

Article

Multi-Class Sleep Stage Analysis and Adaptive Pattern Recognition

Aleš Procházka ^{1,2,*} , Jiří Kuchyňka ³, Oldřich Vyšata ^{1,2,3}, Pavel Cejnar ¹, Martin Vališ ³ and Vladimír Mařík ²

¹ Department of Computing and Control Engineering, University of Chemistry and Technology in Prague, 166 28 Prague, Czech Republic; Vysatao@gmail.com (O.V.); Pavel.Cejnar@vscht.cz (P.C.)

² Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University in Prague, 166 36 Prague, Czech Republic; Vladimír.Marik@cvut.cz

³ Department of Neurology, Faculty of Medicine in Hradec Králové, Charles University, 500 05 Prague, Czech Republic; Kuchynkajir@seznam.cz (J.K.); Valismar@seznam.cz (M.V.)

* Correspondence: A.Prochazka@ieee.org; Tel.: +420-220-444-198

Received: 2 April 2018; Accepted: 27 April 2018; Published: 1 May 2018



Abstract: Multimodal signal analysis based on sophisticated sensors, efficient communication systems and fast parallel processing methods has a rapidly increasing range of multidisciplinary applications. The present paper is devoted to pattern recognition, machine learning, and the analysis of sleep stages in the detection of sleep disorders using polysomnography (PSG) data, including electroencephalography (EEG), breathing (Flow), and electro-oculogram (EOG) signals. The proposed method is based on the classification of selected features by a neural network system with sigmoidal and softmax transfer functions using Bayesian methods for the evaluation of the probabilities of the separate classes. The application is devoted to the analysis of the sleep stages of 184 individuals with different diagnoses, using EEG and further PSG signals. Data analysis points to an average increase of the length of the Wake stage by 2.7% per 10 years and a decrease of the length of the Rapid Eye Movement (REM) stages by 0.8% per 10 years. The mean classification accuracy for given sets of records and single EEG and multimodal features is 88.7% (standard deviation, STD: 2.1) and 89.6% (STD:1.9), respectively. The proposed methods enable the use of adaptive learning processes for the detection and classification of health disorders based on prior specialist experience and man-machine interaction.

Keywords: multimodal signal analysis; pattern recognition; machine learning; computational intelligence; polysomnography; sleep stage classification

1. Introduction

Polysomnography (PSG) represents a diagnostic tool of sleep medicine based on biophysiological changes that occur during the sleep. Multimodal data acquired in sleep laboratories form multichannel records that require wire attachments to the patients in most cases. Selected physiological signals are simultaneously monitored by specific sensors [1–3] and they form time series that are recorded with different sampling frequencies and often combined with videosequences acquired by infra and thermographic imaging cameras. Figure 1(a1,a2) present an example of a multichannel PSG record 5 s long used for diagnosis and classification of sleep disorders either by an experienced neurologist or by an automatic classification model.

The segmentation and pattern recognition process is based either on the analysis of the whole set of multimodal and multichannel PSG data or on the processing of data acquired by selected sensors only. The recorded PSG signals include an electroencephalogram (EEG), electro-oculogram

(EOG), electrocardiogram (ECG), electromyogram (EMG), sound, movement, and breathing records (Flow), among others. Features detected by specific computational methods [4–8] are then used to help neurologists to make a diagnosis and propose the appropriate treatment. Sleep features are very specific and there exist many studies proposing machine learning [9,10] for the automatic detection of sleep stages [11–14]. A specific interest is devoted to the use of hidden Markov models for automatic sleep staging [15–17], to the relation between the adjacent sleep segments, and to deep convolutional neural network [18,19] as well.

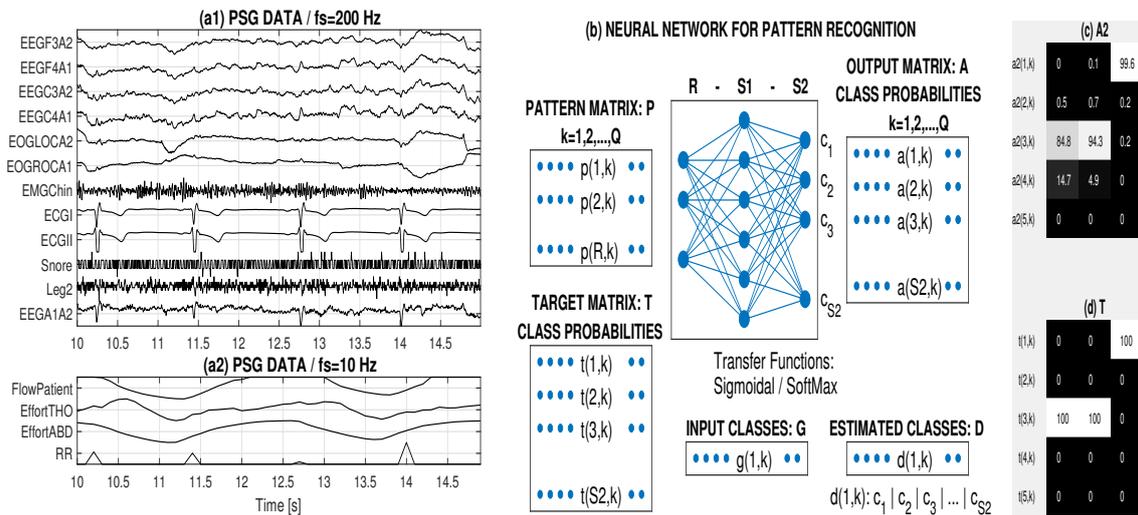


Figure 1. Principle of sleep feature classification using a two-layer neural network with sigmoidal and softmax transfer functions presenting (a1,a2) a sample record of multichannel data 5 s long recorded with the sampling frequency of 200 and 10 Hz, respectively; (b) neural network structure for pattern recognition estimating probabilities of individual classes and evaluating the most probable class for each column vector of input features; and (c,d) evaluated and target probabilities (%) of output classes for selected patterns (detail of Figure 3d and S2 = 5).

Figure 2 presents the rapidly growing interest in the PSG data analysis as illustrated by the number of papers registered in the Web of Science (WoS) database. This increase can be observed both in the area of sleep stages classification (Figure 2a) and machine learning for detection of sleep disorders (Figure 2b). The mostly cited papers (according to the WoS) published recently and devoted to PSG include those of sleep stages classification [20] and machine learning for sleep disorders detection [21,22] with detail comparative reviews of methods and results achieved.

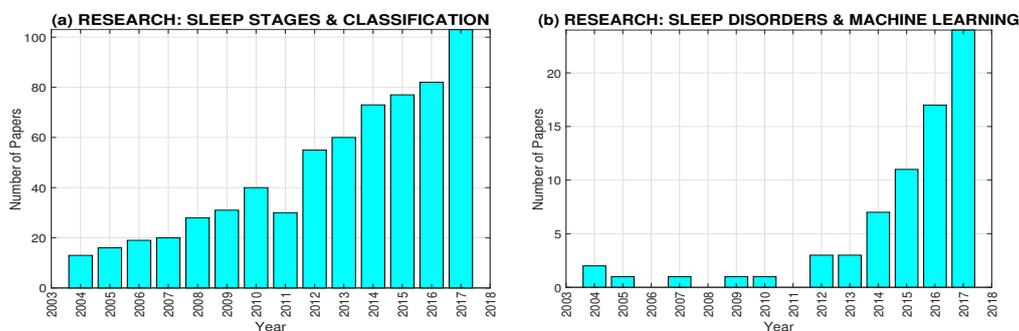


Figure 2. The time evolution of number of journal papers devoted to analysis of (a) sleep stages and their classification and (b) sleep disorders and machine learning according to the Web of Science.

The present paper is devoted to the statistical distribution of sleep stages related to age and diagnosis [23] and to their classification, using neural networks [24–27] with sigmoidal and softmax transfer functions. The mathematical features of each segment are summarized in the pattern matrix and associated with target classes specified by an experienced neurologist. The classification system is presented in Figure 1b. The proposed adaptive process incorporates the prior experience of a neurologist into the decision process using man–machine interaction.

The proposed algorithm was applied for the adaptive recognition and classification of sleep stages [20,28–32] using both EEG data [13,33] and PSG signals, using their spectral components for the estimation of sleep features. Sleep scoring assumes the knowledge of the Flow, EMG, and EOG records in addition to EEG signals but they affect EEG records as well. This fact has motivated some authors to use EEG channels only [21] to classify sleep stages.

The training process analysed in the present paper results in the evaluation of the coefficients of a mathematical model for multiclass pattern recognition [34–37]. The classification accuracy and cross-validation error were then used for comparing the results. The resulting mathematical model can be further updated using expanded databases of individuals and visual sleep scoring as well as the experience of a neurologist in the associated interactive process.

The proposed mathematical model forms a contribution to further automatic systems of sleep stages classification and data analysis for the evaluation of proportion of Wake and REM segments associated with the age of an individual. These information can provide an additional diagnostic tool in the clinical environment.

The present paper is organized as follows. Section 2 presents PSG data acquisition, pattern matrix construction and the methodology of machine learning for pattern recognition of sleep stages. Section 3 describes results of the age analysis of sleep stages and their classification using both multimodal and EEG features of individuals with different diagnosis. The final Section 4 is dedicated to the discussion and conclusion.

2. Methods

2.1. PSG Data Acquisition

The data studied include the own database of 184 polysomnography overnight observations [38–40] of patients recorded, diagnosed and classified by the authors in the sleep neurological laboratory. Table 1 presents a summary of the data for (i) 63 healthy individuals (36 males, 27 females); (ii) 77 individuals with sleep apnea (M = 48, F = 29); (iii) 16 individuals with restless leg syndrome (M = 8, F = 8); and (iv) 28 individuals with the restless leg syndrome combined with sleep apnea (M = 19, F = 9). Each of these records is about 8 hours long, segmented and classified into 5 sleep stages:

- Wake stage (Wake): normal body functions,
- No Rapid Eye Movement 1 (NonREM1): the initial sleep stage with eyes closed,
- No Rapid Eye Movement 2 (NonREM2): the light sleep stage with slower heart rate and body temperature going down,
- No Rapid Eye Movement 3 (NonREM3): the deep sleep stage during which the body repairs and regrows tissues,
- Rapid Eye Movement (REM): the specific sleep period with faster brain activities, faster breathing and heart rate (period of dreams).

These stages were specified by a neurologist specializing in sleep analysis.

The matrix of recorded sleep PSG data of each individual include an information text file and 8 EDF files with subsequent one hour long multichannel records of 22 variables acquired during the night. While EEG, EOG, ECG, sound and movement data were observed with a sampling frequency of 200 Hz, breathing records (Flow) were acquired with a sampling frequency of 10 Hz.

Table 1. Specification of the set of 184 individuals under study (M: number of males, F: number of females) with their age distribution, including (i) healthy individuals; (ii) individuals with sleep apnea (SA); (iii) individuals with the restless leg (RL) syndrome; and (iv) individuals with RL syndrome and sleep apnea syndrome

Diagnosis		M-Male		
		Number	Mean Age (Year)	STD
1	Healthy individuals	36	40.0	15.9
2	Sleep Apnea (SA)	48	50.3	14.2
3	Restless Leg (RL) Syndrome	8	59.9	9.1
4	SA and RL Syndrome	19	58.3	9.0

Diagnosis		F: Females		
		Number	Mean Age (Year)	STD
1	Healthy individuals	27	40.8	14.7
2	Sleep Apnea (SA)	29	54.7	14.4
3	RL Syndrome	8	48.9	12.6
4	RL Syndrome and SA	9	52.7	20.4

A selected EEG data segment 2 h long of a chosen individual together with its manual classification into 5 classes is presented in Figure 3a. The time localization of these classes specified by an experienced neurologist is shown in Figure 3c.

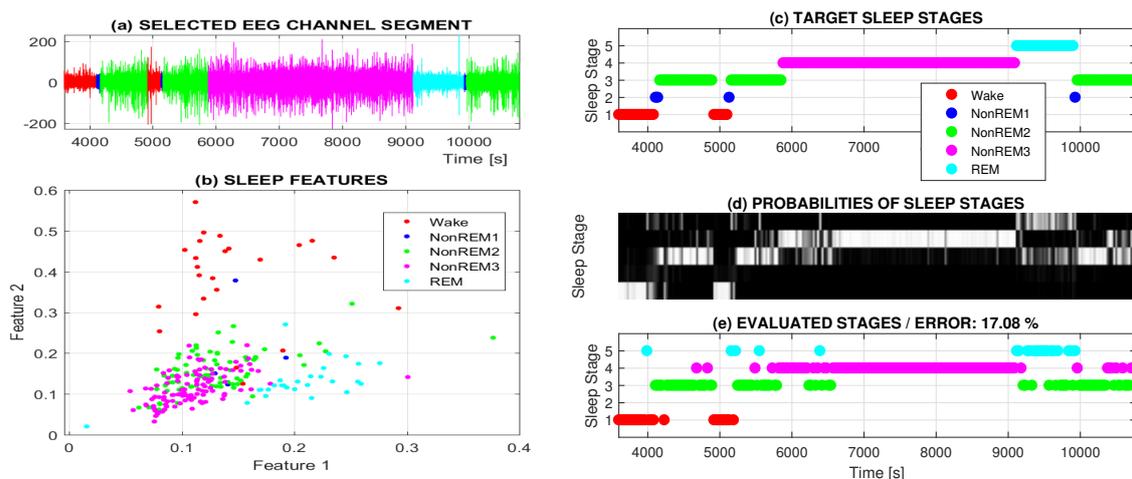


Figure 3. Two-hour long analysis presenting (a) EEG signal segmented into 5 classes and (b) features standing for the mean power in selected frequency ranges (1–4 Hz, 4–8 Hz) in subsequent 30 s long windows forming neural network pattern values; (c) target sleep classes forming desired neural network output; (d) probabilities of evaluated classes by the two layer neural network; and (e) resulting classes with the highest classification probability.

2.2. Pattern Matrix Construction

The classification of sleep stages was based on time segments 30 s long with the corresponding target classes specified by the neurologist during the learning process after the preceded signal preprocessing stage. This initial process included digital filtering for noise rejection and artefact removal [41,42].

Common preprocessing methods include the rejection of noise components present in each individual channel. Denoting a channel signal by $\{x(n)\}_{n=0}^{N-1}$, it is possible to use a selected digital filter with specified cutoff frequencies to define a new sequence $\{y(n)\}_{n=0}^{N-1}$. Finite impulse response (FIR) band pass filters of the 30th order were used in this case. Their cutoff frequencies corresponded

with the frequency ranges of specific PSG channels used for classification. The selection of these frequencies was related to the physiological background of these signals and it included frequency ranges $\langle 0.5, 30 \rangle$ Hz for EEG channels, $\langle 0.05, 0.8 \rangle$ Hz for the flow channel, and $\langle 4, 40 \rangle$ Hz for the EOG channel.

A selected spectral feature FS of a signal segment was evaluated by the discrete Fourier transform as a relative power in the specified frequency band $\langle f_{c1}, f_{c2} \rangle$ by relation

$$FS = \frac{\sum_{k \in \Phi} |Y(k)|^2}{\sum_{k=0}^{N/2} |Y(k)|^2}, \quad Y(k) = \sum_{n=0}^{N-1} y(n) e^{-jkn \frac{2\pi}{N}} \quad (1)$$

where Φ is the set of indices for which frequency values $f_k = \frac{k}{N} f_s \in \langle f_{c1}, f_{c2} \rangle$.

The construction of the pattern matrix $\mathbf{P}_{R,Q}$ presented in Figure 1b assumes the evaluation of R features for each signal segment for $k = 1, 2, \dots, Q$, forming Q column vectors $\mathbf{p}(:, k)$ of R values. As the length of each PSG segment is 30 s, the whole overnight record about 8 h (480 min) long includes on average $Q = 960$ segments for each individual. Each segment is described by R features that represent characteristics evaluated mostly in the time or frequency domains for one or more observed variables.

Figure 3b presents an example of a distribution of EEG features evaluated from one EEG channel only, specifying the average power in selected frequency bands (1–4 and 4–8 Hz). For each feature vector, the associated sleep stage (class) is specified by the neurologist, thus forming the target class for the following classification process as specified in Figure 3c.

2.3. Machine Learning for Pattern Recognition

The pattern matrix $\mathbf{P}_{R,Q}$ formed an input for the two-layer neural network presented in Figure 1b having R inputs, $S1$ neurons in the first layer and $S2$ neurons in the second layer. The outputs of individual layers included the values

$$\mathbf{A1}_{S1,Q} = F1(\mathbf{N1}), \quad \mathbf{A2}_{S2,Q} = F2(\mathbf{N2}) \quad (2)$$

where

$$\mathbf{N1} = \mathbf{W1}_{S1,R} \mathbf{P}_{R,Q} + \mathbf{b1}_{S1,1} \mathbf{ones}(1, Q), \quad \mathbf{N2} = \mathbf{W2}_{S2,S1} \mathbf{A1}_{S1,Q} + \mathbf{b2}_{S2,1} \mathbf{ones}(1, S1) \quad (3)$$

define the arguments of the transfer functions $F1, F2$.

An associated matrix of target values $\mathbf{T}_{S2,Q}$ with $S2 = 5$ rows was defined by the corresponding sleep stages specified by a neurologist in the learning stage (as presented in Figure 1b). For each column vector in the pattern matrix, the corresponding target vector has one unit element in the row pointing to the correct target value. The network coefficients included elements of the matrices $\mathbf{W1}_{S1,R}$, $\mathbf{W2}_{S2,S1}$ and associated vectors $\mathbf{b1}_{S1,1}$, $\mathbf{b2}_{S2,1}$. The proposed model used the sigmoidal transfer function $F1$ in the first layer and the probabilistic softmax transfer function $F2$ in the second layer. The values of the output layer, based on the Bayes theorem [43], using the function

$$F2(\mathbf{N2}) = \frac{\exp(\mathbf{N2})}{\text{sum}(\exp(\mathbf{N2}))} \quad (4)$$

provided the probabilities of each class. Figure 1c,d present an example of estimated and target probabilities for selected sleep pattern vectors. Each column of the output matrix presents class membership to individual classes.

The training of the neural network was performed in the MATLAB (R2017b, The MathWorks, Inc., Natick, MA, United States, 2017) environment using 70% of pattern vectors for training, 15% for validation, 15% for testing, and random initial conditions. Model coefficients were updated for each new sleep record that decreased (with the upper 3% limit) the performance value calculated as the average squared error between the network outputs and target values. The optimization process was done by a single PC with the i7-6500U CPU (2.5 GHz), 16 GB RAM and the 64-bit operating system (Windows 10 Pro).

The evaluation of the classification results was performed by the analysis of the multi-class Receiver Operating Characteristic (ROC) [44,45] to illustrate the performance of the classifier system, and by a confusion matrix presented in Table 2 for classification into S classes. The associated common performance metrics are summarized in Table 3. The ROC analysis was performed on the basis of a pairwise comparison of one class against all other classes [46].

Table 2. The confusion matrix for the evaluation of a model classifying pattern vectors belonging to true (target) classes and estimating their (predicted) classes for the multi-class classification into S categories.

		True (target) Class			False Pos. (out-diag. row sum)
		C(1)	C(k)	C(S)	
Output (predicted) class	C(1)	$x(1,1)$ True Pos. (TP(1))	$x(1,k)$	$x(1,S)$	FP(1)
	
	C(k)	$x(k,1)$	$x(k,k)$ True Pos. (TP(k))	$x(k,S)$	FP(k)
...		
C(S)	$x(S,1)$	$x(S,k)$	$x(S,S)$ True Pos. (TP(S))	FP(S)	
False Neg. (out-diag. col. sum)		FN(1)	FN(k)	FN(S)	

Table 3. Common performance metrics of the confusion matrix for the multi-class classification.

Characteristics	Definition	Comment
PPV Precision of class k (pos. predict. value)	$\frac{TP(k)}{TP(k) + FP(k)}$	Probability of correct classification of class k related to the number of instants classified to class k
TPR Sensitivity of class k (True positive rate, recall)	$\frac{TP(k)}{TP(k) + FN(k)}$	Probability of correct classification of class k related to the number of instants belonging to class k
TNR Specificity of class k (True negative rate)	$\frac{TNK}{TNK + FPK}$ <small>$TNK = \sum TP(j), FPK = \sum FP(j), j \neq k$</small>	Probability of incorrect classification of class k related to the number of instants not classified to class k
FPR False positive rate	$\frac{FPK}{TNK + FPK}$	Probability of positive classification for the negative set (1-specificity)
ACCU Accuracy	$\frac{\sum_k TP(k)}{\sum_k (TP(k) + FP(k))}$	Probability of global correct classification

3. Results

Figure 3a presents the results for a two hour long sleep record selected from the observation of an entire night in the sleep laboratory. The target sleep stages (input classes) associated with 30 s long EEG data windows diagnosed by the experienced neurologist are presented in Figure 3c. The adaptive system was constructed as a two layer neural network with the sigmoidal and softmax transfer functions. Figure 3d presents the evaluated probabilities of the individual classes. The estimated sleep stages pointing to sleep stages with the highest probability are presented in Figure 3e.

The classification results obtained by the neural network system were compared with those achieved with the k -nearest neighbour and decision tree methods (Table 4). The process of class

selection and evaluation of class boundaries of the sleep stages is illustrated in Figure 4 for the two features and their classification.

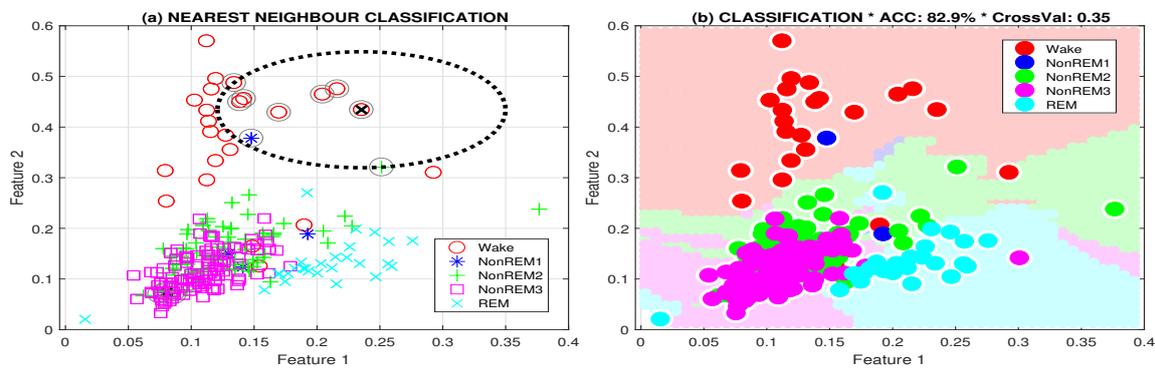


Figure 4. Classification of sleep stages presenting (a) class selection for the 5-nearest neighbour method and (b) class boundaries of the computational model and two features evaluated as the mean power in selected frequency ranges (1–4 Hz, 4–8 Hz).

Table 4. Comparison of the accuracy and cross-validation for classification of sleep features evaluated during the two-hour long observation and 30 s long time windows by the neural network, nearest neighbour, and decision tree models.

Method	Parameters	Accuracy (%)	Cross-Valid.
$R-S1-S2$	$S1 = 3$	79.58	0.13
Neural Net	$S1 = 6$	87.08	0.10
$(R = 4, S2 = 5)$	$S1 = 10$	88.33	0.12
k -Nearest	$k = 3$	82.9	0.35
Neighbour	$k = 5$	77.9	0.29
Decision Tree		85.7	0.14

The classification accuracies and cross-validation errors evaluated by the leave-one-out method for a sample two-hour long EEG record are presented in Table 4. Input features include relative power in 4 frequency bands (1–4, 4–8, 8–12, 12–20 Hz). For the given set of features, the neural network model provide classification results with higher accuracy and lower cross-validation error than did the k -nearest neighbour and decision tree methods.

The methodology of sleep stages processing verified for short records was then applied for an extensive set of 184 individuals specified in Table 1. The results of its analysis with statistical distribution of sleep stages related to age is presented in Figure 5. Linear approximation of these values together with estimation of 95% confidence bounds were evaluated. Resulting regression coefficients point to an average increase of the length of the Wake stage by 2.7% per 10 years and a decrease of the length of the REM stage by 0.8% per 10 years.

Figure 6 presents more detail analysis of the resulting lengths of the sleep stages [47] related to the age for the whole set of 184 individuals divided according to their diagnosis as determined by experienced neurologists. A well documented increase of the Wake stages and decrease of REM stages related to age is confirmed for healthy individuals and patients with sleep apnea. For patients with restless leg syndrome, no such distribution was observed. The summary of associated regression coefficients with their 95% confidence bounds are presented in Table 5.

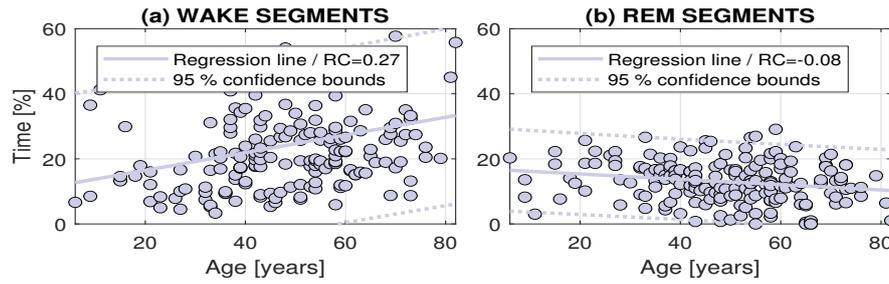


Figure 5. The distributions of the lengths of the Wake and REM sleep stages related to age for the whole set of 184 individuals under study.

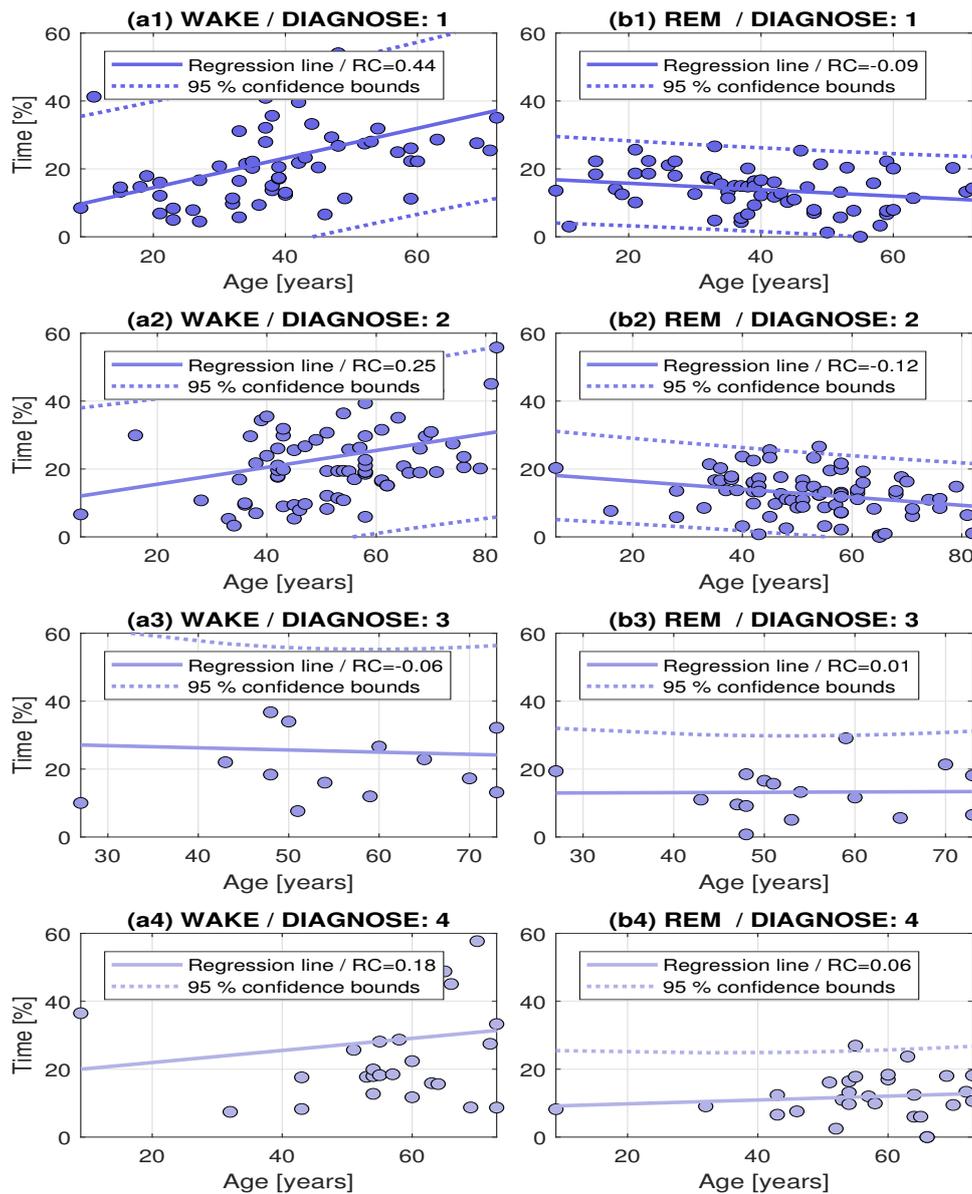


Figure 6. Distributions of Wake and REM sleep stages related to age for (a1,b1) healthy individuals and patients with (a2,b2) sleep apnea; (a3,b3) restless leg syndrome; and (a4,b4) restless leg syndrome and sleep apnea.

Table 5. The summary of regression coefficients (RC) of the age dependence of Wake and REM sleep stages with their 95% confidence bounds.

Diagnosis	Wake Segments		REM Segments	
	RC	Confidence Bounds	RC	Confidence Bounds
All	0.272	(0.145, 0.399)	−0.082	(−0.141, −0.023)
1	0.437	(0.227, 0.648)	−0.094	(−0.198, 0.010)
2	0.248	(0.055, 0.442)	−0.119	(−0.216, −0.022)
3	−0.064	(−0.687, 0.559)	0.009	(−0.334, 0.353)
4	0.178	(−0.377, 0.734)	0.057	(−0.133, 0.246)

Table 6 presents the summary of the results for the individuals specified in Table 1 with 4 different diagnoses and their classification into three classes (C1: Wake, C2: NonREM, C3: REM). The learning performance was evaluated by the cross-entropy for given targets and network outputs. The last column of Table 6 presents cross-validation errors evaluated by the K-fold cross-validation method (for K = 10). A comparison of the classification results was performed using

1. the EEG channels only, with patterns evaluated as mean energies in 5 selected frequency bands (1–4, 4–8, 8–12, 12–16, 16–20 Hz),
2. additional features evaluated from the 4 frequency bands (0.05–0.2, 0.2–0.3, 0.3–0.4, 0.4–0.8 Hz) of the flow channel and 3 frequency bands (4–15, 15–30, 30–40 Hz) of the electro-oculogram channel.

The values in Table 6 show nearly no difference between males and females and a slight improvement of the performance and accuracy of the neural network model using multimodal features consisting of 12 items instead of 5 EEG features associated with each segment.

The results presented in Table 6 are based on the manual classification of the PSG records performed by an experienced neurologist incorporating his experience. The adaptive optimization of the network coefficients can be further performed for newly acquired records.

Table 6. Evaluation of sleep stage classification (Classes: C1—Wake, C2—NonREM, C3—REM) of sets of (i) healthy individuals; (ii) individuals with sleep apnea (SA); (iii) individuals with the restless leg (RL) syndrome, and (iv) individuals with sleep apnea and RL syndrome using the single EEG channel for 5 (F5) mean EEG power features (in bands 1–4, 4–8, 8–12, 12–16, and 16–20 Hz), and 12 (F12) features with additional 4 features evaluated from the flow channel (for frequency bands 0.05–0.2, 0.2–0.3, 0.3–0.4, 0.4–0.8 Hz) and 3 features related to the electro-oculogram channels (4–15, 15–30, 30–40 Hz).

Diagnosis (Male)		Precision (%)			Recall (%)			Acc. (%)	Perf.	10-Fold Cross-Valid.
		C1	C2	C3	C1	C2	C3			
1 Healthy Ind.	- F5	89.4	90.1	81.8	89.9	92.6	71.7	88.9	0.0977	0.15
	- F12	89.8	90.7	79.5	91.6	92.2	70.3	89.1	0.0941	0.12
2 Sleep Apnea	- F5	92.9	94.8	79.7	94.8	95.4	72.5	92.7	0.0632	0.11
	- F12	93.8	95.4	81.4	95.3	95.6	76.5	93.5	0.0598	0.07
3 RL Syndrome	- F5	84.5	94.7	81.7	88.8	92.7	82.5	90.1	0.0845	0.13
	- F12	86.0	95.0	84.3	89.3	93.5	85.1	91.1	0.0758	0.10
4 SA and RL Syn.	- F5	91.5	88.1	76.7	86.8	94.0	62.5	87.8	0.0996	0.19
	- F12	90.9	89.3	83.2	88.3	94.2	67.6	89.1	0.0973	0.15
Diagnosis (Female)		Precision (%)			Recall (%)			Acc. (%)	Perf.	10-Fold Cross-Valid.
		C1	C2	C3	C1	C2	C3			
1 Healthy Ind.	- F5	82.9	90.6	72.5	77.5	92.1	73.8	86.0	0.1186	0.16
	- F12	83.0	91.9	76.9	81.8	92.5	76.2	87.6	0.1055	0.14
2 Sleep Apnea	- F5	86.0	91.7	73.5	90.6	93.3	62.4	88.0	0.1061	0.17
	- F12	87.5	92.2	74.7	92.4	93.4	64.1	88.7	0.1015	0.15
3 RL Syndrome	- F5	85.9	92.0	81.3	86.3	93.1	75.9	89.3	0.1006	0.15
	- F12	87.5	92.2	80.0	86.3	93.5	76.7	89.6	0.0953	0.12
4 SA and RL Syn.	- F5	83.0	89.7	81.9	84.0	91.4	72.6	86.9	0.1063	0.21
	- F12	84.0	90.9	83.9	86.1	91.9	75.3	88.1	0.1049	0.14

An analysis of the features in each 30 s long PSG segment was then performed for spectral components in selected regions for the EEG, Flow and OCG channels for all 184 individuals. This resulted in the specification of the pattern matrix for the adaptive classification by the two layer neural network after the learning process based on associated target values. Results for the training, validation and test sets for healthy individuals (after about 15,000 training epochs) are presented in associated confusion matrices presented in Table 7 for classification into three classes using 12 (F12) multimodal (EEG, Flow and ECG) features. The highest true-positive rate was achieved for the second (NonREM) class and all these sets.

Table 7. Confusion matrix of the classification neural network model for the training, validation, and test sets of sleep stages (for healthy individuals) with true positive values on the matrix diagonal (in the bold).

<i>Training Set</i>		<i>Target Class</i>				
	<i>k</i>	1	2	3	<i>FP (k)</i>	<i>PPV (k)</i>
<i>Output Class</i>	1	4644	601	236	837	84.7%
	2	368	10,971	534	902	92.4%
	3	152	405	1958	557	77.9%
<i>FN (k)</i>		520	1006	770	<i>ACC: 88.4%</i>	
<i>TPR (k)</i>		89.9%	91.6%	71.8%	<i>Error: 11.6%</i>	
<i>Validation Set</i>		<i>Target Class</i>				
	<i>k</i>	1	2	3	<i>FP (k)</i>	<i>PPV (k)</i>
<i>Output Class</i>	1	1007	144	38	182	84.7%
	2	85	2304	114	199	92.0%
	3	44	95	421	139	75.2%
<i>FN (k)</i>		129	239	152	<i>ACC: 87.8%</i>	
<i>TPR (k)</i>		88.6%	90.6%	73.5%	<i>Error: 12.2%</i>	
<i>Test Set</i>		<i>Target Class</i>				
	<i>k</i>	1	2	3	<i>FP (k)</i>	<i>PPV (k)</i>
<i>Output Class</i>	1	1007	155	49	204	83.2%
	2	76	2337	106	182	92.8%
	3	34	79	416	113	78.6%
<i>FN (k)</i>		110	234	155	<i>ACC: 88.3%</i>	
<i>TPR (k)</i>		90.2%	90.9%	72.9%	<i>Error: 11.7%</i>	

Figure 7 presents the target and predicted classes of three selected patients with sleep apnea classified into three classes, including (a) class 1 (Wake), (b) class 2 (NonREM), and (c) class 3 (REM). The associated ROC characteristics in Figure 7d show the true positive rate of each class against the false positive rate of all other classes.

The final results including the accuracy and the performance of neural network models for different sleep disorders and male/female sets of individuals are presented in Figure 8. The results show the similarity in the accuracy for features evaluated from the single EEG channel (F5) and different multimodal (EEG, Flow and OCG) channels (F12). The slightly lower accuracy achieved for females in this case can be explained by the higher deviations of the associated features. The mean classification accuracy is 88.7% (STD: 2.1) and 89.6% (STD: 1.9), respectively, for the given sets of records and classification using single (F5) and multimodal (F12) features. This result can be explained

by the presence of additional attributes in the EEG channel. The accuracy achieved corresponds with that published in [25] for a smaller group of individuals and different feature sets.

Table 8 presents the summary of both the achieved global classification accuracy, the *K*-fold cross-validation errors (for *K* = 10) and the precision of classification into three classes over all sleep disorders with associated detailed values summarized in Table 6.

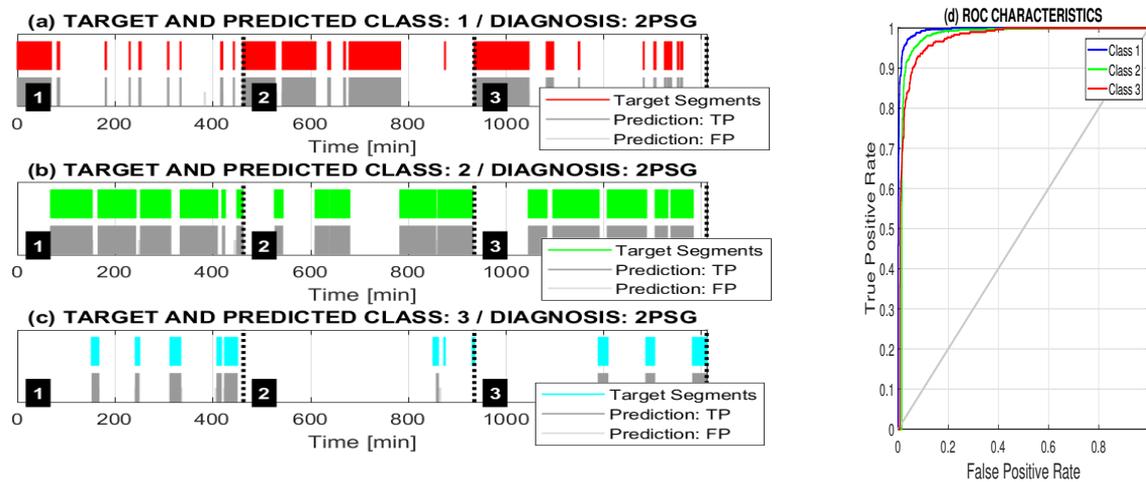


Figure 7. Target and predicted classes of three selected patients with sleep apnea classified by the adaptive neural network model into three classes including (a) class 1 (Wake); (b) class 2 (NonREM); and (c) class 3 (REM) together with (d) associated ROC characteristics.

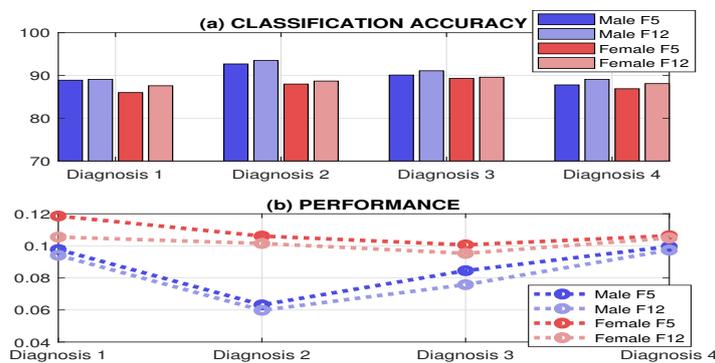


Figure 8. Results of the training process presenting (a) the accuracy and (b) the performance of the resulting model for different sleep disorders for sets of male and female individuals using EEG (F5) and multimodal EEG, Flow and OCG (F12) features.

Table 8. Classification accuracy and the 10-fold cross-validation errors of the sets of males and females and precision of individual classes over all sleep disorders for 5 (F5) mean EEG power features and 12 (F12) features with an additional 4 features evaluated from the flow and electro-oculogram channels.

Sex	Accuracy (%)		10-Fold Cross-Valid.		Class	Precision (%)	
	F5	F12	F5	F12		F5	F12
Male	89.9	90.7	0.15	0.11	Class 1: Wake	87.0	87.8
Female	87.6	88.5	0.17	0.14	Class 2: NonREM	91.5	92.2
Mean	88.7	89.6	0.16	0.12	Class 3: REM	78.6	80.5

4. Conclusions

This paper presents a method involving multi-class classification applied to the recognition of sleep stages using a two layer neural network with the sigmoidal and the softmax transfer functions to

evaluate the probabilities of the output classes. The training set includes overnight PSG records of 184 individuals forming an own database acquired and classified by the authors.

The proposed method allows the construction of a neural network and its training to recognize sleep stages with a mean classification accuracy of 88.7% (STD: 2.1) specified in Tables 6 and 8 for selected spectral features of the single EEG channel. The accuracy increased to 89.6% (STD: 1.9) for multimodal features obtained from EEG, flow and EOG channels was not too high, probably owing to the presence of corresponding artifacts in the EEG channels. The mean 10-fold cross-validation error decreased from 0.16 (for 5 EEG features) to 0.12 (for 12 multimodal features).

The novel approach of the proposed method is the possibility of constructing a mathematical model incorporating the individual experience of a neurologist, allowing a data classification close to his previous strategy and an adaptive modification of model coefficients. Results obtained are compared with selected existing methods in Table 9. The accuracy of sleep stages classification is generally between 70% and 92% [22] and the accuracy of the current model is close to the upper limit of this range.

Table 9. A comparison of the proposed method with further published results.

Reference	Sleep Stages	Signals and Features	Model	Dataset	Accuracy
Fraiwan et al. [8] 2010	Wake, REM NonREM1-4	EEG (21 time-freq. features)	wavelet entropy and discriminant analysis	32 subjects	84.0 %
Pan et al. [15] 2012	Wake, REM, SWS NonREM1-2	EEG, EOG, EMG (13 energy features)	discrete hidden Markov model	20 subjects	85.3 %
Hsu et al. [27] 2013	Wake, REM NonREM1-2	EEG (6 energy features)	Elman recurrent neural classifier	8 subjects	87.2 %
Peker [22] 2016	Wake, REM NonREM1-4	EEG (8 complex features)	complex-valued neural networks	8 subjects	91.6 %
Sors et al. [18] 2018	Wake, REM NonREM1-4	raw EEG (no features selected)	deep convolutional neural network	5728 segments	87.0 %
Current Study	Wake, REM NonREM	EEG, Flow, EOG (12 energy features)	Bayesian neural network classifier	184 subjects (own data)	89.6 %
Current Study	Wake, REM NonREM	EEG (5 energy features)	Bayesian neural network classifier	184 subjects (own data)	88.7 %

It is projected that the neural network model will be further used for more extensive datasets and adaptation of the results to the experience of the neurologist to reduce the time necessary for manual classification. A special attention will be paid to deep convolutional neural networks [17,18] that allow row data processing without specification of separate features.

Author Contributions: A.P. was responsible for the mathematical and algorithmic tools allowing the data processing, J.K. was responsible for PSG data acquisition, O.V. interpreted the acquired data and results obtained from the neurological point of view, P.C. contributed to the data preprocessing, M.V. contributed to interpretation of results from the neurology point of view and V.M. was responsible for computational processing and visualization. All authors have read and approved the final manuscript.

Acknowledgments: Real data were kindly provided by the Department of Neurology of the Faculty Hospital of the Charles University in Hradec Kralove. All procedures involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments. Informed consent was obtained from all individual participants included in the study. This study was partially supported by the grant projects of the Ministry of Health of the Czech Republic (FN HK 00179906) and of the Charles University in Prague, Czech Republic (PROGRES Q40).

Conflicts of Interest: The authors declare that there is no conflict of interest.

References

1. Looney, D.; Goverdovsky, V.; Rosenzweig, I.; Morrell, M.; Mandic, D. Wearable In-Ear Encephalography Sensor for Monitoring Sleep. *Ann. Am. Thorac. Soc.* **2016**, *13*, 2230–2233.
2. Procházka, A.; Schätz, M.; Centonze, F.; Kuchyňka, J.; Vyšata, O.; Vališ, M. Extraction of Breathing Features Using MS Kinect for Sleep Stage Detection. *Signal Image Video Process.* **2016**, *10*, 1278–1286.
3. Procházka, A.; Schätz, M.; Vyšata, O.; Vališ, M. Microsoft Kinect Visual and Depth Sensors for Breathing and Heart Rate Analysis. *Sensors* **2016**, *16*, 996.
4. Liaw, S.; Chen, J. Characterizing sleep stages by the fractal dimensions of electroencephalograms. *Biostat. Biom.* **2017**, *2*, 555584:1–555584:7.
5. Kianzad, R.; Kordy, H. Automatic Sleep Stages Detection Based on EEG Signals Using Combination of Classifiers. *J. Electr. Comput. Eng. Innov.* **2013**, *1*, 99–105.
6. Metsis, V.; Kosmopoulos, D.; Athitsos, V.; Makedon, F. Non-invasive analysis of sleep patterns via multimodal sensor input. *Pers. Ubiquitous Comput.* **2014**, *18*, 19–26.
7. Diykh, M.; Li, Y.; Wen, P. EEG Sleep Stages Classification Based on Time Domain Features and Structural Graph Similarity. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2016**, *24*, 1159–1168.
8. Fraiwan, L.; Lweesy, K.; Khasawneh, N.; Fraiwan, M.; Wenz, H.; Dickhaus, H. Classification of Sleep Stages Using Multi-wavelet Time Frequency Entropy and LDA. *Methods Inf. Med.* **2010**, *49*, 230–237.
9. Lajnef, T.; Chaibi, S.; Ruby, P.; Aguera, P.; Eichenlaub, J.; Samet, M.; Kachouri, A.; Jerbi, K. Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines. *J. Neurosci. Methods* **2015**, *250*, 94–105.
10. Fonseca, P.; Long, X.; Radha, M.; Haakma, R.; Aarts, R.; Rolink, J. Sleep stage classification with ECG and respiratory effort. *Physiol. Meas.* **2015**, *36*, 2027–2040.
11. Rodríguez-Sotelo, J.; Osorio-Forero, A.; Jiménez-Rodríguez, A.; Cuesta-Frau, D.; Cirugeda-Roldán, E.; Peluffo, D. Automatic Sleep Stages Classification Using EEG Entropy Features and Unsupervised Pattern Analysis Techniques. *Entropy* **2014**, *16*, 6573–6589.
12. Sharma, R.; Pachori, R.; Upadhyay, A. Automatic sleep stages classification based on iterative filtering of electroencephalogram signals. *Neural Comput. Appl.* **2017**, *28*, 2959–2978.
13. Huang, C.S.; Lin, C.L.; Ko, L.W.; Liu, S.Y.; Sua, T.P.; Lin, C.T. A hierarchical classification system for sleep stage scoring via forehead EEG signals. In Proceedings of the IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), Singapore, 16–19 April 2013 ; pp. 1–5.
14. Zhu, G.; Li, Y.; Wen, P.P. Analysis and Classification of Sleep Stages Based on Difference Visibility Graphs From a Single-Channel EEG Signal. *IEEE J. Biomed. Health Inform.* **2014**, *18*, 1813–1821.
15. Pan, S.-T.; Kuo, C.-E.; Zeng, J.-H.; Liang, S.-F. A transition-constrained discrete hidden Markov model for automatic sleep staging. *Biomed. Eng. Online* **2012**, *11*, 52.
16. Song, C.; Liu, K.; Zhang, X.; Chen, L.; Xian, X. An Obstructive Sleep Apnea Detection Approach Using a Discriminative Hidden Markov Model From ECG Signals. *IEEE Trans. Biomed. Eng.* **2016**, *63*, 1532–1542.
17. Li, K.; Pan, W.; Li, Y.; Jiang, Q.; Liu, G. A method to detect sleep apnea based on deep neural network and hidden Markov model using single-lead ECG signal. *Neurocomputing* **2018**, *294*, 94–101.
18. Sors, A.; Bonnet, S.; Mirek, S.; Vercueil, L.; Payen, J.-F. A convolutional neural network for sleep stage scoring from raw single-channel EEG. *Biomed. Signal Process. Control* **2018**, *42*, 107–114.
19. Chambon, S.; Galtier, M.N.; Arnal, P.J.; Wainrib, G.; Gramfort, A. A Deep Learning Architecture for Temporal Sleep Stage Classification Using Multivariate and Multimodal Time Series. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2018**, *26*, 758–769.
20. Boostani, R.; Karimzadeh, F.; Nami, M. A comparative review on sleep stage classification methods in patients and healthy individuals. *Comput. Methods Programs Biomed.* **2017**, *140*, 71–91.
21. Aboalayon, K.; Faezipour, M.; Almuhammadi, W.; Saeid Moslehpour, S. Sleep Stage Classification Using EEG Signal Analysis: A Comprehensive Survey and New Investigation. *Entropy* **2016**, *18*, 272.
22. Peker, M. An efficient sleep scoring system based on EEG signal using complex-valued machine learning algorithms. *Neurocomputing* **2016**, *207*, 165–177.
23. Krefting, D.; Jansen, C.; Penzel, T.; Han, F.; Kantelhardt, J.W. Age and gender dependency of physiological networks in sleep. *Physiol. Meas.* **2017**, *38*, 959–975.

24. Lee, J.; Yoo, S. Electroencephalography Analysis Using Neural Network and Support Vector Machine during Sleep. *Engineering* **2013**, *5*, 88–92.
25. Ozsen, S. Classification of sleep stages using class-dependent sequential feature selection and artificial neural network. *Neural Comput. Appl.* **2013**, *23*, 1239–1250.
26. Langkvist, A.; Karlsson, A.; Loutfi, A. Sleep Stage Classification Using Unsupervised Feature Learning. *Adv. Artif. Neural Syst.* **2012**, *12*, 1–9.
27. Hsu, Y.L.; Yang, Y.T.; Wang, J.S.; Hsu, C.Y. Automatic sleep stage recurrent neural classifier using energy features of EEG signals. *Neurocomputing* **2013**, *104*, 105–114.
28. Hashizaki, M.; Nakajima, H.; Kume, K. Monitoring of Weekly Sleep Pattern Variations at Home with a Contactless Biomotion Sensor. *Sensors* **2014**, *15*, 18950–18964.
29. Sen, B.; Peker, M.; Cavusoglu, A.; Celebi, F. A Comparative Study on Classification of Sleep Stage Based on EEG Signals Using Feature Selection and Classification Algorithms. *J. Med. Syst.* **2014**, *38*, 1–21.
30. Bušková, J.; Ibarburu, V.; Šonka, K.; Ružička, E. Screening for REM sleep behavior disorder in the general population. *Sleep Med.* **2016**, *24*, 147–147.
31. Tagluk, E.; Sezgin, N.; Akin, M. Estimation of Sleep Stages by an Artificial Neural Network Employing EEG, EMG and EOG. *J. Med. Syst.* **2010**, *34*, 717–725.
32. Kouchaki, K.S.; Eftaxias, K.; Sanei, S. An Adaptive Filtering Approach Using Supervised SSA for Identification of Sleep Stages From EEG. *Front. Biomed. Technol.* **2014**, *1*, 233–239.
33. Venkatesh, K.; Poonguzhali, S.; Mohanavelu, K.; Adalarasu, K. Sleep Stages Classification Using Neural Network with Single Channel EEG. *Int. J. Recent Adv. Eng. Technol.* **2014**, *2*, 1–8.
34. Ahuja, Y.; Yadav, S. Multiclass classification and support vector machine. *Glob. J. Comput. Sci. Technol.* **2012**, *12*, 15–19.
35. Ou, G.; Murphey, Y. Multi-class pattern classification using neural networks. *Pattern Recognit.* **2007**, *40*, 4–8.
36. Procházka, A.; Vyšata, O.; Ťupa, O.; Mareš, J.; Vališ, M. Discrimination of Axonal Neuropathy Using Sensitivity and Specificity Statistical Measures. *Neural Comput. Appl.* **2014**, *25*, 1349–1358.
37. Procházka, A.; Kuchyňka, J.; Suarez Araujo, C.P.; Vyšata, O. Adaptive Segmentation of Multimodal Polysomnography Data for Sleep Stages Detection. In Proceedings of the 22nd International Conference on Digital Signal Processing, London, UK, 23–25 August 2017; pp. 1–4.
38. Dursun, M.; Ozsen, S.; Yucelbas, C.; Yucelbas, S.; Tezel, G.; Kuccukturk, S.; Yosunkaya, S. A new approach to eliminating EOG artifacts from the sleep EEG signals for the automatic sleep stage classification. *Neural Comput. Appl.* **2017**, *28*, 3095–3112.
39. Saeid, S. *Adaptive Processing of Brain Signals*; John Wiley & Sons Ltd.: Chichester, UK, 2013.
40. Saeid, S.; Chambers, J. *EEG Signal Processing*; John Wiley & Sons Ltd.: Chichester, UK, 2007.
41. Chaparro-Vargas, R.; Cvetkovic, D. A single-trial toolbox for advanced sleep polysomnographic preprocessing. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Osaka, Japan, 3–7 July 2013; pp. 5829–5832.
42. Mares, J.; Vysata, O.; Prochazka, A.; Valis, M. Age-Dependent Complex Noise Fluctuation in the Brain. *Physiol. Meas.* **2013**, *34*, 1269–1279.
43. Procházka, A.; Vyšata, O.; Vališ, M.; Ťupa, O.; Schatz, M.; Mařík, V. Bayesian classification and analysis of gait disorders using image and depth sensors of Microsoft Kinect. *Digit. Signal Process.* **2015**, *47*, 169–177.
44. Everson, R.; Fieldsend, J. Multi-class ROC analysis from a multi-objective optimisation perspective. *Pattern Recognit. Lett.* **2006**, *27*, 918–927.
45. Carlos, J.; Caballero, F.; Martinez, F.; Hervas, C.; Gutierrez, P. Sensitivity versus Accuracy in Multiclass Problems Using Memetic Pareto Evolutionary Neural Network. *IEEE Trans. Neural Netw.* **2010**, *21*, 750–770.
46. Landgrebe, T.; Duin, R. Approximating the multiclass ROS by pairwise analysis. *Pattern Recognit. Lett.* **2007**, *28*, 1747–1758.
47. Espiritu, J. Aging-Related Sleep Changes. *Clin. Geriatr. Med.* **2008**, *24*, 1–14.

