AIP Publishing AIP Conference Proceedings

Volume 3056, Issue 1
7 April 2025

SUSTAINABLE AND INTEGRATED ENGINEERING INTERNATIONAL CONFERENCE: SIE2022
12–13 December 2022
Langkawi, Malaysia

RESEARCH ARTICLE | APRIL 07 2025

Ensemble filter based feature selection technique for classification of human activity recognition

Norfadzlan Bin Yusup ; Azlan Mohd Zain

Author & Article Information
AIP Conf. Proc. 3056, 060006 (2025)
https://doi.org/10.1063/5.0208808

Through the advancement of wearable sensors, wireless communication, and machine learning techniques, Assistive Technologies (AT) which endorse autonomous, active, and healthy lifestyles are emerging in recent years. Among these advances, Human Activity Recognition (HAR) is one of the most innovative means to support or monitor human activities. However, misclassifications such as intra-class variation and inter-class overlap in similar activities degrade classification accuracy in HAR. To improve the recognition of daily human activities, handcrafted features of time-domain and frequency-domain are combined. However, several extracted features may not be significant in describing the activities. Therefore, this research aims to propose a feature selection technique for optimal human activities recognition. The methodology proposed for this research is the Ensemble Filter (Relief-F and mRMR) to select the most relevant and less redundant features. Although a filter feature ranking approach is commonly used in related studies, most works fail to consider the threshold limit to exclude unnecessary and redundant features. An ensemble Random Forest (RF) was used as the base classifier to evaluate the performance of the hybrid algorithm. The results demonstrate that the proposed ensemble filter selection was beneficial in reducing the total number of features while improving overall classification accuracy.

Topics
Machine learning, Wireless communications

REFERENCES

Ensemble Filter Based Feature Selection Technique for Classification of Human Activity Recognition

Norfadzlan bin Yusup[1,a)] and Azlan Mohd Zain[2,b)]

[1]Data Science Centre, Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, 94300, Kota Samarahan, Sarawak, Malaysia
[2]UTM Big Data Centre, Universiti Teknologi Malaysia, 81310, Johor Bahru, Johor, Malaysia

[a)]*corresponding author: ynorfadzlan@unimas.my*
[b)]azlanmz@utm.my

Abstract. Through the advancement of wearable sensors, wireless communication, and machine learning techniques, Assistive Technologies (AT) which endorse autonomous, active, and healthy lifestyles are emerging in recent years. Among these advances, Human Activity Recognition (HAR) is one of the most innovative means to support or monitor human activities. However, misclassifications such as intra-class variation and inter-class overlap in similar activities degrade classification accuracy in HAR. To improve the recognition of daily human activities, handcrafted features of time-domain and frequency-domain are combined. However, several extracted features may not be significant in describing the activities. Therefore, this research aims to propose a feature selection technique for optimal human activities recognition. The methodology proposed for this research is the Ensemble Filter (Relief-F and mRMR) to select the most relevant and less redundant features. Although a filter feature ranking approach is commonly used in related studies, most works fail to consider the threshold limit to exclude unnecessary and redundant features. An ensemble Random Forest (RF) was used as the base classifier to evaluate the performance of the hybrid algorithm. The results demonstrate that the proposed ensemble filter selection was beneficial in reducing the total number of features while improving overall classification accuracy.

INTRODUCTION

The Human Activity Recognition (HAR) is used to enhance the quality of life of individuals in many applications. The objective of HAR is to acknowledge the user's actions and behaviours based on the setting and the sensed movement of the user. In particular, inertial HAR approaches use the Inertial Measuring Unit (IMU) to capture the motion of the user at a particular sensor location such as the wrist, ankle, or foot. With the development of sensor technology for micro-machine mechanical systems (MEMs), inertial sensors such as accelerometer, gyroscope, and magnetometers are developed and tested to perform activity recognition. Human Activity Recognition (HAR) is a vast area of study that seeks to distinguish a person's specific gesture or action based on sensor data (Wang and Meng, 2018). This includes predicting a person's behaviour, which typically requires profound knowledge and techniques from signal processing to the required technical application from the raw data in order to conform to a machine learning model (Wang *et al.*, 2017). The identification of human activity is considered as a challenging classification task in time series data. The misclassification of daily human activities remains one of the challenges in HAR (Wang *et al.*, 2018). A major contributing factor to this issue is due to different people behave differently for the same human daily activities such as 'walking' or 'running'. In HAR, intra-class variation happens when one activity category involves different styles of human motion. For example, in 'walking' activity, a person can walk very fast or slow.

Furthermore, human activity similarities exist not only in different categories (intra-class variation) but also in the same categories (inter-class similarities). Inter-class similarities happened when the activity is physically different (e.g., walking downstairs and walking upstairs) but shows similar characteristics in sensor signal forms (Bulling *et al.*, 2014). These similarities would be challenging for intelligent machines to differentiate, thus contributes to the misclassifications.

The selection of features presents still another critical problem due to the fact that some features are less relevant than others and may not be helpful in capturing the activities. Time-domain features are often used to define static activities as a more generic, but useful one such as standing, and sitting (Zubair *et al.*, 2016; Arif *et al.*, 2017) Despite that, using these features alone to define the dynamic activities (e.g., walking and running) may not be precise (Shoaib *et al*., 2016). While the average accuracy obtained is good, an optimum number of features still have to be chosen to prevent dimensionality curses. Generally, selection strategies for features can be categorised into filter and wrappers. The filter feature ranking approach was employed due to their less complicated nature and is capable of handling a vast number of instances (Bolón-Canedo *et al.*, 2014). Most of the research work, however, only based on selecting top-*k* features and did not define the threshold which discards the low-ranking feature (Doewes *et al.*, 2017; Amezzane *et al.*, 2018; Chandra, 2018; Abo El-Maaty and Wassal, 2019). Clearly, this method is not optimized since it is likely to overestimate or degrade the pruned features. There may also be some feature redundancies and a possible increase in the false detection rate.

METHODOLOGY

In general, the proposed activity recognition is divided into four main stages: pre-processing stage, feature extraction stage, feature selection stage, and classification stage. In this paper, we focus on discussing the proposed ensemble feature selection technique of Relief-F and Maximum Relevancy Minimum Redundancy (mRMR). The Relief-F approach is used to measure the consistency of features utilizing a distance-related weight method of assessment. Since the Relief-F is straightforward and well-performed in analysing the feature subsets, it can generalise noisy data and incomplete data. If the missing information attribute remains continuous, the average function value is determined in the training set to replace the missing data. If not, the maximum frequency value is replaced in the training set. The top ranks would be used as an input in an indicator later. The nearest hit (data point of the same class), as well as the nearest miss (data point of a different class), is determined by using Equation (1) based on their weight of scoring. The original instances are randomly selected from the input data, and the functionality is then re-ranked according to its importance.

$$w_i = \sum_{j=1}^{N} (x_i^j - nearmiss(x^i)_j)^2 - (x_i^j - nearhit(x^i)_j)^2 \quad (1)$$

Where w is the weight of ith feature, x_i^j is the value of ith feature for point x, and N is the total number of data points. Near hit x^j and near-miss x^j are the nearest data point to x^j in the same and different class, respectively (Capela *et al.*, 2015). In this work, an optimum threshold value (τ) is added to decide the limit of the feature. The ranking features produced are evaluated, and the weight features below the defined threshold are pruned or eliminated. The features are then eliminated to generate the most important ranking feature in Equation (2)

$$w = \begin{cases} f & if\ w_i \geq \tau \\ prune & otherwise \end{cases} \quad (2)$$

Where f is the highly ranked i_{th} feature, w is the scoring weight of i_{th} feature, and n is the number of features. If w is greater or equal with the threshold r, the i_{th} features are selected; otherwise, i_{th} features are 'pruned'. Based on the experiment, the value of τ selected in this research is 0.05. The selected features are accumulated, and the ranking of non-eliminated feature subsets, R(f_i), is later served as an input to the mRMR algorithm to generate the optimal feature subsets further. The mRMR aims to find the most important features on the basis of their association with the class label and to minimise the redundancy of the features. This method of filtering shows the highest relevance and minimal redundancy features. Mutual information is used to measure both relevance and redundancies to estimate the mutual dependence of two variables (Alomari *et al.*, 2018). The concept of mutual information is given in Equation (3).

$$I(x_{i},c) = \sum P(x_{i},c)\log\frac{P(x_{i},c)}{P(x_i)P(c)} \tag{3}$$

Where (x_i,c) is the mutual information between feature x_i, and class label c. $P(x_i)$ and $P(c)$ are marginal probability functions and $P(x_i,c)$ is the joint probability distribution. The mutual information value for two completely independent variable random variable is 0. Given x_i, which represents the feature *i*, and the class label *c*, the maximum relevance method selects the best *m* features in descent order of (I,*c*) i.e the best *m* individual features that are relevant to the class labels. The mRMR algorithm aims to find an optimal set of features, *S*, for which relevancy is maximised and redundancy is minimized,

$$Relevancy = \max_{x} \frac{1}{|S|} \sum_{x_i \in S} I(x_i, c) \tag{4}$$

For the complete set of features *x*, the subset of *m* features that has maximal relevance criterion is the subset that satisfies the maximal mean value of all mutual information values between individual features x_i and class *c*. To eliminate the redundancy, the formula is given in Equation (5)

$$Redundancy = \min_{x} \frac{1}{|S|^2} \sum_{x_i,x_j \in S} I(x_i, x_j) \tag{5}$$

mRMR is an algorithm that maximizes the mutual information in feature-class label. In this scenario, this is basically choosing the relevant features. To minimize redundancy, mRMR measures association between each selected feature and makes a calculation of the degree of association. The mRMR criterion eliminates features that are highly correlated with the output class, but weakly correlated with each other. Subjectively, it is best measuredby the objective function.

$$(X_{s}, y) = \max_{x} \frac{1}{|S|} \sum_{x_i \in S} I(x_i, c) - \min_{x} \frac{1}{|S|^2} \sum_{x_i,x_j \in S} I(x_i, x_j) \tag{6}$$

The subset *S* of *m* features that has the minimal redundancy criterion is subset that satisfies the minimal mean value of all mutual information values between all pairs of features x_i and x_j. In mRMR, the subset *S* of *m* best features is grown iteratively using greedy approach. Here, features, one feature at a time, are chosen iteratively. Greedy algorithm selects the features that result in the most massive increase in joint mutual information at each stage. The following criterion is used to add x_j feature to the previous subset of m-1 features:

$$\max_{x_j \in X - S_{m-1}} \left[I(x_j,c) - \frac{1}{m-1} \sum_{x_j \in X - S_{m-1}} I(x_i, x_j) \right] \tag{7}$$

In summary, for this phase, the feature weight will be determined using the Relief-F filter approach. To determine the best number of feature subsets, a threshold value ranging from 0 to 1 was employed. Relief-F is a good feature ranking score, but it cannot detect redundant features, which limits its uses. Then, another filtering technique, mRMR, will be employed to eliminate these redundant features. mRMR seeks to select a subset of attributes that correlate most strongly with a class (relevance) and least strongly with each other (redundancy). The feature space for the mRMR algorithm will be the selected feature subset based on the feature weight computed from Relief-F. Then, mRMR will employ a greedy search to locate near-optimal features. Figure 1 depicts the process flows for the proposed Ensemble Filter method,

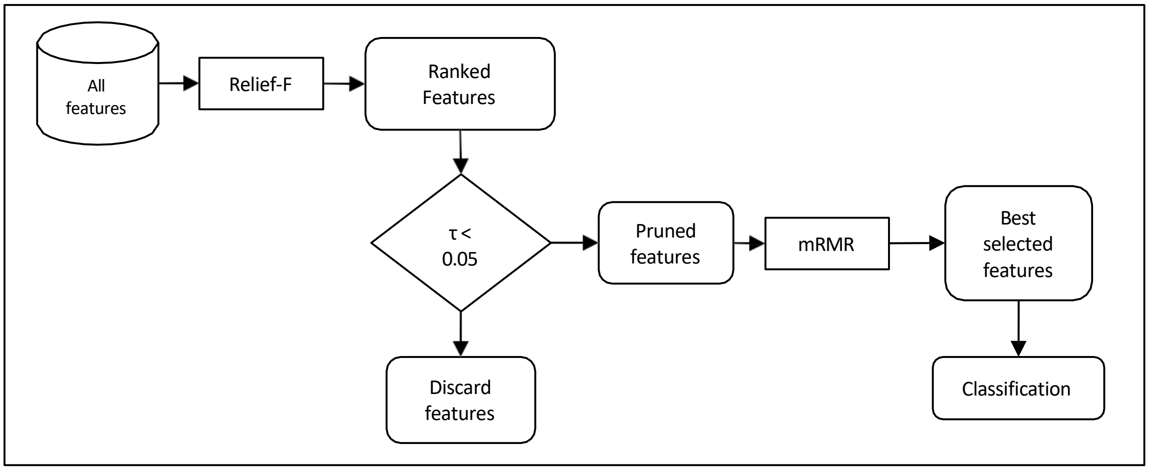

FIGURE 1. Proposed Ensemble Filter Based Feature Selection Technique

EXPERIMENTAL AND RESULTS

This section outlines the experimental study relevant to the proposed filter subset evaluation. In this research, two publicly available physical activities accelerometer sensor datasets are utilized: SBHAR and USC-HAD. Table 1 briefly described the general information of the datasets. Both datasets in this research employed accelerometers and gyroscope sensors, which denotes as (A_x, A_y, A_z) and. (G_x, G_y, G_z) where A and G are referring to Accelerometer and Gyroscope, while *x*,*y*,*z* referring to the three-dimensional (3D) axes of both sensors.

TABLE 1. Dataset

Dataset	Wearable sensors	Sensor Position	No of Activity	No. of Subject
SBHAR	Accelerometers (A_x, A_y, A_z) Gyroscope (G_x, G_y, G_z)	Waist	6	30
USC-HAD			10	14

For SBHAR dataset, a constant rate of 50 Hz is captured for each operation at three-axial linear accelerations (accelerometer) and three-axial angular velocity (gyroscopes). Data labelling is performed manually with a video taken each time during the activity. SBHAR is mobile data collection based on human behaviour identification. The subjects' movements include walking, upstairs walking, walking downstairs and upstairs, laying, sitting, and standing. For USC-HAD dataset, the sensing unit MotionNode was mounted on the subject's upper thigh for 14 subjects, and the frequency of sampling used is 100Hz. This device is embedded with three-axial accelerometer and three-axial gyroscope sensors. The subjects were carrying out 12 different styles of activities: walking forward, walking left, walking right and upstairs, walking downwards, walking up and down, sitting, standing, sleeping, elevator up and downwards. There is a total of six activities for the SBHAR dataset and ten activities for the USC-HAD dataset employed in this research, with the number of subjects 30 and 14, respectively. These two datasets were chosen because they offer information about fundamental human behaviours, including static and dynamic activities.

This research employs a few filter feature selection model parameters. A threshold value between 0.01 and 1 was evaluated. Several well-known filter subset evaluation procedures, such as Relief-F, Chi-Square, Gain Ratio, and Symmetrical Uncertainty, were compared to Relief-F in terms of both average accuracy and total time consumption. Ensemble Random Forest, Naive Bayesian, Support Vector Machine, and k-Nearest Neighbor were utilised as classifiers. A 10-fold cross validation served as the basis for this validation.

In this experiment, a classification model of ensemble random forest is used as an evaluation method. The default parameter value is extended to other methods. The efficacy of each feature subset measurement method is calculated without excluding any features from the subset. Table 2 shows the performance of different filter subset evaluation method. The Relief-F, Chi-Square and Symmetrical Uncertainty received an average of 98.29%, 98.25%, and 98.24% respectively of their results for classification accuracy in all activities. The lowest classification result is from Gain Ratio, where average accuracy is just 81.38%. In comparison, Relief-F reported the longest training time (70.89 seconds) while building the training model and Gain Ratio recorded the shortest training time (19.04 seconds). While Chi-Square and Symmetrical Uncertainty have recorded 51.34 seconds and 36.23 seconds respectively in a reasonable training duration. Based on accuracy the performance, Relief-F outperforms the other filter feature selection technique.

TABLE 2. Performance of different filter subset evaluation methods

Performance	Relief-F	Chi-Square (X^2)	Gain Ratio (GR)	Symmetrical Uncertainty (SU)
Accuracy	**98.29**	98.25	81.38	98.24
Precision	**0.98**	**0.98**	0.81	**0.98**
Recall	**0.98**	**0.98**	0.81	**0.98**
F-Measure	**0.98**	**0.98**	0.81	**0.98**
MCC	**0.98**	**0.98**	0.77	**0.98**
Training time	70.89	51.34	**19.04**	36.23
Testing time	0.24	0.16	0.15	**0.12**

Table 3 displays the complete results of the proposed Ensemble Filter in SBHAR dataset. Developing the proposed Ensemble Filter has allowed the mRMR algorithm to further reduce 79% of the number of features from 311 to 64 features. In all 10 runs, the Precision, Recall, F-Measure have an average score of 0.98 while MCC scored 0.9825.

TABLE 3. Results of proposed Ensemble filter (SBHAR)

Run	**Accuracy**	**Precision**	**Recall**	**F-Measure**	**MCC**
1	98.596	0.986	0.986	0.986	0.983
2	98.558	0.986	0.986	0.986	0.983
3	98.558	0.986	0.986	0.986	0.983
4	98.510	0.985	0.985	0.985	0.982
5	98.596	0.986	0.986	0.986	0.983
6	98.586	0.986	0.986	0.986	0.983
7	98.577	0.986	0.986	0.986	0.983
8	98.558	0.986	0.986	0.986	0.983
9	98.481	0.985	0.985	0.985	0.982
10	98.538	0.985	0.985	0.985	0.982
Avg	**98.560**	**0.986**	**0.986**	**0.986**	**0.983**

Table 4 displays the complete results obtained by using the ensemble Random Forest. The development of the proposed ensemble filter further reduced the number of features from 227 to 62 with the mRMR algorithm. In all 10 trials, the Accuracy, Recall, F- Measure, and MCC all attained an average calculated score that is 0.99.

TABLE 4. Results of proposed Ensemble filter (USCHAD)

Run	Accuracy	Precision	Recall	F-Measure	MCC
1	98.860	0.989	0.989	0.989	0.988
2	98.783	0.988	0.988	0.988	0.987
3	98.838	0.989	0.988	0.988	0.988
4	98.838	0.989	0.988	0.988	0.988
5	98.794	0.988	0.988	0.988	0.987
6	98.838	0.989	0.988	0.988	0.988
7	**98.893**	0.989	0.989	0.989	0.988
8	98.827	0.989	0.988	0.988	0.988
9	98.882	0.989	0.989	0.989	0.988
10	98.838	0.989	0.988	0.988	0.988
Avg	**98.840**	0.988	0.988	0.988	0.988

The summary outcomes of the proposed Ensemble Filter are presented in Table 5. For SBHAR dataset, in terms of accuracy performance, the proposed Ensemble Filter based on ensemble Random Forest classification outperforms Relief-F feature selection with 98.56% accuracy compared to Relief-F with 98.29%.

TABLE 5. Summary of proposed Ensemble Filter

Performance	**Relief-F (SBHAR)**	**Relief-F (USCHAD)**	**Proposed Ensemble Filter (SBHAR)**	**Proposed Ensemble Filter (USCHAD)**
Accuracy	98.29	98.79	**98.56**	**98.84**
Precision	**0.98**	**0.99**	**0.98**	**0.99**
Recall	**0.98**	**0.99**	**0.98**	**0.99**
F-Measure	**0.98**	**0.99**	**0.98**	**0.99**
MCC	**0.98**	**0.99**	**0.98**	**0.99**
Training time	70.89	98.77	**43.23**	**43.41**
Testing time	0.24	0.23	**0.12**	**0.13**

In terms of computation time, the proposed Ensemble Filter took 43.23 seconds for training time, and a small fraction of the testing time (0.12 seconds). The time to train the model has been greatly reduced since the number of features has been reduced by the proposed Ensemble Filter from 311 to 64 features. For USCHAD dataset, the proposed Ensemble Filter based on ensemble Random Forest performed better than Relief-F classification with 98.84% relative accuracy compared to Relief-F with 98.79% accuracy. The proposed Ensemble Filter takes 43.41 seconds for training, and only a small fraction of this time is spent on testing (0.13 seconds). The model's training time has been significantly reduced, as the model's number of features has been drastically reduced by 73% from 227 to 62 features.

CONCLUSION

This study demonstrated that the model complexity can be minimised by only using relevant features. The Relief-F algorithm is a method which has the power in handling noisy and incomplete data. Most of the works conducted did not consider the margin threshold to define the feature boundary. Therefore, an optimal threshold must be identified when selecting the highest-ranking feature. Then mRMR algorithm is used to further reduced the features from Relief-F. Not only is Relief-F easier to use but it also performed exceptionally well compared to the other feature-ranking methods such as Chi Square, Gain Ratio, and Symmetrical Uncertainty. The proposed Ensemble Filter of Relief-F and mRMR outperforms other filter function selection strategies, as shown in Table 5. The proposed ensemble filter improved the accuracy performance of the SBHAR dataset by 0.27% over a single Relief-F algorithm. The proposed ensemble filter enhanced the USC-HAD dataset by 0.05%. In both datasets, the proposed ensemble filter decreased the number of features by 79% and 72% compared to the single Relief-F algorithm. The proposed ensemble filter also decreased training time by 39% and 56% for the SBHAR and USC- HAD datasets, respectively.

ACKNOWLEDGMENTS

This research work is fully funded by Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak.

.

REFERENCES

1. Abo El-Maaty, A. M. and Wassal, A. G. (2019) 'Hybrid GA-PCA Feature Selection Approach for Inertial Human Activity Recognition', *Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence, SSCI 2018*.IEEE, pp. 1027–1032.
2. Alomari, O. A., Tajudin, A., Mohammed, K. and Mohammed, A. A. (2018) 'A novel gene selection method using modified MRMR and hybrid bat-inspired algorithm with β-hill climbing', *Applied Intelligence*. Applied Intelligence, (June).
3. Amezzane, I., Fakhri, Y., El Aroussi, M. and Bakhouya, M. (2018) 'Analysis and Effect of Feature Selection Over Smartphone-Based Dataset for Human Activity Recognition', *Lecture Notes of the Institute for Computer Sciences,Social-Informatics and Telecommunications Engineering, LNICST*, 206, pp. 214–219.
4. Arif, M., Kattan, A. and Ahamed, S. I. (2017) 'Classification of physical activities using wearable sensors',
5. *Intelligent Automation and Soft Computing*, 23(1), pp. 21–30.
6. Bolón-Canedo, V., Sánchez-Maroño, N. and Alonso-Betanzos, A. (2014) 'Data classification using an ensemble offilters', *Neurocomputing*, 135, pp. 13–20.
7. Bulling, A., Blanke, U. L. F. and Schiele, B. (2014) 'A Tutorial on Human Activity Recognition Using Body-Worn',46(3), pp. 1–33.
8. Capela, N. A., Lemaire, E. D. and Baddour, N. (2015) 'Feature selection for wearable smartphone-based humanactivity recognition with able bodied, elderly, and stroke patients', *PLoS ONE*, 10(4), pp. 1–18.
9. Chandra, B. (2018) 'An efficient feature selection technique for gene expression data', *2018 IEEE Conference onComputational Intelligence in Bioinformatics and Computational Biology, CIBCB 2018*. IEEE, pp. 1–6.
10. D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," *ESANN 2013 proceedings, 21st Eur. Symp. Artif. Neural Networks, Comput. Intell. Mach. Learn.*, no. April, pp. 437–442, 2013.
11. Doewes, A., Swasono, S. E. and Harjito, B. (2017) 'Feature selection on Human Activity Recognition dataset usingMinimum Redundancy Maximum Relevance', *2017 IEEE International Conference on Consumer Electronics - Taiwan, ICCE-TW 2017*. IEEE, (1), pp. 171–172.
12. M. Zhang and A. A. Sawchuk, "USC-HAD: A Daily Activity Dataset for Ubiquitous Activity Recognition Using Wearable Sensors," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 2012, pp. 1036–1043, doi: 10.1145/2370216.2370438.

13. Shoaib, M., Bosch, S., Incel, O. D., Scholten, H. and Havinga, P. J. M. (2016) 'Complex human activity recognitionusing smartphone and wrist-worn motion sensors', *Sensors (Switzerland)*, 16(4).
14. Wang, Y., Vasilakos, A., Jin, Q. and Zhu, H. (2017) 'Smartphone-Based Human Activity Recognition', *Device-to-Device based Proximity Service*, (2012), pp. 199–239.
15. Wang, Z. and Meng, F. (2018) 'An overview of human activity recognition based on smartphone'.
16. Wang, Z., Meng, F., Yuan, G., Yan, Q. and Xia, S. (2018) 'An overview of human activity recognition based onsmartphone', *Sensor Review*.
17. Zubair, M., Song, K. and Yoon, C. (2016) 'Human activity recognition using wearable sensors', *Advances inIntelligent Systems and Computing*. IEEE, 1177, pp. 527–538.