Adapted Binary Particle Swarm Optimization for Efficient Features Selection in the Case of Imbalanced Sensor Data

Dorin Moldovan, Ionut Anghel *, Tudor Cioara and Ioan Salomie

Computer Science Department, Technical University of Cluj-Napoca, Memorandumului 28, 400114 Cluj-Napoca, Romania; dorin.moldovan@cs.utcluj.ro (D.M.); tudor.cioara@cs.utcluj.ro (T.C.); ioan.salomie@cs.utcluj.ro (I.S.)

*Correspondence: ionut.anghel@cs.utcluj.ro

Received: 15 January 2020; Accepted: 18 February 2020; Published: 21 February 2020

Abstract: Daily living activities (DLAs) classification using data collected from wearable monitoring sensors is very challenging due to the imbalance characteristics of the monitored data. A major research challenge is to determine the best combination of features that returns the best accuracy results using minimal computational resources, when the data is heterogeneous and not fitted for classical algorithms that are designed for balanced low-dimensional datasets. This research article: (1) presents a modification of the classical version of the binary particle swarm optimization (BPSO) algorithm that introduces a particular type of particles called sensor particles, (2) describes the adaptation of this algorithm for data generated by sensors that monitor DLAs to determine the best positions and features of the monitoring sensors that lead to the best classification results, and (3) evaluates and validates the proposed approach using a machine learning methodology that integrates the modified version of the algorithm. The methodology is tested and validated on the Daily Life Activities (DaLiAc) dataset.

Keywords: machine learning; data streams; daily living activities; classification; wearable devices

1. Introduction

The development and large scale deployment of IoT monitoring devices has made significant amounts of data available, which may be used for improving the care of elders. In this context, the identification of daily living activities out of monitoring data is very important for understanding the elders behaviour and assessing deviations from daily routines, which may signal illness progression and timely planning of intervention processes. For example, in the project ReMIND [1] the enhancement of the quality of life of the elders with mild neurocognitive impairments is considered, and one major challenge addressed by the ReMIND project is the stimulating of the physical activity through music using James robots.

Therefore, the challenge of the daily living activities (DLAs) classification in real-time using low computational resources is critical as those elders might suffer unexpected falls caused by their critical health condition or sudden switches from one DLA to another DLA caused by memory problems. One way to increase the classification speed is to reduce the number of features of the data samples; however, that challenge is even more complex when the data is imbalanced and heterogeneous for each monitored person. In [2] the reduction of the number of features was addressed by proposing a novel bio-inspired algorithm. However, in that approach, the experimental DLAs datasets were balanced and normalized prior to the application of the feature selection.

The challenges of the aging population [3] have been addressed by various types of information and communication technology (ICT) systems in the last years, but few of them evaluated the challenges related to the acceptance and to the usability of the ambient assisted living (AAL) systems [4].
Those challenges are even more complex in particular cases, such as the ones where the elders have symptoms of dementia [5] or symptoms of Parkinson’s disease [6].

The recent developments for the internet of things (IoT) [7] and wearable computing led to the development of various types of health monitoring systems, which use data collected from various sources [8]. Moreover, these systems enable the continuous monitoring of elders; however, there are still challenges that must be considered, such as the need for high processing power and the low latency. This challenge is even more complex in the case of imbalanced data with many features and samples.

The classification of the DLAs using data that is generated by wearable devices has various applications in healthcare, such as falls detection [9], falls prediction [10], the classification of the activities of the elders that live alone [11], the classification of the DLAs of the elders that live in free-living conditions [12], the detection of anomalies in DLAs [13] and behaviour recognition for robust healthcare [14].

In the literature there are many datasets generated by sensors [15], which are used in the prediction of DLAs; however, the identification of the most appropriate techniques that return the best results, in various contexts, is still challenging due to the many dimensions that must be considered, such as the selection of the features, the balancing of the samples and the selection of the classifiers applied in the classification of the DLAs.

Real time data can be collected using various types of monitoring sensors, such as gyroscope sensors, orientation sensors, accelerometer sensors, and pressure sensors. The data generated by monitoring sensors are very useful in research, especially in the health research domain, as a thorough and complex analysis of this data can reveal a great deal of useful information, such as patterns in the daily living behaviour of the persons that are monitored, which might indicate particular health conditions, adherence to a particular type of routine, and anomalies in the normal daily behaviour.

The representative challenges related to the analysis of the data generated by the sensors, which monitor the DLAs are: (1) the determination of the best positions of the monitoring sensors that return the best classification results, (2) the determination of the features of the data generated by each monitoring sensor, which have the most significant influence on the final classification results, and (3) the determination of the combinations of features that correspond to each monitoring sensor, which return the best classification results for each monitored activity. Compared to the classical algorithms from the literature for feature selection that do not consider the particularities of the DLAs data, the approach presented in this article considers the fact that groups of features belong to different monitoring sensors, which are placed on different parts of the bodies of the monitored subjects.

The main contributions of this research article are:

(1) The development of a modified version of the standard binary particle swarm optimization (BPSO) algorithm, which introduces sensors particles characterized by weights proportional to the importance of the monitoring sensors and used further in the equations for the updating of the velocities of the standard particles;
(2) The adaptation of the proposed algorithm for DLAs data reflecting these adaptations in the way in which the objective function is defined and in the ranking of the features returned by the proposed algorithm;
(3) The evaluation and the validation of the adapted version of the BPSO algorithm using a machine learning methodology developed in-house, which compares the proposed algorithm with other feature selection approaches, and which uses as experimental support Daily Life Activities (DaLiAc) dataset [16];

The primary reasons to adapt BPSO for DLAs data are: (1) the flexibility to modify the classical BPSO, either by altering the mathematical equations or by introducing novel concepts in the original version, (2) the fact that the standard classical algorithms, in particular the standard versions of BPSO, do not exploit the nature of the DLAs datasets at maximum, and (3) the fact that in the case of the DLAs
the data is generated by several monitoring sensors described by groups of features—information that might lead to better results if it is integrated in the feature selection algorithms used in the processing of this type of data.

The remaining sections of the article are organized as follows: Section 2 illustrates the research background, Section 3 presents the materials and methods used in this article, Section 4 describes the most representative results obtained in the experiments, Section 5 presents a discussion of those results, and Section 6 concludes the article.

2. Background

The first subsection addresses feature selection challenges related to imbalanced data with many features generated by monitoring sensors. The second subsection presents literature approaches that consider feature selection based on the particle swarm optimization (PSO) algorithm [17]. The third subsection considers various approaches from the literature that use the DaLiAc dataset for experimental support.

2.1. Feature Selection Challenges for Imbalanced Data Generated by Monitoring Sensors

In [18] the high dimensionality and the class imbalance properties of the datasets are tackled. Compared to the approach presented in this article, the authors of that article are motivated by medical applications from real-world systems. Their feature selection method aims to achieve a better separability between the minority class and the majority class. In the approach presented in this article the majority or the minority properties of the classes of the samples are not considered when the features are selected. However, the approach presented in this article considers the variability of the monitored data for each class for each feature, and it also considers the relations between features.

The authors of [19] address one of the major challenges of the feature selection in the case of datasets with many features, namely the ignorance of the class imbalance problem and thus in the majority of the cases, the selected features are biased towards majority classes. In the approach presented in this article the imbalance is addressed by the way in which the objective function of the adapted version of the BPSO algorithm is defined, as all the classes are treated equally.

In [20], the authors deal with the following two major challenges in machine learning, namely the high dimensionality and the class-imbalance. The cardinality of the set of features is penalized by a scaling factors technique, while the class-imbalance problem is addressed using a method based on a support vector machine (SVM) [21]. In the approach presented in this article, each class is treated equally when the objective function is defined, and moreover, the number of features is penalized in the objective function, increasing its value when the number of features is small.

The selection of features for imbalanced datasets is challenging as an effective learning model should be constructed, reducing the memory consumption and the time consumption—two objectives, which are addressed in the approach presented in this article. In [22] that challenge is addressed by rough-set-based feature selection algorithm for imbalanced data (RSFSAID), a novel algorithm for feature selection. Similar to the approach from this article, those authors use PSO. However, in their approach, PSO is applied in the determination of the optimized parameters for the proposed algorithm. They also study the lower and the upper boundary regions, when they define the significance of the features. In the approach from this article the minimum and the maximum values of the ranges of variability are considered in the definition of the matrix of variability.

The authors of [23] address the challenge of feature selection in the case of high dimensional imbalanced datasets from the perspective of the neglecting of the features interaction. Many traditional approaches applied in the case of datasets with imbalanced classes are biased towards the majority classes and the authors of that article propose a method for feature selection based on the interaction information (II). In this article the interactions between the features are considered, defining a matrix that approaches the features from the perspective of the range of variability. Moreover, this article
considers particular characteristics of the DLAs datasets that are not approached in that article, and proposes a more generic algorithm for feature selection.

2.2. Feature Selection Approaches Based on Particle Swarm Optimization

The authors of [24] apply PSO in feature selection, considering three goals, which are also targeted by the approach presented in this research article: (1) the maximization of the classification performance, (2) the minimization of the number of the selected features and (3) the minimization of the computational time. However, they propose, as future research work, the development of an approach based on PSO, which is multi-objective and which simultaneously minimizes the number of the selected features and maximizes the performance of the classification. In the approach presented in this article, both of those two objectives are considered in the definition of the objective function.

In [25], a method is presented in which the features are selected using a hybrid version of PSO, which also integrates a local search strategy. That method is called hybrid particle swarm optimization with local search (HPSO-LS) and determines efficiently the discriminative features with reduced correlations. The objective function of that approach uses the k-nearest neighbor (k-NN) [26] method, and, between two solutions that return the same accuracy, it considers that solution with the smaller number of features. However, that approach is expensive in terms of computational resources, as it requires the training of a k-NN classifier for each particle in each iteration of the algorithm.

The authors of [27] considered an algorithm for feature selection based on PSO with learning memory (PSO-LM). Compared to the approach from this article, that approach introduces a memory learning strategy with the objective to balance the global exploration and the local exploitation in the algorithm. Moreover, their objective function is similar to the objective function of the adapted version of the BPSO algorithm described in this article, as it uses two parts as follows: The first part corresponds to the accuracy of the prediction model trained using the selected features and the second part corresponds to the ratio between the number of the selected features and the total number of features. However, the approach presented in this article is different, as, instead of using a classifier for the first part of the objective function as in the case of that approach, which considers the k-NN classifier, this article introduces an algorithm, which considers the variability of the monitored data, and which also considers the fact that the data is imbalanced.

In [28], an approach is considered for the joint moment prediction, in which the features are selected using a method based on BPSO and the objective function is based on the variance accounted for (VAF). Similar to the approach from this article, their approach is tested and validated on DLAs data. However, that approach does not consider the minimization of the features as a distinctive part of the objective function and the data considered in their experiments for the validation of the approach contains only gait patterns of running, therefore it can not be generalized very easily.

Another approach for feature selection based on PSO is presented in [29]. The method presented in that article is called hybrid particle swarm optimization with a spiral-shaped mechanism (HPSO-SSM) and is based on the following three improvements: (1) the enhancement of the diversity in the searching process using a logistic map sequence, (2) the introduction of two new parameters in the original formulas, used in the updating of the positions, and (3) the adopting of a spiral-shaped mechanism as a local search operator. Their method presents broad applications, as it is validated using twenty classic benchmark classification datasets from the University of California, Irvine (UCI) machine learning repository. However, compared to the method presented in this article, it is not very focused, as it does not exploit the nature of those datasets at maximum.

2.3. Machine Learning Approaches from the Literature that Consider the Daily Life Activities (DaLiAc) Dataset

The Daily Life Activities (DaLiAc) dataset was introduced in [16] as a benchmark dataset that is analyzed in the context of a hierarchical classification system for DLAs. Compared to the method presented in this article, which evaluates the performance of the DLAs classification using 10-fold cross validation, the method presented in that article considers a leave-one-subject-out procedure.
The authors also address the challenge of a high number of features, which leads to a very high computational complexity in the case of embedded systems or real-time applications and they even suggest the reduction of the features, giving, as an illustrative example, the sequential feature selection.

The authors of [30] use the DaLiAc dataset as experimental support for a method based on a classifier ensemble of randomized trees, and their objective is to maximize the recognition performance, both in the replacement scenario and in the relocation scenario. Compared to the approach presented in this article, they used more $K$-folds cross-validation schemes with $K$ from $\{2, 3, 5, 10, 20\}$ and they also used the accuracy to measure the performance. They also included junk data in their experiments and therefore their results are not directly comparable with the results obtained in this article.

The human activity recognition is approached in [31] by combining a number of classifiers, and the method presented in that article, which is based on a combination of models outperforms other types of classifiers combination models. However, the challenge that still remains is how to increase the robustness of that method against interperson variability. The method based on BPSO described in this article addresses that challenge partially by assigning a weight to each monitoring sensor and consequently the weights are different and adapted for each monitored subject.

The algorithms applied in [32] for improving the accuracy of activity recognition are k-NN, logistic regression (LR), naive bayes (NB), random forest (RF), extremely randomized trees (ERT) and SVM. The method described in that article consists of four phases, namely segmentation, feature extraction, feature selection and evaluation. In the case of the DaLiAc dataset, the overall mean classification is 89.6%. The work described in this article is inspired from the work from that article as follows: (1) our work considers the selection of the features according to the positions of the sensors that generate them, (2) it considers the RF algorithm, and (3) it applies the proposed research method on several datasets. However, one weakness of that work is that it does not address the importance of each monitoring sensor, even though it addresses the optimal number of sensors, the positions of the sensors, and the optimal number of features for each sensor.

The method presented in [33] considers the DaLiAc dataset from a different perspective than the other articles, which consider it as experimental support, converting the inertial sensor signals to images and then applying a convolutional neural network (CNN) [34] for images classification. Even if that method leads to better classification results than the traditional classification methods, it is also subject to the limitations of CNNs, such as the class imbalance, the overfitting and poor performance when the input data is small. Moreover, CNNs are also computationally expensive and, in this article, that challenge is addressed by trying to reduce the number of features so that the classifiers can be trained faster and the classification results are optimized.

3. Materials and Methods

This section presents the main concepts and the pseudo-code of the adapted version of BPSO.

3.1. Mathematical Formulation of the Optimization Problem Approached Using an Adapted Variant of the BPSO Algorithm

The objective of the optimization problem is to select an optimal number of features in the case of data generated by sensors that monitor various types of DLAs in order to: (1) minimize the number of selected features and to (2) maximize the value of a mathematical formula based on a heuristic introduced by us, which describes the range of variability of the DLAs using the selected features. Those two sub-objectives are reflected in the objective function of the adapted version of BPSO algorithm.

The optimization variables are represented by arrays of zeros and ones of the form $X = (x_1, ..., x_{24})$, such that $x_i \in \{0, 1\}$ for any $i \in \{1, ..., 24\}$. A value equal to 0 means that a feature is not selected and a value equal to 1 means that a feature is selected. The search space is thus represented by $2^{24} = 16777216$ possible variables.
The optimization constraint is the maximum value of the selected features, which is equal to six as six is also the number of features generated by each monitoring sensor in the data used as experimental support. The search space is thus limited to fewer variables, more specifically \( \sum_{i=0}^{6} \binom{24}{i} = 190051. \) That constraint is considered in this article using the ranking of the features according to an heuristic developed in-house and is applied to the best results obtained using the adapted version of BPSO, when they are compared to the results obtained using other methods in Section 4.

3.2. Matrix of Variability of DLAs Sensors Data for Each Monitored Subject

The matrix of variability \( M \) has the form:

\[
M = \begin{bmatrix}
a_{1,1} & \ldots & a_{1,N_{DLAs}} \\
\vdots & \ddots & \vdots \\
a_{D,1} & \ldots & a_{D,N_{DLAs}}
\end{bmatrix}
\] (1)

where \( D \) is the number of features of the data generated by the monitoring sensors, which is also the dimension of the search space in the adapted version of BPSO, and \( N_{DLAs} \) is the number of daily living activities. Each line \( i \) of the matrix \( M \), where \( i = 1, \ldots, D \), is a permutation of the set \( \{1, 2, \ldots, N_{DLAs}\} \), such that the order of the elements of that set is given by the increasing order of the ranges of variability of each monitored DLA. The range of variability for a DLA \( a \) and the corresponding feature \( f \) is given by the formula:

\[
r_{\text{var}}_{a,f} = v_{\text{max}}_{a,f} - v_{\text{min}}_{a,f}
\] (2)

3.3. Metric for the Evaluation of the Matrix of Variability of DLAs Sensors Data for Each Monitored Subject

The value of the metric \( M_\Delta \) used in the evaluation of the matrix of variability \( M \) of DLAs sensors data is given by:

\[
M_\Delta = \sum_{i=1}^{N_{DLAs}} \left( \left| \max\{j|j \in \mathbb{N}, 1 \leq j \leq N_{DLAs} \land \exists k \in \mathbb{N}, 1 \leq k \leq D, M_{k,j} = i\} - \min\{j|j \in \mathbb{N}, 1 \leq j \leq N_{DLAs} \land \exists k \in \mathbb{N}, 1 \leq k \leq D, M_{k,j} = i\} + 1 \right| \right).
\] (3)

3.4. Heuristic for Features Ranking in the Optimal Solution Returned by the Adapted BPSO Algorithm for Each Monitored Subject

In order to restrict the number of selected features to a maximum threshold, a matrix \( R \) of variability is proposed. The matrix \( R \) is defined by the set of all \( r_{\text{var}}_{a,f} \) such that \( a \) takes values from the set \( \{1, \ldots, N_{DLAs}\} \) and \( f \) takes values from the set \( \{1, \ldots, D\} \):

\[
R = \begin{bmatrix}
r_{\text{var}}_{1,1} & \ldots & r_{\text{var}}_{1,D} \\
\vdots & \ddots & \vdots \\
r_{\text{var}}_{N_{DLAs},1} & \ldots & r_{\text{var}}_{N_{DLAs},D}
\end{bmatrix}
\] (4)

The feature \( f_1 \) is considered better than the feature \( f_2 \) if the value returned by the \( \text{Comparator}(f_1, f_2) \) described next is greater than or equal to 0:

\[
\text{Comparator}(f_1, f_2) = \sum_{i=1}^{N_{DLAs}} s(r_{\text{var}}_{i,f_1} - r_{\text{var}}_{i,f_2}),
\] (5)
such that for any real number x the value of s(x) is 1 if x ≥ 0 or −1 otherwise.

3.5. Mathematical Description of the Objective Function of the Adapted Version of the BPSO Algorithm

Algorithm 1: The objective function of the adapted BPSO algorithm.

| Data: p, D, M, M∆, DLAs |
|--------------------------|
| Result: r               |
1 r = 0;                   |
2 x = position of particle p; |
3 Fselected = { };        |
4 for d ∈ D do            |
5 if xd == 1 then          |
6   Fselected = Fselected + fd; |
7 end                      |
8 end                      |
9 ∆ = 0;                   |
10 for a ∈ DLAs do         |
11   mina = +∞;            |
12   maxa = −∞;           |
13 for f ∈ Fselected do    |
14   if Mfa < mina then    |
15     mina = Mfa;         |
16   end                   |
17   if Mfa > maxa then    |
18     maxa = Mfa;         |
19 end                     |
20 if Fselected ≠ ∅ then   |
21     ∆ = ∆ + |maxa − mina + 1|; |
22 end                     |
23 end                     |
24 r = \( \frac{1}{2} \times \left( \frac{\Delta}{M} + \frac{|D| - |F_{selected}|}{|D|} \right) \); |
25 return r;               |

The inputs of Algorithm 26 are: p—The particle for which the fitness value is computed, D—The number of dimensions of the search space, M—A precomputed matrix, which is presented in more details next, M∆—A metric used for the evaluation of that matrix, and DLAs—The set of daily living activities monitored by the monitoring sensors. The output of the algorithm is r—The fitness value of the particle p.

Initially, the value of r is 0 (line 1), the value of x is equal to the position of the particle p (line 2) and Fselected is an empty set (line 3). For each d from the set \( \{1, \ldots, D\} \), if the value of xd is equal to 1 then the feature fd is included in the set of selected features Fselected (lines 4–8). The initial value of ∆ is 0 (line 9). For each activity a from the set of daily living activities DLAs the steps from (lines 11–23) are repeated. The initial values of mina and maxa are equal to +∞ (line 11) and −∞ (line 12). For each feature f from the set of features Fselected, the values of mina and maxa are updated in (lines 14–19) as follows: If the value of Mfa is less than the value of mina (line 14) then mina is initialized to the value of Mfa (line 15) and if the value of Mfa is greater than the value of maxa (line 17) then maxa is initialized to the value of Mfa (line 18). If Fselected ≠ ∅ (line 21) then
the value of $\Delta$ is incremented with $|\text{max}_a - \text{min}_a + 1|$ (line 22). In (line 25) the value of $r$ is updated using the formula:

$$r = \frac{1}{2} \times \left( \frac{\Delta}{M_\Delta} + \frac{|D| - |F_{\text{selected}}|}{|D|} \right).$$ (6)

The ideal value of $r$ is 1 and corresponds to the case when $\Delta = M_\Delta$ and the number of selected features is 0, and the worst value of $r$ is 0 and corresponds to the case when $\Delta = 0$ and all features are selected. Consequently, $r$ takes values from the interval $[0, 1]$ and the objective of the algorithm is to maximize that value. The algorithm returns the final value of $r$ in (line 26).

3.6. Adapted BPSO Algorithm for Feature Selection in the case of DLAs Sensors Data

**Algorithm 2: Adapted BPSO Algorithm.**

**Data:** $N_{\text{particles}}$, $N_{\text{iterations}}$, $c_1$, $c_2$, $w_{\text{min}}$, $w_{\text{max}}$, $V_{\text{min}}$, $V_{\text{max}}$, $D$, $N_{\text{sensors}}$

**Result:** $g_{\text{best}}$

1. Initialize the particles that correspond to each monitoring sensor $S = \{s_1, \ldots, s_{N_{\text{sensors}}}\}$;
2. for $s \in S$ do
   - $\text{Fitness}(s) = \text{ObjectiveFunction}(s)$;
3. end
4. for $k = 1$ to $N_{\text{sensors}}$ do
   - $t_k = 2 \times \frac{\text{Fitness}(s_k)}{\sum_{i=1}^{N_{\text{sensors}}} \text{Fitness}(s_i)}$;
7. end
8. Initialize the swarm of particles $\text{Swarm} = \{p_1, \ldots, p_{N_{\text{particles}}}\}$;
9. $i = 0$;
10. while $i < N_{\text{iterations}}$ do
11.   - for $p \in \text{Swarm}$ do
12.     - $\text{Fitness}(p) = \text{ObjectiveFunction}(p)$;
13.     - $\text{UpdateLocalBest}(p, p_{\text{best}})$;
14.     - $\text{UpdateGlobalBest}(p_{\text{best}}, g_{\text{best}})$;
15.   end
16.   - $\text{UpdateInertia}(w, w_{\text{min}}, w_{\text{max}}, i)$;
17.   for $j = 1$ to $|\text{Swarm}|$ do
18.     - for $d \in D$ do
19.       - $v_{\text{new}}^d = w \times v_{\text{old}}^d + c_1 \times r_1 \times (p_{\text{best},d} - x_{\text{old}}^d) + c_2 \times r_2 \times (g_{\text{best},d} - x_{\text{old}}^d) + \sum_{k=1}^{N_{\text{sensors}}} t_k \times$
20.       - $s_{k,d} - x_{\text{old}}^d$;
21.       - if $v_{\text{new}}^d \notin (V_{\text{min}}, V_{\text{max}})$ then
22.          - $v_{\text{new}}^d = \max(\min(V_{\text{max}}, v_{\text{new}}^d), V_{\text{min}})$;
23.       - end
24.       - $S(v_{\text{new}}^d) = \frac{1}{1 + e^{-v_{\text{new}}^d}}$;
25.       - if $r_0 < S(v_{\text{new}}^d)$ then
26.          - $x_{\text{new}}^d = 1$;
27.       - else
28.          - $x_{\text{new}}^d = 0$;
29.       - end
30.     end
31.     $i = i + 1$;
32.   end
33. return $g_{\text{best}}$;
The BPSO algorithm is adapted after the binary version of the PSO algorithm [35]. A method similar to the one presented in [36] is considered in order to restrict the values of the positions of the particles to take values from the set \( \{0, 1\} \).

The inputs of the adapted BPSO algorithm are:
- \( N_{\text{particles}} \) — The number of particles,
- \( N_{\text{iterations}} \) — The maximum number of iterations,
- \( c_1 \) — The acceleration coefficient for the cognitive component,
- \( c_2 \) — The acceleration coefficient for the social component,
- \( w_{\text{min}} \) — The minimum value for the inertia,
- \( w_{\text{max}} \) — The maximum value for the inertia,
- \( V_{\text{min}} \) — The minimum possible value for the velocity,
- \( V_{\text{max}} \) — The maximum possible value for the velocity,
- \( D \) — The number of dimensions of the search space, and
- \( N_{\text{sensors}} \) — The number of monitoring sensors. The output \( g_{\text{best}} \) describes the position of the best particle after the maximum number of iterations \( N_{\text{iterations}} \).

The adapted BPSO algorithm starts with the initialization of the particles that correspond to each monitoring sensor (line 1) and then for each particle that corresponds to a monitoring sensor, the corresponding fitness value is calculated (lines 2–4). Next, the weights of each sensor particle are computed (lines 5–7) in such a way that the weights are proportional to the fitness values of the sensor particles.

The particles swarm is initialized randomly (line 8) and the value of the initial iteration \( i \) is set to 0 in (line 9). The algorithm is then executed for a maximum number of iterations \( N_{\text{iterations}} \) given as input.

In each iteration for each particle from the swarm (lines 11–15), the fitness value is computed in (line 12) using the `ObjectiveFunction` and the functions `UpdatedLocalBest` and `UpdateGlobalBest` update the values of \( p_{\text{best}} \) and \( g_{\text{best}} \) as follows:

\[
p_{\text{best}} = \begin{cases} p & \text{if } \text{Fitness}(p) > \text{Fitness}(p_{\text{best}}) \\ p_{\text{best}} & \text{otherwise} \end{cases} \quad (7)
\]

\[
g_{\text{best}} = \begin{cases} p_{\text{best}} & \text{if } \text{Fitness}(p_{\text{best}}) > \text{Fitness}(g_{\text{best}}) \\ g_{\text{best}} & \text{otherwise} \end{cases} \quad (8)
\]

The inertia is updated (line 16) considering the values of \( i, w_{\text{min}} \) and \( w_{\text{max}} \) as follows:

\[
w = \begin{cases} w_{\text{max}} & \text{if } i = 0 \\ \max(w - i, w_{\text{min}}) & \text{if } i > 0 \end{cases} \quad (9)
\]

Next, for each particle from the swarm (lines 17–30) and for each dimension of the particle (lines 18–29) the new value of the velocity of the particle for dimension \( d \) is computed, taking into consideration the following components:

- the inertia component: \( w \times v_{\text{old}}^d \),
- the cognitive component: \( c_1 \times r_1 \times (p_{\text{best},d} - x_{d}^{\text{old}}) \),
- the social component: \( c_2 \times r_2 \times (g_{\text{best},d} - x_{d}^{\text{old}}) \),
- the sensors component: \( \sum_{k=1}^{N_{\text{sensors}}} t_k \times r_{k+2} \times (s_{k,d} - x_{d}^{\text{old}}) \).

If the computed value is not in the interval \( (V_{\text{min}}, V_{\text{max}}) \), then that value is updated either to \( V_{\text{min}} \) or to \( V_{\text{max}} \) (lines 20–22) as follows:

\[
v_{\text{new}}^d = \begin{cases} V_{\text{min}} & \text{if } v_{\text{new}}^d < V_{\text{min}} \\ V_{\text{max}} & \text{if } v_{\text{new}}^d > V_{\text{max}} \end{cases} \quad (10)
\]

The velocity is converted into a probability value (line 23) using the exponential:

\[
S(v_{\text{new}}^d) = \frac{1}{1 + e^{-v_{\text{new}}^d}}. \quad (11)
\]

In (lines 24–28) the value of a random value \( r_0 \) from \([0, 1]\) is applied in order to update the position of the particle as follows:
\[ x_d^{\text{new}} = \begin{cases} 1 & \text{if } r_0 < S(x_d^{\text{new}}) \\ 0 & \text{otherwise.} \end{cases} \] (12)

In (line 31) the current iteration \( i \) is incremented with 1 and finally in (line 33) the algorithm returns the value of \( g_\text{best} \).

4. Results

The adapted version of the BPSO algorithm was written and evaluated in Java in IntelliJ IDEA. The feature selection experiments for the adapted BPSO algorithm were performed in IntelliJ IDEA. The experiments for feature selection using FFS and BFE were performed in Konstanz Information Miner (KNIME) [37]. The experiments for feature selection using RF were performed in JetBrains PyCharm, using the sklearn library from Python. The experiments for feature selection based on the adapted versions of GA and DE were performed in IntelliJ IDEA. The properties of the system in which the experiments were performed are:

1. processor properties: Intel(R) Core(TM) i5-7600K CPU @ 3.80GHz 3.80 GHz;
2. installed memory (RAM) properties: 16.0 GB;
3. system type properties: 64-bit Operating System, x64-based processor.

4.1. Machine Learning Methodology for the Classification of DLAs Based on Adapted BPSO

The steps of the machine learning methodology that is applied in the classification of the DLAs using data generated by monitoring sensors are described in more details in the next subsections. In the Feature Selection step, the methodology integrates the algorithm that is presented in this article, namely the adapted BPSO algorithm. The methodology is also used as support in order to validate and to test this algorithm.

4.1.1. DLAs Sensors Data

In Figure 1, the placement of the monitoring sensors for the DaLiAc dataset is illustrated.

Figure 1. Sensors positions for the Daily Life Activities (DaLiAc) dataset: (1) chest (C), (2) right wrist (RW), (3) left ankle (LA), (4) right hip (RH).
The main characteristics of the DaLiAc dataset are presented in Table 1.

| Characteristic                          | Value  |
|----------------------------------------|--------|
| number of daily living activities (DLAs) | 13     |
| number of subjects                     | 19     |
| number of monitoring sensors           | 4      |
| total number of features               | 24     |
| sampling frequency                     | 200 Hz |

In the case of the DaLiAc dataset the following activities are monitored:

1. DLA1—Sitting;
2. DLA2—Lying;
3. DLA3—Standing;
4. DLA4—Washing dishes;
5. DLA5—Vacuuming;
6. DLA6—Sweeping;
7. DLA7—Walking outside;
8. DLA8—Ascending stairs;
9. DLA9—Descending stairs;
10. DLA10—Treadmill running;
11. DLA11—Bicycling (50 watt);
12. DLA12—Bicycling (100 watt);
13. DLA13—Rope jumping;

The DaLiAc dataset contains information about 19 subjects (eight females and eleven males) aged 26 ± 8 years. The data was acquired using four SHIMMER (Shimmer Research, Dublin, Ireland) sensors. Each of those sensors is equipped with a triaxial accelerometer (three features) and a triaxial gyroscope (three features). The accelerometer measures the proper acceleration while the gyroscope measures the velocity and the angular velocity. The sensors were placed on the left ankle, on the right hip, on the chest and on the right ankle of the monitored subjects. The frequency of sampling of those sensors is 200 Hz.

Figure 2 summarizes the number of samples of each daily living activity for each monitored subject. As can be seen in the figure, the walking outside activity (DLA7) contains the biggest number of samples for each monitored subject and moreover the samples for each activity are not balanced.

4.1.2. Feature Selection

In this step, various combinations of feature selection techniques are investigated. Four different types of feature selection techniques are applied:

1. The first technique considers the features that correspond to the monitoring sensors that are used for collecting the data from the monitoring sensors: six features (chest sensor), six features (right wrist sensor), six features (left ankle sensor), six features (right hip sensor);
2. The second technique considers the following three feature selection algorithms from literature: the forward feature selection (FFS) algorithm [38], the backward features elimination (BFE) algorithm [39] and the random forest (RF) [40];
3. The third technique considers the BPSO algorithm adapted for data generated by monitoring sensors placed on the bodies of the monitored subjects;
4. The fourth technique considers adapted versions of genetic algorithm (GA) [41] and differential evolution (DE) [42] for feature selection using the same objective function as in the case of the adapted BPSO algorithm;
4.1.3. Cross Validation

The approach presented in this article is based on 10-fold cross validation. Thus, the data is split randomly in 10 folds of approximately equal size and the prediction model is run 10 times, such that, in each run, the testing data is represented by a different fold from those 10 folds and the training data is represented by the remaining nine folds.

4.1.4. Machine Learning Classification Model

The machine learning classification model is used in the classification of the daily living activities performed by the monitored subjects based on the data from the monitoring sensors. The machine learning classification model that is applied in this article is RF. RF is an ensemble classification approach that is based on the development of a collection of decision trees and has applications in many domains, such as medicine, ecology, bioinformatics, astronomy, and agriculture, and, moreover, it is a very good option for imbalanced datasets.

4.1.5. DLAs Classification

The final output of the machine learning methodology is represented by the classified DLAs, considering the raw sensors data generated during the monitoring of the DLAs. The metrics that are applied in the evaluation of the models used for the classification of the DLAs are the recall, the precision, and the F-measure.

4.2. Feature Selection Results for DLAs Data Generated by Monitoring Sensors Using the Adapted BPSO Algorithm

The configuration of the parameters of the adapted BPSO algorithm applied in the experiments is summarized in Table 2.

The mappings between the sensor particles weights and the sensor positions for the DaLiAc dataset are: $t_1$—C, $t_2$—RW, $t_3$—LA and $t_4$—RH.

Table 3 summarizes the weights of the sensor particles for the DaLiAc Dataset, for each monitored subject.

Figure 2. Distribution of DLAs samples for each monitored subject of the DaLiAc dataset.
Table 2. Configuration of the adapted binary particle swarm optimization (BPSO) algorithm parameters.

| Parameter       | Significance     | Value |
|-----------------|------------------|-------|
| \( N_{\text{iterations}} \) | number of iterations | 30 |
| \( N_{\text{particles}} \) | number of particles | 20 |
| \( c_1 \) | cognitive component value | 2 |
| \( c_2 \) | social component value | 2 |
| \( V_{\text{min}} \) | minimum value of velocity | 0.5 |
| \( V_{\text{max}} \) | maximum value of velocity | 0.5 |
| \( w_{\text{min}} \) | minimum value of inertia | 0.500 |
| \( w_{\text{max}} \) | maximum value of inertia | 0.901 |

Table 3. Sensor particle weights for the DaLiAc dataset for each monitored subject.

| Monitored Subject/ Sensor Particle Weight | \( t_1 \) | \( t_2 \) | \( t_3 \) | \( t_4 \) |
|------------------------------------------|--------|--------|--------|--------|
| \( S_1 \) | 0.478 | 0.567 | 0.448 | 0.507 |
| \( S_2 \) | 0.544 | 0.490 | 0.466 | 0.500 |
| \( S_3 \) | 0.570 | 0.535 | 0.418 | 0.477 |
| \( S_4 \) | 0.498 | 0.559 | 0.470 | 0.473 |
| \( S_5 \) | 0.565 | 0.492 | 0.465 | 0.478 |
| \( S_6 \) | 0.482 | 0.529 | 0.465 | 0.524 |
| \( S_7 \) | 0.472 | 0.502 | 0.528 | 0.498 |
| \( S_8 \) | 0.564 | 0.520 | 0.435 | 0.481 |
| \( S_9 \) | 0.537 | 0.554 | 0.429 | 0.480 |
| \( S_{10} \) | 0.509 | 0.505 | 0.472 | 0.514 |
| \( S_{11} \) | 0.508 | 0.494 | 0.461 | 0.537 |
| \( S_{12} \) | 0.536 | 0.520 | 0.474 | 0.470 |
| \( S_{13} \) | 0.594 | 0.510 | 0.441 | 0.455 |
| \( S_{14} \) | 0.555 | 0.480 | 0.475 | 0.490 |
| \( S_{15} \) | 0.497 | 0.525 | 0.489 | 0.489 |
| \( S_{16} \) | 0.573 | 0.490 | 0.450 | 0.487 |
| \( S_{17} \) | 0.555 | 0.527 | 0.451 | 0.467 |
| \( S_{18} \) | 0.571 | 0.484 | 0.453 | 0.492 |
| \( S_{19} \) | 0.532 | 0.524 | 0.472 | 0.472 |

In Table 4 is presented as a summary of the experiments for each monitored subject \( S \), as follows: the mean running time in milliseconds \( t_m \), the mean fitness value \( f_m \), the mean number of returned features \( NRF_m \), and the mean accuracy \( a_m \). For each monitored subject, five experiments were performed. The value of \( NRF_m \) is from \{5, 6, 7, 8\}.

In Figure 3, the evolution of the fitness value in each iteration of the adapted BPSO algorithm for each subject is presented. The fitness value is improved after 30 iterations for each monitored subject.

Table 4. Summary of the adapted BPSO algorithm mean results for each monitored subject.

| Subject | \( t_m \) | \( f_m \) | \( NRF_m \) | \( a_m \) |
|---------|----------|----------|-----------|--------|
| \( S_1 \) | 684476 | 0.818 | 8 | 91.0% |
| \( S_2 \) | 910364 | 0.826 | 6 | 93.7% |
| \( S_3 \) | 739985 | 0.795 | 7 | 92.4% |
| \( S_4 \) | 718689 | 0.799 | 7 | 90.0% |
| \( S_5 \) | 725932 | 0.811 | 7 | 90.4% |
| \( S_6 \) | 614417 | 0.814 | 7 | 94.3% |
| \( S_7 \) | 733000 | 0.796 | 5 | 86.0% |
| \( S_8 \) | 625166 | 0.808 | 7 | 91.1% |
| \( S_9 \) | 665925 | 0.800 | 7 | 93.3% |
| \( S_{10} \) | 65937 | 0.804 | 6 | 89.7% |
| \( S_{11} \) | 761430 | 0.802 | 7 | 88.2% |
Table 4. Cont.

| Subject | $t_m$ | $f_m$ | NRF | $n_m$ |
|---------|-------|-------|-----|-------|
| S12     | 793377| 0.803 | 7   | 92.8% |
| S13     | 802259| 0.798 | 7   | 92.1% |
| S14     | 760730| 0.803 | 6   | 90.9% |
| S15     | 813828| 0.811 | 7   | 90.7% |
| S16     | 885900| 0.805 | 7   | 91.0% |
| S17     | 875175| 0.798 | 6   | 92.7% |
| S18     | 774223| 0.793 | 7   | 91.1% |
| S19     | 753161| 0.798 | 6   | 91.7% |

Figure 3. Fitness value evolution in each iteration of the adapted BPSO algorithm for the best results.
4.3. Comparison of the Results Obtained Using the Adapted Version of BPSO for feature selection with the Results Obtained Using Other Methods

Both for the adapted versions of GA and DE for feature selection, the same objective function was applied as in the case of the adapted version of BPSO. Moreover, the number of returned features was restricted to six in the case of GA, DE, and BPSO, using the heuristic for features ranking introduced in this article, which returns the results presented in Appendix A. In the case of the adapted version of BPSO the configuration parameters described in the previous section were applied.

The following configurations are applied in the case of the other methods:

1. FFS and BFE—The standard configurations from KNIME, a threshold for the number of features equal to six and a random drawing strategy;
2. RF—The standard configurations from sklearn, a number of estimators equal to 1000, a maximum number of features equal to six;
3. GA—20 chromosomes, 30 iterations, $CR$ (crossover rate) = 0.5, $MR$ (mutation rate) = 0.5;
4. DE—20 agents, 30 iterations, $CR$ (crossover probability) = 0.5, $F$ (differential weight) = 1.0.

Both for the adapted versions of GA and DE for feature selection, the same objective function was applied, as in the case of the adapted version of BPSO. Moreover, the number of returned features was restricted to six in the case of GA, DE, and BPSO, using the heuristic for features ranking introduced in this article that returns the results presented in Appendix A.

In Table 5, the running times for each feature selection approach are presented. The running time of BPSO is much better than the running time of BFE, GA and DE, but slightly worse than the running time of FFS and comparable to the running time of RF.

Table 5. The running time for forward feature selection (FFS), backward features elimination (BFE), random forest (RF), binary particle swarm optimization (BPSO), genetic algorithm (GA) and differential evolution (DE) for each monitored subject, in milliseconds.

| Subject/Features Selection | FFS  | BFE  | RF   | BPSO | GA   | DE   |
|----------------------------|------|------|------|------|------|------|
| S1                         | 333842 | 2517009 | 914742 | 636930 | 1930886 | 1282519 |
| S2                         | 409865 | 5253360 | 1117351 | 953376 | 2213781 | 1519955 |
| S3                         | 333872 | 3661512 | 800075 | 799151 | 1766370 | 1271330 |
| S4                         | 337070 | 4481319 | 751205 | 716961 | 1787177 | 1269005 |
| S5                         | 348089 | 1632155 | 866558 | 703276 | 1836603 | 1316106 |
| S6                         | 309679 | 1569169 | 680149 | 658894 | 1613077 | 1146485 |
| S7                         | 368906 | 1704399 | 827109 | 705606 | 1882546 | 1373507 |
| S8                         | 326191 | 1540248 | 863979 | 588280 | 1663440 | 1207692 |
| S9                         | 349503 | 4865604 | 723893 | 652314 | 1791708 | 1282526 |
| S10                        | 345890 | 3366558 | 695436 | 575586 | 1750748 | 1285746 |
| S11                        | 329980 | 2669384 | 814888 | 822193 | 1739393 | 1255230 |
| S12                        | 342886 | 1580970 | 830055 | 868976 | 1832432 | 1319197 |
| S13                        | 336975 | 4311050 | 905903 | 849842 | 1855374 | 1294919 |
| S14                        | 333746 | 3019609 | 902391 | 892880 | 1761730 | 1253549 |
| S15                        | 326056 | 3081598 | 873733 | 915939 | 1648931 | 1233069 |
| S16                        | 349341 | 3455577 | 1109170 | 931524 | 1838868 | 1303958 |
| S17                        | 322297 | 1534359 | 730817 | 664468 | 1715616 | 1287910 |
| S18                        | 318742 | 4578390 | 715003 | 875047 | 1644804 | 1157746 |
| S19                        | 319628 | 4515086 | 782075 | 675779 | 1662081 | 1141452 |
In Figure 4, the selected features for each feature selection approach are presented, namely FFS, BFE, RF, BPSO, GA, and DE, for each monitored subject.

4.4. Comparison of the DLAs Classification Results Obtained Using the Adapted Version of BPSO with the Results Obtained Using Other Methods

In Table 6 the DLAs classification results for each feature selection approach are presented, when the applied classification algorithm is RF. For eight out of the 19 monitored subjects, BPSO returns classification results that are from the top three results. Moreover, BPSO returns the best classification result in the case of the ninth monitored subject and this result is very promising, showing that further improvements of the algorithm might lead to even better classification results. Another promising observation is that GA returns the best classification result in the case of the fifteenth monitored subject and DE returns the best classification result in the case of the eighteenth monitored subject.
Therefore, further improvements of the objective function and the application of other bio-inspired algorithms might lead to better classification results.

Table 6. The random forest accuracy classification results for each feature selection approach for the DaLiAc dataset for each monitored subject.

| Subject/Features Selection | C   | RW  | LA  | RH  | FFS | BFE | RF  | BPSO | GA  | DE  |
|----------------------------|-----|-----|-----|-----|-----|-----|-----|------|-----|-----|
| S1                         | 90.9% | 93.7% | 87.1% | 89.2% | 91.7% | 92.8% | 91.5% | 92.1% | 91.5% | 90.1% |
| S2                         | 94.1% | 97.1% | 88.8% | 90.6% | 97.5% | 97.5% | 97.0% | 95.8% | 96.8% | 94.2% |
| S3                         | 91.9% | 94.0% | 86.7% | 90.3% | 93.0% | 93.4% | 94.0% | 93.4% | 93.4% | 92.7% |
| S4                         | 90.5% | 96.2% | 87.5% | 89.5% | 95.9% | 95.9% | 95.4% | 95.7% | 93.5% | 91.2% |
| S5                         | 92.1% | 95.4% | 82.3% | 90.7% | 95.3% | 95.3% | 94.9% | 94.7% | 88.8% | 93.6% |
| S6                         | 95.5% | 94.1% | 88.3% | 93.3% | 95.1% | 96.8% | 95.4% | 96.6% | 93.2% | 94.4% |
| S7                         | 94.6% | 95.0% | 87.1% | 93.0% | 96.1% | 95.2% | 94.5% | 92.5% | 91.9% | 91.2% |
| S8                         | 92.1% | 93.9% | 88.3% | 94.9% | 96.9% | 96.9% | 98.5% | 95.1% | 93.8% | 88.7% |
| S9                         | 92.3% | 93.4% | 83.6% | 88.2% | 94.4% | 94.3% | 94.0% | 94.7% | 89.6% | 94.0% |
| S10                        | 94.5% | 94.1% | 87.0% | 88.4% | 95.0% | 95.0% | 94.5% | 92.4% | 93.7% | 90.0% |
| S11                        | 91.7% | 91.4% | 84.7% | 88.6% | 92.1% | 91.9% | 92.7% | 90.4% | 88.9% | 89.3% |
| S12                        | 94.3% | 94.1% | 85.1% | 90.5% | 94.6% | 95.6% | 94.5% | 94.0% | 93.3% | 94.5% |
| S13                        | 92.6% | 94.1% | 86.0% | 89.8% | 93.2% | 92.8% | 92.8% | 91.9% | 94.0% | 91.4% |
| S14                        | 91.4% | 95.6% | 86.3% | 91.3% | 95.4% | 95.9% | 95.1% | 95.0% | 90.9% | 94.1% |
| S15                        | 91.6% | 93.2% | 88.8% | 91.7% | 93.0% | 93.3% | 93.2% | 92.5% | 93.3% | 90.0% |
| S16                        | 91.5% | 91.6% | 81.3% | 90.1% | 94.2% | 94.2% | 91.5% | 92.8% | 91.2% | 89.3% |
| S17                        | 95.1% | 95.3% | 89.2% | 94.5% | 95.8% | 95.9% | 94.4% | 94.5% | 92.8% | 95.0% |
| S18                        | 89.0% | 93.3% | 83.6% | 88.1% | 93.5% | 93.5% | 93.5% | 92.6% | 88.7% | 93.9% |
| S19                        | 92.1% | 96.0% | 88.4% | 92.2% | 96.2% | 97.0% | 96.2% | 93.4% | 92.9% | 91.2% |

Average: 92.5% | 94.2% | 86.1% | 90.4% | 94.6% | 94.9% | 93.9% | 93.8% | 92.0% | 92.0%

1 (1) The best results are underlined with one line, (2) the second best results are underlined with two lines, (3) the third best results are underlined with three lines, (4) the worst results are underlined with one dash line.

5. Discussion

In this section, a critical discussion and the main focus is on the following aspects is presented: (1) the results obtained when the adapted BPSO algorithm is applied for each type of DLA and (2) the comparison of the performance of the machine learning approach based on the BPSO algorithm with the performance of other approaches from the literature that consider the same benchmark dataset.

5.1. Application of Adapted BPSO Algorithm for DLAs Classification

The experiments were performed in KNIME as follows: the input data for each feature selection approach is read from an Excel Reader (XLS) node, the X-Partitioner node represents the beginning of the cross validation and it applies a number of validations equal to 10 and random sampling, two Number To String nodes are used for converting the data types of the labels of the samples from the training data and from the testing data from number to string, in order to prepare the data for classification, the Random Forest Learner node is used for creating a classification model and the split criterion is Information Gain Ration, the Random Forest Predictor node is used for predicting the results, considering the data resulted from the Random Forest Learner node and the testing data as input,
one **String to Number** node is used for converting the data type of the predicted labels from string to number, the **X-Aggregator** node represents the end of the cross-validation, and finally the **Scorer** node is used for computing the accuracy.

In Figure 5, the classification results for the DaLiAc Dataset for each monitored subject are presented. The best results were obtained in the case of the activities sitting (DLA1) and lying (DLA2) and the worst results were obtained in the case of the activities bicycling (50 watt) (DLA11) and bicycling (100 watt) (DLA12).

| DLA1 | DLA2 | DLA3 | DLA4 | DLA5 | DLA6 | DLA7 | DLA8 | DLA9 | DLA10 | DLA11 | DLA12 | DLA13 |
|------|------|------|------|------|------|------|------|------|-------|-------|-------|-------|
| S1   | 85.4% | 99.9% | 85.2% | 92.4% | 89.4% | 92.6% | 93.3% | 93.7% | 95.1% | 98.3% | 86.0% | 87.8% | 99.2% |
| S2   | 99.9% | 100.0% | 99.2% | 90.8% | 82.2% | 88.5% | 96.5% | 91.7% | 96.2% | 99.8% | 99.6% | 99.5% | 98.2% |
| S3   | 99.7% | 100.0% | 99.6% | 96.1% | 86.0% | 92.1% | 91.1% | 93.5% | 90.0% | 94.8% | 91.6% | 91.2% | 94.5% |
| S4   | 99.9% | 100.0% | 99.5% | 97.2% | 91.6% | 94.3% | 93.3% | 93.2% | 95.8% | 98.9% | 90.4% | 91.4% | 97.5% |
| S5   | 100.0% | 100.0% | 96.8% | 98.2% | 92.7% | 95.3% | 92.5% | 95.0% | 94.9% | 98.1% | 89.0% | 93.6% | 97.4% |
| S6   | 99.9% | 100.0% | 99.5% | 97.5% | 94.7% | 97.3% | 95.5% | 94.1% | 90.8% | 99.2% | 93.8% | 94.4% | 99.1% |
| S7   | 100.0% | 100.0% | 95.2% | 95.6% | 88.4% | 93.8% | 91.6% | 89.2% | 91.5% | 99.2% | 85.6% | 85.2% | 98.4% |
| S8   | 99.7% | 99.9% | 98.4% | 95.6% | 88.3% | 95.1% | 91.9% | 90.5% | 93.1% | 96.7% | 95.6% | 96.0% | 95.8% |
| S9   | 100.0% | 100.0% | 100.0% | 98.0% | 92.3% | 97.8% | 94.3% | 93.6% | 95.4% | 96.7% | 89.1% | 86.4% | 95.4% |
| S10  | 100.0% | 100.0% | 97.8% | 93.7% | 85.1% | 90.4% | 86.0% | 85.1% | 87.5% | 95.9% | 96.0% | 94.6% | 97.0% |
| S11  | 100.0% | 100.0% | 92.9% | 92.9% | 85.4% | 89.5% | 91.1% | 89.3% | 87.6% | 96.9% | 80.9% | 81.8% | 94.7% |
| S12  | 100.0% | 100.0% | 98.6% | 97.2% | 88.1% | 95.6% | 92.3% | 89.5% | 89.1% | 96.2% | 89.9% | 91.7% | 93.4% |
| S13  | 99.9% | 99.9% | 96.9% | 96.0% | 88.9% | 96.2% | 94.2% | 93.8% | 94.0% | 97.3% | 87.0% | 90.5% | 96.9% |
| S14  | 100.0% | 100.0% | 100.0% | 94.8% | 91.2% | 94.9% | 93.6% | 90.7% | 91.9% | 95.8% | 94.6% | 94.7% | 94.7% |
| S15  | 99.9% | 99.9% | 94.8% | 94.8% | 90.2% | 94.1% | 95.8% | 92.7% | 95.2% | 98.0% | 77.7% | 78.6% | 99.2% |
| S16  | 98.0% | 100.0% | 96.7% | 95.5% | 92.1% | 91.4% | 89.9% | 91.0% | 93.3% | 97.8% | 89.9% | 97.9% | 95.3% |
| S17  | 100.0% | 100.0% | 99.6% | 96.6% | 92.0% | 94.2% | 94.9% | 94.9% | 91.9% | 97.7% | 88.4% | 90.1% | 100.0% |
| S18  | 99.6% | 100.0% | 99.6% | 93.9% | 89.4% | 93.1% | 90.7% | 88.1% | 91.3% | 94.7% | 89.4% | 87.5% | 95.3% |
| S19  | 100.0% | 100.0% | 95.2% | 93.5% | 87.4% | 93.0% | 88.2% | 86.5% | 87.7% | 94.2% | 96.3% | 97.1% | 94.3% |

**Figure 5.** The DLAs precision results for the DaLiAc dataset using adapted BPSO and RF for each monitored subject.

5.2. Comparison of the Performance of the BPSO Based Approach with the Performance of Literature Approaches

In this subsection, the results obtained using the machine learning methodology based on BPSO are compared with some of the best literature results. Table 7 presents a comparison of the results obtained, using the adapted version of the BPSO algorithm with the results obtained in the literature.

| Approach                             | Result     | Method                                                                 |
|--------------------------------------|------------|------------------------------------------------------------------------|
| Our approach                         | 93.8% accuracy | The features are selected using an adapted version of BPSO and the data is classified using RF |
| (Zdravevski et al. [32])             | 93.4% accuracy | A method that applies the score-drift feature selection, based on an algorithm for feature extraction, selection, and classification |
| (Leutheuser et al. [16])             | 89.6% accuracy | A hierarchical multi-sensor based classification system |
| (Hur et al. [33])                    | 96.4% accuracy | A method that is based on Iss2Image-UCNet6 |

**Table 7.** Comparison of the results obtained using the adapted version of the BPSO algorithm with the results obtained in the literature using the DaLiAc dataset as the experimental support.
As can be seen in the table, the approach proposed in this article returns results comparable to the ones from the literature. The method presented in [33] returns much better results; however, in our approach we consider a maximum of six features and, in that approach, all features are used, therefore the two approaches are not directly comparable.

6. Conclusions

In this article, we presented a novel approach for feature selection in the case of data that is generated by sensors that monitor various types of DLAs. The method was tested and validated on the DaLiAc dataset. The experimental results show that the adapted version of BPSO presented in this article is comparable to other classical algorithms such as FFS, BFE, and RF, in terms of running time and classification performance.

The running time of the adapted version of BPSO is better than the running time of BFE and the developed objective function also returns promising results in combination with other bio-inspired algorithms, such as GA and DE. Moreover, for some monitored subjects the approach based on BPSO returns results that are much better than the ones obtained using classical algorithms for feature selection. Even if the proposed method based on BPSO returns results, which are not as good as the ones obtained using classical methods for some monitored subjects, further improvements of the objective function might lead to better results.

For example, the method presented in [33] performs significantly better; however, it requires complex computational resources and image processing transformations.

As future research work the following research directions are proposed: (1) The testing and the validation of the proposed methodology using other validation types, such as the leave-one-subject validation type or the leave-two-subjects validation type, (2) the proposal of another heuristic for the ranking of the features returned by the adapted version of BPSO and (3) the comparison of the performance of the adapted version of BPSO with the performance of similar approaches based on other types of bio-inspired algorithms.

Author Contributions: Conceptualization, D.M., I.A. and T.C.; methodology, D.M. and T.C.; software, D.M. and I.A.; validation, D.M., I.A. and T.C.; formal analysis, I.S.; resources, I.A., T.C. and I.S.; writing–original draft preparation, D.M.; writing–review and editing, D.M., I.A. and T.C.; supervision, I.S.; project administration, I.A. and T.C.; funding acquisition, I.A., T.C. and I.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Romanian National Authority for Scientific Research and Innovation, CCCDI–UEFISCDI and of the AAL Programme with co-funding from the European Union’s Horizon 2020 research and innovation programme grant number AAL59/2018 ReMIND within PNCDI III. The APC was funded by the Technical University of Cluj-Napoca through the grants for scientific research support programme.

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

Abbreviations

The following abbreviations are used in this manuscript:
Appl. Sci. 2020, 10, 1496

- AAL: ambient assisted living
- BFE: backward features elimination
- BPSO: binary particle swarm optimization
- C: chest
- CNN: convolutional neural network
- CR: crossover rate or crossover probability
- DaLiAc: Daily Life Activities
- DE: differential evolution
- DLA: daily living activity
- ERT: extremely randomized trees
- F: differential weight
- FFS: forward feature selection
- GA: genetic algorithm
- k-NN: k-nearest neighbor
- KNIME: Konstanz Information Miner
- HPSO-LS: hybrid particle swarm optimization with local search
- HPSO-SSM: hybrid particle swarm optimization with a spiral-shaped mechanism
- ICT: information and communication technology
- II: interaction information
- IoT: internet of things
- LA: left ankle
- LR: logistic regression
- MR: mutation rate
- NB: naive bayes
- PSO: particle swarm optimization
- PSO-LM: particle swarm optimization with learning memory
- RF: random forest
- RH: right hip
- RW: right wrist
- SVM: support vector machine
- UCI: University of California, Irvine
- VAF: variance accounted for

Appendix A

The ranking of the features for each monitored subject $S$ of the DaLiAc dataset in the adapted version of BPSO introduced in this article is presented next:

1. $S_1 = CA_1, CA_3, WA_3, CA_2, WA_1, WA_2, HA_3, HA_1, HA_2, AA_3, AA_2, AA_1, CG_3, HG_3, WG_2, AG_1, AG_2, CG_2, HG_2, HG_1, CG_1, WG_3, WG_1, AG_3$;
2. $S_2 = CA_1, CA_3, HA_3, CA_2, HA_1, WA_3, HA_2, WA_1, WA_2, AA_2, AA_3, AA_1, CG_3, HG_3, HG_1, CG_1, CG_2, HG_2, WG_2, AG_1, WG_3, AG_2, AG_3, WG_1$;
3. $S_3 = CA_1, HA_3, HA_2, CA_2, HA_1, CA_3, WA_1, WA_3, WA_2, AA_3, AA_2, AA_1, CG_3, HG_3, HG_1, HG_2, CG_2, WG_2, CG_1, WG_3, AG_1, AG_2, AG_3, WG_1$;
4. $S_4 = CA_1, HA_2, CA_3, HA_1, HA_3, CA_2, WA_1, WA_3, WA_2, AA_3, AA_2, AA_1, CG_3, HG_3, HG_1, AG_1, AG_3, CG_2, HG_2, WG_2, WG_3, AG_2, CG_1, WG_1$;
5. $S_5 = CA_1, CA_3, CA_2, WA_3, HA_3, HA_2, WA_1, HA_1, WA_2, AA_3, AA_2, AA_1, CG_3, HG_3, CG_2, WG_2, AG_1, HG_2, HG_1, CG_1, WG_3, AG_2, AG_3, WG_1$;
6. $S_6 = HA_3, CA_1, WA_3, HA_2, CA_3, WA_1, CA_2, HA_1, WA_2, AA_3, AA_2, AA_1, CG_3, HG_3, WG_2, HG_2, HG_1, CG_2, WG_1, AG_3, AG_1, AG_2, WG_1, AG_3$;
7. $S_7 = CA_1, CA_3, CA_2, WA_1, WA_3, HA_1, HA_3, HA_2, WA_2, AA_3, AA_2, AA_1, CG_3, CG_1, HG_3, CG_2, WG_2, HG_1, WG_3, AG_1, AG_2, AG_3, HG_2, WG_1$;
8. $S_8 = CA_1, CA_3, CA_2, HA_2, WA_3, HA_1, WA_1, HA_3, WA_2, AA_3, AA_2, AA_1, CG_3, AG_1, HG_3, CG_2, WG_2, HG_1, WG_3, AG_2, HG_2, CG_1, AG_3, WG_1$;
References

1. ReMIND. Available online: https://www.aalremind.eu/ (accessed on 15 January 2020).
2. Moldovan, D.; Anghel, I.; Cioara, T.; Salomie, I.; Chifu, V.; Pop, C. Kangaroo mob heuristic for optimizing features selection in learning the daily living activities of people with Alzheimer’s. In Proceedings of the 22nd International Conference on Control Systems and Computer Science (CSCS), Bucharest, Romania, 28–30 May 2019; pp. 236–243. [CrossRef]
3. Schneider, C.; Trukeschitz, B.; Rieser, H. Measuring the use of the active and assisted living prototype CARMO for home care service users: Evaluation framework and results. Appl. Sci. 2020, 10, 38. [CrossRef]
4. Maskelius, R.; Damasevicius, R.; Segal, S. A review of internet of things technologies for ambient assisted living environments. Future Internet 2019, 11, 1–23. [CrossRef]
5. Dziak, D.; Jachimczyk, B.; Kulesza, W. J. IoT-based information system for healthcare application: Design methodology approach. Appl. Sci. 2017, 7, 996. [CrossRef]
6. Terashi, H.; Mitoma, H.; Yoneyama, M.; Aizawa, H. Relationship between amount of daily movement measured by a triaxial accelerometer and motor symptoms in patients with Parkinson’s disease. Appl. Sci. 2017, 7, 486. [CrossRef]
7. Samie, F.; Bauer, L.; Henkel, J. From cloud down to things: An overview of machine learning in internet of things. IEEE Internet Things J. 2019, 6, 4921–4934. [CrossRef]
8. Vitabile, S.; Marks, M.; Stojanovic, D.; Plana, S.; Molina, J.M.; Krzyżton, M.; Sikora, A.; Jarynowski, A.; Hosseinpour, E.; Jakobik, A.; et al. Medical data processing and analysis for remote health and activities monitoring. In High-Performance Modelling and Simulation for Big Data Applications: Selected Results of the COST Action IC1406 chilPSet; Kolodziej, J., Gonzalez-Velez, H., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 186–220. [CrossRef]
9. Chelli, A.; Patzold, M. A machine learning approach for fall detection and daily living activity recognition. IEEE Access 2019, 7, 38670–38687. [CrossRef]
10. Saadeh, W.; Butt, S.A.; Alfat, M.A.B. A patient-specific single sensor IoT-based wearable fall prediction and detection system. IEEE Trans. Neural Syst. Rehabil. Eng. 2019, 5, 995–1003. [CrossRef] [PubMed]
11. Yatbaz, H.Y.; Eraslan, S.; Yesilada, Y.; Ever, E. Activity recognition using binary sensors for elderly people living alone: Scanpath trend analysis approach. IEEE Sens. J. 2019, 19, 7575–7582. [CrossRef]
12. Awais, M.; Chiari, L.; Ihlen, E.A.F.; Helbostad, J.; Palmerini, L. Physical activity classification for elderly people in free-living conditions. *IEEE J. Biomed. Health* 2019, 23, 197–207. [CrossRef] [PubMed]

13. Yahaya, S.W.; Lotfi, A.; Mahmud, M. A consensus novelty detection ensemble approach for anomaly detection in activities of daily living. *Appl. Soft Comput.* 2019, 83, 105613. [CrossRef]

14. Uddin, M.Z.; Hassan, M.M.; Alsanad, A.; Savaglio, C. A body sensor data fusion and deep recurrent neural network-based behavior recognition approach for robust healthcare. *Inform. Fusion* 2020, 55, 105–115. [CrossRef]

15. De-La-Hoz-Franco, E.; Ariza-Colpas, P.; Quero, J.M.; Espinilla, M. Sensor-based datasets for human activity recognition—A systematic review of literature. *IEEE Access* 2018, 6, 59192–59210. [CrossRef]

16. Leutheuser, M.; Schludhaus, D.; Eskofier, B.M. Hierarchical, multi-sensor based classification of daily life activities: Comparison with state-of-the-art algorithms using a benchmark dataset. *PLoS ONE* 2013, 8, 1–11. [CrossRef] [PubMed]

17. Lin, Q.; Liu, S.; Zhu, Q.; Tang, C.; Song, R.; Chen, J.; Coello, C.A.; Wong, K.-C.; Zhang, J. Particle swarm optimization with a balanceable fitness estimation for many-objective optimization problems. *IEEE T. Evolut. Comput.* 2018, 22, 32–46. [CrossRef]

18. Zhou, P.; Hu, X.; Li, P.; Wu, X. Online feature selection for high-dimensional class-imbalanced data. *Knowl. Based Syst.* 2017, 136, 187–199. [CrossRef]

19. Liu, M.; Xu, C.; Luo, Y.; Xu, C.; Wen, Y.; Tao, D. Cost-sensitive feature selection by optimizing F-measures. *IEEE T. Image Process.* 2018, 27, 1323–1335. [CrossRef] [PubMed]

20. Maldonado, S.; Lopez, J. Dealing with high-dimensional class-imbalanced datasets: Embedded feature selection for SVM classification. *Appl. Soft Comput.* 2018, 67, 94–105. [CrossRef]

21. Xu, Y. Maximum margin of twin spheres support vector machine for imbalanced data classification. *IEEE T. Cybern.* 2017, 47, 1540–1550. [CrossRef] [PubMed]

22. Chen, H.; Li, T.; Fan, X.; Luo, C. Feature selection for imbalanced data based on neighborhood rough sets. *Inform. Sciences* 2019, 483, 1–20. [CrossRef]

23. Hosseini, E.S.; Moattar, M.H. Evolutionary feature subsets selection based on interaction information for high dimensional imbalanced data classification. *Appl. Soft Comput.* 2019, 82, 105581. [CrossRef]

24. Xue, B.; Zhang, M.; Browne, W.N. Particle swarm optimisation for feature selection in classification: Novel initialisation and updating mechanisms. *Appl. Soft Comput.* 2014, 18, 261–276. [CrossRef]

25. Moradi, P.; Gholampour, M. A hybrid particle swarm optimization for feature subset selection by integrating a novel local search strategy. *Appl. Soft Comput.* 2016, 43, 117–130. [CrossRef]

26. Samanthula, B.K.; Elmehdwi, Y.; Jiang, W. K-nearest neighbor classification over semantically secure encrypted relational data. *IEEE T Knowl. Data En.* 2015, 27, 1261–1273. [CrossRef]

27. Wei, B.; Zhang, W.; Xia, X.; Zhang, Y.; Yu, F.; Zhu, Z. Efficient feature selection algorithm based on particle swarm optimization with learning memory. *IEEE Access* 2019, 7, 166066–166078. [CrossRef]

28. Xiong, B.; Li, Y.; Huang, M.; Shi, W.; Du, M.; Yang, Y. Feature selection of input variables for intelligence joint moment prediction based on binary particle swarm optimization. *IEEE Access* 2019, 7, 182289–182295. [CrossRef]

29. Chen, K.; Zhou, F.-Y.; Yuan, X.-F. Hybrid particle swarm optimization with spiral-shaped mechanism for feature selection. *Expert Syst. Appl.* 2019, 128, 140–156. [CrossRef]

30. Casale, P.; Altini, M.; Amft, O. Transfer learning in body sensor networks using ensembles of randomized trees. *IEEE Internet Things* 2015, 2, 33–40. [CrossRef]

31. Nazabal, A.; García-Moreno, P.; Artes-Rodriguez, A.; Ghahramani, Z. Human activity recognition by combining a small number of classifiers. *IEEE J. Biomed. Health* 2016, 20, 1342–1351. [CrossRef] [PubMed]

32. Zdravevski, E.; Lameski, P.; Trajkovik, V.; Kulakov, A.; Chorbev, I.; Goleva, R.; Pombo, N.; Garcia, N. Improving activity recognition accuracy in ambient-assisted living systems by automated feature engineering. *IEEE Access* 2017, 5, 5262–5280. [CrossRef]

33. Hur, T.; Bang, J.; Huynh-The, T.; Lee, J.; Kim, J.-I.; Lee, S. Iss2Image: A novel signal-encoding technique for CNN-based human activity recognition. *Sensors* 2018, 18, 3910. [CrossRef] [PubMed]

34. Kumar, A.; Kim, J.; Lyndon, D.; Fulhorn, M.; Feng, D. An ensemble of fine-tuned convolutional neural networks for medical image classification. *IEEE J. Biomed. Health* 2017, 21, 31–40. [CrossRef] [PubMed]

35. Liu, J.; Mei, Y.; Li, X. An analysis of the inertia weight parameter for binary particle swarm optimization. *IEEE T Evolut. Comput.* 2016, 20, 666–681. [CrossRef]
36. Too, J.; Abdullah, A.R.; Saad, N.M. A new co-evolution binary particle swarm optimization with multiple inertia weight strategy for feature selection. *Informatics* 2019, 6, 21. [CrossRef]

37. Feltrin, L. KNIME an open source solution for predictive analytics in the geosciences [Software and Data Sets]. *IEEE Geosci. Rem. Sens. M.* 2015, 3, 28–38. [CrossRef]

38. Macedo, F.; Oliveira, M.R.; Pacheco, A.; Valadas, R. Theoretical foundations of forward feature selection methods based on mutual information. *Neurocomputing* 2019, 325, 67–89. [CrossRef]

39. Maldonado, S.; Weber, R.; Famili, F. Feature selection for high-dimensional class-imbalanced data sets using support vector machines. *Inform. Sci.* 2014, 286, 228–246. [CrossRef]

40. Bader-El-Den, M.; Teitei, E.; Perry, T. Biased random forest for dealing with the class imbalance problem. *IEEE Trans. Neural Netw. Learn. Syst.* 2019, 30, 2163–2172. [CrossRef] [PubMed]

41. Alirezazadeh, P.; Fathi, A.; Abdali-Mohammadi, F. A genetic algorithm-based feature selection for kinship verification. *IEEE Signal Process. Lett.* 2015, 22, 2459–2463. [CrossRef]

42. Zhang, Y.; Gong, D.-W.; Gao, X.-Z.; Tian, T.; Sun, X.-Y. Binary differential evolution with self-learning for multi-objective feature selection. *Inform. Sci.* 2020, 507, 67–85. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).