A Novel Selective Ensemble Learning Method for Smartphone Sensor-Based Human Activity Recognition Based on Hybrid Diversity Enhancement and Improved Binary Glowworm Swarm Optimization

YIMING TIAN, JIE ZHANG, (Senior Member, IEEE), QI CHEN, AND ZUOJUN LIU

1College of Information Engineering, Tianjin University of Commerce, Tianjin 300134, China
2School of Engineering, Mezz Court, Newcastle University, NE1 7RU Newcastle upon Tyne, U.K.
3School of Artificial Intelligence and Data Science, Hebei University of Technology, Tianjin 300130, China

Corresponding author: Yiming Tian (yimingtian@tjcu.edu.cn)

This work was supported in part by the Tianjin Science and Technology Ombudsman Project 22YDTPJC00290, in part by the National Natural Science Foundation of China under Grant 61803143 and Grant 61703135, and in part by the National Key Technology Research and Development Program of the Ministry of Science and Technology of China under Grant 2015BAI06B03.

ABSTRACT
Human activity recognition (HAR) is gaining interest with many important applications including ubiquitous computing, health-care services and detection of diseases. Smartphone sensors have high acceptance and adherence in daily life and they provide an alternative and economic way for activity recognition. To improve the performance of smartphone sensor-based HAR, a novel smartphone sensor-based HAR method (hybrid diversity enhancement with selective ensemble learning, HDESEN) that utilizes selective ensemble learning with differentiated extreme learning machines (ELMs) is proposed, where hybrid diversity enhancement is proposed to boost the diversity of base models and an improved binary glowworm swarm optimization (IBGSO) is employed to effectively enhance the learning process by choosing a superior subset for ensemble instead of all. Firstly, statistical features in the time domain and frequency domain are extracted and integrated from smartphone sensors and then three filter-based feature selection methods are utilized for desirable base models. Secondly, to enhance the diversity of the base models, three types of diversities are introduced to construct different base models, respectively. Among them, Bootstrap is introduced to design distinctive training data subsets for differential base models, random subspace and optimized subspace are proposed to obtain different feature spaces for constructing base models. Thirdly, a pruning method based on glowworm swarm optimization (GSO) is proposed to find the optimal sub-ensemble from the pool of models from all diverse types to implement selective ensemble learning. The experimental results on tow publicly available datasets (UCI-HAR and WISDM) demonstrate the proposed HDESEN can reliably improve the performance of HAR and outperforms the relevant state-of-the-art approaches.

INDEX TERMS
Human activity recognition, smartphone sensors, selective ensemble learning, wearable sensor, diversity enhancement.

I. INTRODUCTION

With the recent development of smart devices, wireless communication and machine learning, HAR has become an active
research area and will benefit various applications including ubiquitous computing [1], sports activity [2] and smart building [3]. Besides, HAR can be a promising solution to aging problems and many achievements have been gained in this area, for example, HAR system can provide information and intelligent services to related agencies so that the elderly people can remain healthily and safely at home as long as possible [4]. Recognizing deviations from preceding normal patterns can help us to identify whether any emergency has occurred [5]. A variety of fall detection systems [6], [7] with functions of monitoring and raising alarms have been proposed and developed for elderly care, which would be advantageous to emergency treatment after falling.

According to the sensing technologies used, the HAR approaches can be divided into two categories: 1) computer vision-based HAR and 2) Inertial sensor-based HAR [8]. The vision-based HAR has first attracted the attention of researchers from videos and images and shows satisfactory performance on HAR. However, this approach is impractical when everyday life is considered because it will bring privacy concerns. The systems can only work in the specific area where the camera is located. Most studies on HAR are based on wearable inertial sensors. Although this approach has no privacy issues and is insensitive to the environment, wearing the sensors for a long time will bring inconvenience to our daily life and the sensor system will increase additional living expenses. Nowadays, smartphones have been widely utilized in our daily lives and our lives are highly dependent on smartphones. The inertial sensors embedded in smartphones is a good alternative for sensor-based HAR [9], [10], which can release us from wearing additional sensing components. Therefore, using the sensors from smartphones for HAR is a promising research area.

Although previous HAR studies have made some progress by using sensors from smartphones, there are still two crucial aspects that are closely related to the performance of the smartphone-based HAR system. First, the smartphone is usually put in the pocket and the signals of smartphone sensors would contain a lot of noise, which requires more robust and distinguished features obtained by using feature engineering for the final identification of human activities. Second, the original feature set contains some redundant features which will reduce the performance of the HAR system. Besides, the quality and the large size of the original feature set will bring high computational load hindering the real-time response of the smartphone-based HAR system.

Feature selection aims to optimize the HAR performance by selecting the discriminative features from the original feature set, which can benefit the recognition and generalization performance of a classifier, reduce the computational load and energy consumption for smartphone-based HAR system. In the previous studies, many feature selection methods have been proposed to deal with the feature selection problem in HAR [11]. These methods can be grouped into two main categories, the filter method and the wrapper method [12]. The wrapper methods utilize the performance index of the classification algorithm as the evaluation criterion of the feature subset. It is usually more effective compared with the filter method, however, if the dimensionality of the feature set is large, the iterative process of evaluating features using classification algorithms will be time-consuming. Contrary to the wrapper method, the filter methods utilize evaluation criteria to measure the relationship between the features and category labels and remove the features that have low thresholds. The filter methods such as ReliefF, maximal relevant minimal redundant (mRMR) and information gain have been widely applied to many fields, including remote sensing [13], EEG signals [14] and fault diagnosis [15] as they have many advantages including effectiveness, low computational load and independence of the machine learning classifier. However, most of the methods in this technique rank the features based on their evaluation criteria and usually ignore the dependency among the selected features, which may lead to select redundant features and results in low classification accuracy. To overcome the shortcomings above, multiple filter-based feature selection methods are adopted in this paper to obtain the optimized feature sets respectively which are further utilized for establishing the ensemble recognition system.

After feature optimization, various machine learning algorithms can be applied to recognize human activities. Recently ensemble learning algorithms have attracted much attention from scholars for HAR. These methods can acquire higher accuracy and generalization ability of the system with a small dataset by combining several diverse models. It has been demonstrated effective and robust in many researches on HAR [16], [17]. Although deep learning has achieved remarkable performance for HAR. With its stacking structure, it can even automatically learn representative features. However, deep learning algorithms essentially require a huge dataset for model training. Besides, the high computational load of deep learning makes it unsuitable for real-time smartphone-based HAR issue. Ensemble learning has been successfully applied in many real applications. Currently, the most common ensemble learning methods for HAR are bagging [18] and boosting [19]. However, some studies demonstrated the methods that utilize all features like bagging and boosting are not always superior to the methods that utilize a portion of the features such as random subspace [20]. This may be due to the overfitting problem caused by high-dimensional feature vectors or a small number of training samples. Inspired by this, we propose an ensemble learning method for HAR considering the optimized feature subsets generated from different filter-based feature selection methods, which aims to reduce the low conformity between the size of the training samples and the length of the feature set, hence reduce computation load and also acquire more robust and generalized performance.

The diversity among the base models is an important condition to ensure optimal performance of the ensemble. Therefore, we proposed a method called optimized subspace for generating base models. Unlike the randomly
constructed subspace method, the optimized subspace method utilized different optimized feature subsets from ReliefF, mRMR, and maximum relevancy maximum complementary (MRMC) for obtaining base models. Besides, the random subspace and optimized subspace methods are used simultaneously as manipulation on the feature set for diverse base models. What’s more, manipulation of the data set for the diversity of base models is also considered. We apply Bootstrap which is the most commonly utilized and successful ensemble learning method to create diverse training data subsets for each base model. Eventually, the diverse base models trained based on manipulations on the data set and feature set constitute the entire ensemble system.

Based on intuition and experience, the larger the ensemble, the better its performance. However, many studies [21, 22] demonstrated that the performance of the ensemble has no positive relationship with the number of base models. Besides, a very large number of base classifiers will require large memory storage and computational resources, which may not suitable for smartphone-based HAR. Thus, this paper proposes a novel selective ensemble learning method HDESEN which applies hybrid diversity enhancement for smartphone-based HAR. The proposed method HDESEN contains a hybrid diversity enhancement method which utilizes three types of diversities (Bootstrap, random subspace and optimized subspace) for constructing the base models to boost the performance of ensemble learning. Furthermore, an optimized subspace algorithm is proposed which utilizes three kinds of filter-based feature selection methods for base model construction. In the pruning phase, the IBGSO is proposed to conduct selective ensemble and optimal sub-ensemble with excellent performance and huge diversity is selected. The main contributions are as follows:

1. A novel selective ensemble approach HDESEN is proposed for smartphone sensor-based human activity recognition. The novelty of the HDESEN lies in its novel hybrid diversity enhancement method and the proposed IBGSO, which can boost the diversity and optimize the subset for optimal recognition accuracy.

2. The hybrid diversity enhancement method is proposed for constructing the base models in the HDESEN, which is the first method to consider training data subsets as well as feature space aspects to boost diversity and brings better generalization performance for selective ensemble learning.

3. The proposed HDESEN method could select superior base models for selective ensemble learning based on the proposed IBGSO, which is different from the previous studies in terms of its searching process. This makes it more appropriate to deal with smartphone sensor-based human activity recognition issues.

The rest of this paper is organized as follows. In Section II, related works about smartphone-based HAR and feature selection methods for HAR are introduced. In Section III, the proposed method is demonstrated in detail. The experimental setup and dataset are introduced in Section IV. In Section V, experiments are carried out to verify the effectiveness of the proposed method HDESEN and results are analyzed. Section VI concludes the paper.

II. RELATED WORKS

Various scholars have carried out research on the smartphone-based HAR. Li et al. [23] proposed a PSDRNN scheme for smartphone-based HAR to improve the efficiency in training and recognition. Compared with the most accurate DRNN-based HAR scheme, the PSDRNN reduces the average recognition and training time by 56% and 80% respectively. Wang et al. [24] compared the effectiveness of triaxial accelerometer and gyroscope in a smartphone-based HAR system when they were used simultaneously or separately and demonstrated that fusion of two kinds of sensors can easily achieve better recognition performance. Chen et al. [16] proposed a novel ensemble ELM algorithm for smartphone-based HAR system. The input weight of the basic ELM is initialized with Gaussian random projection, which can enhance the diversity of the ensemble system. The proposed method has been proven effective on two datasets. Chen et al. [9] proposed a robust smartphone-based HAR system based on CT-PCA and OISVM to deal with the poor performance caused by orientation, placement, and subject variations. Experiments have demonstrated the generalization ability of CT-PCA scheme on the data from unseen orientation and the effectiveness of OISVM on placement and subject variations.

In the past few years, a large number of feature selection methods have been proposed for HAR. Turker et al. [25] utilized ensemble residual networks as a feature extractor to obtain 3000 features for HAR and applied ReliefF to select 1000 most discriminative features from the original feature set in the feature selection phase. A game theory-based feature selection method which is based on entropy and mutual information theory is proposed by Wang et al. [26] to evaluate the acceleration features from waist and ankle in HAR. The proposed feature selection method has been demonstrated to select fewer features and provide higher accuracy compared with the ReliefF and mRMR. Zdravevski et al. [27] propose a two-phase feature selection approach which firstly considers the importance and drift sensitivity of the features and then utilizes diversified forward-backward method to evaluate the quality of a feature set. Application to five publicly available data sets demonstrated that it yielded better accuracy than when using hand-tailored features. Ghasemzadeh et al. [28] proposed the notion of power-aware feature selection for minimizing energy consumption in mobile-based HAR. The method utilizes a graph model to represent the computing complexity of the features and a greedy approximation approach is applied to select the features that can have low computation complexity. Experimental results demonstrate that the features selected by the proposed approach can significantly reduce energy consumption for the wearable sensor network-based HAR.
FIGURE 1. The flowchart of the proposed HDESEN approach.

III. PROPOSED METHODOLOGY
A. OVERVIEW OF THE PROPOSED FRAMEWORK
As mentioned before, the smartphone-based HAR system requires not only high accuracy and good generalization ability but also low computation load. All these facts would have a crucial impact on the user experience. In general, ensemble learning algorithms showed optimal performance when compared with the traditional machine learning algorithms [16], [25]. The generalization performance of ensemble learning algorithms is highly related to the effectiveness and diversity of base models. Therefore, we propose an optimized subspace method which utilizes the optimized feature subsets to obtain diverse base models. The pool of base models is established on manipulation of feature vectors including optimized subspace and random subspace methods and manipulation of training sample using Bootstrap re-sampling method. Besides, eliminating the redundant base models and finding the most optimal sub-ensemble will benefit the compact recognition system. Thus, IBGSO is proposed in this paper as a modified GSO pruning method to select the superior sub-ensemble. The architecture of the proposed HDESEN approach is shown in Figure 1. The HDESEN approach consists of four parts: data preprocessing, hybrid diversity enhancement, IBGSO-based pruning and activity recognition. In the following subsections, the main components of the proposed approach are described.

B. FEATURE EXTRACTION
In the smartphone-based activity recognition, features are extracted from the triaxial accelerometer and triaxial gyroscope embedded in a smartphone. For a triaxial accelerometer or a triaxial gyroscope, their readings are composed of data corresponding to the three axes of \( x \), \( y \), and \( z \). The readings are divided into vectors of length \( N \) through the sliding window, from which we can extract the various features that are previously demonstrated to be effective for activity recognition. For each axis of the triaxial accelerometer and triaxial gyroscope, the mean value, standard deviation, maximal value, minimal value, median absolute deviation, signal magnitude area, energy, variance, interquartile range, and autoregression coefficients are extracted from the time-domain signals of the \( N \) readings in a sliding window. Furthermore, frequency features are also extracted using the Fast Fourier Transform (FFT). These extracted frequency domain features include skewness of the signal, kurtosis of the signal, and the signal energy and entropy. The features extracted in this work are shown in Table 1. Particularly, we merged the features for both acceleration and gyroscope into a single vector after features extraction.

| Domain     | Features                                |
|------------|-----------------------------------------|
| Time       | Mean value, Standard deviation, Maximal value, Minimal value, Median absolute deviation, Signal magnitude area, Energy, Variance, Interquartile range, Autoregression coefficients |
| Frequency  | Fast Fourier Transform (FFT), Skewness, Kurtosis, Energy, Entropy |

C. FEATURE SELECTION
1) A BRIEF INTRODUCTION TO MRMC
The evaluation criterion is the core of the filter-based feature selection method. Most filter-based feature selection methods ignore how a feature complements the already selected features, which results that the number of selected features is always larger than actually required. MRMC is a novel filter-based feature selection method that considers the relevance and complementary among features, which has been demonstrated to be effective for sensor-based HAR [11].
Mutual information (MI) is based on information theory, measures the dependency between two variables. The MI value is zero if and only if the variables are independent. Compared with other commonly used feature selection methods based on MI, such as MRMR and NMIFS, MRMC utilizes complementary measurement to detect the feature redundancy so that the redundant features can be represented by a low complementary score. It shows many advantages in dealing with data with a large number of features.

\[ \text{Rel}_{fi} = 0.8 \] means that the network’s performance will decrease by 80% if the feature \( f_i \) is not selected.

**c: COMPLEMENTARY SCORE**

The relevance measurement only focuses on the relationship between the feature and the classes and neglects the relationship among features such as redundancy and complementarity. Therefore, relevance measurement tends to select features that have strong discriminatory as individuals but are weak as a group. The complementary measurement takes the relationship between the features and the selected feature set into consideration so that the redundancy features which do not benefit the performance will be assigned a low complementary score. The baseline performance is based on the network trained on the selected feature set \( S \) and a new feature \( f_i \) is added to the network which is referred as the complementary network.

Given an already selected feature set \( S \), the complementary score of feature \( f_i \) to \( S \) can be calculated as:

\[ \text{Com}_{fi} = \frac{P(S \cup f_i)}{P(S)} - 1 \] (3)

where \( P(S \cup f_i) \) denotes the generalized performance of the complementary network and \( P(S) \) is the generalized performance of the baseline network. The complementary score indicates how much the new feature \( f_i \) contributes to the baseline network. For example, \( \text{Com}_{fi} = 0.3 \) implies that selecting the feature \( f_i \) can improve the performance of the baseline network by 30%.

**d: THE RELEVANCE-COMPLEMENTARY (RC) SCORE**

The MRMC ranks the features based on their relevance and complementary scores. According to the relevance and complementary scores, the RC score can be computed as follows:

\[ \text{RC}_{fi} = \text{Rel}_{fi} + \text{Com}_{fi} \] (4)

The feature is then selected based on the maximum RC score. In the MRMC algorithm, the complementary measurement can reduce the chance of selecting overlapping or redundant features and the RC evaluation criteria can effectively improve the discrimination of feature subsets and reduce redundant features. The detail steps of the MRMC method can be found in [11].

2) A BRIEF INTRODUCTION TO ReliefF

ReliefF which was proposed by Kononenko in 1994 [30] is an extension of Relief to handle multiclass problems. ReliefF selects the optimal feature subset based on the relevance between features and classification. Each feature will be assigned a weight and ReliefF ranks the features according to the descending order of the weights. The weight is the core of the ReliefF which is calculated based on the correlation between the feature and the category label. The correlation between the feature and the category label reflects the ability of the feature to distinguish samples. The features whose weights are greater than a certain threshold will be selected.
Relieff can deal with multiclass problems with more robustness and be able to deal with incomplete and noisy data.

When dealing with multiclass problems, the class label of training set is $C = \{c_1, c_2, \ldots, c_l\}$. Firstly, Relieff randomly selects a sample $R_k$ from the training dataset. Then, it searches for $k$ nearest neighbors (denoted by set $H_j$) of $R_k$ from the same class and also $k$ nearest neighbors (denoted by set $M(c)$) from each of different classes. Lastly, the Relieff algorithm repeats these two steps $m$ times. The weight of the feature $t$ is updated by the following equation:

$$ W_t = W_t - \sum_{j=1}^{k} \left[ \text{diff}(t, R_i, H_j)/(m \cdot k) \right] $$

$$ + \sum_{c \neq \text{class}(R_k)} \left[ \frac{p(c)}{1 - p(\text{class}(R_i))} \times \sum_{j=1}^{k} \text{diff}(t, R_i, M_j(c)) \right] $$

where $p(c)$ is the proportion of samples belonging to category $c$ in the training samples, $p(\text{class}(R_i))$ is the proportion of samples of the same class as $R_i$ to the total samples, $\text{diff}(t, R_1, R_2)$ represents the difference between the sample $R_1$ and the sample $R_2$ in the feature $A$ and can be calculated as:

$$ \text{diff}(t, R_1, R_2) = \left| \frac{R_{gt} - R_{jt}}{\max(t) - \min(t)} \right| $$

where $R_{gt}$ and $R_{jt}$ are the $t$th feature of samples $R_1$ and $R_2$ respectively, and $\max(t)$ and $\min(t)$ are the maximum and minimum values of the corresponding feature $t$ in all the samples.

3) A BRIEF INTRODUCTION TO mRMR

The mRMR is a filter-based method, which originates from MI. It utilizes relevance and redundancy as the evaluation criteria to optimize the feature set. Specifically, the algorithm firstly utilizes MI to measure the relevance between features and classes and then calculate the redundancy between every two features. The features will be ranked according to the minimal redundancy and maximal relevance criterion. The mRMR algorithm is described as follows:

Given two discretized variables $f_i$ and $f_j$ with probability density $p(f_i)$ and $p(f_j)$ respectively, and their joint probability density is $p(f_i, f_j)$, then the mutual information between $f_i$ and $f_j$ can be expressed as:

$$ MI(f_i, f_j) = \sum_{i} \sum_{j} p(f_i, f_j) \log \frac{p(f_i, f_j)}{p(f_i)p(f_j)} $$

Given set $S = \{\}$ which contains the selected features and feature set $F = \{f_1, f_2, f_3, f_j, \ldots, f_N\}$ contains $N$ features. The mRMR selects the feature $f_j$ in $F$ according to the maximal relevance criterion as shown in Eq. (8). The selected features in feature set $S$ have the largest dependency of the target class $c$.

$$ \max D(S, c), \quad D = \frac{1}{|S|} \sum_{f_i \in S} MI(f_i, c) $$

where $|S|$ is the number of the selected features. $D$ represents the mean value of MI between each feature in the feature set $S$ and the target class $c$. Minimal redundancy criterion which is utilized to reduce the redundancy of the feature set can be expressed as Eq. (9).

$$ \min R(S), \quad R = \frac{1}{|S|^2} \sum_{f_i, f_j \in S} MI(f_i, f_j) $$

where $R$ is the MI value between the features in $S$, which represents the redundancy between the features. Lastly, based on the evaluation criteria mentioned above. The mRMR selects the features in $F$ according to the Eq. (10) which ensures the maximum correlation between the feature set and the sample class and the minimum redundancy between the features.

$$ \max \Phi(D, R), \quad \Phi = D - R $$

D. THE PROPOSED HYBRID DIVERSITY ENHANCEMENT METHOD

In this paper, we generate diverse base models based on the three methods which originate from random samples, randomly selected features, and different optimized feature sets. To take advantages of these methods, the ELM is chosen as the base model due to its fast learning, high diversity and good generalization performance which make it suitable for ensemble learning.

Firstly, bootstrapping is utilized to generate base models with random samples. It creates multiple sets of samples by random sampling from the original training data. Then, the base models are trained on the above multiple sets of samples. Therefore, we can obtain multiple base models with diversity based on different training subsets. The procedure of establishing base models with bootstrap is illustrated in Algorithm 1.

### Algorithm 1 Framework of Bootstrap

**Input:** $T$: Training set; $L$: Base model; $N^B$: Number of iterations

**Output:** Multiple base models $\{CB n|n = 1, \ldots, N^B\}$

1. for $n = 1$ to $N^B$
2. $S_n = \text{bootstrap sample from } S$.
3. $CB n = L(S_n)$.
4. end.

Secondly, the diverse base models are trained on randomly selected features. We utilize the random subspace [31] method which has the advantages of bootstrapping and aggregation to generate base models. Different from the bootstrap method of bootstrapping training samples, the random subspace performs bootstrapping on the feature space. We can utilize random subspace to obtain a new feature subset generated randomly from the original feature set, and then a base model can be trained on the new feature subset. The base models with diversity can be established by repeating the random sampling applied to the feature space. The procedure of
establishing base models with random subspace is illustrated in Algorithm 2.

Algorithm 2 Framework of Random Subspace

**Input:** $F$: Feature set; $L$: Base model; $N^R$: Number of iterations

**Output:** Multiple base models \{ $CR n | n = 1, \ldots, N^R$ \}

1. for $n = 1$ to $N^R$
2. $F_n =$ bootstrap feature from $F$.
3. $CR n = L(F_n)$.
4. end.

Thirdly, inspired by random subspace, this paper proposes a new method to generate diverse models based on different optimized feature sets. The feature subset is not generated by random sampling applied to the feature space like the random subspace, it is obtained through the ranking result of the filter-based feature selection methods that three methods that mentioned above. After using the ranking results of the filter-based feature selection methods to sort the features in descending order of weight, we can train a base model using the first $m$ important features of a filter-based feature selection method. Using new subsets obtained by setting different values of $m$, we can derive multiple base models with diversity. According to the name of the random subspace, we call this new method optimized subspace, because each subset is the optimization result of a filter-based feature selection method. The procedure of establishing base models with optimized subspace is illustrated in Algorithm 3.

Algorithm 3 Framework of Optimized Subspace

**Input:** $O$: Optimized feature set from feature selection method; $m$: number of important features; $L$: Base model; $N^O$: Number of iterations

**Output:** Multiple base models \{ $CO n | n = 1, \ldots, N^O$ \}

1. for $n = 1$ to $N^O$
2. $O_n =$ select first $m$ features from $O$.
3. $CO n = L(O_n)$.
4. end.

E. ENSEMBLE PRUNING USING THE PROPOSED IBGSO METHOD

The GSO algorithm [32] inspired by the lighting behavior of glowworms has the advantages of easy implementation, strong robustness, and strong search ability. GSO searches for the optimal solution utilizing a population composed of randomly generated glowworms in a solution space. In the population, each glowworm represents a feasible solution to the optimization problem. The luciferin value of the glowworm is updated according to its fitness value. After several iterations, we can have the optimal glowworm in the population, which is the optimal solution of the optimization problem. The basic steps of GSO can be found in [32]. However, it still has some drawbacks. For example, it is easy to get in the local optimum and has a slow convergence rate. To overcome the above weaknesses and improve the searching efficiency, we propose the IBGSO by modifying its searching processes of the basic GSO. Firstly, the moving way of glowworms is improved so that GSO can search in a binary discrete space. Secondly, the search behavior of GSO is modified, which can increase the randomness of the algorithm and ensure that the algorithm avoids falling into the local optimum. Finally, crossover operation and mutation operation are introduced to improve the search efficiency of the algorithm. These improvements are detailed as follows:

1) IMPROVEMENT OF GLOWWORM MOVING

GSO cannot handle the discrete combinatorial optimization problem, it is necessary to modify the moving way of glowworms. In this paper, the position of glowworms is changed by using probability. In the $rth$ iteration of the algorithm, let $X_i(t) = [x_{i1}(t), x_{i2}(t), \ldots, x_{in}(t)]$ be the position of the current glowworm and $X_i(t) = [x_{i1}(t), x_{i2}(t), \ldots, x_{in}(t)]$ be the position of the target glowworm that $X_i(t)$ will move to. The location update process can be expressed by the following formula.

$$x_{id}(t+1) = \begin{cases} x_{id}(t), & r \leq p_1 \\ x_{id}(t), & p_1 < r < p_2 \\ r_0, & r \geq p_2 \end{cases}$$

(11)

where $p_1$ and $p_2 \in [0, 1]$ are both selected parameters for the update formula, $r$ is a random number uniformly distributed between $(0, 1)$ and $r_0$ randomly takes the value of 0 or 1, $d = 1, 2, \ldots, N$, $N$ is the number of glowworms.

2) IMPROVEMENT OF SEARCH BEHAVIOR

In order to improve the convergence speed and the performance of the algorithm, this paper improves the search behavior as follows. In the $rth$ iteration, the glowworm $X_i(t)$ respectively moves to the best position in the bulletin board, the optimal position of glowworm in the decision domain and a random position in the decision domain. These positions are marked as $X_i'(t+1), X_i''(t+1)$ and $X_i'''(t+1)$. Then, the best one of the $X_i'(t+1), X_i''(t+1)$ and $X_i'''(t+1)$ will be the position of $X_i(t+1)$.

3) CROSSOVER OPERATION

In order to increase the population diversity in the iterative process, crossover operations is added. The crossover operation for the $dth$ dimension of the $ith$ glowworm is:

$$x_{id} = \begin{cases} x_{jd}, & \text{if rand} \geq C_j \\ x_{id}, & \text{otherwise} \end{cases}$$

(12)

where $j = 1, 2, \ldots, N$, but $j \neq i$, the rand is a random number uniformly distributed between $(0, 1)$, the crossover probability $C_j$ is dynamically changing and is defined as:

$$C_j = 0.3 \times F_{i, opt}$$

(13)

$$F_{i, opt} = \frac{F_i - F_{opt}}{F_{worst} - F_{opt}}$$

(14)
where $F_i$ is the fitness value of the $i$th glowworm, $F_{\text{worst}}$ and $F_{\text{opt}}$ are the worst fitness and optimal fitness value of the current iteration. Through formula (12)-formula (14), the proposed IBGSO can ensure that the current optimal individual will not change, and the crossover probability is inversely proportional to the fitness value.

4) MUTATION OPERATION
In order to further improve the diversity of the population and prevent the algorithm from falling into the local optimum, the mutation operation is added after the crossover operation. By performing mutation operations, the glowworms can jump out of the local optimal solution, which helps the algorithm expand the search space and improve the possibility of finding a better value. The crossover operation for the $d$th dimension of the $i$th glowworm is:

$$x_{id} = \begin{cases} x_{g\text{best},d} + r_3(x_{q1,d} - x_{q2,d}) & \text{rand} \geq \eta \\ x_{id} & \text{otherwise} \end{cases}$$

(15)

And

$$\eta = 0.2 \times F_{i,\text{opt}}$$

(16)

where $q_1, q_2 = 1, 2, \ldots, N$, but $q_1 \neq q_2 \neq i$; $x_{g\text{best},d}$ is the optimal glowworm of $d$ dimension in the entire iteration process so far, the $\text{rand}$ is a random number uniformly distributed between $(0, 1)$ and $\eta$ is the mutation probability. If the fitness value of the current individual is greater than the global optimal individual, it will replace the global optimal individual. According to formula (15)-(16), the mutation probability of the global optimal individual is 0, and the mutation probability of the worst individual is 0.2. Based on the above analysis, the proposed IBGSO algorithm is presented as Algorithm 4, and the architecture of the IBGSO is shown in Figure 2.

Algorithm 4 IBGSO

**Input:** the parameters of IBGSO

**Output:** the sub-ensemble $X_{\text{opt}}$ and its fitness value $F_{\text{opt}}$

1. Initialize the parameters of IBGSO
2. Generate $N$ glowworms by random method and compute their fitness values
3. $X_{\text{opt}} \leftarrow \max \{F(1), F(2), \ldots, F(N)\}; F_{\text{opt}} \leftarrow \max \{F(1), F(2), \ldots, F(N)\}$
4. $t = 1$
5. While $t \leq T_{\text{max}}$ do
6. Update the fitness value of each glowworm
7. Construct the neighborhood set $N(t)$
8. Calculate the global optimal solution $X_{\text{opt}}$
9. Determine moving targets based on improved search behavior
10. Update the position of $X_{i}$ according to formula (11) and assign the optimal position to the current glowworm
11. Update the dynamic decision radius
12. Implement the crossover operation to create new glowworms and update the current glowworm $X_{i}$
13. Implement the mutation operation to jump out of the local optimal solution, and update the current glowworm $X_{i}$
14. end for
15. $X_{\text{opt}} \leftarrow \max \{F(1), F(2), \ldots, F(N)\}$
16. $F_{\text{opt}} \leftarrow \max \{F(1), F(2), \ldots, F(N)\}$
17. $t = t + 1$
18. end while
19. return $X_{\text{opt}}$ and $F_{\text{opt}}$

5) THE COMPLEXITY OF THE IBGSO METHOD
Assuming that the initial population size of IBGSO is $n$, the time complexity of IBGSO is analyzed as follows.

1) Initialization of population and parameters of IBGSO, and its cost is $O(n)$;
2) Searching process of glowworms, and its cost at each iteration is $O(n^2)$;
3) The crossover operation of glowworms, and its cost at each iteration is $O(n)$;
4) The mutation operation of glowworms, and its cost at each iteration is $O(n)$;
5) Iteration process of IBGSO, and its total computational cost after $T_{\text{max}}$ iterations is $O(T_{\text{max}} \times n^2)$.

In general, the total time complexity of IBGSO is $O(T_{\text{max}} \times n^2)$.

IV. EXPERIMENTAL SETUP

A. EXPERIMENTAL SETUP AND EXPERIMENTAL DATASET
In order to evaluate the performance of our proposed HAR approach, we have carried out extensive experiments on public dataset collected using a Samsung Galaxy SII smartphone with 3-D accelerometer and gyroscope sensor [33]. The 3-D
accelerometer is STMicroelectronics K3DH, which is accurate to ±2G with resolution = 0.0625 (G is the gravitational constant) and the gyroscope sensor is K3G with the maximum range of 8.72665 and resolution = 0.000305433. They all work at a sampling frequency of 50Hz. Thirty subjects are involved in the data collection and their ages range from 19 to 48. They performed activities including walking (W), upstairs (US), downstairs (DS), sitting (S), standing (ST) and lying (LY) with a smartphone attached to the waist. The dataset was pre-processed and segmented with a sliding window of 2.56s with an overlap of 50%. Therefore, there were 128 data points in a single window for feature extraction. The statistical features in Table 1 are extracted. In total, 10 299 samples were obtained, where 5885 of them are used for training, 1471 of them are used for validation and the remaining for testing.

In addition, we also utilized the WISDM dataset [34] to test the performance of the proposed HDESEN HAR method. The WISDM dataset contains smartphone sensor data from 51 subjects, each of whom was asked to perform 18 daily activities, including 5 simple activities and 13 complex activities. The WISDM dataset was collected using three-axis accelerometer and gyroscope sensors to capture sensor data at a constant rate of 20 Hz while each subject performed 3 minutes of activity.

### B. PERFORMANCE MEASURES

To evaluate the effectiveness of the proposed method and show its superiority over the comparative methods, we compare them in terms of the number of selected features and obtained recognition performance which will be measured by the following four measures:

The accuracy measure is used to evaluate the performance of the proposed method, which can be expressed as:

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]

where \(TP, TN, FP,\) and \(FN\), respectively, represent the number of true positive, true negative, false positive, and false negative outcomes in a given experiment.

Precision and recall are defined as measuring the recognition rate of records correctly classified from a class of total positive records and the recognition rate of records correctly classified from a class of total true samples in a class, respectively.

\[
\text{precision} = \frac{TP}{TP + FP} \quad (18)
\]

\[
\text{recall} = \frac{TP}{TP + FN} \quad (19)
\]

In addition, F1 evaluation criterion is also considered. F1 is defined as the combination of precision and the recall, which is defined as follows:

\[
F1 = \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \quad (20)
\]

### V. EXPERIMENTAL RESULTS

In this paper, 45 diverse base models, including 15 models from Adaboost, 15 models from random subspace and 15 models from optimized subspace (5 for each feature selection method) are constructed. The parameters of the IBGSO are set as follows: fluorescein volatilization factor \(\rho = 0.4\), fluorescein update rate \(\gamma = 0.6\), dynamic decision domain update rate \(\beta = 0.08\), threshold \(n_1 = 5\) and the maximum number of iterations \(s_{\text{max}} = 300\), \(p_1 = 0.15\), \(p_2 = 0.85\). The experiments are implemented in Matlab 2014a using a computer with a 2.8 GHz processor and 6 GB memory.

### A. COMPARISON WITH SINGLE DIVERSITY ENHANCEMENT METHOD

In order to ensure the fairness of the comparison, each diversity enhancement method produces 45 different basic models, which is consistent with the number of the proposed hybrid diversity enhancement method in this paper. Table 2 shows the average training accuracy, verification accuracy and test accuracy of different diversity enhancement methods. These results are the average of 10 trials. From Table 2 we can see that the average training accuracy, validation accuracy and testing accuracy of the proposed hybrid diversity enhancement method are 99.25%, 99.04% and 98.92%, respectively, whereas the optimal performance of the single diversity enhancement method achieves 94.67%, 92.98% and 91.15% for training accuracy, validation accuracy and testing accuracy respectively. Besides, the corresponding standard deviations of the proposed hybrid diversity enhancement method are 0.43%, 0.54%, and 0.56%, respectively, whereas the smallest ones of the single diversity enhancement methods are 1.16%, 1.34%, and 1.08%, respectively.

Figure 3 shows the specific testing accuracies of different methods for 10 trials. The following three conclusions can be drawn from Figure 3. Firstly, the proposed hybrid diversity enhancement method achieves the highest test accuracy in most trials. Secondly, while obtaining optimal recognition results, the performance of our proposed method in 10 experiments is relatively stable compared to other methods, so it can also achieve a better balance between the accuracy and robustness compared with other methods. While obtaining optimal recognition results, the performance of our proposed method in 10 experiments is relatively stable compared to...

### TABLE 2. The average results comparison of the proposed hybrid method and other methods for 10 trials.

| Method            | Training accuracy | Validation accuracy | Testing accuracy |
|-------------------|-------------------|---------------------|-----------------|
| Adaboost          | 93.82±1.16        | 92.98±1.34          | 90.52±1.16      |
| Random subspace   | 93.86±1.89        | 91.72±1.95          | 88.56±2.05      |
| Optimized subspace| 94.67±1.38        | 92.95±1.46          | 91.15±1.08      |
| Proposed method   | 99.25±0.43        | 99.04±0.54          | 98.92±0.56      |
other methods. Thirdly, the performance of the random subspace method in 10 trials varies greatly. These results further demonstrate that the proposed hybrid diversity enhancement method can acquire more robust and higher accuracy than the other single diversity enhancement methods.

To further demonstrate the effectiveness of the proposed hybrid diversity enhancement method, the more specific results of different methods for trial 4 are shown as follows. Figures 4-6 show the testing precision rate, recall rate and F1 score comparison of the proposed method and other single diversity enhancement methods for trial 4, respectively. From Figures 4-6, three points can be intuitively acquired. Firstly, the results show that the performances of the proposed hybrid diversity enhancement method are higher than that single diversity enhancement methods in most cases. Secondly, for activity type upstairs and downstairs, the performances of other methods are relatively low, whereas the performances of the proposed hybrid diversity enhancement method are improved to about 90%. Thirdly, the performance of the proposed method is greatly improved than other single diversity enhancement methods, especially for the activity upstairs, downstairs and lying. This confirms that the performance of the proposed hybrid diversity enhancement method is superior to the single diversity enhancement method.

B. COMPARISON WITH OTHER PRUNING METHODS

The superiority of the proposed method HDESEN mainly includes two aspects i.e. hybrid diversity enhancement method for base model construction and selective ensemble through the proposed IBGSO. In order to further demonstrate the superiority of the proposed method HDESEN, the experiment is carried out to compare the proposed pruning method IBGSO with several state-of-the-art pruning approaches including GA [35], BGSO [36] and BAFSA [37]. In this experiment, the hybrid diversity enhancement method produces base models and these four methods including the proposed IBGSO, GA, BGSO and BAFSA are utilized to select the optimal subset for ensemble learning. The population size with these methods and the parameters with BGSO and IBGSO are set to the same. Other parameters are set according to their respective papers.

Table 3 shows the average verification accuracy, test accuracy, and ensemble size of different pruning methods. From Table 3 we can see that the average validation accuracy, testing accuracy and ensemble size of the proposed method IBGSO are 99.04%, 98.92% and 13.7, respectively, whereas the optimal performance of other comparative pruning methods achieves 92.82%, 91.06% and 17.6 for verification accuracy, testing accuracy and ensemble size. Besides, the corresponding standard deviations of the proposed method are 0.54%, 0.56%, and 1.4%, respectively, whereas the
TABLE 3. The average results comparison of different pruning methods for 10 trials.

| Methods  | Validation accuracy | Testing accuracy | Ensemble size | Run time (s) |
|----------|---------------------|------------------|---------------|--------------|
| GA       | 90.32±1.56          | 88.34±1.38       | 33.2±1.9      | 19.43±1.23   |
| BAFSA    | 91.87±1.37          | 89.56±1.49       | 23.9±2.8      | 17.5±0.67    |
| BGSO     | 92.82±1.42          | 91.06±1.45       | 17.6±2.8      | 16.49±0.85   |
| Proposed IBGSO | 99.04±0.54   | 98.92±0.56       | 13.7±1.4      | 12.17±0.75   |

FIGURE 7. The testing accuracy comparison of different methods for 10 trials.

Performances of the optimal comparative pruning method are 1.37%, 1.38%, and 1.9%, respectively. The running time and deviation of the proposed IBGSO are less than the comparative methods, indicating the stability and effectiveness of IBGSO. In summary, the overall performance of IBGSO is significantly superior to that of the other 3 state-of-the-art methods. Therefore, compared with the other pruning methods, the proposed method IBGSO can achieve optimal performance and robustness with the smallest size of the ensemble.

Figure 7 shows the specific testing accuracy comparison of different pruning methods for 10 trials. It can be seen from Figure 7 that the proposed IBGSO method generally achieves the highest testing accuracy in each of the 10 trials. Additionally, Figure 8 compares the number of selected base models of different pruning methods for 10 trials. Obviously, it can be seen from Figures 7-8 that the proposed method usually obtains the highest testing accuracy in each of the 10 trials with the minimum number of base models, which demonstrate the effectiveness of the proposed IBGSO.

Figure 7 shows the specific testing accuracy comparison of different pruning methods for 10 trials. It can be seen from Figure 7 that the proposed IBGSO method generally achieves the highest testing accuracy in each of the 10 trials. Additionally, Figure 8 compares the number of selected base models of different pruning methods for 10 trials. Obviously, it can be seen from Figures 7-8 that the proposed method usually obtains the highest testing accuracy in each of the 10 trials with the minimum number of base models, which demonstrate the effectiveness of the proposed IBGSO.

In order to further show the performance of the proposed method HDESEN, Tables 4–6 present the precision (P), recall (R) and F1 score of the four pruning methods for trial 4, respectively. It can be observed from Tables 4–6 that the proposed pruning method IBGSO generally achieves better performance than the other three methods for the six types of activities. Specifically, in terms of walking, upstairs and downstairs, the performance improvement of the proposed pruning method IBGSO is quite obvious. For example, as for upstairs, the precision rate of the proposed method is improved to 90.34% from 89.98% of BGSO, 86.19% of

TABLE 4. The precision rates comparison of different pruning methods for trial 4.

| Method  | P (%) values of different activity types |
|---------|-----------------------------------------|
| W       | US          | DS          | S           | ST          | LY          |
| GA      | 85.63       | 86.19       | 83.02       | 91.93       | 93.95       | 98.51       |
| BAFSA   | 86.10       | 86.29       | 83.85       | 92.93       | 96.12       | 99.44       |
| BGSO    | 88.74       | 89.98       | 89.18       | 90.41       | 93.95       | 99.07       |
| Proposed | 90.20       | 90.34       | 89.93       | 93.45       | 95.95       | 99.81       |

TABLE 5. The recall rates comparison of different pruning methods for trial 4.

| Method  | R (%) values of different activity types |
|---------|-----------------------------------------|
| W       | US          | DS          | S           | ST          | LY          |
| GA      | 87.88       | 82.34       | 85.00       | 95.11       | 90.58       | 98.51       |
| BAFSA   | 90.10       | 81.70       | 84.05       | 96.33       | 93.22       | 99.25       |
| BGSO    | 90.71       | 87.87       | 88.33       | 94.09       | 90.58       | 99.63       |
| Proposed | 91.11       | 89.57       | 89.29       | 95.93       | 93.79       | 100        |

TABLE 6. The F1 scores comparison of different pruning methods for trial 4.

| Method  | F1 (%) values of different activity types |
|---------|-----------------------------------------|
| W       | US          | DS          | S           | ST          | LY          |
| GA      | 86.74       | 84.22       | 84.00       | 93.49       | 92.23       | 98.51       |
| BAFSA   | 88.05       | 83.93       | 83.95       | 94.58       | 94.65       | 99.34       |
| BGSO    | 89.71       | 88.91       | 88.75       | 92.21       | 92.23       | 99.35       |
| Proposed | 90.65       | 89.95       | 89.61       | 94.67       | 94.86       | 99.90       |
GA and 86.29% of BAFSA. Similarly, as for walking, the recall rate of the proposed method IBGSO is improved to 91.11% from 90.71% of BGSO, 90.10% of BAFSA and 87.88% of GA. For the activity type stand, even though the precision rate of the proposed method is not the highest, its F1 score which gives comprehensive evaluation is still the best among the four methods. These results future demonstrate that the proposed pruning method is superior to the other three pruning methods.

Additionally, in order to gain a better insight into the proposed method HDESEN, the confusion matrixes of different pruning methods for trial 4 are presented in Figure 9. According to the results, it can be seen that the proposed IBGSO contributes to improve the discrimination of similar activities, such as sit and stand, upstairs and downstairs. For example, according to Figure 9, we can see that the GA method makes 59 errors in distinguishing between sitting and standing, the BGSO method makes 63 errors and the proposed method reduces to 50 errors. The activities of sitting and standing were not differentiated highly effectively possibly due to their similarity from the viewpoint of the smartphone sensors. Similarly, we can also see that in differentiating upstairs and downstairs, the proposed method only results in 43 errors in comparison to the 45 errors of the BGSO method, 78 errors of the BAFSA method and 70 errors of the GA. Furthermore, we can also observe that the proposed method can distinguish between dynamic activity (walking, upstairs and downstairs) and static activity (sitting, standing and lying) with higher performance. Specifically, activity recognition using the proposed method makes 6 classification errors out of the 2943 test samples, compared to the 8 errors of using the BGSO method, 15 errors of using the BAFSA method and 22 errors of using the GA method.

### Table 7. The average results comparison of different pruning methods on the WISDM dataset.

| Methods  | Validation accuracy | Testing accuracy | Ensemble size | Run time (s) |
|----------|---------------------|------------------|---------------|--------------|
| GA       | 89.75±1.54          | 88.32±1.78       | 36.7±2.6      | 23.64±3.15   |
| BAFSA    | 91.38±1.64          | 90.79±1.76       | 25.3±1.8      | 20.51±1.67   |
| BGSO     | 94.87±1.27          | 95.32±1.38       | 19.6±2.1      | 19.72±2.73   |
| Proposed IBGSO | 98.62±0.43  | 98.25±0.56       | 15.7±1.3      | 15.38±1.27   |

### Table 8. Performance comparison of different pruning methods on the WISDM dataset.

| Methods  | Acc  | Pre  | Rec  | F1   |
|----------|------|------|------|------|
| GA       | 88.32% | 88.43%  | 88.24%  | 88.33%  |
| BAFSA    | 91.38% | 91.41%  | 91.27%  | 91.34%  |
| BGSO     | 94.87% | 94.86%  | 94.87%  | 94.87%  |
| Proposed IBGSO | 98.25% | 97.71%  | 98.17%  | 97.94%  |
C. ADDITIONAL EXPERIMENT ON THE WISDM DATASET

In order to comprehensively verify the performance of the proposed HDESEN method, the WISDM dataset was utilized in the experiment. Table 7 shows the average verification accuracy, test accuracy and ensemble size of different pruning methods on the WISDM dataset. From Table 7 we can see that the average validation accuracy, testing accuracy and ensemble size of the proposed method IBGSO are 98.62%, 98.25%, and 15.7, respectively, whereas the optimal performance of other comparative pruning methods achieves 94.87%, 95.32% and 19.6 for validation accuracy, testing accuracy, and ensemble size. Besides, the corresponding standard deviations of the proposed method are 0.43%, 0.56%, and 1.3%, respectively, whereas the standard deviations of the optimal comparative pruning method are 1.27%, 1.38%, and 1.8%, respectively. The running time and deviation of the proposed IBGSO are 15.38 and 1.27, respectively, which are the optimal among the comparative methods. This demonstrated the stability and effectiveness of the IBGSO on the WISDM dataset.

Table 8 presents the performance comparison of different pruning methods on the WISDM dataset. It can be observed from Table 8 that the proposed pruning method IBGSO achieves 98.25% accuracy, 97.71% precision, 98.17% recall and 97.94% F1. The proposed pruning method IBGSO generally achieves better performance than the other comparative methods, which shows the effectiveness of the IBGSO on the WISDM dataset.

Figure 10 shows the corresponding confusion matrix of our proposed method HDESEN on the unbalanced WISDM dataset. The activities of sitting and jogging were not differentiated highly effectively as we have observed in Figure 10, possibly due to their similarity in data and it is difficult to mine deeper information only by smartphone-based sensor data.

D. COMPARISON WITH OTHER HAR METHODS

In this section, we have compared the proposed HDESEN with some state-of-the-art approaches in the literature. Table 9 lists the performance comparison of some studies using the two datasets. In relation to the WISDM dataset, our proposed HDESEN method performed better than other works, while in relation to the UCI-HAR dataset, the proposed method attained accuracy of 98.9%. By comparison, the maximum accuracy achieved for the UCI-HAR dataset was 97.4% via using ensemble ELM. Therefore, the proposed HDESEN can achieve an optimal result with respect to smartphone sensor-based HAR and displays better performance compared with the earlier research on WISDM and UCI-HAR datasets.

VI. CONCLUSION

This paper proposes a novel smartphone sensor-based HAR method HDESEN which is based on hybrid diversity enhancement and selective ensemble learning. The proposed method HDESEN is different from the ensemble learning algorithm that relies on an individual method to build a set of models and a combination of all base models. To create more diversity for boosting the performance of ensemble learning, we propose to utilize three filter-based feature selection methods for initialization of the optimized subspace algorithm. Then, bootstrap, random subspace algorithm and optimized subspace algorithm are applied to design distinctive training subsets for base models with diversities. The IBGSO is proposed to implement selective ensemble and find the optimal sub-ensemble. Comparative experiments and analysis are conducted to verify the effectiveness of the proposed method. The results show that the proposed hybrid diversity enhancement method is beneficial for recognizing different types of activity accurately and reliably than the individual diversity enhancement method. Also, the proposed IBGSO is more effective than the other comparative pruning methods, which can achieve superior performance over other
comparative pruning methods with minimal ensemble scales. The proposed method shows better performance compared with the earlier research on UCI-HAR and WISDM. This demonstrates our proposed method HDSEEN is superior in HAR tasks with respect to stronger generalization ability and learning efficiency. In addition, compared with some state-of-the-art approaches for smartphone sensor-based HAR in the literature on UCI-HAR and WISDM, our proposed method HDSEEN is superior to these latest methods.

In future work, we will attempt to use recent swarm intelligence algorithms such as hunter-prey optimizer, slime mould algorithm, hunter-prey optimizer and aquila optimizer to search the optimal sub-ensemble for constructing a selective ensemble-based HAR system. Furthermore, we will test the effectiveness of the proposed approach using public datasets with more complex human activities.

REFERENCES

[1] M. Cornacchia, K. Oztan, Y. Zheng, and S. Velipasalar, “A survey on activity detection and classification using wearable sensors,” IEEE Sensors J., vol. 17, no. 2, pp. 386–403, Jan. 2017.

[2] R. Mooney, G. Corley, A. Godfrey, L. R. Quinlan, and G. ÓLaighin, “A survey on UCI-HAR and WISDM, our proposed method

[3] D. Thakur and S. Biswas, “Smartphone based human activity monitoring effectiveness of the proposed approach using public datasets

[4] X. Zhang, Q. Zhang, M. Chen, Y. Sun, X. Qin, and H. Li, “A two-stage feature selection and intelligent fault diagnosis method for rotating machinery using hybrid filter and wrapper method,” Neurocomputing, vol. 275, pp. 2426–2439, Jan. 2018.

[5] Z. Chen, C. Jiang, and L. Xie, “A novel ensemble ELM for human activity recognition using smartphone sensors,” IEEE Trans. Ind. Informat., vol. 15, no. 5, pp. 2691–2699, May 2019.

[6] S. S. Khan and B. Taati, “Detecting unseen falls from wearable devices using channel-wise ensemble of autencoders,” Expert Syst. Appl., vol. 87, pp. 280–290, Nov. 2017.

[7] L. Breiman, “Bagging predictors,” Mach. Learn., vol. 24, no. 2, pp. 123–140, 1996.

[8] Y. Freund and R. E. Schapire, “A decision-theoretic generalization of online learning and an application to boosting,” J. Comput. Syst. Sci., vol. 55, no. 1, pp. 119–139, Apr. 1997.

[9] K. Gu, Z. Xia, and J. Qiao, “Stacked selective ensemble for PM2.5 forecast,” IEEE Trans. Instrum. Meas., vol. 69, no. 3, pp. 660–671, Mar. 2020.

[10] G. D. C. Cavalcanti, L. S. Oliveira, T. J. M. Moura, and G. V. Carvalho, “Combining diversity measures for ensemble pruning,” Pattern Recognit. Lett., vol. 74, pp. 38–45, Apr. 2016.

[11] H. Cao, H. Liu, R. Li, C. Wu, Y. Gao, and M. Xu, “Margin diversity based ordering ensemble pruning,” Neurocomputing, vol. 275, pp. 237–246, Jan. 2018.

[12] X. Li, Y. Wang, B. Zhang, and J. Ma, “PSDRNN: An efficient and effective HAR scheme based on feature extraction and deep learning,” IEEE Trans. Ind. Informat., vol. 16, no. 10, pp. 6703–6713, Oct. 2020.

[13] A. Wang, G. Chen, J. Yang, S. Zhao, and C.-Y. Chang, “A comparative study on human activity recognition using inertial sensors in a smartphone,” IEEE Sensors J., vol. 16, no. 11, pp. 4566–4578, Jun. 2016.

[14] T. Tuncer, F. Ertam, S. Dogan, E. Aydinem, and P. Plawiak, “Ensemble residual network-based gender and activity recognition method with signals,” J. Supercomput., vol. 76, no. 3, pp. 2119–2138, Mar. 2020.

[15] Z. Wang, D. Wu, J. Chen, A. Ghomei, and M. A. Hossain, “A triaxial accelerometer-based human activity recognition using EEMD-based features and game-theory-based feature selection,” IEEE Sensors J., vol. 16, no. 9, pp. 3189–3207, May 2016.

[16] E. Zdravevski, P. Lameski, V. Trajkovik, A. Kulakov, I. Chorbey, R. Goleva, N. Ponomo, and N. Garcia, “Improving activity recognition accuracy in ambient-assisted living systems by automated feature engineering,” IEEE Access, vol. 5, pp. 5262–5280, 2017.

[17] H. Ghasemzadeh, N. Anini, R. Saeedi, and M. Sarrafzadeh, “Power-aware computing in wearable sensor networks: An optimal feature selection,” IEEE Trans. Mobile Comput., vol. 14, no. 4, pp. 101–112, Apr. 2015.

[18] W. Wang, P. Jones, and D. Partridge, “Assessing the impact of input features in a feedforward neural network,” Neural Comput. Appl., vol. 9, no. 2, pp. 101–112, Jul. 2000.

[19] I. Kononenko, “Estimating attributes: Analysis and extensions of RELIEF,” in Proc. Eur. Conf. Mach. Learn. ECML Mach. Learn., (ECML), 1994, pp. 171–182.

[20] T. T. Alemayoh, J. H. Lee, and S. Okamoto, “New sensor data structuring pipeline,” in Proc. 10th Int. Conf. Natural Comput. (ICNC), 2014, pp. 1697–1707.

[21] Z. Chen, Q. Zhu, A. Ghoneim, and M. A. Hossain, “A triaxial accelerometer-based human activity recognition via EEMD-based features and game-theory-based feature selection,” IEEE Sensors J., vol. 16, no. 9, pp. 3189–3207, May 2016.

[22] H. Guo, H. Liu, R. Li, C. Wu, Y. Gao, and M. Xu, “Margin diversity based ordering ensemble pruning,” Neurocomputing, vol. 275, pp. 237–246, Jan. 2018.

[23] X. Li, Y. Wang, B. Zhang, and J. Ma, “PSDRNN: An efficient and effective HAR scheme based on feature extraction and deep learning,” IEEE Trans. Ind. Informat., vol. 16, no. 10, pp. 6703–6713, Oct. 2020.

[24] A. Wang, G. Chen, J. Yang, S. Zhao, and C.-Y. Chang, “A comparative study on human activity recognition using inertial sensors in a smartphone,” IEEE Sensors J., vol. 16, no. 11, pp. 4566–4578, Jun. 2016.

[25] T. Tuncer, F. Ertam, S. Dogan, E. Aydinem, and P. Plawiak, “Ensemble residual network-based gender and activity recognition method with signals,” J. Supercomput., vol. 76, no. 3, pp. 2119–2138, Mar. 2020.

[26] Z. Wang, D. Wu, J. Chen, A. Ghomei, and M. A. Hossain, “A triaxial accelerometer-based human activity recognition using EEMD-based features and game-theory-based feature selection,” IEEE Sensors J., vol. 16, no. 9, pp. 3189–3207, May 2016.

[27] E. Zdravevski, P. Lameski, V. Trajkovik, A. Kulakov, I. Chorbey, R. Goleva, N. Ponomo, and N. Garcia, “Improving activity recognition accuracy in ambient-assisted living systems by automated feature engineering,” IEEE Access, vol. 5, pp. 5262–5280, 2017.

[28] H. Ghasemzadeh, N. Anini, R. Saeedi, and M. Sarrafzadeh, “Power-aware computing in wearable sensor networks: An optimal feature selection,” IEEE Trans. Mobile Comput., vol. 14, no. 4, pp. 101–112, Apr. 2015.

[29] W. Wang, P. Jones, and D. Partridge, “Assessing the impact of input features in a feedforward neural network,” Neural Comput. Appl., vol. 9, no. 2, pp. 101–112, Jul. 2000.

[30] I. Kononenko, “Estimating attributes: Analysis and extensions of RELIEF,” in Proc. Eur. Conf. Mach. Learn. ECML Mach. Learn., (ECML), 1994, pp. 171–182.

[31] T. T. Alemayoh, J. H. Lee, and S. Okamoto, “New sensor data structuring pipeline,” in Proc. 10th Int. Conf. Natural Comput. (ICNC), 2014, pp. 1697–1707.
[41] Y. Li and L. Wang, “Human activity recognition based on residual network and BiLSTM,” *Sensors*, vol. 22, no. 2, p. 635, Jan. 2022.

[42] S. Mekruksavanich, A. Jitpattanakul, K. Sitthithakerngkit, P. Youplao, and P. Yupapin, “ResNet-SE: Channel attention-based deep residual network for complex activity recognition using wrist-worn wearable sensors,” *IEEE Access*, vol. 10, pp. 51142–51154, 2022.

YIMING TIAN received the B.S. degree in automation from the Hebei University of Technology, Tianjin, China, in 2012, the M.S. degree in control science and control engineering from the University of Science and Technology Liaoning, Anshan, China, in 2015, and the Ph.D. degree in control science and control engineering from the Hebei University of Technology, in 2020. He is currently a Lecturer with the College of Information Engineering, Tianjin University of Commerce. His research interests include neural networks, wearable sensors, pattern recognition, and bioinformatics.

JIE ZHANG (Senior Member, IEEE) received the B.Sc. degree in control engineering from the Hebei University of Technology, Tianjin, China, in 1986, and the Ph.D. degree in control engineering from City University, London, in 1991. He is currently a Reader with the School of Engineering, Newcastle University, Newcastle upon Tyne, U.K. He has published over 180 papers in international journals, books, and conferences. His research interests include neural networks, neuro-fuzzy systems, intelligent control systems, genetic algorithms, optimal control of batch processes, and multivariate statistical process control. Dr. Zhang served as a Reviewer for many prestigious international journals, including IEEE *Transactions on Neural Networks*, IEEE *Transactions on Fuzzy Systems*, *Neural Networks, Automatica, Chemical Engineering Science*, and *AIChE Journal*. He is a member of the IEEE Control Systems Society and IEEE Computational Intelligence Society.

QI CHEN received the Ph.D. degree from Tianjin University, Tianjin, China. She is currently an Associate Professor with the College of Information Engineering, Tianjin University of Commerce. She has published more than ten papers and five were retrieved by SCI and EI. Her research interests include information processing of Industrial Internet of Things and production process control.

ZUOJUN LIU received the M.S. degree in control science and control engineering from the Hebei University of Technology, Tianjin, China, and the Ph.D. degree in control science and control engineering from Nankai University, Tianjin. He is currently a Professor and a Ph.D. Supervisor with the School of Artificial Intelligence and Data Science, Hebei University of Technology. His research interests include rehabilitation aids, intelligent robots, and intelligent control theory and application.