Number | Activity         | Feature Samples |
|--------|------------------|-----------------|
| 1      | Sitting (ST)     | 1380            |
| 2      | Jumping (JP)     | 1380            |
| 3      | Walking (W)      | 1385            |
| 4      | Standing (SD)    | 1985            |
| 5      | Lying (LY)       | 2005            |
| 6      | Running (7km/h) (R7) | 2225         |
| 7      | Running (10km/h) (R10) | 2310          |
| 8      | Ascending stairs (AS) | 1590         |
| 9      | Descending stairs (DS) | 1535         |

Table 2. The activities collected in the experiment.

Figure 4. Human activity data acquisition platform based on acceleration sensor: (a) the data acquisition platform, (b) data collection node containing the triaxial accelerometer, (c) experimental data acquisition process.
signal for feature extraction. The sliding window with 50% overlap is chosen and the window length is 300 samples. The features of the each window are extracted as shown in Table 1 and then the features are normalized. Finally, under the standard position and the deviation position, the number of feature samples for each action is shown in the third column of Table 2.

5. Experimental results and analysis

Since data was acquired from five subjects, the leave one out (LOO) strategy was utilized to train and recognize activities. The feature samples collected under the standard position of the four volunteers were used as training data and the feature samples under the deviation position of another volunteer were used as test data. The experiment was repeated five times until the deviation data of five volunteers were tested and the final results were the average of five experimental values.

5.1. The influence of wavelet energy spectrum features on recognition

Feature information with higher discrimination has great benefits for the performance of the classifier. Figure 6 shows the effect of the wavelet energy spectrum features of the two classifiers on the overall accuracy of the recognition.

Tables 3 and 4 show the effects of wavelet energy spectrum feature on the accuracy of each activity using SVM and KNN classifiers, respectively. It can be observed from Tables 3 and 4 that between these two classifiers, the improvement of wavelet energy spectrum features for different activities is different. The accuracy of KNN classifier for the activities ‘Ascending stairs’, ‘Descending stairs’, ‘Running (7km/h)’ and ‘Running (10km/h)’ are significantly increased. The accuracy of SVM classifier for the activities ‘Ascending stairs’, ‘Descending stairs’, ‘Jumping’, ‘Sitting’, ‘Running (7km/h)’ and ‘Running (10km/h)’ have increased by more than 2%.
Table 3. Comparison of recognition accuracy of SVM classifier using the feature set with and without the proposed features.

| Activity | Feature set without proposed features | Feature set with proposed features |
|----------|----------------------------------------|-----------------------------------|
| ST       | 98.91%                                 | 99.63%                            |
| JP       | 83.70%                                 | 87.68%                            |
| W        | 92.78%                                 | 94.22%                            |
| SD       | 93.45%                                 | 95.47%                            |
| LY       | 88.28%                                 | 89.28%                            |
| R7       | 62.47%                                 | 68.76%                            |
| R10      | 67.53%                                 | 69.70%                            |
| AS       | 79.56%                                 | 86.16%                            |
| DS       | 82.08%                                 | 87.29%                            |
| Total    | 81.77%                                 | 85.01%                            |

Table 4. Comparison of recognition accuracy of KNN classifier using the feature set with and without the proposed features.

| Activity | Feature set without proposed features | Feature set with proposed features |
|----------|----------------------------------------|-----------------------------------|
| ST       | 98.55%                                 | 98.91%                            |
| JP       | 81.16%                                 | 82.61%                            |
| W        | 93.86%                                 | 94.58%                            |
| SD       | 92.69%                                 | 92.95%                            |
| LY       | 90.27%                                 | 90.27%                            |
| R7       | 63.37%                                 | 67.87%                            |
| R10      | 65.15%                                 | 67.97%                            |
| AS       | 79.87%                                 | 82.08%                            |
| DS       | 80.13%                                 | 82.41%                            |
| Total    | 81.32%                                 | 83.06%                            |

Additionally, in order to gain a better insight into the effect of wavelet energy spectrum on distinguishing different activities, Tables 5–8 show the confusion matrix of activity recognition using KNN and SVM classifier with feature set of wavelet energy spectrum features or without wavelet energy spectrum features, which can demonstrate the effectiveness of wavelet energy spectrum features in activity recognition.

It can be seen from Tables 5–6 that the KNN classifier has many misrecognitions of some similar activities when the wavelet energy spectrum features are not considered. For example, activity ‘Running (7km/h)’ is misrecognized as ‘Running (10km/h)’, activity ‘jumping’ is misrecognized as ‘Ascending stairs’ and ‘Running (10km/h)’. After applying the wavelet energy spectrum features, the discrimination of activity is increased. For example, the discriminations between activities ‘jump’ and ‘Ascending stairs’, activity ‘Running (10km/h)’ and ‘Running (7km/h)’ are increased. It can also be seen from Tables 7–8 that after application of wavelet energy spectrum features, it is easier to distinguish the activity ‘jump’ and ‘Running (7km/h)’, activity ‘Ascending stairs’ and ‘Descending stairs’.

5.2. Comparison of feature selection methods

In the previous section, we analyzed the effect of wavelet energy spectrum features on activity recognition performance. Based on this, we further optimized the feature set with robustness to the deviation position by feature selection method to improve recognition performance. First, all the features in Table 1 are extracted as data samples for the data set under the deviation position. Considering the randomness of the samples, the feature ranking results of each feature selection method may be not representative. Therefore, in this paper, each feature selection results of each feature selection method is run five times and the feature ranking result is the average of the five experiments. For convenience of expression, the ranking results of the four feature selection methods are arranged according to the feature IDs in Table 1, as shown in Figure 7.

As can be seen from Figure 7, some features are ranked higher in the selected four feature selection methods. For example, feature ID7 ranks high in the gain ratio, chi-square statistics and ReliefF method and feature ID56 ranks high among the four feature selection methods. This indicates that these features have a higher degree of activity discrimination ability. Similarly, there are some features that are ranked behind by some feature selection methods. For example, feature ID27 ranks lower in the information gain, chi-square statistic, and ReliefF methods, feature ID14 ranks lower in the information gain, gain ratio, and ReliefF methods. This indicates that these features are considered to be less distinguishing or redundant features by some feature selection methods.

According to the results of the first three feature selection methods, some of the proposed 24 new features (ID 36-59) are ranked after 20. Compared with the first three feature selection methods, less proposed features...
Table 6. Confusion matrix for activity recognition using KNN and wavelet energy spectrum features.

|   | ST  | JP  | W   | SD  | LY  | R7  | R10 | AS  | DS  |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ST | 273 | 0   | 2   | 1   | 0   | 0   | 0   | 0   | 0   |
| JP | 0   | 228 | 0   | 0   | 0   | 4   | 11  | 31  | 2   |
| W  | 1   | 0   | 262 | 5   | 0   | 6   | 2   | 1   | 0   |
| SD | 3   | 0   | 10  | 369 | 1   | 9   | 2   | 2   | 1   |
| LY | 0   | 0   | 0   | 2   | 362 | 3   | 22  | 0   | 12  |
| R7 | 0   | 3   | 3   | 12  | 4   | 302 | 118 | 2   | 1   |
| R10| 0   | 7   | 0   | 0   | 18  | 121 | 314 | 0   | 2   |
| AS | 0   | 21  | 0   | 0   | 1   | 0   | 261 | 35  |     |
| DS | 0   | 2   | 0   | 2   | 15  | 0   | 0   | 35  | 253 |

Table 7. Confusion matrix for activity recognition using SVM and feature set without wavelet energy spectrum.

|   | ST  | JP  | W   | SD  | LY  | R7  | R10 | AS  | DS  |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ST | 273 | 0   | 1   | 1   | 1   | 0   | 0   | 0   | 0   |
| JP | 0   | 231 | 0   | 1   | 0   | 16  | 8   | 18  | 2   |
| W  | 1   | 0   | 259 | 2   | 5   | 9   | 0   | 0   | 1   |
| SD | 1   | 0   | 8   | 371 | 3   | 6   | 3   | 3   | 2   |
| LY | 0   | 1   | 6   | 2   | 354 | 8   | 15  | 4   | 11  |
| R7 | 0   | 12  | 10  | 6   | 20  | 278 | 113 | 4   | 2   |
| R10| 0   | 11  | 1   | 2   | 15  | 118 | 312 | 2   | 1   |
| AS | 1   | 22  | 1   | 5   | 5   | 8   | 5   | 253 | 18  |
| DS | 0   | 5   | 3   | 4   | 15  | 3   | 2   | 23  | 252 |

Table 8. Confusion matrix for activity recognition using SVM and wavelet energy spectrum features.

|   | ST  | JP  | W   | SD  | LY  | R7  | R10 | AS  | DS  |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ST | 275 | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   |
| JP | 0   | 242 | 0   | 1   | 0   | 11  | 5   | 15  | 2   |
| W  | 0   | 0   | 261 | 2   | 4   | 8   | 0   | 1   | 1   |
| SD | 0   | 0   | 4   | 379 | 2   | 6   | 3   | 2   | 1   |
| LY | 0   | 1   | 5   | 1   | 358 | 9   | 12  | 4   | 11  |
| R7 | 0   | 9   | 6   | 12  | 4   | 306 | 105 | 2   | 1   |
| R10| 0   | 4   | 0   | 0   | 15  | 120 | 322 | 0   | 1   |
| AS | 1   | 11  | 0   | 2   | 3   | 2   | 1   | 274 | 24  |
| DS | 0   | 1   | 1   | 1   | 14  | 0   | 3   | 19  | 268 |

Figure 7. The ranking of the selected features using the four feature selection methods.
are ranked after 20 by ReileF. This indicates that different evaluation criteria have different rankings for the proposed features. Some of the proposed features show poor ability to distinguish in the first three methods. However, more proposed features show better ability to distinguish in ReliefF. Therefore, the feature rankings obtained by the four feature selection methods are used as a reference for EFFS method. This enables us to eliminate the features that are irrelevant or redundant.

In order to verify the EFFS method experimentally, several comparative experiments are performed. The EFFS is compared with four traditional filter feature selection methods. In addition to the feature selection approaches, we also test the EFFS method under the classifiers of different classification principles. And the SVM and KNN classifiers are utilized to measure the effectiveness of the proposed feature selection method. Figures 8 and 9 show the recognition accuracy of the proposed EFFS and four comparison feature selection methods using KNN and SVM classifiers respectively with feature selection dimensions of 1–50 dimensions. In Figures 8 and 9, the number x on the X-axis refers to the top x selected features with descending order of importance using different methods. The parameters of the EFFS $\alpha$, $\beta$, $\gamma$, and $\epsilon$ are set as 2, 1, 2, and 1 respectively. This combination of parameters was chosen randomly in the experiment.

The purpose of feature selection is to replace the redundant features with features that are as distinguishing as possible. Too many features may contain some redundant features, which does not have much significance for improving the accuracy of the classification algorithm and may increase the computational complexity of the algorithm. Therefore, selecting the appropriate number of features can greatly improve the computational efficiency of the classification algorithm. However, few features tend to reduce the recognition accuracy of the algorithm to a great extent. It can be seen from Figures 8–9 that few features cannot guarantee high recognition accuracy. For example, no matter which method is adopted, if the feature set contains less than 8 features, both classifiers SVM and KNN achieve less than 65% accuracy. To a certain extent, as the number of selected features increases, the recognition performance gradually gets better. The accuracy can reach 80.4% by using EFFS with 24 features when using KNN and can reach 81.7% by using EFFS with 20 features when using SVM. It can also be observed from Figures 8–9 that the proposed EFFS outperforms other four filter-based feature selection methods. The method proposed in this paper can obtain higher accuracy under the same feature dimension, which can show its superiority.

### 5.3. The performance of EFFS under different parameters

In the proposed method, the weights $\alpha$, $\beta$, $\gamma$ and $\epsilon$ are used to coordinate the relationship between the four independent feature selection methods in the EFFS method. Since the ranking criteria of each feature selection method is different, weight configurations can reflect the diversity effects of the EFFS. To better illustrate the EFFS method, we utilized SVM and KNN to evaluate the effect of the EFFS method on the recognition accuracy against different weight configurations. Three sets of parameter combinations were randomly selected for the experiment. Figures 10 and 11 respectively shows the recognition accuracy of SVM and KNN against three weight configurations of $\alpha$, $\beta$, $\gamma$ and $\epsilon$ using the proposed EFFS. RelieF is chosen as the benchmark method for comparison.

It can be seen from Figures 10–11 that when the respective weights of the proposed method are (2, 1, 2, 1), the accuracy of the classifier can achieves its highest value. This may be due to the fact that this set of weights can balance the four filter-based feature selection methods and obtain the most distinguishing feature set. It can also be found from Figures 10–11 that the accuracies of the classifiers are different against different
weighted configurations, but the overall average level is higher than the results obtained by the baseline method ReliefF, which again proves the validity of the proposed EFFS method in this paper.

6. Conclusions

In order to reduce the influence of minor deviation from the standard position of accelerometer on HAR, the wavelet energy spectrum features and EFFS method have been proposed in this paper. We utilized two kinds of classifier KNN and SVM to test the feature set with and without our proposed features. The experiments show that compared with the feature set without introduced features, the feature set with wavelet energy spectrum features can improve the discrimination of activities, which is beneficial to improve the accuracy of classifier. Several comparison experiments with four filter-based feature selection methods are employed to verify the effectiveness of the proposed EFFS method in human activity recognition. The results show that the proposed method is more robust and can get better recognition performance compared with other approaches under the same number of features. In future work, we will conduct experiments using the proposed EFFS method to analyze the performance of wavelet energy spectrum features in HAR. Besides, we will utilize data sets from other body positions such as chest or ankle and more classifiers such as extreme learning machine or deep learning methods to verify the significance of wavelet energy spectrum features and the proposed feature selection method in HAR.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Natural Science Foundation of China [grant numbers 61703134 and 61803143], National Key Technology Research and Development Program of the Ministry of Science and Technology of China [grant number 2015BAI06B03], the Natural Science Foundation of Tianjin [grant number 17JCQNJQ04400] and the Natural Science Foundation of Hebei Province [grant number F2019202369].

References

Abeel, T., Helleputte, T., Van de Peer, Y., Dupont, P., & Saesys, Y. (2010). Robust biomarker identification for cancer diagnosis with ensemble feature selection methods. Bioinformatics, 26, 392–398.

Agarwal, B., & Mittal, N. (2013). Optimal feature selection for sentiment analysis. Springer Berlin Heidelberg, 7817, 13–24.

Apolloni, J., Leguizamón, G., & Alba, E. (2016). Two hybrid wrapper-filter feature selection algorithms applied to high-dimensional microarray experiments. Applied Soft Computing, 38, 922–932.

Bouten, C. V. C., Koekkoek, K. T. M., Verduin, M., Kodde, R., & Janssen, J. D. (1997). A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. IEEE Transactions on Biomedical Engineering, 44(3), 136–147.

Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. Computers and Electrical Engineering, 40(1), 16–28.

Chen, Z., Zhu, Q., Chai, S. Y., & Zhang, L. (2017). Robust human activity recognition using smartphone sensors via CT-PCA and online SVM. IEEE Transactions on Industrial Informatics, 13(6), 3070–3080.

Cheng, W.-C., & Jhan, D.-M. (2013). Triaxial accelerometer-based fall detection method using a self-constructing cascade-adaboost-SVM classifier. IEEE Journal of Biomedical & Health Informatics, 17(2), 411–419.

Cherubumroong, S., Shuang, C., & Yu, H. (2015). Genetic algorithm-based classifiers fusion for multisensor activity recognition of elderly people. IEEE Journal of Biomedical and Health Informatics, 19, 282–289.

Cornachia, M., Ozcan, K., Zheng, Y., & Velipasalar, S. (2017). A Survey on activity detection and classification using wearable sensors. IEEE Sensors Journal, 17, 386–403.
Gibson, R. M., Amira, A., Ramzan, N., Casaseca-De-La-Higuera, P., & Pervuz, Z. (2016). Multiple comparator classifier framework for accelerometer-based fall detection and diagnostic. Applied Soft Computing, 39, 94–103.

Gravina, R., Alinia, P., Ghasemzadeh, H., & Fortino, G. (2017). Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges. Information Fusion, 35, 68–80.

Haury, A. C., Gestraud, P., & Vert, J. P. (2011). The influence of feature selection methods on accuracy, stability and interpretability of molecular signatures. PloS ONE, 6(6), e28210.

Hua, Z., Bao, Y., Xiong, T., & Chiong, R. (2015). Hybrid filter–wrapper feature selection for short-term load forecasting. Engineering Applications of Artificial Intelligence, 40, 17–27.

Huynh-The, T., Hua, C.-H., Tu, N. A., Hur, T., Bang, J., Kim, D., … Lee, S. (2018). Hierarchical topic modeling with pose-transition feature for action recognition using 3D skeleton data. Information Sciences, 444, 20–35.

Ibrahim, H., Badr, S., & Shaheen, M. (2012). Adaptive layered approach using machine learning techniques with gain ratio for intrusion detection systems. International Journal of Computer Applications, 56(7), 10–16.

Javed, A., Larijani, H., Ahmadinia, A., & Gibson, D. (2017). Smart random neural network controller for hvac using cloud computing technology. IEEE Transactions on Industry Applications, 13(1), 351–360.

Khan, A. M., Lee, Y.-K., Lee, S. Y., & Kim, T.-S. (2010). A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. IEEE Transactions on Information Technology in Biomedicine, 14(5), 1166–1172.

Kononenko, I. (1994). Estimating attributes: Analysis and extensions of RELIEF(C). European Conference on Machine Learning ECML 1994: Machine Learning: ECML-94. pp. 171–182.

Kwolek, B., & Kepski, M. (2015). Improving fall detection by the use of depth sensor and accelerometer. Neurocomputing, 168, 637–645.

Lai, C. F., Chang, S. Y., Chao, H. C., & Huang, Y. M. (2011). Detection of cognitive injured body region using multiple triaxial accelerometers for elderly falling. IEEE Sensors Journal, 11(3), 763–770.

Lai, X., Liu, Q., Wei, X., Wang, W., Zhou, G., & Han, G. (2013). A survey of body sensor networks. Sensors, 13(5), 5406.

Li, W. F., & Yao, B. F. (2016). Recognition of human activities based on single triaxial accelerometer. Information Technology in Biomedicine, IEEE Transactions on, 44(4), 58–62.

Liu, S., Gao, R. X., John, D., Staudemayer, J. W., & Freedson, P. S. (2012). Multisensor data fusion for physical activity assessment. IEEE Transactions on Biomedical Engineering, 59(3), 687–696.

Ma, X., Wang, H., Xue, B., Zhou, M., Ji, B., & Li, Y. B. (2014). Depth-based human fall detection via shape features and improved extreme learning machine. IEEE Journal of Biomedical & Health Informatics, 18(6), 1915–1922.

Margarito, J., Helouai, R., Bianchi, A. M., Sartor, F., & Bonomi, A. G. (2016). User-independent recognition of sports activities from a single wrist-worn accelerometer: A template-matching-based approach. IEEE Transactions on Biomedical Engineering, 63(4), 788–796.

Mo, L., Liu, S., Gao, R. X., & Freedson, P. S. (2013). Energy-efficient and data synchronized body sensor network for physical activity measurement. IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 1120–1124.

Mooney, R., Corley, G., Godfrey, A., Quinlan, L. R., & Ó Laighin, G. (2016). Inertial sensor technology for elite swimming performance analysis: A systematic review. Sensors, 16(1), 18.

Nissim, N., Moskovitch, R., Rokach, L., & Elovich, Y. (2012). Detecting unknown computer worm activity via support vector machines and active learning. Pattern Analysis and Applications, 15(4), 459–475.

Patidar, S., Pachori, R. B., & Acharya, U. R. (2015). Automated diagnosis of coronary artery disease using tunable-Q wavelet transform applied on heart rate signals. Knowledge-Based Systems, 82, 1–10.

Reyes, O., Morell, C., & Ventura, S. (2015). Scalable extensions of the Relieff algorithm for weighting and selecting features on the multi-label learning context. Neurocomputing, 161(C), 168–182.

Riou, O., & Vetterli, M. (1991). Wavelets and signal processing. IEEE Signal Processing Magazine, 8(4), 14–38.

Rodriguez-Martín, D., Samà, A., Pérez López, C., Cabestany, J., Catalá, A., & Rodríguez-Molina, A. (2015). Posture transition identification on pd patients through a SVM-based technique and a single waist-worn accelerometer. Neurocomputing, 164(C), 144–153.

Shoaib, M., Bosch, S., Incel, O. D., Scholten, H., & Havinga, P. J. (2015). A survey of online activity recognition using mobile phones. Sensors, 15(11), 2059–2085.

Shriram, R., Sundhararajan, M., Shete, S., & Daimiwal, N. (2017). Statistical features based comparison of analysis and synthesis of normal and epileptic electroencephalogram for various wavelets. Turkish Journal of Electrical Engineering and Computer Sciences, 25(3), 1795–1806.

Solorio-Fernández, S., Carrasco-Ochoa, J. A., & Martínez-Trinidad, J. F. (2016). A new hybrid filter-wraper feature selection method for clustering based on ranking. Neurocomputing, 214, 866–880.

Song, B., Kamal, A. T., Soto, C., Ding, C., Farrell, J. A., & Roy-Chowdhury, A. K. (2010). Tracking and activity recognition through consensus in distributed camera network. IEEE Transactions on Image Processing, 19(10), 2564–2579.

Tachibana, A., Noah, J. A., Bronner, S., Ono, Y., & Onozuka, M. (2011). Parietal and temporal activity during a multimodal dance video game: An fNIRS study. Neuroscience Letters, 503(2), 125–130.

Vergara, J. R., & Estévez, P. A. (2014). A review of feature selection methods based on mutual information. Neural Computing and Applications, 24(1), 175–186.

Wang, A. G., Chen, G. L., Yang, J., Zhao, S. H., & Chang, C. Y. (2016). A comparative study on human activity recognition using inertial sensors in a smartphone. IEEE Sensors Journal, 16(11), 4566–4578.

Wang, Z. L., Guo, M., & Zhao, C. (2016). Badminton Stroke recognition based on body sensor Networks. IEEE Transactions on Human-Machine Systems, 46(5), C1–C1.

Wang, Z. L., Qiu, S., Cao, Z. K., & Jiang, M. (2013). Quantitative assessment of dual gait analysis based on inertial sensors with body sensor network. Sensor Review, 33(1), 48–56.

Wang, Z., Wu, D., Chen, J., Ghoneim, A., & Hossain, M. A. (2016). A triaxial accelerometer-based human activity recognition via
EEMD-based features and game-theory-based feature selection. *IEEE Sensors Journal*, 16(9), 3198–3207.

Yan, R., Gao, R. X., & Chen, X. (2014). Wavelets for fault diagnosis of rotary machines: A review with applications. *Signal Processing*, 96(5), 1–15.

Zhu, C., & Sheng, W. (2009). Human daily activity recognition in robot-assisted living using multi-sensor fusion. *IEEE International Conference on Robotics and Automation in: Robotics and Automation*, 2154–2159.