

RESEARCH

Open Access



Sleep stage classification using extreme learning machine and particle swarm optimization for healthcare big data

Nico Surantha* , Tri Fennia Lesmana and Sani Muhamad Isa

*Correspondence:
nico.surantha@binus.ac.id
Computer Science
Department, BINUS
Graduate Program-Master
of Computer Science, Bina
Nusantara University, Jl. K. H.
Syahdan No. 9, Kemanggisian,
Palmerah, Jakarta 11480,
Indonesia

Abstract

Recent developments of portable sensor devices, cloud computing, and machine learning algorithms have led to the emergence of big data analytics in healthcare. The condition of the human body, e.g. the ECG signal, can be monitored regularly by means of a portable sensor device. The use of the machine learning algorithm would then provide an overview of a patient's current health on a regular basis compared to a medical doctor's diagnosis that can only be made during a hospital visit. This work aimed to develop an accurate model for classifying sleep stages by features of Heart Rate Variability (HRV) extracted from Electrocardiogram (ECG). The sleep stages classification can be utilized to predict the sleep stages proportion. Where sleep stages proportion information can provide an insight of human sleep quality. The integration of Extreme Learning Machine (ELM) and Particle Swarm Optimization (PSO) was utilized for selecting features and determining the number of hidden nodes. The results were compared to Support Vector Machine (SVM) and ELM methods which are lower than the integration of ELM with PSO. The results of accuracy tests for the combined ELM and PSO were 62.66%, 71.52%, 76.77%, and 82.1% respectively for 6, 4, 3, and 2 classes. To sum up, the classification accuracy can be improved by deploying PSO algorithm for feature selection.

Keywords: Electrocardiogram, Extreme learning machine, Heart rate variability, Particle swarm optimization, Sleep stages

Introduction

During good sleep, muscle and tissue rejuvenate. Memory consolidation also occurs. Therefore, maintaining the quality of sleep is crucial for human beings. Based on the survey of the National Sleep Foundation on 1000 participants from USA [1], 13% of the participants did not have enough sleep on non-workdays. The percentage was higher on workdays, in which 30% of the participants did not have enough sleep. The findings of this survey show that many people lack sleep, particularly on workdays.

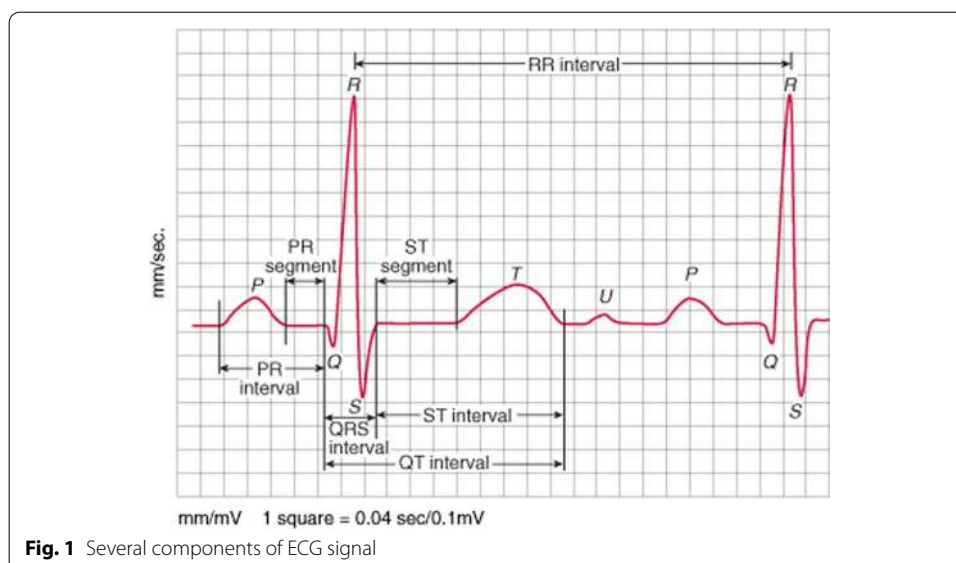
The lack of sleep may disturb individuals' life by reducing productivity due to energy loss [2]. Early detection of sleep related disorders may prevent further sleep disorders in which evaluating sleep stages can be a good method for sleep study and early indication. According to the gold standard of sleep study, several stages of sleep can be

classified as awake, REM, and NREM [3]. The NREM contains several stages such as light sleep (stage 1 and 2) and deep sleep (stage 3 and 4).

Classification of sleep stages can be done by visiting a sleep specialist. An experiment will be conducted to patients in the specific room to monitor their psychological conditions during sleep. Common devices such as electroencephalography and ECG can be employed for detecting brain wave and heart rate, respectively. For this work, ECG was chosen rather than EEG because it provides more comfort, where in EEG installation, EEG electrodes need to be placed on the patients [4]. For ECG case, a non-contact ECG sensor already exists [5]. The ECG produces signal as the result, as shown in Fig. 1. Each data generated can be identified by sleep specialists. However, manual identification is very tedious and exhausting. Automatic identification of sleep stages will be very helpful for sleep research as the initial diagnosis.

Recent developments of portable sensors, embedded system, cloud computing, and machine learning algorithms have led to the emergence of big data analytics for healthcare. Sleep monitoring system is developed using these technologies to allow patients to monitor their sleep condition on a regular basis. The ECG signal collected using a portable ECG sensor can be aggregated in the data aggregator before being sent to cloud computing for signal processing [6]. Sleep stage can be classified automatically by machine learning algorithm for early diagnostic. The sleep stages classification can be used to predict the proportion of each sleep stage. Where sleep stages proportion information can provide an insight of human sleep quality. For example, when it is found that the patient wakes up more than once during the night sleep, it is one indication of poor sleep quality.

In the field of machine learning, artificial neural network has been known as a promising and emerging method which can be applied in many fields. One of the advantages of applying neural network in sleep study is that it has tolerance for some undesirable data or events [7] with the ability to learn non-linear and complex



relations [8]. On the other hand, neural network has some drawbacks, such as slow backpropagation learning which can lead to local optima solution [9].

Considering the time required for learning algorithm, neural network algorithm such as ELM provides a faster processing ELM can be considered as a feedforward neural network algorithm which usually has input, hidden, and output layer learned only in a single step. According to the previous implementation of ELM algorithm, it gives the best accuracy compared to SVM and BPNN with faster processing time [10]. The faster processing time of ELM is due to a better generalization ability where ELM chooses the input weights randomly and uses simple matrix computations to determine the output weights.

SVM is another commonly used machine learning algorithm. It is the algorithm that operates using an optimal hyperplane and utilizes it for separating data based on the classes. Using this concept, a reliable classification tool can be achieved, in which it is widely applied in signal processing [11]. Due to SVM popularity and well-known performances, SVM was deployed in this work.

The data features must be extracted in order to use ELM and SVM algorithm for identifying sleep stages. Features can be extracted based on heart rate variability from ECG. The HRV is one of physiological phenomena showing the continuous beating of the heart which varies during sleeping. To extract the features of HRV, several methods can be employed, such as geometrical [12], time domain [12], poincare [13] as well as frequency domain [14]. Some features can be irrelevant and redundant for the classification model, especially when using a large number of features. Hence, the optimum features should be carefully selected to obtain low processing time for the model. To aid the feature selection process and determining hidden node number, PSO is chosen as algorithm in this work. Previous study has shown a promising result in using PSO for selecting features [15]. PSO showed minor accuracy reduction with significantly faster performance compared to other algorithms such as GA [16].

The study was conducted using several methods. The classification of sleep stages was conducted using ELM and SVM. PSO was later integrated with ELM for finding optimum features and determining the number of hidden nodes. The performance of ELM, SVM, and the combination of ELM and PSO were evaluated to obtain the optimum algorithm by calculating the accuracy percentage.

The main contributions of this study are:

1. Building a classifier model using the combination of ELM-PSO to achieve optimal accuracy in terms of sleep stages classification using HRV features obtained from ECG signal;
2. Evaluating the performance of the proposed ELM-PSO model in 6, 4, 3, and 2 sleep stage classes.

The rest of the paper is arranged as follows: Sect. “[Related Works](#)” discusses related works. Sect. “[Methods](#)” presents the dataset and the algorithm used in this study. Finally, in sect. “[Results and discussion](#)”, the steps, results, and analysis are presented. Section “[Conclusions](#)” provides the conclusion of the paper.

Related works

Since the invention of standardized sleep scoring in 1968, it has become a gold standard for studies on sleep [3]. Along with the standard, algorithm for automatic sleep stage classification has been developed since 1969 using EEG signal [17]. It was found that sleep stages can be identified using several physiologic parameters which can be extracted from EEG, ECG, electro-oculogram, electromyogram, pulse oximetry, airflow and respiratory effort [18]. Studies on sleep stage classification use ECG signal, rather than multi signal to reduce the complexity and increase comfort because the use of ECG signal could produce high accuracy [19]. Several methods of machine learning have been used to analyze the data from ECG to identify sleep stages (e.g., sleep, NREM, REM, awake) [20–25] and sleep disorders [26–28]. The sleep disorders are not covered in this work since the study is focused on sleep stage classification.

Adnane et al. used MITBPD for signal input and categorized sleep stages into two classes, which are sleep and awake [20]. SVM was utilized as machine learning algorithm, and MITBPD was deployed to record physiologic signal in that study. Twelve features were extracted based on HRV, the detrended fluctuation analysis (DFA), and windowed DFA; 20% of the data were utilized for training, while the remaining 80% were dedicated for classification. The testing was divided into two sets of features, which are 12 and optimum 10 features, in which the optimum ten features were determined using SVM-RFE. After 5 repeated testing for each set, the generated accuracies were $79.31 \pm 4.52\%$ for 12 features (Cohen's kappa value, $k=0.41$) and $79.99 \pm 4.64\%$ for 10 features ($k=0.43$). Also, it was concluded that HRV performed better than DFA and windowed DFA.

Xiao et al. classified sleep stages into three categories, which are awake, REM, and NREM sleep [21]. RF was used for the machine learning method. The data were gathered from SHRSV, resulting in 41 HRV features in which 25 features were novel. Using RF method, two scenarios were conducted, which were subject specific and subject independent classifier. In subject specific classifier scenario, data were divided into 80% training and 20% testing for each recording, followed by 10 repetitions of randomized data set to minimize bias. In the subject independent classifier scenario, 44 recordings were grouped as training data while one recording was set as testing data. The process was repeated 45 times, and the resulting data were averaged. The resulting accuracies were 88.67% for subject specific classifier ($k=0.7393$) and 72.58% for subject independent classifier ($k=0.4627$). Xiao et al. suggest that the proposed method could be used as an alternative or aiding technique for solving complex usage of PSG system and convenient sleep stages classification.

Werteni et al. defined two classes of sleep stages, namely sleep and awake [22]. MITBPD was used in this study, and the features were extracted using HRV and DFA based on QRS detection times. ELM was used as algorithm and compared to SVM and BPNN. One-third of the data were assigned as training data while the remaining 2/3 were testing data. The resulting accuracies were 78.33%, 76.74%, 78.12% respectively for ELM, BPNN, and SVM. One must note that the ELM algorithm exhibited significantly faster detection time compared to BPNN and SVM. Werteni et al. presents a simple automatic sleep–wake stages classifier using only RR series obtained from ECG. Seven features were extracted from the RR series by three methods, HRV, DFA and W DFA.

Rahimi et al. utilized SVM with extension of ECOC and RBF as kernel algorithm for sleep stage classification [25]. Two sets of classes were used in the study, which were sleep–wake and wake-REM-NREM. The methods were HRV, and ECG-derived respiration (EDR) extracted from ECG signal using neural principal component analysis (PCA). The features were extracted by nonparametric Kruskal–Wallis test and optimized using the mRMR algorithm. The resulting accuracies were 81.76% for two sleep stages and 76% for three sleep stages. Rahimi et al. present a sleep stage classification using HRV and EDR features with multi-class SVM classification.

Radha et al. conducted a massive study involving 292 participants within the period of 1997 to 2000 in 7 labs from 5 European countries [23]. The sleep stages were classified into four stages as wake, REM, N1/N2, N3 in which N1, N2, N3 are the stages of NREM sleep. To incorporate long-term temporal dependencies, Long Short-Term Memory (LSTM) network was used in this work as algorithm. The study was performed based on HRV data and achieved $77.00 \pm 8.90\%$ for four classes of sleep stages. Our previous study combined SVM with PSO in which SVM was deployed for classification while PSO was utilized for feature selection [24]. By using HRV as a method to classify 2 sleep stages, PSO was able to increase accuracy from $\approx 72\%$ to 78.41%. In light of those related works, the rapid detection algorithm such as ELM was combined with PSO for feature selection to improve the resulting accuracy. Radha et al. present HRV-based sleep stage classification using LSTM. However, there are still some issues until the system can reach the expected accuracy of EEG-based sleep stage classification.

Yücelbaş et al. conducted research on sleep stage classification with Morphological method using MITBPD public dataset [29]. The classifier model used 15 features extracted based on HRV resulting into 3 sleep stage classes and achieved 77.02% accuracy for the model performance. Yücelbaş et al. present an automatic sleep staging system using morphological method based on ECG signal.

On the other hand, Wei et al. carried out research to classified sleep stages into 3 sleep stages such as awake, REM and NREM [30]. By using 11 HRV features extracted from MITBPD, Wei et al. classified the sleep stages with 77% accuracy. Wei et al. suggest that the values of RR interval can be used as the features and can reduce the number of wearable devices for sleep tracking.

In the previous research, there is a room for improvement for the model performance in 2 and 3 classes. Also, the models implemented on the health monitoring system are required to perform fast and precisely. Therefore, in this study, we propose the sleep stage classification using ELM-PSO. The proposed algorithm was tested in terms of accuracy and processing time. The summary of related works on sleep stage classifications using ECG is presented in Table 1.

Methods

The methodology includes: (1) data collection, (2) pre-processing of data, (3) feature extraction of HRV, and (4) model building and evaluation. The algorithms used in this work were SVM, ELM, and the integration of ELM and PSO. Performance comparison was made. The phases of the methodology is presented as shown in Fig. 2.

Table 1 Summary of related works on sleep stage classifications using ECG

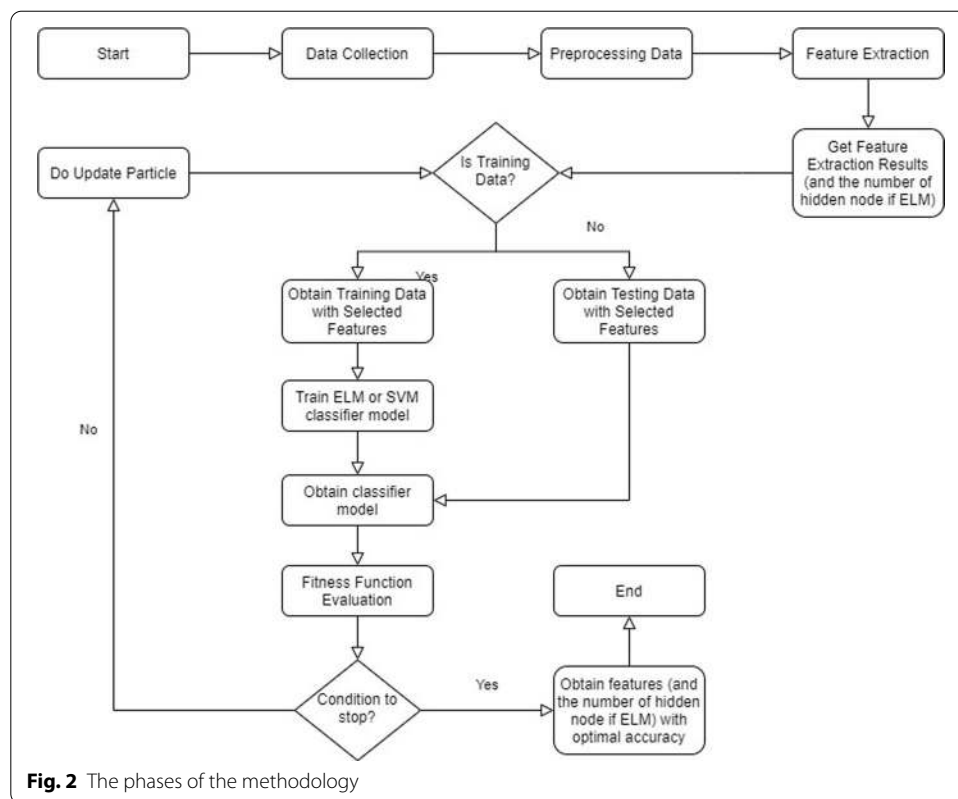
Algorithm [Ref.]	Dataset	Number of classes	Methods	Accuracy
SVM [20]	MITBPD	2	HRV, DFA, windowed DFA	79.31 ± 4.52% for 12 features 79.99 ± 4.64% for 10 features
RF [21]	SHRSV	3	HRV	88.67% for subject specific classifier 72.58% for subject independent classifier
ELM [22]	MITBPD	2	HRV and DFA	78.33% for both HRV and DFA; 76.29% for HRV; 73.48% for DFA
BPNN [22]	MITBPD	2	HRV and DFA	76.74% for both HRV and DFA; 73.81% for HRV; 71.07% for DFA
SVM [22]	MITBPD	2	HRV and DFA	78.12% for both HRV and DFA; 76.03% for HRV; 73.21% for DFA
SVM with ECOC extension and RBF [25]	MITBPD	2 and 3	HRV and EDR	81.76% for 2 classes 76% for 3 classes
LSTM network [23]	Siesta	4	HRV	77.00 ± 8.90%
Combination of SVM and PSO [24]	MITBPD	2	HRV	78.41% for RBF kernel 77.08% for linear kernel
Morphological [29]	MITBPD	3	HRV	77.02%
Deep Neural Network [30]	MITBPD	3	HRV	77%
Combination of ELM and PSO (Proposed Method)	MITBPD	2	HRV	81.96% for 2 classes 76.59% for 3 classes 71.44% for 4 classes 63.18% for 6 classes

Data collection of MIT-BIH polysomnographic dataset

The data obtained [31] contained recordings of physiological signal of the human body during sleeping. ECG, EEG, blood pressure, and respiration were included in MITBPD. The reason for using MITBPD dataset is because it is one of public ECG datasets which has labelled annotations and has been used by many previous research to make sure that the results of comparison carried out are fair. For simplicity, only ECG signal was used, which was digitized with 12 bit per sample and sampling frequency of 250 Hz. The eighteen files recorded during the sleeping period of 2 female and 16 male subjects were utilized. The participants' ages ranged from 32 to 56 years old ($\bar{x}=43$) with weight 89–152 kg ($\bar{x}=119$ kg). Prior to pre-processing, 10274 samples were retrieved in which each sample recording has a duration of 30 s. The MITBPD data distribution are 17.79%, 38.28%, 4.76%, 1.78%, 6.89% and 30.5% for NREM1, NREM2, NREM3, NREM4, REM and awake sleep stage, respectively.

Pre-processing of the data

Data annotation and RR interval were recorded and pre-processed. RR interval is a time delay between the current and previous heart beats [32]. By calculating the time which elapsed between two consecutive R waves of ECG signal, RR interval was



obtained. Data annotation was labelled by a sleep specialist confirming the sleep stage for every 30 s.

Firstly, problematic data were removed. In this step, the annotation data and RR interval were synchronized. The problematic data refer to the missing timestamp that makes both data do not synchronize. The data will be removed. Each annotation data may contain many RR intervals. If the annotation data did not match the scope of the research (such as awake, NREM, REM), then the data were removed. Then, feature extraction was performed to those set of RR interval from each annotation.

Feature extraction of heart rate variability

To track and interpret the electric activity of the heart, ECG can be used [33]. There is a constant variation in human heart beat, which is recorded by ECG. The nervous systems such as parasympathetic and sympathetic nerves control the heart rate fluctuation. In this work, those variabilities were utilized to extract the 18 features as shown in Table 2.

First algorithm: extreme learning machine

Compared to BPNN, ELM provides faster processing time due to better generalization; hence it is easier to implement the algorithm [9]. ELM is a feedforward neural network algorithm which usually has input, hidden, and output layer learned in only one step and does not need iterative training. ELM randomly chooses hidden nodes and analytically determines the output weights of hidden layer and all the parameters are tuned once [34].

Table 2 The list of features extracted from HRV

No.	Name of feature	Method of extraction	Explanation
1	AVNN	Time domain	AVNN is mean of all normal sinus to normal sinus of NN intervals. $AVNN = \frac{1}{N} \sum_{j=1}^N RR_j$ where j is interval index and N is number of total intervals
2	SDNN	Time domain	SDNN is the standard deviation of NN intervals $SDNN = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (RR_j - \overline{RR})^2}$
3	RMSSD	Time domain	Root Mean Square of the Successive Differences (RMSSD) $RMSSD = \sqrt{\frac{1}{N-1} \sum_{j=1}^{N-1} (RR_{j+1} - RR_j)^2}$
4	SDSD	Time domain	The standard deviation of the successive difference between adjacent R-R intervals
5	NN50	Time domain	The number of pairs of adjacent NN intervals differing by more than 50 ms $Numberof(RR_{j+1} - RR_j) > 50$
6	pNN50	Time domain	Percentage of adjacent NN intervals differing by more than 50 ms $pNN50 = \frac{NN50}{N-1} \times 100$
7	HRV Triangular Index	Geometrical	(The total number of RR intervals)/(the height of histogram of all RR intervals with 7.8125 ms bin)
8	SD1	Poincare	$SD1^2 = \frac{1}{2}SDSD^2$ where SD1 = the standard deviation of perpendicular points to the line-of-identity
9	SD2	Poincare	$SD2^2 = 2SDNN^2 - \frac{1}{2}SDSD^2$ where SD2 = the standard deviation of parallel points along the line-of-identity
10	SD1SD2 Ratio	Poincare	$SD1SD2Ratio = \frac{SD1}{SD2}$
11	S	Poincare	$S = \pi \times SD1 \times SD2$
12	TP	Frequency domain	The total power (TP) of the HRV
13	VLF	Frequency domain	Very low frequency (VLF) is defined as the total power (TP) ranging from 0–0.04 Hz
14	LF	Frequency domain	Low frequency is defined as the total power (TP) ranging from 0.04–0.15 Hz
15	HF	Frequency domain	High frequency is defined as the total power (TP) ranging from 0.15–0.4 Hz
16	LFHF Ratio	Frequency domain	The ratio of LF to HF indicates the sympathovagal balance.
17	LFnorm	Frequency domain	$LFnorm = \frac{LF}{TP-VLF}$
18	HFnorm	Frequency domain	$HFnorm = \frac{HF}{TP-VLF}$

Second algorithm: support vector machine

The optimum decision boundary, which separates data, can be found with SVM algorithm. Usually, high value of margin indicates the best decision boundary. Herein, the gap between adjacent data to decision boundary is denoted as margin. For the most real-world problems, the data can be categorized as non-linearly separable in which kernel function can be used to solve the problem. Using kernel function, the data can be separated linearly by mapping the data from input to feature space with higher dimension. Two types of kernel functions can be used in SVM which is linear kernel function and RBF kernel. RBF was chosen in this work, which provided better accuracy in our previous work [24].

Table 3 Particle representation of ELM and PSO integration

Feature mask			Total hidden nodes		
p_i^1	$\dots p_i^k \dots$	p_i^n	p_i^1	$\dots p_i^j \dots$	$p_i^{n_c}$

Feature selection by particle swarm optimization

PSO is a population-based optimization technique inspired by the social behaviour of certain animals in colonies such as bird swarm that work together to achieve goals by moving together, spreading out of the group, and re-group [35]. In PSO, each position of a bird is denoted as particle in which the fitness value will be optimized. A particle with the best fitness value is denoted as g-best while p-best is indicated by the best fitness value for each particle. All of the particles have velocity which is used to evaluate the particle movement to the next position. PSO algorithm is as follows [36]:

- (a) Randomly initializing velocity and particle population;
- (b) Starting new iteration;
- (c) Evaluating the fitness function for all particles;
- (d) Determining p-best of each particle and replacing the p-best when it is better than the previous one;
- (e) Determining g-best value;
- (f) Updating each particle velocity by Eq. (15).

$$V_{id}^{t+1} = WV_{id}^t + c_1 \cdot r_1 \cdot (pbest_{id}^t - x_{id}^t) + c_2 \cdot r_2 \cdot (p_{gd}^t - x_{id}^t). \tag{15}$$

where V_{id}^t is the i-th particle velocity of the t-th iteration, W is inertia weight, and d-th variable [31]. Meanwhile, c_1 is each particle tendency to follow its p-best. r_1 and r_2 is a random factor of 0 to 1. $pbest_{id}^t$ is denoted as the best i-th particle position. c_2 is each particle tendency for following its g-best while p_{gd}^t is g-best. x_{id}^t is the position of i-th particle.

- (g) Updating the position of each particle by Eq. (16).

$$x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1} \tag{16}$$

- (h) The iteration must be stopped when the termination conditions are met. Otherwise, the process must be repeated from point c).

Third algorithm: the integration of ELM with particle swarm optimization

In this work, ELM was integrated with PSO where PSO was used for selecting features and determining hidden node number. Table 3 shows each particle representation in binary containing total hidden nodes and feature mask [37].

The feature selection was represented in the part of feature mask where n is the total features to be selected, where “1” means that the feature has been chosen or else the value is “0”. As an example, the binary of “101011” means the features of 1, 3, 5, and 6 have been selected with a total of 6 features.

The determination of hidden node number was performed in the part of total hidden nodes where n_c is the maximum bit number representing the maximum hidden node number. As an example, the binary of “00111” indicates that 7 hidden nodes were used by the particle.

To define the fitness function, the model was evaluated using the total selected features and the accuracy of ELM. The formula for fitness function is shown in Eq. (17) where W_F and W_A are weighting factors indicating the significance of high accuracy with less selected features. Meanwhile, f_j is j -th bit of feature mask and n_f is total number of features. In the integration of ELM with PSO, the accuracy of the particle was evaluated by ELM algorithm.

$$fitness_i = W_A \times acc_i + W_F \times \left[1 - \frac{\left(\sum_{j=1}^{n_f} f_j \right)}{n_f} \right] \quad (17)$$

Results and discussion

The algorithms were used to classify 4 sets of data, which are 6, 4, 3, and 2 number of sleep stages in which the sets can be categorized as follows (1) 6 classes: awake, stage 1, stage 2, stage 3, stage 4, and REM, (2) 4 classes: awake, deep sleep, light sleep, and REM, (3) 3 classes: awake, NREM, and REM, (4) 2 classes: awake and sleep. The MITBPD dataset distribution are 17.79%, 38.28%, 4.76%, 1.78%, 6.89% and 30.5% for NREM1, NREM2, NREM3, NREM4, REM and awake sleep stage, respectively. To validate the data, the 70% training and 30% testing data were implemented using stratified sampling where the training and testing data were used further for accuracy evaluation. The number of retrieved data was adjusted proportionally. The training and testing results data from train-test split are different from data that used in previous works.

Pre-processing

The invalid data without annotation or irrelevant annotation as well as incomplete RR interval were taken out. 1.17% of total data removed amounted to 120 data. The rest of the 98.83% data which is 10154 samples were synchronized with annotation data, data normalization, feature extraction, and RR interval.

First algorithm: extreme learning machine

Twenty-five experiments consisting of 100 ELM iterations for each experiment were performed to evaluate ELM algorithm. All 18 features were executed without a prior selection process. As shown in Fig. 3, the mean accuracy of training data for 6, 4, 3, and 2 classes were respectively 74.86%, 83.02%, 86.97%, and 92.85%. Consecutively, the mean accuracies of testing data for 6, 4, 3, and 2 classes were 57.62%, 66.83%, 72.2%, and 78.12%. According to the accuracy evaluation, the difference of training and testing accuracy was significant in which the differences were 17.24%, 16.19%, 14.77%, and

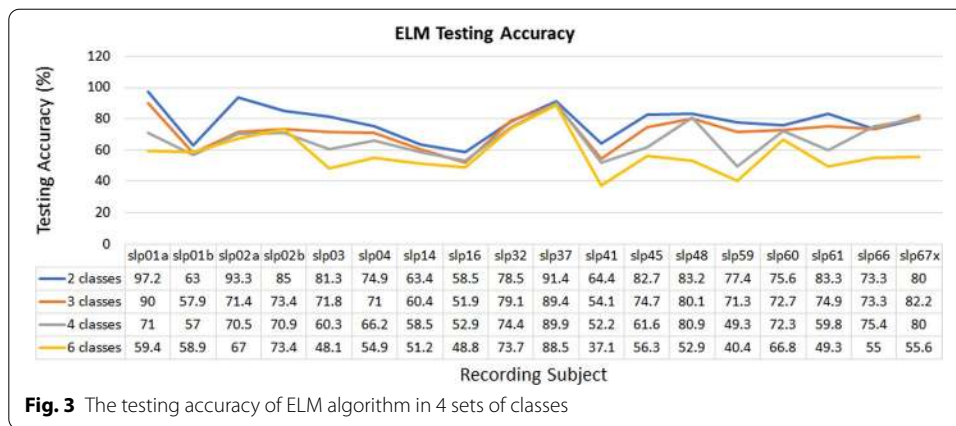


Fig. 3 The testing accuracy of ELM algorithm in 4 sets of classes

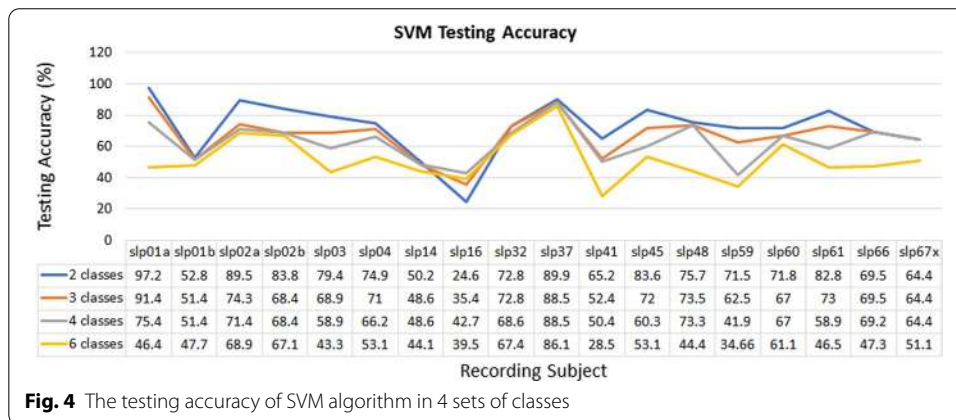


Fig. 4 The testing accuracy of SVM algorithm in 4 sets of classes

14.73% for 6, 4, 3, and 2 classes. The large differences may lead to overfitting. The large differences caused by ELM does not do well with multiple class classification where ELM needs a large number of hidden nodes to ensure generalization performance, but also easy to lead to overfitting. To solve the overfitting problem, we propose the ELM with PSO, which is explained in Sect. [Third algorithm: the integration of extreme learning machine with particle swarm optimization](#).

Second algorithm: support vector machine

All 18 features were used for SVM study without feature selection. RBF was utilized as kernel function. As shown in Fig. 4, the mean accuracy of training data for 6, 4, 3, and 2 classes by SVM were 68.72%, 77.25%, 81.8%, 85.93% respectively and the mean accuracy of testing were 51.66%, 62.52%, 66.94%, and 72.2%. Similar to ELM, the training and testing accuracy which showed significant differences were 17.06%, 14.73%, 14.86%, 13.73% for 6, 4, 3, and 2 classes, which can also lead to overfitting. The RBF kernel more likely exposed to overfitting than the linear kernel.

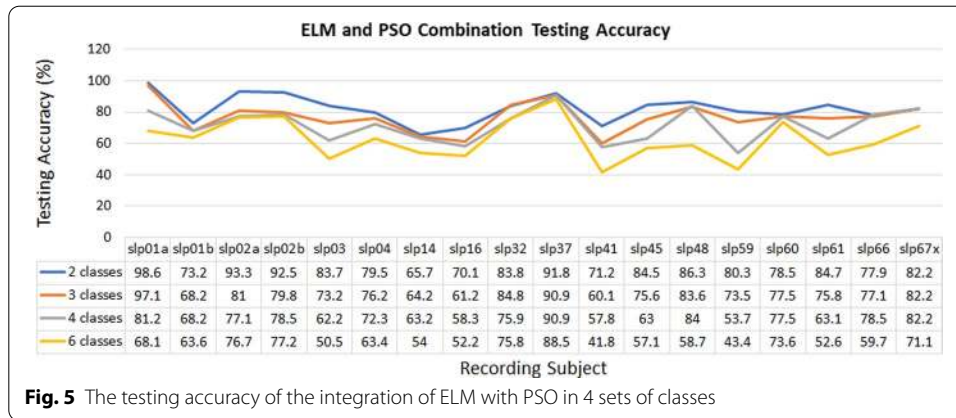


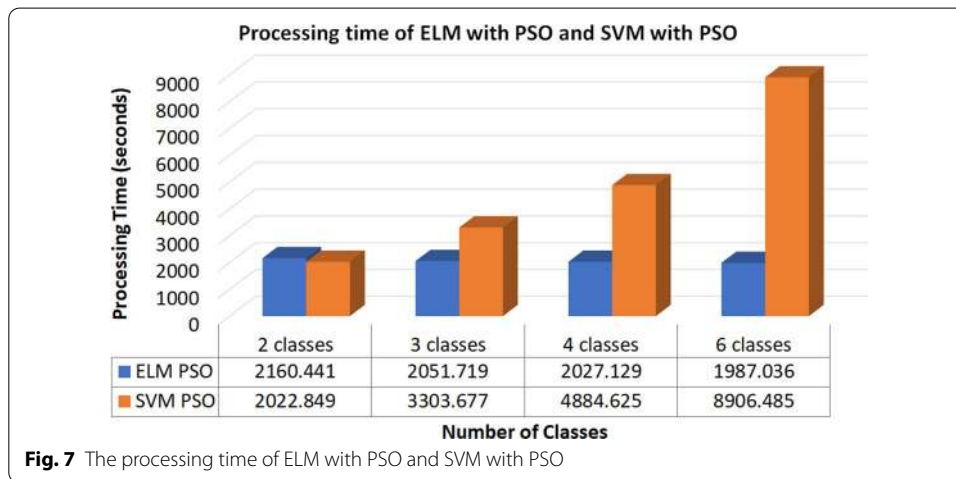
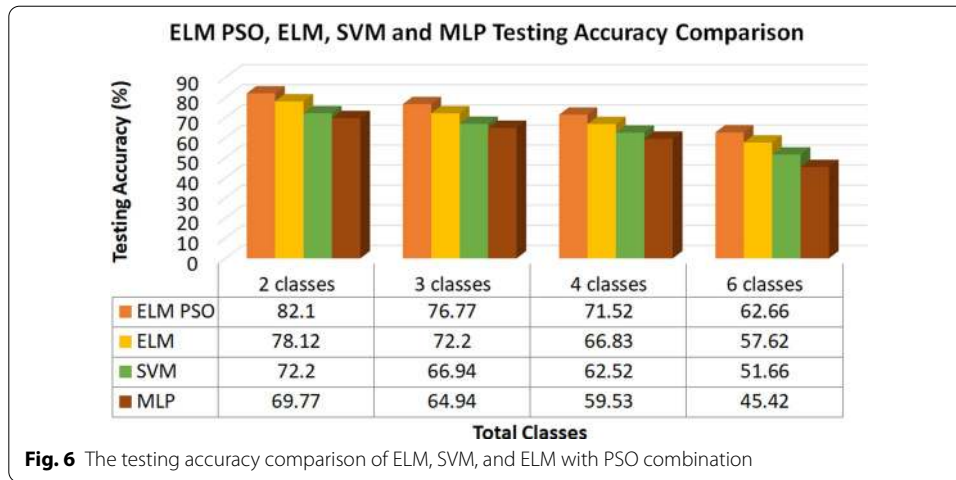
Fig. 5 The testing accuracy of the integration of ELM with PSO in 4 sets of classes

Table 4 The selected features for 6-classes classification

Recording subject	Selected features
slp01a	1 2 4 15 17
slp01b	1 3 4 6 8 9 13 14 15 16 18
slp02a	2 3 6 7 8 11 14 15 16 18
slp02b	2 3 6 8 9 10 11 14
slp03	1 7 9 10 12 14 15
slp04	6 7 8 11 13 14 15 16 18
slp14	1 2 3 10 12 13 14 16 17 18
slp16	1 2 4 8 9 10 11 12 14 15 17 18
slp32	3 4 6 7 10 14 16 18
slp37	1 3 5 7 8 11 14 16
slp41	2 3 5 6 8 10 13 14 17
slp45	2 4 6 7 8 9 10 11 13 14 15 17 18
slp48	1 2 3 4 6 8 10 12 13 14 15 16
slp59	2 3 4 8 12 13 17 18
slp60	1 4 7 8 9 13
slp61	3 5 6 8 17 18
slp66	2 4 7 11 12 13 14 16 18
slp67x	2 4 5 6 9 12 15

Third algorithm: the integration of extreme learning machine with particle swarm optimization

According to ELM and SVM study, the accuracy differences of training and testing data from ELM and SVM algorithm led to overfitting. Hence, the integration of ELM with PSO was proposed to overcome the overfitting problem. Twenty-five experiments were performed with 100 PSO iterations applied on each experiment for feature selection. The parameter of PSO was outlined as follows: W was 0.6, $c1$ and $c2$ values were 1.2, 20 particles, W_F and W_A and were 0.05 and 0.95. With those set of parameter combination, the small difference of the training and testing accuracy was expected. In Fig. 5, ELM with PSO feature selection exhibited mean accuracy of testing data for 6, 4, 3, and 2 classes of 62.66%, 71.52%, 76.77%, 82.1% while the mean accuracy of training data were 69.09%, 72.73%, 77.97%, 84.9%. The differences were shown to be less significant compared to



ELM and SVM, in which those are 6.43%, 1.21%, 1.2%, 2.8% for 6, 4, 3, and 2 classes. It occurs because the feature selection by PSO can reduce redundancy in the the information given by the selected features. Therefore, it improves the generalization of the model. Table 4 shows the selected features for each recorded subject in 6 class classification. It shows that different best subsets of features are selected for each recorded subject using PSO.

The comparison of three algorithms

The new data were deployed to evaluate the generalization ability and the accuracy of testing data from ELM, SVM, and the integration of ELM with PSO, as shown in Fig. 6. There is also a simple MLP experiment carried out as another comparison to the three algorithms. The integration of ELM with PSO exhibited the best testing accuracy compared to ELM, SVM and MLP in which PSO improved the testing accuracy of ELM methods by 5.04%, 4.69%, 4.57%, 3.98% for 6, 4, 3, and 2 classes. It is clear from the results that the selection of the best subset of features and the right number of hidden neurons will increase the classification accuracy of the proposed ELM-PSO method.

The processing time of ELM with PSO was compared with SVM with PSO in Fig. 7, and it can be concluded that the model using SVM with PSO was not effective for multi-class calculation due to the significant increment of processing time with the increase of the number of classes. In ELM with PSO, there was no significant difference in processing time for 2, 3, 4, and 6 classes. More time was required in SVM to process the training data, in which the number of SVM to be trained can be described as $k(k - 1)/2$, where k is the number of classes. In ELM, the processing time correlated with the time spent to generalize inverse matrices of the hidden layer output matrix.

The comparison with previous works

In this section, the result of the proposed method is compared with previous works in 2 class classification. Compared with the previous research presented in Table 5 conducted by Werteni et al. [22], the proposed combination of ELM and PSO has an accuracy rate of 5.81%, 8.29%, 6.07%, and 3.72% higher than the ELM method, BPNN and SVM. Meanwhile, compared with research conducted by Adnane et al. [20], the combination of ELM and PSO has an accuracy rate of 2.79% and 2.11%, respectively higher than the SVM method (12 features) and SVM (10 features).

Essentially, ELM has a limitation in calculating the appropriate number of hidden nodes to achieve sufficient accuracy. In addition, a comparison of the two previous studies [20, 22] indicates that the selection of features plays a very important role in defining the sleep cycle in order to achieve a better accuracy. With the combination of ELM and PSO for the selection of features and the determination of the number of hidden nodes, the particles in the PSO will help each other to achieve one objective function even if they originate from different positions. It is proven that PSO is able to boost the ability of the ELM to make stronger generalizations. Therefore it can achieve the best performance in 2 classes classification compared to the references in Table 5, even though the other references [20, 22] use more than HRV signal to predict the sleep stage. The comparison of our algorithm with the previous research is shown in Table 5.

The result of proposed method is also compared to previous works in multi-class classification using HRV signal only. Compared with the previous research in Table 6 conducted by Werteni et al. [22], the proposed combination of ELM and PSO has an accuracy rate of 3.77%, 5.36% and 3.98% higher than the ELM method, BPNN and SVM. Meanwhile, when compared with research conducted by Adnane et al. [20], the combination of ELM and PSO has an accuracy rate of 2.79% and 2.11%, respectively higher

Table 5 Comparison with previous research

Paper	Dataset	Method	Accuracy			
			HRV	DFA	HRV + DFA	HRV + DFA + W DFA
Werteni et al. [22]	MITBPD	ELM	76.29%	73.48%	78.33%	–
		BPNN	73.81%	71.07%	76.74%	–
		SVM	76.03%	73.21%	78.12%	–
Adnane et al. [20]	MITBPD	SVM (12 features)	–	–	–	79.31%
		SVM (10 features)	–	–	–	79.99%
Our paper	MITBPD	ELM and PSO	82.1%	–	–	–

Table 6 Comparison with Previous Research

Paper	Dataset	Method	Accuracy			
			2 class	3 class	4 class	6 class
Werteni et al. [22]	MITBPD	ELM	78.33%	–	–	–
		BPNN	76.74%	–	–	–
		SVM	78.12%	–	–	–
Adnane et al. [20]	MITBPD	SVM (12 features)	79.31%	–	–	–
		SVM (10 features)	79.99%	–	–	–
Xiao et al. [21]	SHRSV	Random Forest (41 features)	–	88.67%	–	–
Yücelbaş et al. [29]	SHRSV	Morphological	–	77.02%	–	–
Wei et al. [30]	PSG	DNN	–	77%	–	–
Our paper	MITBPD	ELM and PSO	82.1%	76.77%	71.52%	62.66%

than the SVM method (12 features) and SVM (10 features). Compared with research conducted by Xiao et al. [21], the RF method (41 features) has an accuracy rate of 11.9% higher than proposed combination ELM and PSO. Compared with research conducted by Yücelbaş et al. [29], the morphological has an accuracy rate of 11.9% higher than proposed combination ELM and PSO. However, both Xiao et al. [21] and Yücelbaş et al. used SHRSV dataset, not MITBPD. Meanwhile, when compared with research conducted by Wei et al. that used PSG dataset [30], the combination of ELM and PSO has an accuracy rate of 0.23% lower than the DNN method. The comparison of our algorithm with the previous research is shown in Table 6.

Conclusions

In this study, sleep stage classification from Heart Rate Variability of ECG signal was performed using 3 types of algorithm, 4 set of classes, and 2 set of features. Meanwhile, ELM, SVM, and the integration of ELM with PSO were used as algorithms. The set of classes are 6, 4, 3, 2 sleep stages in which the accuracy was compared. The 6 out of 18 features were selected by PSO algorithm.

The study shows that the integration of ELM and PSO exhibited the highest accuracy, followed by ELM and SVM, respectively. It can be concluded that PSO incorporation improved the accuracy of ELM and SVM algorithm. The probability of model overfitting was also decreased by PSO incorporation. There was no significant difference in the processing time of ELM with PSO for 2, 3, 4, and 6 classes.

Abbreviations

BPNN: Backpropagation neural network; DFA: Detrended fluctuation analysis; ECOC: Error correcting output codes; ECG: Electrocardiogram; EDR: ECG-derived respiration; EEG: Electroencephalogram; ELM: Extreme learning machine; GA: Genetic algorithm; HF: High frequency; HRV: Heart rate variability; LSTM: Long short-term memory; LF: Low frequency; mRMR: Minimum redundancy maximum relevance; MITBPD: MIT/BIH polysomnographic database; MLP: Multi-layer perceptron; NREM: Non-rapid eye movement; PSO: Particle swarm optimization; PCA: Principal component analysis; REM: Rapid eye movement; RF: Random forest; RBF: Radial basis function; RMSSD: Root mean square of the successive differences; SVM: Support vector machine; SVM-RFE: Support vector machine recursive feature elimination; SHRSV: Sleep heart rate and stroke volume data bank; SDNN: The Standard Deviation of NN Intervals; TP: Total power; VLF: Very low frequency.

Acknowledgements

Not applicable.

Authors' contributions

NS is the main author of this article. TFL worked on the research under the supervision of NS and SMI. All authors read and approved the final manuscript.

Authors' information

NS is an assistant professor at Binus Graduate Program-Master of Computer Science, Bina Nusantara (Binus) University, Indonesia. NS received his B.Eng (2007) and M.Eng. (2009) from Institut Teknologi Bandung, Indonesia. He received his PhD degree from Kyushu Institute of Technology, Japan, in 2013.

TFL received her bachelor degree in computer science from Bina Nusantara University. TFL currently is student at Master of Computer Science. Bina Nusantara University.

SMI is an associate professor in Computer Science Department, BINUS Graduate Program-Master of Computer Science. SMI has some experiences in teaching and research in remote sensing and biomedical engineering areas. SMI received his doctoral degree in Computer Science from University of Indonesia. SMI also received his master degree in Computer Science from University of Indonesia as well as bachelor degree from Padjadjaran University, Bandung, Indonesia.

Funding

This work is supported by the Directorate General of Strengthening for Research and Development, Ministry of Research, Technology, and Higher Education, Republic of Indonesia as part of Penelitian Terapan Unggulan Perguruan Tinggi Research Grant to Binus University entitled "Prototipe dan Aplikasi Monitoring Kualitas Tidur Portabel berbasis Teknologi Cloud Computing dan Machine Learning" or "Portable Sleep Quality Monitoring Prototype and Application based on Cloud Computing Technology and Machine Learning" with contract number: 039/VR.RTT/IV/2019 and contract date: 29 April 2019.

Availability of data and materials

The datasets used during the current study are available in <https://physionet.org/>.

Competing interest

The authors declare that they have no competing interests.

Received: 13 May 2020 Accepted: 26 December 2020

Published online: 07 January 2021

References

1. Hirshkowitz M. National Sleep Foundation 2013 poll. In: *Sleep Heal*; 2013.
2. Tibbitts GM. Sleep disorders: causes, effects, and solutions. *Prim Care Clin Off Pract*. 2008;35:817–37. <https://doi.org/10.1016/j.pop.2008.07.006>.
3. Kales A, Rechtschaffen A. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Neurological Information Network: U. S. National Institute of Neurological Diseases and Blindness; 1968.
4. J. Nedoma et al., "Comparison of BCG, PCG and ECG signals in application of heart rate monitoring of the human body," 2017 40th Int. Conf. Telecommun. Signal Process. TSP 2017, vol. 2017-Janua, pp. 420–424, 2017, <https://doi.org/10.1109/tsp.2017.8076019>.
5. Majumder S, Chen L, Marinov O, Chen CH, Mondal T, Jamal Deen M. Noncontact Wearable Wireless ECG Systems for Long-Term Monitoring. *IEEE Rev Biomed Eng*. 2018;11:306–21. <https://doi.org/10.1109/RBME.2018.2840336>.
6. Surantha N, Kusuma GP, Isa SM. Internet of things for sleep quality monitoring system: A survey. In: *Proceedings - 11th 2016 International Conference on Knowledge, Information and Creativity Support Systems (KICSS)*, IEEE, Yogyakarta; 2016.
7. Robert C, Guilpin C, Limoge A. Review of neural network applications in sleep research. *J Neurosci Methods*. 1998;79:187–93. [https://doi.org/10.1016/S0165-0270\(97\)00178-7](https://doi.org/10.1016/S0165-0270(97)00178-7).
8. Erguzel TT, Ozekes S, Tan O, Gultekin S. Feature selection and classification of electroencephalographic signals: an artificial neural network and genetic algorithm based approach. *Clin EEG Neurosci*. 2015;46:321–6. <https://doi.org/10.1177/1550059414523764>.
9. Huang GB, Zhu QY, Siew CK. Extreme learning machine: theory and applications. *Neurocomputing*. 2006;70:489–501. <https://doi.org/10.1016/j.neucom.2005.12.126>.
10. Song Y, Crowcroft J, Zhang J. Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine. *J Neurosci Methods*. 2012;210:132–46. <https://doi.org/10.1016/j.jneumeth.2012.07.003>.
11. Jafari A. Sleep apnoea detection from ECG using features extracted from reconstructed phase space and frequency domain. *Biomed Signal Process Control*. 2013;8:551–8. <https://doi.org/10.1016/j.bspc.2013.05.007>.
12. Malik M, Camm AJ, Bigger JT, et al. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. *Eur Heart J* 1996;17:381. <https://doi.org/10.1111/j.1542-474X.1996.tb00275.x>.
13. Karim AHMZ, Rhaman MM, Haque MA. Identification of premature ventricular contractions of heart using poincare plot parameters and sample entropy. *Eng e Transaction*. 2010;5:1–5.
14. Stein PK, Pu Y. Heart rate variability, sleep and sleep disorders. *Sleep Med Rev*. 2012;16:47–66. <https://doi.org/10.1016/j.smr.2011.02.005>.
15. N N, Vashishtha J. Particle swarm optimization based feature selection. *Int J Comput Appl* 2016;146:11–17. <https://doi.org/10.5120/ijca2016910789>.
16. Syarif I. Feature selection of network intrusion data using genetic algorithm and particle swarm optimization. *Emit Int J Eng Technol*. 2016;4:277–90. <https://doi.org/10.24003/emitter.v4i2.149>.

17. Itil TM, Shapiro DM, Fink M, Kassebaum D. Digital computer classifications of EEG sleep stages. *Electroencephalogr Clin Neurophysiol.* 1969;27:76–83. [https://doi.org/10.1016/0013-4694\(69\)90112-6](https://doi.org/10.1016/0013-4694(69)90112-6).
18. Rundo JV, Downey R. Polysomnography. In: *Handbook of Clinical Neurology*. Elsevier B.V.; 2019. p. 381–392.
19. Utomo OK, Surantha N, Isa SM, Soewito B. Automatic sleep stage classification using weighted ELM and PSO on imbalanced data from single lead ECG. *Procedia Comput Sci.* 2019;157:321–8. <https://doi.org/10.1016/j.procs.2019.08.173>.
20. Adnane M, Jiang Z, Yan Z. Sleep-wake stages classification and sleep efficiency estimation using single-lead electrocardiogram. *Expert Syst Appl.* 2012;39:1401–13. <https://doi.org/10.1016/j.eswa.2011.08.022>.
21. Xiao M, Yan H, Song J, et al. Sleep stages classification based on heart rate variability and random forest. *Biomed Signal Process Control.* 2013;8:624–33. <https://doi.org/10.1016/j.bspc.2013.06.001>.
22. Werteni H, Yacoub S, Ellouze N. An automatic sleep-wake classifier using ECG signals. *Int J Comput Sci Issues.* 2014;11:84.
23. Radha M, Fonseca P, Moreau A, et al. Sleep stage classification from heart-rate variability using long short-term memory neural networks. *Sci Rep.* 2019;9:1–11. <https://doi.org/10.1038/s41598-019-49703-y>.
24. Surantha N, Isa SM, Lesmana TF, Setiawan IMA. Sleep stage classification using the combination of SVM and PSO. In: *2017 1st International Conference on Informatics and Computational Sciences (ICICoS)*; Institute of Electrical and Electronics Engineers (IEEE), Semarang; 2017. p. 177–182.
25. Rahimi A, Safari A, Mohebbi M. Sleep stage classification based on ECG-derived respiration and heart rate variability of single-lead ECG signal. In: *2019 26th National and 4th International Iranian Conference on Biomedical Engineering (ICBME)*. Institute of Electrical and Electronics Engineers (IEEE), Tehran; 2020. p. 158–163.
26. Widasari ER, Tanno K, Tamura H. Automatic sleep disorders classification using ensemble of bagged tree based on sleep quality features. *Electronics.* 2020;9:512. <https://doi.org/10.3390/electronics9030512>.
27. Wang T, Lu C, Shen G, Hong F. Sleep apnea detection from a single-lead ECG signal with automatic feature-extraction through a modified LeNet-5 convolutional neural network. *PeerJ.* 2019;7:e7731. <https://doi.org/10.7717/peerj.7731>.
28. Dey D, Chaudhuri S, Munshi S. Obstructive sleep apnoea detection using convolutional neural network based deep learning framework. *Biomed Eng Lett.* 2018;8:95–100. <https://doi.org/10.1007/s13534-017-0055-y>.
29. Yücelbaş Ş, Yücelbaş C, Tezel G, Özşen S, Yosunkaya Ş. Automatic sleep staging based on SVD, VMD, HHT and morphological features of single-lead ECG signal. *Expert Syst Appl.* 2018;102:193–206. <https://doi.org/10.1016/J.ESWA.2018.02.034>.
30. Wei R, Zhang X, Wang J, Dang X. The research of sleep staging based on single-lead electrocardiogram and deep neural network. *Biomed Engineering Lett.* 2018. <https://doi.org/10.1007/s13534-017-0044-1>.
31. Goldberger AL, Amaral LA, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 2000;101:E215–20. <https://doi.org/10.1161/01.CIR.101.23.E215>. Available: <https://physionet.org>.
32. F. I. Alarsan and M. Younes, "Analysis and classification of heart diseases using heartbeat features and machine learning algorithms," *J. Big Data*, vol. 6, no. 1, 2019, <https://doi.org/10.1186/s40537-019-0244-x>.
33. Kumar Joshi A, Tomar A, Tomar M. A review paper on analysis of electrocardiograph (ECG) signal for the detection of arrhythmia abnormalities. *Int J Adv Res Electr Electron Instrum Eng* 2007;3:12466–12475. <https://doi.org/10.15662/ijareeie.2014.0310028>.
34. G. Bin Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006, <https://doi.org/10.1016/j.neucom.2005.12.126>.
35. Kennedy J, Eberhart R. Particle swarm optimization. In: *Proceedings of ICNN'95 - International Conference on Neural Networks*. IEEE, Perth; 1995. p. 1942–1948.
36. Poli R, Kennedy J, Blackwell T. Particle swarm optimization: an overview. *Swarm Intell.* 2007;1:33–57. <https://doi.org/10.1007/s11721-007-0002-0>.
37. Ahila R, Sadasivam V, Manimala K. An integrated PSO for parameter determination and feature selection of ELM and its application in classification of power system disturbances. *Appl Soft Comput.* 2015;32:23–37. <https://doi.org/10.1016/j.asoc.2015.03.036>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.