www.czasopisma.pan.pl PAN www.journals.pan.pl

Metrol. Meas. Syst., Vol. 30 (2023) No. 2, pp. 323–336
DOI: 10.24425/mms.2023.144866

METROLOGY AND MEASUREMENT SYSTEMS
Index 330930, ISSN 0860-8229
www.metrology.wat.edu.pl

SINGLE-CHANNEL EEG PROCESSING FOR SLEEP APNEA DETECTION AND DIFFERENTIATION

Monika A. Prucnal, Adam G. Polak

Department of Electronic and Photonic Metrology, Faculty of Electronics, Photonics and Microsystems, Wroclaw University of Science and Technology, Wroclaw, Poland (✉ *monika.a.prucnal@pwr.edu.pl*, *adam.polak@pwr.edu.pl*)

Abstract

Sleep apnea syndrome is a common sleep disorder. Detection of apnea and differentiation of its type: obstructive (OSA), central (CSA) or mixed is important in the context of treatment methods, however, it typically requires a great deal of technical and human resources. The aim of this research was to propose a quasi-optimal procedure for processing single-channel electroencephalograms (EEG) from overnight recordings, maximizing the accuracy of automatic apnea or hypopnea detection, as well as distinguishing between the OSA and CSA types. The proposed methodology consisted in processing the EEG signals divided into epochs, with the selection of the best methods at the stages of preprocessing, extraction and selection of features, and classification. Normal breathing was unmistakably distinguished from apnea by the k-nearest neighbors (kNN) and an artificial neural network (ANN), and with 99.98% accuracy by the support vector machine (SVM). The average accuracy of multinomial classification was: 82.29%, 83.26%, and 82.25% for the kNN, SVM and ANN, respectively. The sensitivity and precision of OSA and CSA detection ranged from 55 to 66%, and the misclassification cases concerned only the apnea type.

Keywords: single-channel EEG, sleep apnea detection, optimization of signal processing, medical decision support.

1. Introduction

Sleep is one of the basic states of the human body. Being closely related to the central nervous system, it affects its relaxation and is necessary for the normal functioning the entire human being. *Sleep apnea syndrome* (SAS) is one of the most common sleep disorders that involves the temporary reduction (hypopnea) or complete cessation (apnea) of breathing multiple times during the night, which causes sleep fragmentation and has an influence on health, safety, and the quality of human life. Generally, there are three types of apnea: *obstructive* (OSA), *central* (CSA), and *mixed* (MSA) one, differing in factors influencing their occurrence. Counting only the respiratory events (apnea or hypopnea) enables the determination of apnea severity: mild, moderate, severe,

Article history: received January 1, 2023;; revised February 7, 2023; accepted February 10, 2023; available online June 19, 2023.

or extreme, whereas the differentiation of the event type is important in the context of therapeutic options and the selection of treatment methods [1].

The gold standard in sleep disorder diagnostics is *polysomnography* (PSG) [1,2]. PSG consists in recording various physiological signals during the night, including EEG, to evaluate the length and quality of sleep and breathing patterns. PSG is complex, expensive, time-consuming and based on recordings realized in a sleep laboratory and then annotated by experts.

The EEG represents the electrical activity of the brain and enables evaluation of the function of the central nervous system by observing changes in the five brainwave activities. Typically, two symmetrical EEG channels are recorded during PSG from the recommended and backup electrode locations. Then, only one of them is used for further analysis [3]. The observed effect of apnea on the EEG signal is the modification of the brainwaves [4, 5] and arousals occurring after its episodes [6]. The above documents that SAS manifests in the EEG signal. Therefore, automatic diagnostics with the use of portable EEG recorders can improve the effectiveness of detecting apnea, including home surveillance. On the other hand, the optimized processing of the single-channel EEG is of particular interest because it may be helpful also in the multi-channel EEG signals analysis or EEG combination with other signals.

In recent years, a lot of studies have been published on the detection of apnea events based on the use of single signals recorded during PSG, but including few with the single-channel EEG [2, 7–27]. Most of the work applied binary classification to detect *normal breathing* (NB) and apnea events (usually of the obstructive type only) [2, 7–11, 13, 16–20, 22–26, 28, 29]. The studies, using solely the single EEG channel, achieved accuracy ranging from 76.70% [25] to 99.53% [11]. Differentiation of the apnea type into OSA, CSA, or MSA, using the single-channel EEG signal, has been reported in only a few studies [14, 15, 21, 27]. There are also works on the classification of the apnea type based on two symmetrical EEG channels [15], [28, 30–32]. The accuracy of NB, OSA, and CSA classification obtained in these studies ranged from 63.80% to 64.30% with the use of the single-channel EEG signal [15] and from 70.30% to 99.68% with the combination of features from two symmetrical EEG channels [15, 28, 31, 32]. Moreover, two works proposed to distinguish also between apnea and hypopnea with the accuracy of 75.90% [21], and within six classes: apnea or hypopnea of each type at the level of 48.24% [21] and 98.82% [27].

The aim of this research was to propose an optimized procedure for a single-channel EEG signal processing from overnight recordings, maximizing the accuracy of both automatic differentiation between sleep apnea and normal breathing, as well as the distinguishing between the apnea types. The main contribution of this work is to demonstrate that processing only one EEG channel allows for the correct detection of NB and apnea events, as well as to show the possibility of distinguishing the type (OSA or CSA) of respiratory events. In order to achieve the goal above, the present research used the classic approach to multinomial classification: signal preprocessing, its division into 30-second epochs, extraction and selection of features, classification, and the final evaluation of classification accuracy.

2. Materials and methods

2.1. EEG data from St. Vincent's University Hospital database

Data from the PhysioBank were used in this study [33]. This database contains EEG signals from two symmetric locations (C3-A2 and C4-A1), sampled at 128 Hz from, 25 patients (4 women and 21 men), diagnosed with SAS or primary snoring. In this study, the C3-A2 channel was

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 323–336
DOI: 10.24425/mms.2023.144866

chosen according to the American Academy of Sleep Medicine recommendations [1, 3]. The database is annotated by one expert about the sleep stages and respiratory events classified as three (obstructive, central, and mixed) types of apnea and hypopnea, all scored with a resolution of 1 second.

2.2. Methodology for selecting the best methods

The proposed methodology concerns the selection of the best possible methods at subsequent stages of EEG processing (Fig. 1). To this end, the approaches reported in the literature on the use of EEG in sleep research that yielded high accuracy of classification were first identified. The selection of optimal methods for preprocessing as well as feature selection was made by maximizing the performance of the knearest neighbors (kNN) algorithm (Section 2.3) [14, 15, 30]. At individual stages of the analysis, the best method was ascertained by applying two- and one-sided Student's t-tests ($\alpha = 0.05$), which, respectively, made it possible to determine whether the differences between the methods were statistically significant, and if any of the approaches was better. Then, the chosen methods were used at the succeeding stages. All procedures were implemented in MATLAB R2020b (*The MathWorks*, USA).

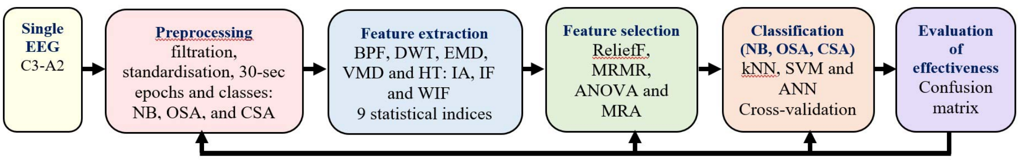

Fig. 1. The applied methodology for processing a single-channel EEG (see text for abbreviations used)

2.3. Preprocessing

The following preprocessing methods were compared: removing too high magnitudes and effects of saturation, removal of disturbances from the power grid by low-pass and notch filtration, and standardization.

Too high magnitudes and saturations are characterized by significant deviations in the signal amplitude. Detection of the samples with excessive magnitudes was made by calculating the z-score for each sample and then specifying thresholds for the outliers [34]. In addition, the removed sequences had to be longer than 10 samples. Two methods for removing disturbances from the power grid were compared i.e. low-pass filtering with a cut-off frequency of 45 Hz [9, 35], and notch filtering at 50 Hz (removing interference from the power grid) and 60 Hz (a component of an unknown source observed in some EEG recordings). Low-pass filtering was performed with a 17th-order type II Chebyshev zero-phase filter [36]. Then a 2nd order zero-phase notch filter with an attenuation width of 1 Hz was applied to the same data.

Because the EEG signals collected in the database come from tests performed at different times and on different subjects, they vary in signal energy. Therefore, detrending and standardization were performed to unify the average signal powers.

Finally, the signals were divided into 30-second epochs based on expert description. Next, the respiratory events with a duration of at least 10 seconds were found and placed in the middle of the epochs created (if the event was shorter than 30 s, it was supplemented with adjacent samples from normal breathing). Similarly, the signal fragments with normal breathing were also split into 30-second epochs. Then the epochs were gathered in three classes representing: NB, obstructive apnea/hypopnea (OSA), and central apnea/hypopnea (CSA). The three classes were balanced taking into account the number of epochs corresponding to the least numerous set. The comparison of the preprocessing methods was made with the features obtained with the *discrete wavelet transform* (DWT) and *Hilbert transform* (HT) combined with scalar metrics (see Section 2.4), and then using the kNN classifier [14, 15].

2.4. Feature extraction

Feature extraction was performed in two steps, i.e., signal decomposition (one- or two-stage) and the calculation of scalar metrics of the obtained components. Among the one-stage decomposition methods chosen were *band-pass filtering* (BPF) [8, 10], DWT [37], *empirical mode decomposition* (EMD) [10], and *variational mode decomposition* (VMD) [12]. After that, the HT was applied to components from EMD (*the Hilbert–Huang transform*, HHT), as well as from the DWT [14–16, 30] and VMD.

BPF was performed with a cascade combination of constant-phase low-pass and high-pass Chebyshev Type II filters with the minimum order, considering the frequency ranges characteristic for the five brainwaves: *delta* (0–4 Hz), *theta* (4–8 Hz), *alpha* (8–16 Hz), *beta* (16–32 Hz), and *gamma* (32–64 Hz) [1]. The DWT is a time-scale signal analysis method returning detailed and approximation coefficients at a given level. The decomposition of epochs was performed using the Daubechies 3 wavelet, whose shape is similar to EEG signal fragments [37]. Taking into account the sampling frequency (128 Hz) and the frequency ranges of the brainwaves, the number of decomposition levels was set at 4 to obtain 5 sub-signals matching these subbands. EMD is an iterative algorithm that decomposes a signal into a finite number of oscillating *intrinsic mode functions* (IMFs), having two properties: the number of extremes is the same as the number of zero crossings (± 1), and their envelopes are symmetrical in relation to the baseline [10]. The number of IMFs depends on the information contained in the signal and as such may be different for each epoch. Therefore, in order to unify the amount of obtained features for all epochs, the first 13 IMFs were used in further analyses. VMD allows to decompose a signal into a predetermined number of IMFs with limited frequency bands located around the adaptive center frequencies, calculated by solving a limited variational problem [38]. Each epoch was decomposed into 5 IMFs, as there are 5 brainwaves. The HT returns an analytic signal. For the analytical signals of the components from the first stage of decomposition, their *instantaneous amplitudes* (IA) and *frequencies* (IF) were calculated, as well as the *weighted frequencies* (WIF) [14].

At the end, 9 scalar features were calculated for all waveforms characterizing the EEG epochs, such as statistical indexes: skewness, kurtosis, median [7,8,10,16,35]; Hjorth parameters: activity, mobility, complexity and the ratio of the activity to the sum of the activities of all signals [35], Shannon entropy [12, 16], and maximum amplitude [35].

2.5. Feature selection

Three feature selection methods based on feature filtering were deduced from the literature: the *minimum redundancy maximum relevance* (MRMR) approach [35], the ReliefF algorithm [40] and the *analysis of variance* (ANOVA) [14,15,30], as well as their combinations with *multivariate regression analysis* (MRA) [14, 15, 30]. The effectiveness of these methods was tested for the largest set of features (351), extracted with the HHT method.

In the MRMR method, the feature ranking was made based on the *mutual information quotient* (MIQ) with values indicating the significant features [39]. Then, based on a decrease in MIQs,

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 323–336
DOI: 10.24425/mms.2023.144866

two thresholds were selected: 0.015 and 0. ReliefF is a ranking algorithm finding the closest k neighbors for each feature and determining which of them have different values for different classes and similar for the same class. It assigns weights according to their relevance. As a result, the features are sorted according to importance related to their weights [40]. The selection of features was made for the k values from 3 to half the size of the feature vector and different values of the weight thresholds th (set for $k = 10$, which is the default value [40]) to reject some of the features and then to test new, smaller vectors. ANOVA allows to determine whether the observed variability of a feature is related to its belonging to a class. Therefore, it was used to reject those features that did not differentiate between the classes ($\alpha = 0.05$). On the other hand, MRA backward selection made it possible to discard features that were almost linearly dependent on others, using the adjusted coefficient of determination $\overline{R}^2 > 0.95$ [30].

After finding the best method, it was used to select features separately for vectors from each extraction method. Then these features were combined into one overall set and the final features were reselected using the same method.

2.6. Classification

In order to test the effectiveness of classification within the three classes (NB, OSA and CSA), three classifiers were compared: kNN [8, 9, 12, 16, 31, 32], the *support vector machine* (SVM) [5, 7, 10, 11, 16, 31, 32], and two architectures of the *artificial neural network* (ANN) [7, 11, 16, 37].

The kNN method is a nonlinear classifier that enables assigning a new element to a class based on knowledge of its closest neighbors belonging. Results of this multinomial automatic classification were tested for 10 metrics: Hamming, Jaccard, Spearman, Chebyshev, Minkowski, Euclidean, correlation, cosine, city-block, and standardized Euclidean, the number of neighbors from 3 to 150, and the majority decision taking method [41].

The SVM is a supervised learning algorithm that allows to create a hyperplane dividing the transformed feature space into classes and minimize the classification error by obtaining the maximum geometric distance between them [32]. The classification within three classes was made by decomposing the problem into three binary classifications. The search for the optimal SVM structure was performed for the most popular kernel functions: linear, radial ($2\sigma^2$ from 10^{-5} to 10^7) and polynomial (degree of 2 and 3), as well as for different values of box constraint parameter C (from 10^{-5} to 10^7), which enables to avoid overfitting, and then, with smaller step nearly the optimal value, by scaling C and $2\sigma^2$ by 0.2; 0.4; 0.6; 0.8; 1; 2; 4; 6; 8 [42].

The ANN is a nonlinear model inspired by the nervous system. The search for suboptimal network architecture was performed for two ANN structures: with one hidden layer with a radial function and two hidden layers with log-sigmoidal activation functions, which are sufficient to approximate any mapping [43]. In the output layer, the *SoftMax* function was used, which allows to estimate the probability of belonging to the classes. The selection of the number of neurons in the hidden layers was carried out in two stages. First, the number of neurons in one or both layers were set from 10 to 100, in increments of 10. Then, in the proximity of the best performance, the number of neurons was changed in increments of 1. Neural network weights and thresholds were randomly initialized before each network training. The search for suboptimal ANN architecture was performed using the scaled conjugate gradient training algorithm because for this algorithm the network training time was about 700 times shorter than for other tried ones. Then, for the ANN selected as optimal, training was also performed with the use of two additional training algorithms: gradient steepest descent and *Levenberg–Marquardt* (LM).

2.7. Assessment of efficiency of classifiers

The performance evaluation of the classifiers was performed using 32-fold cross-validation since testing the model 32 times allows to obtain a sufficiently good statistical averaging. The obtained subsets were of roughly equal numbers: 128 or 129 epochs. The average accuracy with the standard deviation of classification for the validation data was used as the primary performance indicator for a classifier with specific hyperparameters. The obtained results were then used to optimize the model hyperparameters, which allowed to achieve the maximum average accuracy (A) of classification. Moreover, the average precision (P) and sensitivity (S) were also used to evaluate preprocessing and the feature selection methods [44].

3. Results

Various types of combinations of the preprocessing methods were tested. The best classification result of 54.43%, with a statistically significant difference to the signal without preprocessing (52.36%, p-value = 8.14×10^{-6}), was observed for a combination of removing too high magnitudes and signal saturations, low-pass filtering, and signal standardization. After dividing the preprocessed signals into 30-second segments, the apneas covered 335 central, 208 obstructive, and 128 mixed epochs. Hypopneas included 1401, 1038, and 102 of the obstructive, central and mixed types. After merging them, the CSA proved to be the least numerous class. For this reason, 1373 epochs representing also the other two classes were randomly drawn. Thus, the total number of epochs in the three balanced classes was 4119.

During the first stage of feature extraction, EEG epochs were decomposed into multiple waveforms: each 5 by BPF, DWT and VMD; first 13 by EMD; 15 (5×3) by DWT or VMD combined with HT; and 39 (13×3) by HHT. Then, after calculating the 9 scalar features for each component, the successive feature vectors contained: 45 features for each BPF, DWT and VMD; 117 for EMD; 135 for DWT+HT and VMD+HT, and 351 for HHT. Combining all subsets into one gave a total of 873 features (Table 1). The results of the comparison between the feature selection methods for the HTT set, including the Student t-tests, are shown in Table 2. The ReliefF method reduced the number of features from 351 to 292 and improved the classification precision from 68.04 ± 2.25 to 68.65 ± 2.34%, however, without statistically significant difference

Table 1. Number of features before and after two-stage selection using the ReliefF algorithm shown for individual extraction methods.

Extraction method	**Number of features**			**Percentage of selected features (%)**
	Before selection	After first selection	After final selection	
BPF	45	11	11	24.4
DWT	45	18	15	33.3
EMD	117	48	22	18.8
VMD	45	39	17	37.8
HHT	351	292	257	**73.2**
DWT+HT	135	105	58	43.0
VMD+HT	135	7	7	5.2
Overall set	873	520	387	44.3

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 323–336
DOI: 10.24425/mms.2023.144866

($p-$value = 0.059). Nevertheless, it was chosen for further stages of research. This final set was used to optimize the three classifiers by tuning their hyperparameters. Tables 3–5 demonstrate the best classification results, respectively. The average accuracy of the multinomial classification, assessed by 32-fold cross-validation, was 82.29 ± 2.29%, 83.26 ± 2.62%, and 82.25 ± 2.80% for the kNN with the standardized Euclidean metric and $k = 47$; the SVM with a radial basis function, $2\sigma^2 = 40$ and $C = 20$; and the feedforward ANN with 14 + 14 neurons in two hidden layers, log-sigmoid activation function, and the LM learning algorithm, respectively.

Table 2. Comparison of the precision of classification for different methods of feature selection (features extracted by the HTT; classified by the kNN: $k = 17$, city metrics).

	P ±SD (%)	**Number of features**
Original set	68.04 ± 2.25	351
MRMR	67.85 ± 2.32	323
ANOVA	68.08 ± 2.27	329
ReliefF	**68.65 ± 2.34**	**292**
ANOVA +MRA	66.52 ± 2.34*	217
ReliefF+MRA	66.79 ± 2.4*	238

The results of statistical tests (p-value = 0.119 for kNN and SVM, 0.938 for kNN and ANN, and 0.137 for SVM and ANN compared) indicate that there is no significant difference between the average accuracy of individual classifiers, and therefore the results obtained using these methods are comparable.

Based on the confusion matrices of the tested classifiers (Tables 3–5), it can be seen that the NB epochs were flawlessly distinguished from the apnea epochs by kNN and ANN, and with an

Table 3. Confusion matrix (expert qualification in columns, prediction in rows) and performance of the kNN model for $k = 47$ and standardized Euclidean metric.

Confusion matrix				**Classifier performance [%]**					
				In classes			Averaged		
	NB	OSA	CSA	A_i	P_i	S_i	A	P	S
NB	1373	0	0	**100.00**	**100.00**	**100.00**	82.29	73.44	73.44
OSA	0	1084	805	73.44	57.38	78.95			
CSA	0	289	568	73.44	66.28	41.37			

Table 4. Confusion matrix (expert qualification in columns, prediction in rows) and performance of the SVM model for the radial function, $C = 20$ and $2\sigma^2 = 40$.

Confusion matrix				**Classifier performance [%]**					
				In classes			Averaged		
	NB	OSA	CSA	A_i	P_i	S_i	A	P	S
NB	1372	0	0	99.98	**100.00**	99.93	83.26	74.90	74.90
OSA	0	859	519	74.92	62.34	62.56			
CSA	1	514	854	74.90	62.38	62.20			

accuracy of 99.98% by SVM (with one confused element). On the other hand, the sensitivity and precision of the detection of OSA and CSA epochs range from 55 to 66%, and the cases of misclassification concern only the apnea type. The average sensitivity and precision for these classifiers were between 73 and 75%.

Table 5. Confusion matrix (expert qualification in columns, prediction in rows) and performance of the ANN model with 14 + 14 neurons in hidden layers, after LM learning.

Confusion matrix				**Classifier performance [%]**					
				In classes			Averaged		
	NB	OSA	CSA	A_i	P_i	S_i	A	P	S
NB	1373	0	0	**100.00**	**100.00**	**100.00**	82.25	73.37	73.37
OSA	0	807	531	73.37	60.31	58.78			
CSA	0	566	842	73.37	59.80	61.33			

4. Discussion

The aim of this research was to propose a quasi-optimal procedure for single-channel EEG signal processing, maximizing the accuracy of both automatic sleep apnea detection and the differentiation between the types of apnea. This procedure assumed the classic approach to decision making based on machine learning with features extraction and selection and consisted in selecting the most effective methods at the subsequent stages of signal processing.

First, the tested EEG preprocessing methods showed that removing too high magnitudes and saturations visible in the signal did significantly decrease the precision of sleep apnea detection. Nevertheless, their removal seems reasonable because they certainly do not represent the electrical activity of the brain and therefore might have an undefined effect on the interpretation of the EEG signal, including its classification. In addition, 98.16% of the detected too high magnitudes came from signal samples annotated as NB and only 1.84% as OSA. There were no too high magnitudes in the CSA class. These results shows also that notch filtering did not affect the classification performance. The observed positive effect of limiting the frequency band to 45 Hz was confirmed in the literature [12]. Additionally, such an upper limit from 35 to 45 Hz was also applied in other works on the detection of sleep apnea on the basis of the EEG signal [9, 12, 26, 35]. Moreover, there are studies that investigated changes in EEG signal power before, during and after sleep apnea episodes, which did not include frequencies above 30 Hz [4]. In contrast, there is research in which the set of selected features contained about 20% elements in the high frequency range of 32–64 Hz [14, 30]. Although frequencies below 30 Hz are dominant during sleep, and therefore this range was usually analyzed in sleep apnea studies, it seems worthwhile to conduct further research on the influence of the *beta* and *gamma* brainwaves (16–64 Hz) on the accuracy of sleep apnea detection. Summarizing, standardization of the EEG signals has proven to be the most effective preprocessing method.

In the studies on the EEG use to detect apnea, features were usually extracted by decomposing the signal into waveforms related to the brainwaves and then calculating their scalar measures. In the works with the classification accuracy about 90%, the feature vectors had from 5 to 140 elements. In this study with the extraction methods: BPF, DWT, EMD, VMD, HHT, DWT+HT and VMD+HT, the sets with 45 to 351 features have been computed. However, after combining all the subsets, a total of 873 features have been obtained, which is a much larger set than in the

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 323–336
DOI: 10.24425/mms.2023.144866

quoted literature. This large set allowed to finally select the features that distinguished the classes very effectively.

The selection of features was performed only in a few of relevant studies: [11,14–17,26,30,35], by applying the MRMR algorithm [35], Fisher's method [11], the ANOVA and MRA [14,15,30], the p-values of Student's t-test [26] and SVM recursive elimination [17]. In other works, the feature vector dimension was not reduced, and the discriminant ability of its elements was not checked, but their number was usually small. The high accuracy of sleep apnea classification, from 89.01% to 99.00%, was obtained in studies without feature selection [10, 12, 13, 16–18, 20, 23, 24] and from 89.90% to 99.53% when they were selected [11, 26, 35]. In this study, the MRMR, ReliefF, ANOVA, ANOVA+MRA and ReliefF+MRA were compared. Each of these methods reduced the number of HHT features (Table 1). On the basis of the Student t-tests, it can be concluded that the average precision without selection is higher than after it for ANOVA+MRA and ReliefF+MRA (however, the greatest reduction to 217 features was obtained with the ANOVA+MRA), whereas the outcomes of ReliefF as well as the MRMR and ANOVA alone did not show a statistically significant difference. The deterioration of the classification precision to 66.52% obtained after additional MRA application was unexpected and should be further investigated. In the end, the ReliefF method was chosen because it allowed the greatest increase in average precision to 68.65% while reasonably reducing the number of HHT features from 351 to 292. Afterwards, the reselection of features from the overall set (520 features yielding the precision 72.91%) did not improve the precision of classification (72.56%). According to the t-test, there was no statistically significant difference between using the overall set and the features after reselection, however, this limited their number to 387 (Table 1) and therefore this vector was considered reasonable to use in the next step. Nevertheless, it is worth noting that the two-step selection reduced the vector size to 43.3% (Table 1), and the HTT yielded the 257 most informative features. The second method in the order, DWT+HT, provided 15% of the differentiating features.

Most of the works published in recent years have dealt with binary classification within the class of NB and apnea, regardless of its type, or only with OSA. In this study, the accuracy of distinguishing between NB and apnea, regardless of its type, is 100% for kNN and ANN (Tables 3–5), so it is better or comparable to those achieved in the literature, that ranged from 78.10% to 99.53% [7–11, 13, 16–18, 18, 20, 22, 24–26, 29]. Such high accuracy results primarily from the appropriately extracted and selected features and, perhaps, from constructing the EEG epochs with the respiratory events placed in their center.

Recently, several papers have been published aimed at distinguishing the type of apnea based on a single EEG channel only. Research depicted in [14, 15] yielded the maximal accuracy 64.30% of the multinomial classification (NB, OSA, and CSA) in contrast to the higher accuracy of 83.26% reached in this work (Table 4). There are works which differentiated both apnea and hypopnea, having six classes corresponding also to the types of events [21, 26]. Alimardani and Moor achieved a mean accuracy of 48.24% and the highest precision within the obstructive type 50.00% for apnea (SVM) and 72.50% for hypopnea (*linear discriminant analysis,* LDA), and for the central type 38.10% for apnea (LDA) and 50.88% for hypopnea (SVM) [21]. In this study, the highest precision for the combined apnea/hypopnea classes was 62.24% (SVM) for obstructive and 66.28% (kNN) for the central type. Chatterjee and Jana attained high mean accuracy within six classes from 81.39% to 98.92% (kNN) depending on the used database [27]. The discrimination between NB and obstructive apnea and hypopnea was realized by Gurrala et al. with the average accuracy 95.90%, but the authors did not present classification assessment for each class separately [23]. The results achieved in the presented work, however, cannot be directly compared with those of Alimardani and Moor, Chatterjee and Jana and Gurrala et. al, since in those studies the apnea and hypopnea classes were considered separately. Nevertheless, it was there noticed that for the central type, the accuracy for apnea/hypopnea classes was higher than for separated classes, while for the combined obstructive type it was higher than for apnea, but lower than for hypopnea alone [21].

There are also studies in which the detection of apnea was made based on features extracted from two symmetrical EEG channels, achieving accuracy from 94.33% to 99.68% [17, 28, 31] for binary classification (NB vs. apnea) and from 70.30% to 88.99% [2, 15] for multinomial classification (NB, OSA, CSA). Using this approach, Zhao et al. obtained a classification accuracy of 88.99% for the same three classes [32], which is higher than accuracy obtained in this study using a single channel (83.26%). This confirms the hypothesis from previous research [15, 30] that exploiting both symmetrical channels (instead of one) improves the accuracy of classifiers applied in automated sleep apnea diagnostics. On the other hand, that research was conducted with unbalanced learning subsets to the benefit of the well-recognized NB class, which had to be translated into an increase in average accuracy. Two of the classifiers analyzed there, kNN and SVM, were tested also in this study, yielding the accuracy 3% lower for kNN and 14% higher for SVM. The accuracy of distinction between apnea and NB epochs in this study (accuracy and precision of 100%, see Tables 3 and 5) is also higher than the accuracy achieved by others using two symmetrical EEG channels: 94.33% [17], 95.10% [31] and 99.68% [28].

Some limitations of the proposed methodology were identified and should be considered in future research. First of all, the study used signals from one public database [33] and only one EEG channel (C3–A2). The size of the analyzed dataset was limited, particularly by the number of epochs identified as apnea or hypopnea of the central type. To be able to obtain a greater generalization of the classifier, it would be necessary to enlarge the dataset of signal segments from various types of apnea and hypopnea, and from subjects with different severity of the SAS. Moreover, during this study, processing was performed for the 30-second epochs in which the episodes of the respiratory events were in the middle of these time intervals, and therefore it is not known how such learned classifiers would work e.g., in the online analysis of consecutively transmitted signal samples. Lastly, the classical scheme of machine learning, with the separate stages of feature extraction and selection, was applied in this work, and it seems worthwhile to extend the study to the use of models enabling deep learning, which made it possible to achieve the accuracy of apnea detection from 76.70% to 95.90% [2, 19, 20, 23, 25, 29].

5. Conclusions

The main attempt at this work was to systematically review and test the methods used in the automatic detection of sleep apnea in order to identify the best of them at each stage of the single-channel EEG processing. The flawless differentiation between normal breathing and sleep apnea/hypopnea as well as the accuracy of differentiation of the apnea type higher than in the previous studies confirm the validity of the proposed methodology. Thus, the main conclusion is that by appropriate extracting and selecting features from 30-second EEG epochs, it is possible to reliably detect sleep apnea epochs when analyzing only a single-channel EEG signal. Among the tested methods, the most effective approaches to EEG processing turned out to be: low-pass filtering and standardization of the overnight signals as preprocessing, the Hilbert–Huang transform (HHT) and the DWT combined with the Hilbert transform at the feature extraction stage, as well as the ReliefF algorithm for features selection. Nevertheless, all the extraction methods, selected primarily from the literature, had a significant impact on these remarkable outcomes. Moreover, the possibility of achieving very accurate results only on the basis of one EEG channel should motivate the technological development of simple and cheap portable

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 323–336
DOI: 10.24425/mms.2023.144866

electroencephalographs with only a few channels. Most importantly, these encouraging results, achieved on a relatively large database (4119 EEG epochs), show the potential of the proposed methodology not only in the diagnostics of sleep apnea, but also of other diseases affecting brain function.

In summary, the highlights of this work are as follows:

- the most effective methods were selected in the single-channel EEG processing chain used to detect obstructive or central sleep apnea;
- normal breathing was distinguished from apnea with 100% accuracy using the kNN method and the ANN, and 99.98% using the SVM;
- the proposed methodology may also be fruitful in the detection of other diseases affecting the brain function, and in general – in other biosignals processing used for decision making.

Further research is needed to improve the accuracy of discriminating the type of apnea. The first idea, which will not significantly increase the complexity of measurements, is to include the second or more additional EEG channels (depending on the available hardware solution), because even symmetrical channels have been shown to contain additional information increasing the accuracy of classification. It seems also beneficial to use a larger database in future research, with a greater number of respiratory events, in particular central type sleep apnea, the number of which (assuming balanced classes) limited the amount of analyzed epochs. Finally, a hierarchical approach to classification is considered: after accurately distinguishing between apnea and normal breathing, to repeat the process of selecting the appropriate features (from the same originally extracted set) that will best differentiate between the apnea types. These aforementioned directions for further research should lead to full automation of apnea detection and differentiation, taking into account a single- or few-channel EEG dedicated device, processing the last segment of the recorded signal in accordance with the procedure developed in this work, and using the learned model.

Acknowledgements

Research supported by the Grant no. 0402/0148/18 from the Ministry of Science and Higher Education, Poland.

References

[1] American Sleep Disorders Association. (2001). *The International Classification of Sleep Disorders, Revised: Diagnostic and Coding Manual*. Chicago, IL: American Academy of Sleep Medicine.

[2] Kokkalas, L., Korompili, G., Tatlas, N. A., Mitilineos, S. A., & Potirakis, S. M. (2021). Severe obstructive sleep apnea event detection from EEG recordings. In *2nd International Electronic Conference on Applied Sciences: Proceedings*, 15, 31.

[3] Berry, R. B., Brooks, R., Gamaldo, C., Harding, S. M., Lloyd, R. M., Quan, S. F., ... & Vaughn, B. V. (2017). AASM scoring manual updates for 2017 (version 2.4). *Journal of Clinical Sleep Medicine*, *13*(5), 665–666. https://doi.org/10.5664/jcsm.6576

[4] Coito, A. L., Belo, D., Paiva, T., & Sanches, J. M. (2010, October 29). EEG spectral power changes before, during and after obstructive sleep apnea [Poster presentation]. *16th Portuguese Conference on Pattern Recognition*.

[5] Zhou, G., Pan, Y., Yang, J., Zhang, X., Guo, X., & Luo, Y. (2020). Sleep electroencephalographic response to respiratory events in patients with moderate sleep apnea-hypopnea syndrome. *Frontiers in Neuroscience*, *14*, 310. https://doi.org/10.3389/fnins.2020.00310

[6] Simms, T., Brijbassi, M., Montemurro, L. T., & Bradley, T. D. (2013). Differential timing of arousals in obstructive and central sleep apnea in patients with heart failure. *Journal of Clinical Sleep Medicine*, *9*(8), 773–779. https://doi.org/10.5664/jcsm.2918

[7] Almuhammadi, W. S., Aboalayon, K. A., & Faezipour, M. (2015, May). Efficient obstructive sleep apnea classification based on EEG signals. In *2015 Long Island Systems, Applications and Technology* (pp. 1–6). IEEE. https://doi.org/10.1109/LISAT.2015.7160186

[8] Ahmed, F., Paromita, P., Bhattacharjee, A., Saha, S., Azad, S., & Fattah, S. A. (2016, December). Detection of sleep apnea using sub-frame based temporal variation of energy in beta band in EEG. In *2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)* (pp. 258–261). IEEE. https://doi.org/10.1109/WIECON-ECE.2016.8009131

[9] Saha, S., Bhattacharjee, A., Ansary, M. A. A., & Fattah, S. A. (2016, November). An approach for automatic sleep apnea detection based on entropy of multi-band EEG signal. In *2016 IEEE Region 10 Conference (TENCON)* (pp. 420–423). IEEE. https://doi.org/10.1109/TENCON.2016.7848033

[10] Shahnaz, C., & Minhaz, A. T. (2016, December). Sleep apnea frame detection based on empirical mode decomposition of delta wave extracted from wavelet of EEG signals. In *2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)* (pp. 233–236). IEEE. https://doi.org/10.1109/WIECON-ECE.2016.8009125

[11] Taran, S., & Bajaj, V. (2019). Sleep apnea detection using artificial bee colony optimize Hermite basis functions for EEG signals. *IEEE Transactions on Instrumentation and Measurement*, *69*(2), 608–616. https://doi.org/10.1109/TIM.2019.2902809

[12] Bhattacharjee, A., Fattah, S. A., Zhu, W. P., & Ahmad, M. O. (2018). VMD-RiM: Rician modeling of temporal feature variation extracted from variational mode decomposed EEG signal for automatic sleep apnea detection. *IEEE Access*, *6*, 77440–77453. https://doi.org/10.1109/ACCESS.2018.2883062

[13] Jiang, D., Ma, Y., & Wang, Y. (2018, October). A multi-scale parallel convolutional neural network for automatic sleep apnea detection using single-channel EEG signals. In *2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)* (pp. 1–5). IEEE. https://doi.org/10.1109/CISP-BMEI.2018.8633132

[14] Prucnal, M. A., & Polak, A. G. (2018, July). Analysis of features extracted from EEG epochs by discrete wavelet decomposition and Hilbert transform for sleep apnea detection. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 287–290). IEEE. https://doi.org/10.1109/EMBC.2018.8512201

[15] Prucnal, M. A., & Polak, A. G. (2019, July). Effectiveness of sleep apnea detection based on one vs. two symmetrical EEG channels. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 4056–4059). IEEE. https://doi.org/10.1109/EMBC.2019.8856632

[16] Vimala, V., Ramar, K., & Ettappan, M. (2019). An intelligent sleep apnea classification system based on EEG signals. *Journal of Medical Systems*, *43*(2), 1–9. https://doi.org/10.1007/s10916-018-1146-8

[17] Wang, Y., Ji, S., Yang, T., Wang, X., Wang, H., & Zhao, X. (2020). An efficient method to detect sleep hypopnea-apnea events based on EEG signals. *IEEE Access*, *9*, 641–650. https://doi.org/10.1109/ACCESS.2020.3038486

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 323–336
DOI: 10.24425/mms.2023.144866

[18] Wang, Y., Yang, T., Ji, S., Wang, X., Wang, H., Wang, J., & Zhao, X. (2021). A robust sleep apnea-hypopnea syndrome automated detection method based on fuzzy entropy using single lead-EEG signals. https://doi.org/10.21203/rs.3.rs-558448/v1

[19] Khan, I. A., Mahmud, T. I., Mahmud, T., & Fattah, S. A. (2020, December). Deep convolutional neural network based sleep apnea detection scheme using spectro-temporal subframes of EEG signal. In *2020 11th International Conference on Electrical and Computer Engineering (ICECE)* (pp. 463–466). IEEE. https://doi.org/10.1109/ICECE51571.2020.9393059

[20] Mahmud, T., Khan, I. A., Mahmud, T. I., Fattah, S. A., Zhu, W. P., & Ahmad, M. O. (2021). Sleep apnea detection from variational mode decomposed EEG signal using a hybrid CNN-BiLSTM. *IEEE Access*, *9*, 102355–102367. https://doi.org/10.1109/ACCESS.2021.3097090

[21] Alimardani, M., & de Moor, G. (2021). Automatic classification of sleep apnea type and severity using EEG signals. In *BIODEVICES* (pp. 121–128). https://doi.org/10.5220/0010288301210128

[22] Bello, S. A., & Alqasemi, U. (2021). Computer aided detection of obstructive sleep apnea from EEG signals. *Available at SSRN 3890660*. https://doi.org//10.2139/ssrn.3890660

[23] Gurrala, V., Yarlagadda, P., & Koppireddi, P. (2021). Detection of sleep apnea based on the analysis of sleep stages data using single channel EEG. *Traitement du Signal*, *38*(2). https://doi.org/10.18280/ts.380221

[24] Jayaraj, R., & Mohan, J. (2021). Classification of sleep apnea based on sub-band decomposition of EEG signals. *Diagnostics*, *11*(9), 1571. https://doi.org/10.3390/diagnostics11091571

[25] Barnes, L. D., Lee, K., Kempa-Liehr, A. W., & Hallum, L. E. (2022). Detection of sleep apnea from single-channel electroencephalogram (EEG) using an explainable convolutional neural network (CNN). *PLOS one*, *17*(9), e0272167. https://doi.org/10.1101/2021.04.11.439385

[26] Bhalerao, S. V., & Pachori, R. B. (2022). Sparse spectrum based swarm decomposition for robust nonstationary signal analysis with application to sleep apnea detection from EEG. *Biomedical Signal Processing and Control*, *77*, 103792. https://doi.org/10.1016/j.bspc.2022.103792

[27] Chatterjee, A., & Jana, N. D. (2022, July). Classification of sleep apnea event type using imbalanced labelled EEG signal. In *2022 IEEE Region 10 Symposium (TENSYMP)* (pp. 1-6). IEEE.

[28] Khursheed, M. D., Parveen, S., & Farooq, O. (2022, September). Machine learning based system for sleep apnea classification using EEG signal. In *2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA)* (pp. 1–5). IEEE.

[29] Cheng, L., Luo, S., Li, B., Liu, R., Zhang, Y., & Zhang, H. (2023). Multiple-instance learning for EEG based OSA event detection. *Biomedical Signal Processing and Control*, *80*, 104358.

[30] Prucnal, M. A., & Polak, A. G. (2019). Comparison of information on sleep apnoea contained in two symmetric EEG recordings. *Metrology and Measurement Systems*, *26*(2). https://doi.org/10.24425/mms.2019.128351

[31] Gupta, R., Zaidi, T. F., & Farooq, O. (2020, November). Automatic detection of sleep apnea using sub-band features from EEG signals. In *2020 3rd International Conference on Signal Processing and Information Security (ICSPIS)* (pp. 1–4). IEEE. https://doi.org/10.1109/ICSPIS51252.2020.9340133

[32] Zhao, X., Wang, X., Yang, T., *et al.* (2021). Classification of sleep apnea based on EEG sub-band signal characteristics. *Scientific Reports*, *11*(1), 1–11. https://doi.org/10.1038/s41598-021-85138-0

[33] Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, *101*(23), e215–e220. http://circ.ahajournals.org/cgi/content/full/101/23/e215

[34] Zhang, X., Yao, L., Zhang, D., Wang, X., Sheng, Q. Z., & Gu, T. (2017, November). Multi-person brain activity recognition via comprehensive EEG signal analysis. In *Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services* (pp. 28–37). https://doi.org/10.1145/3144457.3144477

[35] Bayatfar, S., Seifpour, S., Oskoei, M. A., & Khadem, A. (2019, April). An automated system for diagnosis of sleep apnea syndrome using single-channel EEG signal. In *2019 27th Iranian Conference on Electrical Engineering (ICEE)* (pp. 1829–1833). IEEE. https://doi.org/10.1109/IranianCEE.2019.8786667

[36] Deepa, V. B., & Thangarajs, P. (2011). A study on classification of EEG Data using the Filters. *IJACSA (4)*. https://doi.org/10.14569/IJACSA.2011.020415

[37] Lin, R., Lee, R. G., Tseng, C. L., Zhou, H. K., Chao, C. F., & Jiang, J. A. (2006). A new approach for identifying sleep apnea syndrome using wavelet transform and neural networks. *Biomedical Engineering: Applications, Basis and Communications*, *18*(03), 138–143. https://doi.org/10.4015/S1016237206000233

[38] Dragomiretskiy, K., & Zosso, D. (2013). Variational mode decomposition. *IEEE Transactions on Signal Processing*, *62*(3), 531–544. https://doi.org/10.1109/TSP.2013.2288675

[39] Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(8), 1226–1238. https://doi.org/10.1109/TPAMI.2005.159

[40] Robnik-Šikonja, M., & Kononenko, I. (2003). Theoretical and empirical analysis of ReliefF and RReliefF. *Machine learning*, *53*(1), 23–69. https://doi.org/10.1023/A:1025667309714

[41] Prasath, V. B., Alfeilat, H. A. A., Hassanat, A., Lasassmeh, O., Tarawneh, A. S., Alhasanat, M. B., & Salman, H. S. E. (2017). Distance and similarity measures effect on the performance of k-nearest neighbor classifier – a review. *arXiv preprint arXiv:1708.04321*. https://doi.org/10.1089/big.2018.0175

[42] Saccá, V., Campolo, M., Mirarchi, D., Gambardella, A., Veltri, P., & Morabito, F. C. (2018). On the classification of EEG signal by using an SVM based algorithm. In *Multidisciplinary Approaches to Neural Computing* (pp. 271–278). Springer, Cham. https://doi.org/10.1007/978-3-319-56904-8_26

[43] Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, *5*(4), 455. https://doi.org/10.1007/BF02551274

[44] Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, *45*(4), 427–437. https://doi.org/10.1016/j.ipm.2

Monika A. Prucnal obtained the M.Sc. and Ph.D. degrees in Electronics, Technical Sciences, from Wroclaw University of Science and Technology, Wroclaw, Poland, in 2016 and 2021, respectively. Her doctoral thesis was entitled *Methods for EEG Signal Analysis Aimed at Sleep Apnea Detection*. She is currently an Assistant Professor with Wroclaw University of Science and Technology, Poland. Her research activity focuses on biomedical signal processing, including EEG, ECG and PPG signals, as well as on artificial intelligence methods.

Adam G. Polak received the M.Sc., Ph.D., and D.Sc. degrees in Technical Sciences from Wroclaw University of Science and Technology (WUST), Poland, in 1988, 1994 and 2008, respectively. Currently, he is an Associate Professor at WUST. His research includes: indirect measurements, regularization of parameter estimation and signal restoration, modelling and measurements in physiology, machine learning in disease detection. He was a recipient of the awards of the Foundation for Polish Science (1995), Department IV of the Polish Academy of Sciences (PAS) (2008), and the scholarship "Enhancing Science Linkages between New Zealand and Europe through Poland" (2010). He is a member of the Committee on Metrology and Scientific Instrumentation, PAS, and an IEEE Member.