Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2018) 26: 1378 – 1389

doi:10.3906/elk-1709-138

Two-stage feature selection using ranking self-adaptive differential evolution algorithm for recognition of acceleration activity

Muhammad Noorazlan Shah ZAINUDIN[1,2,*], **Md. Nasir SULAIMAN**[1], **Norwati MUSTAPHA**[1], **Thinagaran PERUMAL**[1], **Raihani MOHAMED**[1]

[1]Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Selangor, Malaysia

[2]Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

Received: 14.09.2017 • **Accepted/Published Online:** 03.04.2018 • **Final Version:** 30.05.2018

Abstract: Widespread research on activity recognition is becoming an imperative topic for improving the quality of human health. The fast development of sensing technology has become a fundamental platform for researchers to implement a system that could fulfill human needs. Due to privacy interests and low cost, wearable sensing technology is used in numerous physical activity monitoring and recognition systems. While these systems have proved to be successful, it is crucial to pay attention to the less relevant features to be classified. In such circumstances, it might happen that some features are less meaningful for describing the activity. Less complex and easy to understand, feature ranking is gaining a lot of attention in most feature dimension problems such as in bioinformatics and hyperspectral images. However, the improvement of ranking features in activity recognition has not yet been achieved. On the other hand, an evolutionary algorithm has proven its effectiveness in searching the best feature subsets. An exhaustive searching process of finding an optimal parameter value is another challenge. Consequently, this paper proposes a ranking self-adaptive differential evolution (rsaDE) feature selection algorithm. The proposed algorithm is capable of selecting the optimal feature subsets while improving the recognition of acceleration activity using a minimum number of features. The experiments employed real-world physical acceleration data sets: WISDM and PAMAP2. As a result, rsaDE performed better than the current methods in terms of model performance and its efficiency in the context of random forest ensemble classifiers.

Key words: Activity recognition, feature ranking, self-adaptive differential evolution, relief-f, random forest

1. Introduction

Sensing technology has become an imperative topic in security surveillance [1], human/computer interaction [2], and activity recognition [3]. Activity recognition establishes the relations between machine learning, artificial intelligence, ubiquitous computing, human/computer interaction, and psychology and sociology . Activity recognition is successfully applied to a variety of populations, such as patients from weight control programs and also for rehabilitation programs [5]. Unfortunately, it is problematic to manually self-record daily physical activity. Thus, use of wearable sensors to monitor physical activity while exercising is receiving great attention. The increase in healthy lifestyle awareness and physical concerns change the daily human routines. Consequently, such physical activities are essential to improve the quality of life [6]. The performances of activity recognition are measured in terms of its effectiveness and efficiency. The selections of features from acceleration sensor data are considered to be challenging, which is interrelated with recognition accuracy. In such circumstances,

*Correspondence: noorazlan@utem.edu.my

some of the extracted features may be less relevant or unusable as well as redundant for portraying the activity [7]. These features are considered less meaningful in the sense that they possibly can increase the incorrect classification rate [8]. On the other hand, differential evolution (DE) is broadly employed to generate the feature subsets [9]. However, a thorough process of finding an optimal parameter in traditional DE is exhausting [10]. Hence, self-adaptive parameter initialization is introduced to tackle this matter. Several contributions are carried out in this paper. Features from the statistical and spectral frequency are introduced to improve the learning capability of physical activity recognition. The highly ranked features of relief-f are examined and pruned according to the boundary threshold. Two adaptive parameters, number of generations and population size, are adaptively defined from the number of high ranking features to minimize the complexity of the search space. On top of that, a self-adaptive scaling factor and crossover probability are proposed to minimize the time of finding an optimal parameter value by selecting the optimum feature subsets. We prove that the proposed ranking self-adaptive differential evolution (rsaDE) algorithm is able to achieve a higher level of accuracy in activity recognition than several state-of-the-art feature selection methods.

2. Materials and methods

The proposed framework as illustrated in Figure 1 consists of four stages, namely preprocessing, feature extraction, feature selection, and classification. The proposed method is mainly found in Stages 2 and 3. Details of each stage are discussed in the following subsections.

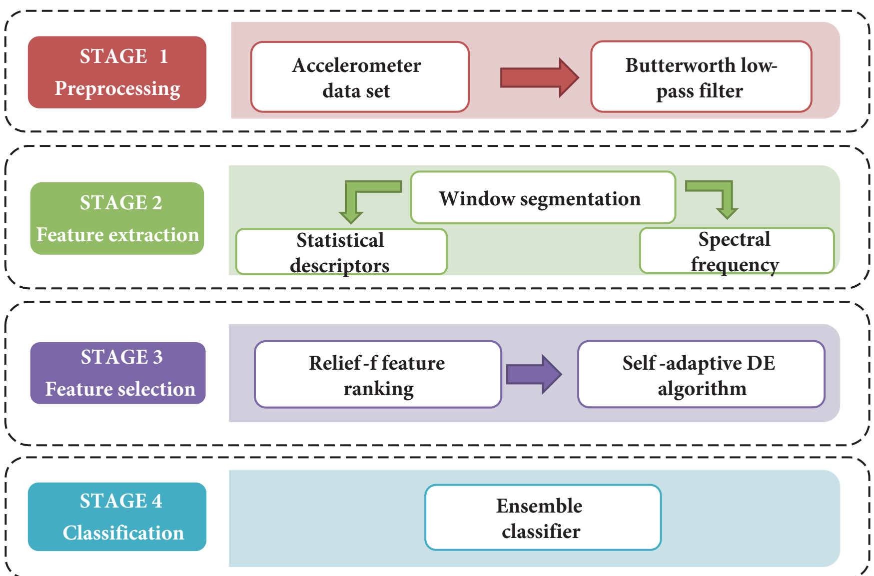

Figure 1. The proposed framework for activity recognition.

2.1. Physical activity data set

Two real-world physical acceleration data sets are utilized: Wireless Sensor Data Mining (WISDM) [11] and Physical Activity Monitoring for Aging People (PAMAP2) [12]. WISDM utilizes a 20-Hz sampling rate through the use of an accelerometer, which is embedded in Android smartphones and attached at the front pocket of pants. Meanwhile, in PAMAP2, three inertial measurement units (IMUs) are used and attached to three different places: the wrist, chest, and ankle. Unlike WISDM, a 100-Hz sampling rate is used in PAMAP2 due to it covering various types of complex activity.

2.2. Signal filtering

The acceleration data are recorded in each direction to capture the sum of gravitational and body acceleration. The acceleration signal is recorded continuously and preprocessed in order to isolate the gravitational and body acceleration [13]. Each acceleration signal consists of two different signal components: low-frequency components and high-frequency components. Low-frequency components are captured by gravity according to the orientation of the sensor with respect to ground data. Meanwhile, high-frequency components are generated due to the noise presented from gravitational forces. The gravitational acceleration signal generated in low-frequency components needs to be removed in order to produce a decent signal presentation for classification purposes [14]. In such states, a Butterworth low-pass filter is applied to eliminate high-frequency components by retaining the low-frequency components [7].

2.3. Feature extraction

It is necessary to extract the original signal for further computational efficiency. The extracted information is referred to as feature vectors and later will be used for classification [15]. The acceleration signal in each dimension is divided into several predetermined windows sizes [16]. A sliding window is extensively applied to solve various time series application problems [17–19]. In order to transform the acceleration signal in a more meaningful way, feature extraction has taken place. In this work, we propose 12 features, where 9 features are extracted from statistical features while 3 features are extracted from frequency features. The list of the extracted features is given in Table 1.

Table 1. List of extracted features.

Statistical features	Frequency features
Minimum and maximum, mean, variance, standard deviation, skewness, kurtosis, correlation coefficient, harmonic mean	Band power, power bandwidth, occupied bandwidth

3. The proposed rsaDE feature selection algorithm

The proposed rsaDE encompasses two stages. The highly relevant features are ranked using relief-f according to the score in the first stage. Afterward, the highly ranked features are evaluated by using the self-adaptive differential evolution algorithm. Simple but capable of generalizing noisy and incomplete data, relief-f is broadly applied in various feature dimension problems. Nearest hits (data points from the same class) and nearest misses (data points from a different class) of each instance are calculated using Eq. (1).

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

Here, w is the weight of the ith feature, x_i^j is the value of the ith feature for point x^j, and N is the total number of data points. Near hit x^j and near miss x^j are the nearest data points to x^j in the same and different classes, respectively. In order to prune the less relevant features, selected features considered as highly ranking are selected using the following equation.

$$f_i = \begin{cases} \sum f_{i+1} & \text{if } w_i \geq 0.01 \\ \text{prune} & \text{otherwise} \end{cases} \tag{2}$$

Here, f is the selected ith features, w is the weight of the ith features, and if w is greater than or equal to 0.01, the ith features are selected. If otherwise, the ith features are pruned.

The DE algorithm is introduced based on the differential reproduction that occurs within a population from environmental pressures that lead to natural selection (survival of the fittest). This algorithm is easy and effective in convergence, capable of handling nondifferentiable, nonlinear, and multimodal cost functions [20], as well as able to perform a parallel search with fast implementation properties. This condition makes DE ideal to be extensively applied to select the best feature subsets. DE is capable of manipulating a target vector using a different vector to generate a trial vector. The population is generated from each D-dimensional real-valued parameter where NP represents population size and D represents the number of parameters to be optimized. A trial vector is generated based on the weight from different vectors between two population members x_{r1} and x_{r2}, which are added to the third member, x_{r0}. For each target vector, a mutant vector $v_{j,i,g}$ is generated according to Eq. (3).

$$v_{j,i,g}{=}x_{j,r0,g}{+}F\left(x_{j,r1,g}{-}x_{j,r2,g}\right) \tag{3}$$

Here, $r_0 r_1 r_2 \in \{0, 2, \ldots, NP-1\}$are randomly chosen integers, $g = NP = f_i$, and population members must be different from each other, which is $r_0 \neq r_1 \neq r_2$. Index g indicates the generation to which a vector belongs. Index i indicates the population index that runs from to $NP-1$. Parameters within the vectors are indexed with j, which runs from 0to $D-1$. The value of F in Eq. (4) is a scaling factor that controls the rate at which the population evolves.

$$F_{i,G+1} = \begin{cases} F_l + rand_2 * F_u & \text{if } \mathrm{rand}_1 < \tau_1 \\ F_{i,G} & \text{otherwise} \end{cases} \tag{4}$$

Here, $rand_j j \in \{1, 2, 3, 4\}$ lies between $[0, 1]$ and $\tau_1 = 0.1$. F_l and F_u are set to 0.1 and 0.9, respectively. In order to increase the diversity of perturbed parameter vectors, a crossover is introduced. Trial vector $u_{j,i,g}$is obtained from an element of mutant vector $v_{j,i,g}$ and target vector $x_{j,i,g}$; $u_{j,i,g}$ is written as shown in Eq. (5).

$$u_{j,i,g} = \left\{ \begin{array}{ll} v_{j,i,g} & \text{if rand } (0,1) \leq CR \\ x_{j,i,g} & \text{otherwise} \end{array} \right\} \tag{5}$$

Here, $j = 0, 2, \ldots D-1$ and crossover probabilityCR is self-adapted by using Eq. (6). $u_{j,i,g}$ is the jth dimension from the jth trial vector along with the current population g.

$$CR_{i,G+1} = \begin{cases} rand_4 & \text{if } \mathrm{rand}_3 < \tau_2 \\ CR_{i,G} & \text{otherwise} \end{cases} \tag{6}$$

Here, $rand_j j \in \{1, 2, 3, 4\}$ lies between $[0, 1]$ and $\tau_2 = 0.1$. F_l and F_u are set to 0.1 and 0.9, respectively. Selection as shown in Eq. (7) is a step to choose the vector between the target vector and trial vector for creating an individual population for the next generation.

$$x_{j,i,g+1}{=}\left\{ \begin{array}{ll} u_{j,i,g} & \text{if } f\left(u_{i,G+1}\right){\geq}f\left(x_{i,G}\right) \\ x_{j,i,g} & \text{otherwise} \end{array} \right\} \tag{7}$$

If the generated vector has a lower objective function value (better fitness) than a predetermined population member, the resulting vector replaces the vector with which it was compared. Otherwise, the predetermined population member remains for the next generation.

4. Evaluation of proposed rsaDE feature selection

In this work, a 5th order Butterworth low-pass filter with 0.3-Hz cut-off frequency is chosen. This amount is considered sufficient to separate high- and low-frequency components [22,23]. A total of 36 features (12 $\times$ 3 dimensions) with an additional one feature are added to the activity label accordingly, which serves the classification stage later. The experimental results are measured using several performance indicators: overall accuracy, precision, recall, and training build time. A random forest ensemble classifier is utilized as a classifier model and validation is conducted in a 10-fold cross-validation classification strategy.

4.1. Analysis of relief-f feature ranking

In this section, the evaluation of the relief-f feature ranking method is described. We conduct the experiment to rank the features (original 36 features) using relief-f for WISDM. Since both data sets, WISDM and PAMAP2, utilize the same features (as explained in Section 2.3), the evaluation has only been conducted with WISDM. For comparison, the experimental result is compared with three different feature ranking methods: gain ratio, information gain, and principal component analysis (PCA).

Table 2. Classification performance using different feature ranking methods.

Indicators	Gain Ratio	Info Gain	PCA	Relief-f
Accuracy	0.997	0.997	0.913	0.997
Precision	0.997	0.997	0.913	0.997
Recall	0.997	0.997	0.913	0.997
Time (in seconds)	13.71	13.76	18.08	13.32

As demonstrated in Table 2, accuracies of gain ratio, information gain, and relief-f reveal the same result of 97%. PCA contributes the lowest accuracy and precision at 91%. PCA has the longest time taken to build the training model at 18.08 s. Even though there is no large difference between times taken to build the training model compared to other feature ranking methods, relief-f has the shortest time recorded at 13.32 s. Figure 2 shows the ranking score of each feature, according to relevant criteria.

The weight scored within the range of $[8^{e-5}, 6^{e-2}]$ is considered as the heaviest weight and the most significant features. The result clearly shows that 70% of features have a score above 0.02, and very few features have a score lower than 0.01. In order to evaluate the classification performance within the selected features range, a few numbers with selected thresholds have been investigated.

As visualized in Figure 3, the average accuracy within the boundary of 0.001 to 0.01 is recorded as stagnant at 99.7%. However, accuracy decreased about 0.1% when the features weight of 0.03 was selected. In such circumstances, it is likely that the number of features might be inadequate or it might possibly be that highly relevant features have been removed. The accuracy drastically drops when features with weights above 0.05 are selected. Consequently, features weights above 0.01 will be selected and applied in the next stage. As a result, 30 features are produced and we believe that these features are the most influential features to portray the class.

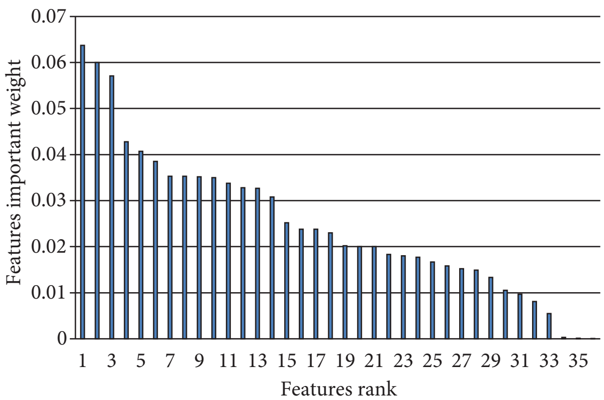

Figure 2. Ranking features according to the relevance score.

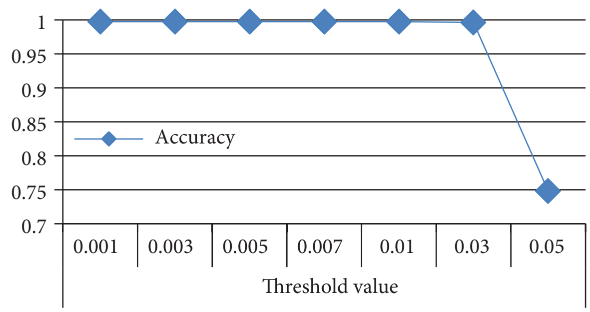

Figure 3. Classification accuracy of the different thresholds.

4.2. Analysis of rsaDE algorithm

In this section, experimental analysis of the rsaDE is assessed. GEN and NP are initialized to 30 and these parameters are kept constant during each run. On the other hand, F and CR have been self-adaptively initialized according to Eqs. (1) and (6), respectively. In contrast with the previous two parameters (GEN and NP), F and CR are dynamically changed within each generation. Random numbers [0.1, 1.0] are used to initialize F and CR in first generation. In order to measure the subset performance, there are some criteria to be considered. The algorithm is terminated when the difference between fitness values of two consecutive generations of the best subsets is negligible (<0.001). The same scenarios prevail for a certain number of generations. However, in our case, GEN is equal to the number of dimensions, D. According to [24], the desired number of features (DNF) is suggested within the range of 10 to 20.

Figure 4 illustrates the classification accuracy for different numbers of DNF. The accuracy increases to 99.5% when 12 and 13 features are chosen. However, the accuracy declines when features are selected within the range of 14 to 20. Lowest accuracy of 99.4% is recorded when 17 random features are chosen. The classification accuracy of rsaDE has also been compared with traditional DE. For verification purposes, an experiment is conducted with an average of 10 runs.

Figure 5 illustrates the average accuracy between the rsaDE and traditional DE. It is clearly shown that the classification accuracy of rsaDE is acceptable compared to traditional DE. On average, rsaDE has recorded accuracy of 99.6%, which is a bit superior to traditional DE by 0.1%. There is no large difference between the methods, but rsaDE is more capable of minimizing the complexity of finding an optimal parameter than traditional approaches.

The performance of rsaDE has also been compared with several state-of-the-art subset generation methods such as particle swarm optimization (PSO), reduced scatter search (RSS), the evolutionary algorithm (EA), and tabu search. For fair comparison, the same value of NP is applied to PSO, RSS, and EA. Besides that, GEN = 30 is applied for both PSO and EA. Crossover probabilities and mutation probabilities are initialized as 0.5 and 0.01, respectively, for EA. Meanwhile, number of neighbors = 1 is chosen for the tabu search.

Figure 6 shows the average classification accuracy of various subset generation methods. It is clearly shown that rsaDE performs better as compared to other methods. PSO, RSS, and tabu search have the same percentage for accuracy while EA performs better with an increase of 0.1% than traditional DE, RSS, and tabu search. In terms of number of selected features, 13 and 11 features are selected for EA and PSO, respectively.

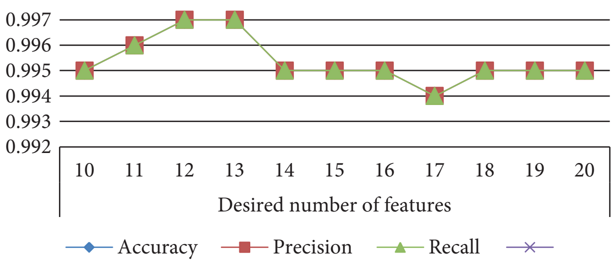

Figure 4. Classification performance using different numbers of DNF.

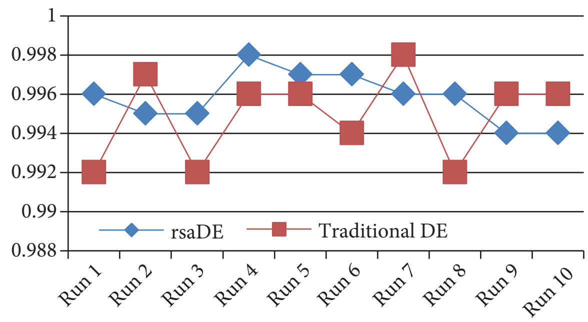

Figure 5. Comparison classification accuracy of rsaDE and traditional DE.

Meanwhile, RSS and tabu search have selected an equal number of features, 11. Even though the result from traditional DE is acceptable, 17 features is still considered as a higher value than that for rsaDE.

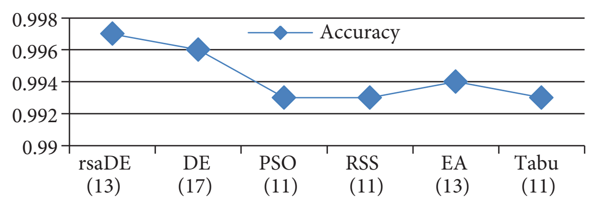

Figure 6. Classification accuracy for different types of subset generation methods.

5. Experimental results and analysis

In this section, we conduct an experiment for both data sets, WISDM and PAMAP2. We also compare our performance with several benchmark works that utilize both data sets. However, different sizes in segmentation and different validations in classification have been applied according to the benchmark work accordingly.

5.1. WISDM

WISDM applies sliding window sizes of 10 s with overlapping 2.5 s between two consecutive window segments. Hence, it will produce about 200 samples per second with 25 samples overlapping with the next following window segment. For validation purposes, 80% of randomly selected samples are used for training and the remaining 20% of subsets are applied for testing.

Table 3 shows the comparison accuracy of rsaDE with traditional DE, EA, and RSS as reported by the authors [22]. The authors also reported the most difficult activities to be those involving the stairs (down stairs and up stairs) and walking, even though data collection was conducted in a different space. They also applied Decremental Reduction Optimization Procedure 2 (DROP2) to reduce the complexity of the search space. However, going down stairs and up stairs are reported to have the worst performances, with about 81% and 88%. Meanwhile, our experiment shows that rsaDE achieved the highest accuracy, even though traditional DE performed somewhat better than EA and RSS. Up stairs, down stairs, and walking could be effectively differentiated with precision of 99.7%, 100%, and 99.9%, accordingly. EA was also slightly better, except for going down stairs, although EA is able to produce the same number of features as rsaDE. In comparison with a number of features, almost 40% of features could be reduced using rsaDE, which is considered as better than RSS and traditional DE.

Table 3. Comparison accuracy of rsaDE, traditional DE, EA, and RSS.

Activity	EA (13)	RSS (30)	DE (17)	rsaDE (13)
Down stairs	0.988	0.816	0.993	0.997
Up stairs	0.996	0.888	0.999	1.000
Walking	0.991	0.963	0.998	0.999
Jogging	0.994	0.986	1.000	1.000
Sitting	1.000	0.996	1.000	1.000
Standing	1.000	0.989	1.000	1.000
Average	0.993	0.953	0.998	1.000

Table 4. Confusion matrix of rsaDE.

Activity	Down stairs	Up stairs	Walking	Jogging	Sitting	Standing
Down stairs	398	0	0	0	0	0
Up stairs	1	481	1	0	0	0
Walking	0	0	1659	0	0	0
Jogging	0	0	0	1329	0	0
Sitting	0	0	0	0	223	0
Standing	0	0	0	0	0	181

Table 4 presents the confusion matrix of rsaDE. As illustrated, 1% of "up stairs" instances have been recognized as going down stairs and walking. Obviously, rsaDE tends to increase the classification accuracy by using a minimum number of features. In this sense, 13 features are considered as adequate to recognize various types of physical activities and have good performance in reducing the feature subsets compared to the previous work (n = 30).

5.2. PAMAP2

In a previous work [5], the authors utilized only twelve types of activities in their experiment. A sliding window of 5 s with overlapping 1 s between two consecutive window segments was applied and 70% of randomly selected instances were used for training while the remaining instances were applied for testing. In order to minimize the number of features to be classified, RSS subset generation has been proposed. Unlike WISDM, PAMAP2 is evaluated according to individual and combinational sensor placements.

Table 5 shows the classification accuracy for each placement (wrist, chest, and ankle). The average accuracy of rsaDE for each placement is 99%. Surprisingly, ascending and descending walk achieves an outstanding accuracy above 99% for all placements. The ankle has the lowest accuracy among all other placements. However, there is a slight drop in ascending walk when the sensor is worn on the wrist. The wrist has a better performance in recognizing Nordic walk and walking. Hence, the wrist and chest have good performance for hand motion activity as compared with the ankle. The ankle is considered as the worst for recognizing activity that involves hand motion activity. In previous work, poor performance was also reported, below 94% for each placement. Standing has been reported to have the worst accuracy for the chest and ankle, while the wrist and ankle are thought the worst to describe ironing activity.

In this experiment, feature subsets from each sensor placement are combined to form the new feature

Table 5. Classification accuracy of each placement using rsaDE.

Activity	Wrist		Chest		Ankle	
	RSS (30)	rsaDE (13)	RSS (30)	rsaDE (13)	RSS (30)	rsaDE (13)
A1-Ascending	0.907	0.986	0.942	1.000	0.951	0.992
A2-Cycling	0.987	0.998	0.958	0.998	0.949	0.994
A3-Descending	0.954	1.000	0.976	0.994	0.965	1.000
A4-Ironing	0.845	1.000	0.902	1.000	0.787	0.990
A5-Jumping	0.963	1.000	0.993	1.000	1.000	1.000
A6-Lying down	0.974	0.995	0.985	1.000	0.995	1.000
A7-Nordic walk	0.989	0.996	0.974	0.989	0.962	0.984
A8-Running	0.996	1.000	0.996	0.993	0.981	1.000
A9-Sitting	0.922	1.000	0.917	1.000	0.963	0.996
A10-Standing	0.855	0.998	0.838	0.998	0.838	0.995
A11-Vacuuming	0.961	0.994	0.928	0.992	0.853	0.987
A12-Walking	0.984	1.000	0.952	0.997	0.968	0.983
Average	0.941	0.997	0.940	0.997	0.925	0.992

Table 6. Classification accuracy of rsaDE, traditional DE, EA, and RSS for sensor combination.

Label	Activity	EA (39)	RSS (46)	DE (51)	rsaDE (39)
A1	Ascending	0.997	0.974	1.000	1.000
A2	Cycling	1.000	1.000	1.000	1.000
A3	Descending	0.994	0.966	1.000	1.000
A4	Ironing	1.000	0.964	1.000	1.000
A5	Jumping	1.000	1.000	1.000	1.000
A6	Lying down	1.000	0.995	1.000	1.000
A7	Nordic walk	0.998	0.996	1.000	1.000
A8	Running	1.000	1.000	1.000	1.000
A9	Sitting	1.000	0.975	1.000	1.000
A10	Standing	1.000	0.949	1.000	1.000
A11	Vacuuming	1.000	0.983	0.996	1.000
A12	Walking	0.999	0.975	1.000	1.000
Average		0.999	0.980	1.000	1.000

subsets. Hence, the new feature subset produced had 39 features (13 features $\times$ 3 placements). Table 6 shows the classification accuracy of the combination from all placements (wrist, chest, and ankle). Referring to previous studies, RSS is able to reduce about 60% (46 features) compared to the original features (117 features). Previously, k-nearest neighbors was chosen as the evaluation model and average accuracy produced was 98%. Running, cycling, and jumping had the highest accuracy, while standing had the poorest. In our work, rsaDE has clearly outperformed the average accuracy of the previous work. All activities had 100% precision when the reduced feature subsets were combined. Meanwhile, DE had similar performance to rsaDE, followed by

EA. Unfortunately, 51 features is still considered higher than rsaDE, which leads to increased learning model complexity. Also, there is slightly decreased accuracy in recognizing the stride activities by EA. Furthermore, it can be observed that at the cost of fractional loss of classification accuracy the number of features would reduce dramatically to 39, which is about 64% size reduction in feature subsets.

Table 7. Confusion matrix of rsaDE for all placements.

Act	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
A1	353	0	0	0	0	0	0	0	0	0	0	0
A2	0	493	0	0	0	0	0	0	0	0	0	0
A3	0	0	316	0	0	0	0	0	0	0	0	0
A4	0	0	0	720	0	0	0	0	0	0	0	0
A5	0	0	0	0	145	0	0	0	0	0	0	0
A6	0	0	0	0	0	581	0	0	0	0	0	0
A7	0	0	0	0	0	0	562	0	0	0	0	0
A8	0	0	0	0	0	0	0	291	0	0	0	0
A9	0	0	0	0	0	0	0	0	558	0	0	0
A10	0	0	0	0	0	0	0	0	0	572	0	0
A11	0	0	0	0	0	0	0	0	0	0	529	0
A12	0	0	0	0	0	0	0	0	0	0	0	702

It is clearly seen that the confusion matrix in Table 7 produces an excellent recognition performance for all types of physical activities. All instances are classified exactly according to their categories. Even though stairs activities (ascending and descending) are considered as somewhat similar and are always confused with walking, these activities can also be effectively differentiated using our feature subsets.

6. Conclusion

This paper discusses the experimental evaluation of enhancing the previous activity recognition for two different acceleration physical activity data sets: WISDM and PAMAP2. Several features from statistics and frequency are introduced to increase the diversity in differentiating between stationary and locomotion activity. The correlations between statistical and frequency features are also measured and evaluated. The boundary threshold of 0.01 is introduced to select the relevant features, which were previously ranked by relief-f. Afterwards, irrelevant features are pruned and eliminated before the selected feature subset is fed as input to rsaDE. In this work, GEN and NP parameters are adaptively initialized from the numbers of input features. The experimental result verifies that adaptive parameter setting tends to produce an optimum accuracy in reducing the searching space complexity. Furthermore, self-adaptive F and CR are applied to minimize the thoroughness of finding the optimal parameter value. This parameter mechanism is to avoid defining different parameters in each algorithm run. Therefore, rsaDE proves better in reducing the number of features and is able to produce a high level of accuracy as compared with other methods. As another way to verify the effectiveness of rsaDE, all possible sensor placements are combined into one collection of feature subsets. This assessment also shows that rsaDE has decent accuracy of 100% in distinguishing 12 physical types effectively compared to previously reported work. For future work, we plan to investigate the effectiveness of rsaDE in other domain areas such as in bioinformatics and text mining to evaluate its applicability and consistency.

References

[1] Lara OD, Labrador MA. A survey on human activity recognition using wearable sensors. IEEE Commun Surv Tut 2013; 15: 1192-1209.

[2] Murthy GRS, Jadon RS. Hand gesture recognition using neural networks. In: IEEE 2nd International Advance Computing Conference; 1 March 2010; Patiala, India. New York, NY, USA: IEEE. pp. 134-138.

[3] Guiry JJ, van De Ven P, Nelson J, Warmerdam L, Riper H. Activity recognition with smartphone support. Med Eng Phys 2014; 36: 670-675.

[4] Khan AM. Human activity recognition using a single tri-axial accelerometer. PhD, Kyung Hee University, Seoul, Korea, 2011.

[5] Arif M, Kattan A. Physical activities monitoring using wearable acceleration sensors attached to the body. PLoS One 2015; 10: e0130851.

[6] Dobbins C, Rawassizadeh R, Momeni E. Detecting physical activity within lifelogs towards preventing obesity and aiding ambient assisted living. Neurocomputing 2013; 230: 110-132.

[7] Machado IP, Gomes AL, Gamboa H, Paixao V, Costa RM. Human activity data discovery from triaxial accelerometer sensor: non-supervised learning sensitivity to feature extraction parametrization. Inf Process Manag 2015; 51: 201-214.

[8] Martinoyić G, Bajer D, Zorić B. A differential evolution approach to dimensionality reduction for classification needs. Int J Appl Math Comput Sci 2014; 24: 111-122.

[9] Olvera-Lopez JA, Carrasco-Ochoa JA, Martinez-Trinidad JF, Kittler J. A review of instance selection methods. Artif Intell Rev 2010; 34: 133-143.

[10] Brown C, Jin Y, Leach M, Hodgson M. μJADE: adaptive differential evolution with a small population. Soft Comput 2016; 20: 4111-4120.

[11] Weiss G, Lockhart J, Pulickal T. Actitracker: A smartphone-based activity recognition system for improving health and well-being. In: IEEE International Conference on Data Science and Advanced Analytics; 26 December 2016; Montreal, Canada. New York, NY, USA: IEEE. pp. 682-688.

[12] Reiss A, Stricker D. Introducing a new benchmarked dataset for activity monitoring. In: 16th International Symposium on Wearable Computers; 23 July 2012; Newcastle, UK. pp. 108-109.

[13] Acharjee D, Mukherjee A, Mandal JK, Mukherjee N. Activity recognition system using inbuilt sensors of smart mobile phone and minimizing feature vectors. Microsyst Technol 2016; 22: 2715-2722.

[14] Xiao W, Lu Y, Cui J, Ji L. Recognition of human stair ascent and descent. Proceedings in Adaptation, Learning and Optimization 2015; 2: 253-265.

[15] Hoseini-Tabatabaei SA, Gluhak A, Tafazolli R. A survey on smartphone-based systems for opportunistic user. ACM Comput Surv 2013; 45: 1-51.

[16] Ignatov AD, Strijov VV. Human activity recognition using quasiperiodic time series collected from a single tri-axial accelerometer. Multimed Tools Appl 2016; 75: 7257-7270.

[17] Fida B, Bernabucci I, Bibbo D, Conforto S, Schmid M. Varying behavior of different window sizes on the classification of static and dynamic physical activities from a single accelerometer. Med Eng Phys 2015; 37: 705-711.

[18] Ronao CA. Human activity recognition using smartphone sensors with two-stage continuous hidden Markov models. In: 10th International Conference on Natural Computation; 8 December 2014; Xiamen, China. pp. 681-686.

[19] Siirtola P, Roning J. Recognizing human activities user-independently on smartphones based on accelerometer data. Int J Artif Intell Interact Multimed 2012; 1: 38-45.

[20] Storn R, Price K. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. J Glob Optim 1997; 11: 341-359.

[21] Bharathi P., Subashini P. Optimal feature subset selection using differential evolution and extreme learning machine. Int J Sci Res 2014; 3: 1898-1905.

[22] Arif M, Bilal M, Kattan A, Ahamed SI. Better physical activity classification using smartphone acceleration sensor. J Med Syst 2014; 38: 1-10.

[23] Anguita D, Ghio A, Oneto L, Parra X, Reyes-Ortiz JL. A public domain dataset for human activity recognition using smartphones. In: European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning; 24–26 April 2013; Bruges, Belgium. pp. 437-442.

[24] Khushaba RN, Al-Ani A, Al-Jumaily A. Feature subset selection using differential evolution and a statistical repair mechanism. Expert Syst Appl 2011; 38: 11515-11526.