Automatic Sleep Staging Employing Convolutional Neural Networks and Cortical Connectivity Images

Panteleimon Chriskos, Christos A. Frantzidis, Polyxeni T. Gkivogkli, Panagiotis D. Bamidis, and Chrysoula Kourtidou-Papadeli

***Abstract*— Understanding of the neuroscientific sleep mechanisms is associated with mental/cognitive and physical well-being and pathological conditions. A prerequisite for further analysis is the identification of the sleep macroarchitecture through manual sleep staging. Several computer-based approaches have been proposed to extract time and/or frequency-domain features with accuracy ranging from 80% to 95% compared with the golden standard of manual staging. However, their acceptability by the medical community is still suboptimal. Recently, utilizing deep learning methodologies increased the research interest in computer-assisted recognition of sleep stages. Aiming to enhance the arsenal of automatic sleep staging, we propose a novel classification framework based on convolutional neural networks. These receive as input synchronizations features derived from cortical interactions within various electroencephalographic rhythms (delta, theta, alpha, and beta) for specific cortical regions which are critical for the sleep deepening. These functional connectivity metrics are then processed as multidimensional images. We also propose to augment the small portion of sleep onset (N1 stage) through the Synthetic Minority Oversampling Technique in order to deal with the great difference in its duration when compared with the remaining sleep stages. Our results (99.85%) indicate the flexibility of deep learning techniques to learn sleep-related neurophysiological patterns.**

***Index Terms*— Automatic sleep staging, convolutional neural networks (CNNs), default mode network (DMN), functional connectivity features, minority class oversampling technique.**

I. Introduction

ONE of the most vital biological processes that is necessary for mental and physical health is sleep [1]. The duration and quality of sleep are associated with cardiovasular diseases, diabetes, cognitive decline, and loss of everyday life quality [2]. Therefore, the understanding and analysis of sleep functions may provide valuable insight to personal health.

Manuscript received September 25, 2018; revised December 20, 2018 and January 8, 2019; accepted February 12, 2019. Date of publication March 15, 2019; date of current version January 3, 2020. This work was supported in part by the European Union's Horizon 2020 Research and Innovation Programme under Grant 681120. *(Corresponding author: Panteleimon Chriskos.)*

P. Chriskos is with the Laboratory of Medical Physics, Medical School, Aristotle University of Thessaloniki, Thessaloniki 541 24, Greece (e-mail: pchriskos@gmail.com).

C. A. Frantzidis, P. T. Gkivogkli, P. D. Bamidis, and C. Kourtidou-Papadeli are with the Laboratory of Medical Physics, Medical School, Aristotle University of Thessaloniki, Thessaloniki 541 24, Greece, and also with the Greek Aerospace Medical Association-Space Research, Kalamaria 551 32, Greece.

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNNLS.2019.2899781

According to the American Sleep Medical Association, American Academy of Sleep Medicine (AASM) [3], [4], sleep is divided into four distinct stages, each of which is characterized by significant fluctuations in brain activity. The sleep stages defined are N1, N2, N3, and rapid eye movement (REM). These stages are grouped into Non REM (NREM), encompassing the N1, N2, and N3 sleep stages and REM sleep. The first step in automatic sleep staging is to extract features that can be subsequently used for analysis. The duration and characteristics of each sleep stage can be used to provide an overview of the sleeping human brain [5]. This process is time-consuming, potentially error-prone, and tedious usually carried out by experienced sleep experts. Aiming to reduce the amount of time required and the number of potential errors, several methodologies have been proposed toward the accurate automatic sleep staging [6]–[8].

Prior studies in the field of automatic sleep staging rely on one or more feature extraction methods and domains, leading to an extensive variety of features used in automatic sleep staging applications. For example, temporal features (time domain) have been utilized in [9], spectral features were extracted in [10], while time-frequency features were used by Hassan and Bhuiyan [11]. Energy- and entropy-related features have been proposed in [12]. In many instances, features were combined from various domains aiming to provide higher utility feature sets as in [13] and [14].

Apart from extracting features from various domains, it is also possible to utilize multiple biological signals for automatic sleep staging purposes, which is advantageous as sleep is a biological process affects multiple body systems. In some cases, sleep staging was based solely on electroencephalographic (EEG) recordings as in [7] and [15]. Other authors used a combination of EEG and electooculographic (EOG) signals [16], whereas Liang *et al.* [17] used EEG and electromyographic (EMG) signals, while more comprehensive approaches presented in [9] and [14] relied on EEG, EOG, and EMG signals.

The accuracy percentage of sleep staging methodologies found in the recent bibliography range between 62% and 99%, rounded to the nearest whole percentage for simplicity. It must be noted, however, that the 99% rate presented in [18] used as ground truth labels that were extracted by an automatic sleep staging algorithm, not labels produced through expert annotation. The highest accuracy attained with expert scored sleep data was 97% in [7]. The classifiers used were usually

supervised in nature, such as the k-nearest neighbor (KNN) classifier used in [19] achieving an accuracy equal to 89%, decision trees [7], [13] achieving a maximum accuracy equal to 97%, and support vector machines [14], [20] reaching 94% accuracy. In the majority of the aforementioned approaches, the features extracted did not retain spatial information regarding the brain's activity, information that may allow more accurate sleep staging.

In order to utilize spatial information, the classifiers evaluated in this paper are convolutional neural networks (CNNs) that are commonly used in image applications such as image classification [21], [22], object detection [23], [24], and image caption generation [25], [26]. In general, CNNs use orthogonal filters specific to the task that should be learned and are therefore more suited for image-based recognition and not for identification of sleep stage patterns. However, a pioneering work from Tsinalis *et al.* [8] employed CNN for automatic sleep staging from single-channel EEG recordings with the first convolutional layer comprised of 1-D filters. The maximum accuracy achieved was equal to 74% possibly limited by the limited number of sensors used and fact that the polysomnographic (PSG) series were directly used as input. A similar approach with a single EEG channel as input was presented in [27], in which an accuracy of 82% was attained. The DeepSleepNet was introduced in [27], which employs raw single-channel EEG sensors to extract time-invariant features focusing on learning transition patterns across different sleep stages. A main advantage of the proposed model is its robustness to different data sets with varying recording properties (e.g., sampling rate) and even heterogeneity of the adopting scoring standard. However, the model could not outperform state-of-the-art methodologies. The most promising approach is that of Chambon *et al.* [28] who proposed a novel multivariate and multimodal approach. They used time series from EEG, EOG, and EMG and provided a novel framework of low computational cost, which exploits the temporal neurophysiological interactions on sensor level. The classification performance they achieved was at 80%, which is acceptable but not surpassing state-of-the art methodologies.

Contrary to the aforementioned works, where the PSG signals were used as input, in our approach, we transform the problem of classifying sleep-related biosingals (time series) into classification task. This, however, could not be achieved on the data level or even employing traditional feature extraction methods. In a previous study, we employed synchronization-based features for performing sleep stage classification [6]. These features quantified the cooperative activity among electrode sensors and, therefore, were represented as a 2-D matrix, which could easily be visualized as an image. It also provided a novel approach of regarding the sleeping brain as an integrative system, which could provide a macroscopic view of sleep physiology and how the brain cooperative activity evolves during the sleep evolution. Despite its novelty, there were also drawbacks [6]. More specifically, the image representation of the cooperative activity during sleep contained redundant information since the functional connectivity estimation as estimated through synchronization likelihood (SL) is not directional. Consequently, the images were symmetrical with regards to their main diagonal. This could be solved either by introducing a directed connectivity metric or by merging two symmetrical matrices.

In this paper, we transform the 19 channel EEG recordings, recorded according to the 10–20 International System, to 18 brain region activations, through an appropriate transformation matrix (inverse solution problem) for cortical activity estimation. These 18 regions are part of the human brain's default mode network (DMN), which has been observed to vary during the sleep-wake cycle [29]. The functional connectivity features extracted from these activations, namely, SL [30] and relative wavelet entropy (RWE) [31] values, along with graph metrics and energy ratio features were treated as image pixels, whose values represent the interactions between human brain regions during sleep and, at the same time, retaining spatial information related to brain region interaction. Our approach allowed the use of 3-D inputs to the CNN, similar in structure to RGB images, which is expected to provide with higher accuracy rates than prior works. Looking back on prior approaches, that fed CNNs with time series as input, it was not surprising that relatively low accuracy rates were achieved compared to other applications of CNNs. This indicates that the proper training of CNNs, with appropriately structured inputs, allows these versatile learners to achieve their maximum potential.

An issue that arises when using CNNs is that underrepresented classes tend to be ignored during training. This is true in the case of N1 sleep stages since the total duration of N1 is [6] approximately 5% of the total sleep time. In order to overcome this limitation, synthetic N1 samples were generated using the Synthetic Minority Oversampling Technique (SMOTE).

Our use of deep learning methods is one of many applications that these methodologies have provided results surpassing other state-of-the-art methods, including face recognition [21], object detection [24], thermal analysis of heat exchanges [32], and the real-time control and synchronization of the Tokamak device [33], [34].

The rest of this paper is organized as follows. Section II presents the data preprocessing steps that were applied on the raw EEG data, along with a description of the process behind calculating the transformation matrix that allows the calculation of the DMN brain region activations. The two feature extraction methodologies SL and RWE are also presented in this section along with the graph metrics calculated on the previous values and the technique used to augment the N1 training set through SMOTE. Section III describes the raw data that were used in this paper and how the feature values were transformed into images acceptable as CNN input along with the CNN architectures that were evaluated. In Section IV, the experimental results are presented, while final conclusions and the discussion are included in Section V.

II. Methodology

In this section, we describe the methodology that was followed during the preparation of the initial PSG data, the transformation of sensorial data into the cortical activations of the DMN, the feature extraction approaches used, as well

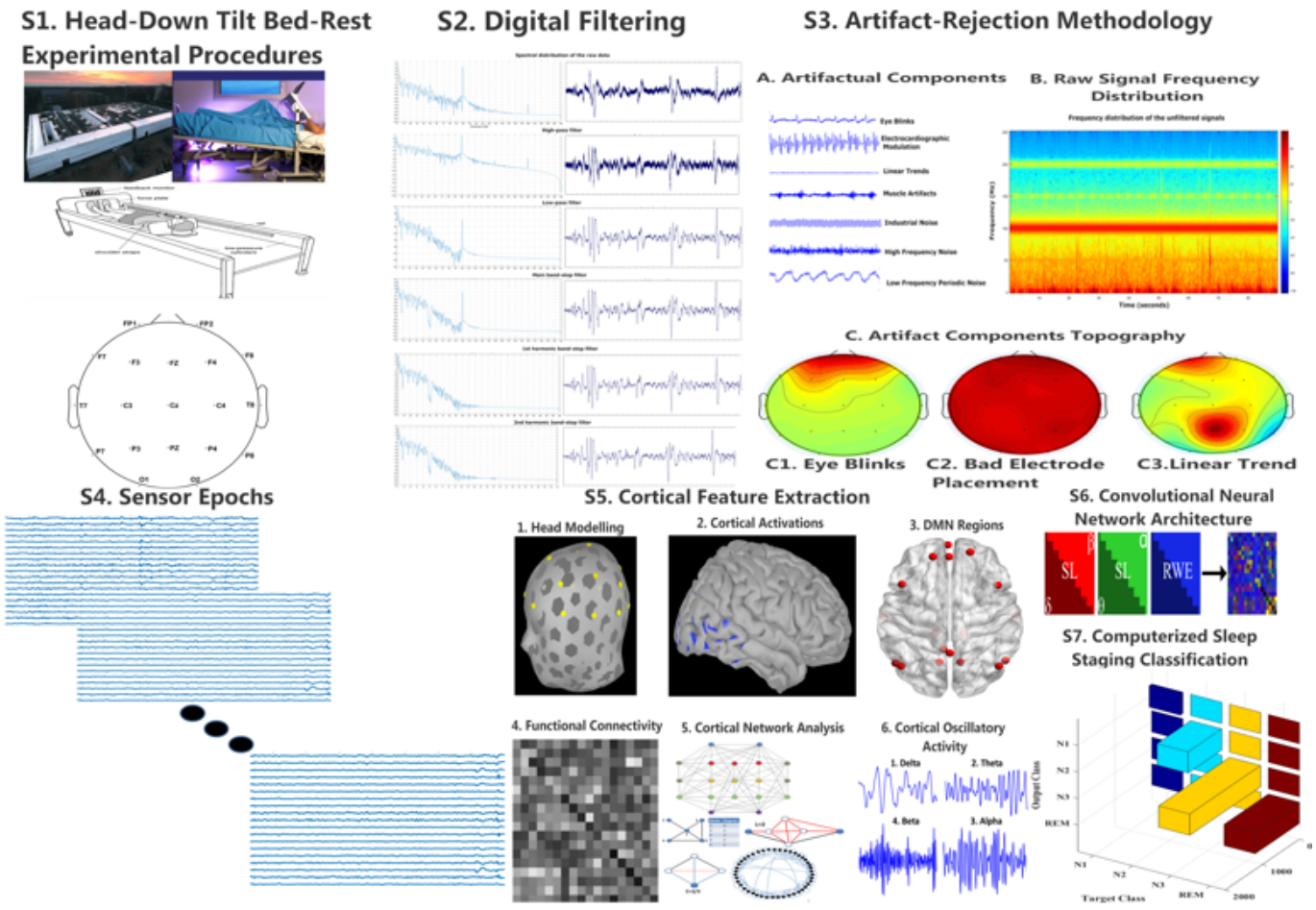

Fig. 1. Overview of the proposed methodology.

as the steps taken in order to increase the number of available N1 sleep stages aiming to maximize the classification accuracy. The methodology is outlined in Fig. 1.

A. Electroencephalographic Signal Preprocessing

In order to achieve automatic sleep staging, features were extracted from EEG recordings [Fig. 1(1)]. Before feature extraction, it was necessary to preprocess the EEG signal in order to eliminate content irrelevant to brain activity. This was achieved through a multistep process as detailed in [35]. In short, the initial EEG recordings were separated into five segments from which the mean value of each signal was subtracted. Afterward, the signals were filtered with third-order Butterworth filters in order to remove unwanted spectral content such as power line interference, noise due to body movements, and noise from other unknown sources [Fig. 1(2)]. After filtering independent component analysis (ICA) was performed as described in [36], and a subset of the resulting components was used to reconstruct the EEG signal, in order to remove noise components related to body movements, interference from other biological signals such as EOG, EMG, and electrocardiographic (ECG) modulation [Fig. 1(3)]. Finally, the EEG signal was split into 30-sec epochs according to the guidelines given by the AASM [3] [Fig. 1(4)]. From each epoch, features were extracted through SL [30] and RWE [31], which are briefly described below.

B. Estimation of Default Mode Network Cortical Activity

The brainstorm interface [37] ran under MATLAB environment was used for estimating the entire cortical activity through the solution of the inverse problem [Fig. 1(5)]. First, we modeled the generic head anatomy through the open M/EEG boundary elements methodology and the use of the magnetic resonance imaging (MRI) volume (Colin 27 stereotaxic T1-weighted MRI volume). This is a commonly used methodology when there are no individual MRI scans for each participant. Therefore, the entire cortex was modeled through 15 000 fixed dipoles. The cortical activity of each dipole was then estimated through the standardized low resolution electromagnetic tomography methodology as implemented in Brainstorm [38]. It estimates through a linear solution of the origin of EEG activity due to the postsynaptic activations of cortical neuronal assemblies. Then, we employed *a priori* knowledge derived from more localization precise neuroimaging techniques (fMRI) in order to define the DMN nodes according to Table I.

C. Synchronization Likelihood From Electroencephalographic Signals

SL [30] [Fig. 1(6)] quantifies the degree of synchronization between a number of time series, in this case, the channels of the EEG signal. The SL for a channel k at time i with respect

TABLE I
DEFINITION OF DMN NODES ACCORDING TO THEIR MONTREAL NEUROLOGICAL INSTITUTE COORDINATES

Region	Label	X	Y	Z
Left Posterior Inferior Parietal Lobe	LpIPL	-51	-67	26
Right Posterior Inferior Parietal Lobe	RpIPL	52	-64	27
Orbitofrontal Cortex	OFC/vACC	-2	62	-17
Brodmann Area 8	dMPFC BA8	-16	49	43
Brodmann Area 9	dMPFC BA9	19	54	36
Left Dorsolateral Prefrontal Cortex L	DLPFC	-45	19	44
Right Dorsolateral Prefrontal Cortex	R DLPFC	45	18	43
Left Parahippocampal Gyrus	L PHG	-13	-35	-5
Right Parahippocampal Gyrus	R PHG	12	-35	-5
Left Inferior Temporal Cortex	L ITC	-62	-16	-20
Right Inferior Temporal Cortex	R ITC	62	-15	-22
Medial Prefrontal Cortex	MPFC	-1	49	-5
Posterior Cingulate Cortex	PCC	-6	-52	-40
Precuneus	PRE	0	-56	28
Left Precuneus/Posterior Cingulate Cortex	3IL	-10	-66	24
Right Precuneus/Posterior Cingulate Cortex	3IR	10	-66	24
Left Lateral Parietal	LLP	-46	-70	36
Right Lateral Parietal	RLP	46	-70	36

to all other channels is calculated as

$$S_{k,i} = \frac{1}{2(w_2 - w_1)} \sum_{\substack{j=1 \\ w_1 < |i-j| < w_2}}^{N} S_{k,i,j} \tag{1}$$

where w_1 is the Theiler correction for autocorrelation and w_2 is the window for time resolution sharpening as described in [39], while $S_(k, i, j)$ is the SL for each channel k at time pair (i, j) calculated through

$$S_{k,i,j} = \begin{cases} \dfrac{V_{i,j} - 1}{N - 1} & \text{if } |\mathbf{x}_{k,i} - \mathbf{x}_{k,j}| < \epsilon_{k,i} \\ 0 & \text{if } |\mathbf{x}_{k,i} - \mathbf{x}_{k,j}| \geq \epsilon_{k,i} \end{cases} \tag{2}$$

where $V_{i,j}$ denotes the number of channels in the current window that is closer than distance $\epsilon_{k,i} = 10^{-5}$ from each other, N is the number of channels in the signal, and $\mathbf{x}_{k,i}$ is the time-delay embedded vector of signal $\mathbf{x}_k$. The SL values are stored in an $N \times N$ matrix $\mathbf{S}$ which is symmetric with reference to the main diagonal.

SL values were calculated on the signals that resulted after bandpass filtering the signal in four EEG rhythm ranges, namely, δ (0.5–4 Hz), θ (4–8 Hz), α (8–12 Hz), and β (12–20 Hz). This filtering allowed the calculation of the interactions between pairs of cortical regions in each band with more accuracy, providing a more comprehensive representation of the human brain during sleep. However, filtering the signals resulted in artifacts during the first few samples of the recording, and therefore, the first 0.7 sec (350 samples) was removed from all filtered EEG band recordings, which was visually confirmed.

D. Relative Wavelet Entropy Calculated on Electroencephalographic Signals

RWE [31], Fig. 1(6), values quantify the extent to which the recordings from each pair of electrodes are similar to one another based on the orthogonal discrete wavelet transform (ODWT), [40]. Given a uniformly sampled signal $\mathbf{x}$, the ODWT, X_t, is calculated as

$$X_t = \sum_{j=-M}^{-1} \sum_{l} C_j(l) \psi_{j,l}(t) \tag{3}$$

where M denotes the decomposition levels of the ODWT and $C_j(l)$ denotes the wavelet coefficients at times l. The types of mother wavelets used are orthogonal cubic spline functions. The number of decomposition levels is equal to five, one for each of δ, θ, α, β, and γ (20–50 Hz) brain rhythms. Coefficient amplitude quantifies the similarity of signal $\mathbf{x}$ compared to the mother wavelet and its sign describes the polarity of said similarity. The total energy, E_{tot}, of the signal is given by

$$E_{\text{tot}} = \sum_{j<0} \sum_{l} |C_j(l)|^2 \tag{4}$$

which is required in order to calculate the relative energies p_j for each level. These ratios are equal to the ratio of the energy at level j over the total energy. The relative entropy between each pair of electrodes is given by the Shannon entropy [41]

$$H(p|q) = -\sum_{j<0} p_j \ln\left(\frac{p_j}{q_j}\right). \tag{5}$$

RWE values are stored in a $N \times N$ matrix $\mathbf{T}$ which is not symmetric with reference to the main diagonal contrary to the case of SL. Apart from the RWE values, the energy ratios of the five EEG brain wave rhythms were also calculated, namely, the energies of δ, θ, α, β, and γ bands over the total energy of the signal.

E. Graph Metrics

The aforementioned matrices $\mathbf{S}$ and $\mathbf{T}$ can be regarded as graph adjacency matrices allowing the calculation of graph metrics which describe the overall connectivity strength between electrode pairs, effectively providing an overview of the graph's organization and structure. The graph metrics calculated [40], [42] are the mean clustering coefficient which quantifies the intensity of graph node connectivity, the characteristic path efficiency and length, the characteristic path being the shortest path that connects all nodes in the graph with all other nodes of the same graph. Other metrics include the graph density, which is the ratio of total connections present in the graph over the total possible connections in the same graph, the mean betweenness centrality, i.e., the number of shortest paths connecting all other pairs of nodes that incorporate the given node, as well as the betweenness centrality of regions posterior cingulate cortex (PCC), orbitofrontal cortex (OFC), and Precuneus (PRE), which are the key regions of the DMN. The last graph metric calculated is the small world property of the graph which quantifies the degree of optimal network organization [43].

F. Augmenting the N1 Training Set Using the Synthetic Minority Oversampling Technique

A common issue that arises in automatic sleep staging is that the number of epochs per sleep stage present in the data set varies significantly due to the physiological differences

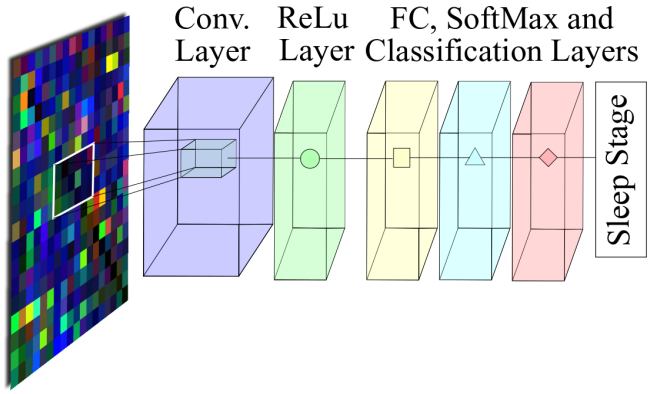

Fig. 2. Diagram of the CNN architecture that provided the highest classification accuracy.

in the duration of each sleep stage. This introduces a challenge, especially in the case of N1 sleep which only lasts approximately 5% of the total sleep duration and is oftentimes polluted by noise due to body movements. Therefore, in order to oversample the N1 minority class, the Synthetic Minority Oversampling Technique (SMOTE) [44] was used. SMOTE augments the underrepresented class with synthetic samples allowing the trained classifiers to achieve accuracies on minority classes comparable to those attained on the other classes in the data set. In short, new samples of the minority class, in this case features extracted from N1 sleep stages, are constructed by combining two samples of the initial data set in order to create a new synthetic sample. As a first step, for each given training sample, s_T, the KNNs are found. In our application of SMOTE, K was equal to 100 and the distance metric used was the squared Euclidean distance. Afterward, a random neighbor, s_R, of the K nearest is selected to be used in constructing the new synthetic sample, s_S, which is given by

$$s_S = s_T + u * (s_T - s_R), u \in U(0, 1) \tag{6}$$

where $U(0, 1)$ denotes the uniform distribution in the interval [0, 1]. When applying SMOTE on the N1 sleep stage features, each training sample was used 10 times, resulting in a training data set containing 11 times the initial number of training samples. It must be noted that the use of SMOTE may provide synthetic training samples that enable the trained classifiers to achieve higher or lower classification accuracies, dependent on the random number generator. Therefore, multiple synthetic training data sets were constructed and multiple classifiers trained, as is described in a later section.

G. Convolutional Network Architectures Evaluated

A total of six different CNN architectures were evaluated for automatic sleep staging using the inputs described in Sections II-C–II-E. The architecture of the CNN that achieved the highest accuracy is presented in Fig. 2. In half of the CNNs trained, the first layer after the input layer was a convolutional layer [45] with square filters of size 3, 6, and 8, followed by a fully connected layer of size 4, equal to the number of classes. Afterward, a softmax layer was utilized calculating the softmax or normalized exponential function on each element k of the input vector $\mathbf{x}$ defined as: $\text{softmax}(\mathbf{x}_k) = \exp \mathbf{x}_k / 1 + \sum_j \exp \mathbf{x}_j$ [45]. The final layer in the architecture is a classification layer with cross entropy loss function [45] defined as $H_C(y, p) = -\sum_i y_i \log(\text{Softmax}(\mathbf{x}_k))$ where y_i is the label of the input. These architectures are referred to as C3, C6, and C8 depending on the size of the convolutional layers' filters. The rest of the networks were similar in structure but after the convolutional layer, a rectified linear unit (ReLu) layer [46] was added implementing the function $\max(x, 0)$, where x is the input value. These networks are referred to as C3R, C6R, and C8R. The total number of layers in the architectures was either 4 or 5, depending on whether a ReLu layer was present or not. Deeper architectures were also tested but did not provide higher accuracy rates, perhaps due to the small size of the input matrices. The sleep stage assigned to each input epoch was selected based on which output neuron had the maximum activation value.

No prior work was found regarding the training of CNNs on synchronization features derived from EEG recordings for automatic sleep staging. Therefore, the CNN architectures, the number of filters employed in the convolutional layer, 150 were used in all networks, as well as the values of the training parameters, learning rate equal to 0.005 and momentum equal to 0.7, were selected empirically based on the accuracy percentage attained on the test set. It must be noted that the above-mentioned architectures and parameters were optimized for the available train and test sets and, thus, may require calibration for other sleep data sets. All CNNs were trained using stochastic gradient descent with momentum [45].

The objective of our automatic sleep staging approach was, apart from achieving an overall high classification accuracy, to facilitate high recall rates for all the sleep stages. This was especially important in the case of N1 sleep epochs which were available in a limited number, and thus required the utilization of SMOTE. However, SMOTE may provide varying results depending on the random number generator. In order to verify that our results were not due to pure chance, 15 different training sets were created through SMOTE and used to train 15 different CNNs for each of the six architectures mentioned above. As such in Section IV, the minimum, maximum, and median accuracy values are presented for the test set data. Training multiple networks were possible first due to the relatively small number of available training data compared to the databases used in [47] and [48] which contain millions of images. Furthermore, the shallow CNN architectures used, once more compared to the networks described in [21] and [22], allowed multiple networks to be trained in an acceptable time period.

III. Experimental Setup

In this section, we describe the initial data and how they were recorded along with the construction of the appropriate data structures that allowed the training of the classifiers.

A. Data Description

The data used in this paper are a subset of the experimental data recorded during the experiment conducted in the ENVIHAB premises of the German Aerospace Agency, funded by

the European Space Agency as described in [49]. During this experiment, a microgravity environment was simulated for a total of 60 days. Throughout the 60 days, the participants followed a head down tilt bed rest protocol. The participants were 23 healthy male adults between the ages of 23 and 45 (mean: 29 $\pm$ 6 years). The purpose of the experiment was to assess the effect of simulated microgravity on the human body including but not limited to cardiac function, hormone levels, muscle activity, body mass, and body composition. The participants were screened in terms of their medical and psychological status and then assigned either to a control or to a training group which practiced through a reactive sledge jumps (RSL) device in order to minimize the effect of microgravity on human physiology. All of them were informed for the RSL study and then they signed a written informed consent form [49]. The ethics committee which approved the study was that of the Northern Rhine Medical Association (Arztekammer Nordrhein) in Duesseldorf, Germany Strahlenschutz. During the same study, the Greek research team (under the leadership of Greek Aerospace Medical Association and the Laboratory of Medical Physics, Medical School of the Aristotle University of Thessaloniki) conducted several entire night PSG recordings in order to study the effect of microgravity on sleep quality as discussed in [50]. PSG recordings took place 2 weeks before the bed rest initialization (BDC-14 phase), 21, 35, and 50 days after the experiment began (HDT-21,35,50 phases) and finally 7 days after the end of the bed rest period (recovery phase). The acquired signals included EEG recordings from 19 Ag/AgCl electrodes positioned according to the International 10–20 System as well as ECG, chin EMG, and vertical and horizontal electrooculogrammic signals. In this paper, automatic sleep staging was achieved by extracting features only from EEG recordings acquired during the BDC-14 phase of the experiments, in order to avoid EEG fluctuations due to the head-down tilt protocol.

From the initial data set, a total of 9766 epochs were manually selected after preprocessing took place, in which noise sources did not distort the signal beyond utility. This set of epochs was divided into training and test sets each containing 5674 and 4092 epochs, respectively. From a total of 22 participants, the PSG recordings of 14 were used in the training set and the recording of the remaining eight was used for testing. After the application of SMOTE, the number of epochs in the training set was equal to 10976. The number of available epochs per training and test sets and sleep stage prior and after applying SMOTE is presented in Table II. From this table, it is evident that the small number of N1 samples initially available may compromise the ability of the trained networks to accurately classify N1 epochs.

B. Feature Data Structure

In order to achieve automatic sleep staging, all the features described above were fused into a single data structure for each sleep epoch. As already stated, SL matrices contain redundant information since these matrices are symmetrical with reference to the main diagonal. Therefore, only the lower or upper triangular of said matrices can be kept without any loss of information. Furthermore, the values of the main diagonal are also redundant since they indicate that each channel is synchronous to itself and, hence, were omitted from the analysis. The same applies to the RWE matrices with regards to the values of the main diagonal, since RWE matrices are not symmetrical. All features extracted from the EEG recordings were normalized, per feature, in the interval [0, 1].

TABLE II
NUMBER OF SLEEP STAGES PER SET AND SLEEP STAGE

Set	Sleep Stage				
	N1	N2	N3	REM	Total
Train	121	1344	2607	1602	5674
Train + SMOTE	1331	1344	2607	1602	6884
Test	77	908	1912	1195	4092

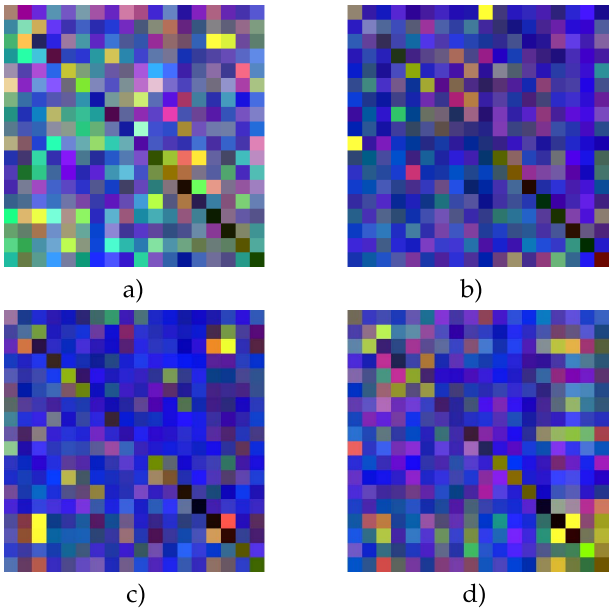

Fig. 3. Sample images constructed from the three synchonization matrix values and metrics for each sleep stage. (a) N1. (b) N2. (c) N3. (d) REM.

In our approach, each of the three synchronization matrices was treated as an image channel resulting in a 3-D matrix of $18 \times 18 \times 3$ dimensions, representing an RGB image of size 18×18. Examples of the resulting images are presented in Fig. 3 for each sleep stage. From this figure, it is possible to pinpoint several differences between the four sleep stages.

The values contained in each matrix were organized as follows. The strictly lower triangular portion of the first matrix (red channel) contained the SL values calculated for the δ EEG band, while the strictly upper triangular was populated with the β EEG band SL values. The diagonal of this matrix contained the graph metrics calculated on the δ and β SL values. The second matrix (green channel) is constructed in a similar fashion populated with the SL values and metrics of the θ and α EEG bands. The final matrix (blue channel) contained the RWE values between DMN region pairs, while the main diagonal was populated with the graph metric values calculated on the RWE values and the five EEG band energy ratios. This approach is summarized in Fig. 4.

Fig. 4. RGB image formation using the three synchronization, graph metric, and energy ratio values. Red: δ and β SL. Green: θ and α SL. Blue: whole signal RWE.

In the case of the SL matrices, the diagonals are populated in value pairs. The first of each pair corresponds to the graph metric of the strictly lower triangular matrix and the second one to the strictly upper triangular matrix. The graph metric pairs were ordered as follows: small world property, mean clustering coefficient, characteristic path length and efficiency, graph density, mean betweenness centrality, and PCC, OFC, and PRE betweenness centralities. The ordering of the graph metrics in the diagonals was arbitrary. In the RWE matrix, the values in the diagonal were the energy ratios of five EEG bands followed by the nine graph metrics as mentioned in the SL matrices. Any remaining diagonal cells were filled with zeros.

In this paper, in order to facilitate the use of CNNs, we first performed digital filtering to extract oscillatory activity on delta, theta, alpha, and beta frequency bands. The logic behind this approach was that the transition to sleep from wakefulness is accompanied by gradual loss of beta and subsequent increase of alpha activity. Further increase in the amplitude of low-frequency activity (δ and θ) enables the transition to deeper sleep during the NREM period. During the REM period, an increase of alpha and beta waves is observed. Thus, we hypothesized that quantifying the cooperation degree of the aforementioned brain rhythms would increase the accuracy of the proposed methodology. This led us to adopt the data structure described above to train the CNNs. The underlying methodology provides a three-layer transformation of oscillatory time series into connectivity images representing distinct sleep stages and physiological properties.

IV. Results

The overall test set classification accuracies are presented in Table III. From the accuracy values presented, it is evident that all network architectures are capable of providing high correct classification percentages as all minimum accuracy percentages are above 99.50%. Also, the difference between the maximum and minimum reported classification accuracies is below 0.2% (a maximum of 0.17% is reported for C3) indicating that the randomness included within SMOTE does not significantly impact the classification ability of the networks. The maximum reported accuracy on the test set was achieved by a C6R network equal to 99.85%, emphasized in bold in Table III. The network's confusion matrix is presented in Table IV. From this confusion matrix, it can be observed that perfect recall is achieved for the N2, N3, and REM sleep stages. Classification errors are only present in the case of N1 sleep epochs, where 6 out of the 77 N1 epochs are missclassified as N2 with the recall value of the N1 class being equal to 92.21%. The maximum accuracy of the C6R network is followed by a value of 99.80% achieved by C6, C8, and C8R networks. The lowest maximum classification accuracies were achieved by C3 and C3R networks equal to 99.71% and 99.78%, respectively.

TABLE III
Overall Test Set Accuracy Percentages

Architecture	Test Set Accuracy		
	Min	Median	Max
C3	99.61 %	99.68 %	99.71 %
C3R	99.66 %	99.73 %	99.78 %
C6	99.68 %	99.76 %	99.80 %
C6R	99.73 %	99.78 %	**99.85 %**
C8	99.73 %	99.80 %	99.80 %
C8R	99.71 %	99.76 %	99.80 %

TABLE IV
Confusion Matrix for C6R Network That Achieved the Maximum Overall Test Set Accuracy

		Target Class				
		N1	N2	N3	REM	
Output Class	N1	71	0	0	0	
	N2	6	908	0	0	
	N3	0	0	1912	0	
	REM	0	0	0	1195	Accuracy
	Recall	92.21 %	100.00 %	100.00 %	100.00 %	99.85 %

Apart from the test set accuracy, the recall percentages for each class attained by the networks that achieved the maximum overall test set accuracy are presented in Table V. Recall is the portion of a class' samples that are correctly classified in said class. In the majority of cases regarding the N3 and REM sleep stages, perfect recall is achieved. This is true for half of the networks in the case of N2 sleep epochs where C3, C3R, and C6R networks achieve perfect recall. For N1 sleep epochs, the highest recall rate is 93.51% achieved by a C8 network, not the C6R network that achieved the maximum test set accuracy, which achieved a N1 recall equal to 92.21%. The difference, however, is not significant since the number of test set N1 epochs is equal to 77. Therefore, the 1.3% difference between the two networks means that the C8 network managed to classify correctly 72 N1 epochs compared to the 71 N1 epochs correctly classified by the C6R network.

It must be noted that, according to Table II, the initial number of available N1 epochs in the training set was equal to 2.1% of the total number of training samples prior to applying SMOTE. This led to the disregard of the N1 class from the classifiers, leading to very poor recall rates for the N1 sleep stage equal to about 11%. After SMOTE, the percentage of N1 epochs in the training set increased to 19.3%, approximately equal to the percentage of N2 sleep stages, allowing higher classification accuracies and recall rates for the N1 epochs as reported above.

TABLE V
CLASS RECALL PERCENTAGES FOR THE NETWORKS THAT ACHIEVED THE MAXIMUM OVERALL TEST SET ACCURACY

Archit.	Accuracy	Class Recall			
		N1	N2	N3	REM
C3	99.78 %	89.61 %	100.00 %	99.95 %	100.00 %
C3R	99.80 %	90.90 %	100.00 %	100.00 %	99.92 %
C6	99.80 %	90.90 %	99.89 %	100.00 %	100.00 %
C6R	**99.85** %	92.21 %	100.00 %	100.00 %	100.00 %
C8	99.80 %	93.51 %	99.67 %	100.00 %	100.00 %
C8R	99.80 %	92.21 %	99.78 %	100.00 %	100.00 %

V. DISCUSSION

In this paper, we evaluated the efficacy of using CNNs in order to achieve accurate automatic sleep staging. Contrary to other approaches, the input to the CNN was not the signal time series, but functional connectivity features structured in an image-like manner, allowing us to achieve very high accuracy rates compared to prior uses of CNNs in sleep staging tasks. This proves that the same learner, fed with the appropriate representation of the data, can be trained to a much more efficient state. The notion of transforming time series into images can be extended to several fields such as sound/voice classification, text classification, and more or less to any application involving time series. Our approach underlines that CNNs are powerful learners that can be applied successfully to a wide range of data and scenarios. However, the upper limit of their performance can be fine-tuned, apart from the selection of the proper architecture and training parameters, if they are trained with the appropriate input.

In short, in the first step of our methodology, we employ a preprocessing pipeline to remove noise content. It must be noted that given a different data set, some preprocessing steps may require revision, such as the filters' cutoff frequencies, or can be omitted, such as the ICA component pruning step. When omitting this step, the whole preprocessing pipeline can be carried out automatically [6]. The next step was to extract functionality connectivity metrics in order to model the state of the human brain during sleep. From the experimental results presented in Section IV, it is evident that our approach of extracting SL and RWE features from EEG signals, augmenting the underrepresented N1 class using SMOTE along with graph metrics and energy ratios is apt for automatic sleep staging using CNNs. The maximum classification accuracy achieved was equal to 99.85% surpassing previous works in the field. Such high-accuracy percentages may have not been achieved in automatic sleep staging but are regularly achieved in much more challenging image classification problems [21], [22] where the number of classes is much greater, in the thousands.

Although previous attempts to use CNNs for automatic sleep staging did not significantly improve the classification performance [8], [27], [28], they provided insights into future research regarding the use of heterogeneous recordings and processing techniques. These may be adopted as preliminary steps for reaching a robust and data invariant classification model. Despite their elaborate computational framework, they focused on the actual time series and how to extract temporal information from these. The latter may have limited their classification accuracy. The reason may be the dimensions of PSG data, which are highly rectangular (much smaller number of sensors than the number of sample points). This may hamper the effectiveness of CNN filters, which seem to be more suitable for images with similar number of dimensions in both rows and columns. Moreover, spatial information is usually neglected and there is no attempt to exploit the inherent properties of CNNs that make them so successful in the field of image processing.

Functional connectivity analysis on sensor (electrode) level faces several methodological limitations due to the poor spatial resolution of the EEG [51], and the use of CNNs is hindered. In that case, the volume conduction problem induces the linear combination of the cortical activity from various activation sources into a single-electrode recording [52]. Thus, we lack precise localization of the recording activity, and therefore, we could not provide spatial information. Aiming to improve our previous approach, we extended the analysis by estimating cortical activations of the DMN model. The DMN is a thoroughly studied resting-state network, which is active during the lack of external stimulation and was demonstrated to reflect intrinsic brain properties [53]. It was also found to be dependent of sleep depth [54], while some of its nodes were also involved in sleep disturbances like insomnia [55]. We hypothesized that the previous neuroscientific findings would fit well with our research aim to identify various sleep stages. This paper results demonstrate that the estimation of DMN cortical activity significantly improved the classification accuracy. Therefore, the selection of DMN nodes as regions of interest for sleep related studies is a suitable one.

CNNs are regularly employed in cases of big data problems. Aiming to maximize the classification accuracy, their design is focused on the identification of the distinct patterns associated with the most frequent clusters/groups. Therefore, in that case, there is the chance that the signature of a minority group (e.g., $\leq$1%–2% of the total data size) would often be neglected. This may be acceptable in the case of computer vision or image recognition problems but not in medical problems. These rare patterns may be associated with life-threatening situations or cases of significant clinical/research interest. The latter happens also in sleep staging with the sleep onset as represented by the N1 stage. That stage lasts approximately 3–10 min during physiological sleep whose duration is approximately 8 h. However, its identification is very important and could provide us with valuable information about sleep quality. Its classification may be neglected in previous big data approaches since it contributes minimally to the overall classification accuracy. Our approach with the utilization of the SMOTE algorithm ensures that the CNN architecture would properly regard N1 patterns and would result in its identification, primarily because it represents a significant portion of the training data set.

Despite the great novelty of this paper, there are several issues that should be addressed regarding the

proposed framework. More specifically, EEG source connectivity analysis was previously demonstrated to improve the localization accuracy and minimize but not neutralize severe limitations of sensor-level analysis [56]. However, the small number of electrodes adopted in order to assure convenient sleeping environment is the minimum one used for source estimation with scientific merit [57], [58]. Another study limitation is the employment of a generic head anatomy model. More advanced techniques should adopt individual anatomy of each participant in order to improve the localization accuracy. Therefore, our results should be interpreted with caution and validated by other neuroimaging techniques. Adoption of cortical connectivity analysis with the standard technological equipment may not be ideal for neuroimaging localization studies but it is suitable for providing precise estimation of sleep macroarchitecture and its disturbances in clinical routine practice and for sleep assessment during extreme conditions like that of microgravity simulation and spaceflights. This was also assessed by the high correlation of the system's outcome with the results of the manual sleep staging performed by two sleep experts. Future work would focus on the validation of our proposed DMN-based framework on sleep data sets derived from high-density EEG or other neuroimaging modalities. We also aim to preserve the same classification accuracy on data recorded either under noisy conditions or from clinical populations with sleep-related breathing disorders. In these cohorts, we aim to provide a system not only detecting the proper sleep stage but also abnormal pulmonary patterns.

To sum up, this paper proposes a novel framework for exploiting the intrinsic properties of CNNs (CNNs) for computer-based sleep staging. CNNs are based on filters in order to extract meaningful, mainly spatial information from images. However, previous approaches for sleep staging employed CNNs during temporal feature extraction from sleep-related time series. Our contribution is to perform the analysis not on the PSG or EEG electrodes but to estimate the activity of a spatially distributed on the cerebral cortex brain network. This network is the DMN which is consisted of nodes placed on both anterior and posterior positions. Then, we exploit both spatial and temporal information as derived from the cortical functional connectivity among the DMN nodes. Since, DMN activity is highly dependent from sleep stages, we did not focus only on the entire EEG spectrum but we computed three different activity images based on the main rhythms associated with transition to sleep and its depth. Our third point of novelty is not to neglect the relatively short duration of sleep onset (N1 stage) for the sake of much longer stages (e.g., N2) but to provide a solution for the proper identification of these minority states based on the SMOTE algorithm. The proposed computational methodology seems to provide great insight into the newly developed area of CNN-based tools for learning meaningful neuroimage patterns. It could also assist sleep experts during their tedious task of manual sleep staging.

REFERENCES

[1] S. Mukherjee *et al.*, "An official American Thoracic Society statement: The importance of healthy sleep. Recommendations and future priorities," *Amer. J. Respiratory Crit. Care Med.*, vol. 191, no. 12, pp. 1450–1458, 2015.

[2] K. J. Reid, K. G. Baron, B. Lu, E. Naylor, L. Wolfe, and P. C. Zee, "Aerobic exercise improves self-reported sleep and quality of life in older adults with insomnia," *Sleep Med.*, vol. 11, no. 9, pp. 934–940, 2010.

[3] R. B. Berry *et al.*, "AASM scoring manual updates for 2017 (version 2.4)," *J. Clin. Sleep Med., JCSM, Off. Publication Amer. Acad. Sleep Med.*, vol. 13, no. 5, p. 665, 2017.

[4] M. H. Silber *et al.*, "The visual scoring of sleep in adults," *J. Clin. Sleep Med.*, vol. 3, no. 2, p. 22, 2007.

[5] E. Olbrich, J. C. Claussen, and P. Achermann, "The multiple time scales of sleep dynamics as a challenge for modelling the sleeping brain," *Philos. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 369, no. 1952, pp. 3884–3901, 2011.

[6] P. Chriskos, C. A. Frantzidis, P. T. Gkivogkli, P. D. Bamidis, and C. Kourtidou-Papadeli, "Achieving accurate automatic sleep staging on manually pre-processed EEG data through synchronization feature extraction and graph metrics," *Frontiers Hum. Neurosci.*, vol. 12, p. 110, Mar. 2018.

[7] K. A. I. Aboalayon, W. S. Almuhammadi, and M. Faezipour, "A comparison of different machine learning algorithms using single channel EEG signal for classifying human sleep stages," in *Proc. Long Island Syst., Appl. Technol. (LISAT)*, 2015, pp. 1–6.

[8] O. Tsinalis, P. M. Matthews, Y. Guo, and S. Zafeiriou. (2016). "Automatic sleep stage scoring with single-channel EEG using convolutional neural networks." [Online]. Available: https://arxiv.org/abs/1610.01683

[9] S. Yucelbas, S. Ozsen, C. Yucelbas, G. Tezel, S. Kuccukturk, and S. Yosunkaya, "Effect of EEG time domain features on the classification of sleep stages," *Indian J. Sci. Technol.*, vol. 9, no. 25, pp. 1–8, 2016.

[10] T. H. Sanders, M. McCurry, and M. A. Clements, "Sleep stage classification with cross frequency coupling," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 4579–4582.

[11] A. R. Hassan and M. I. H. Bhuiyan, "A decision support system for automatic sleep staging from EEG signals using tunable Q-factor wavelet transform and spectral features," *J. Neurosci. Methods*, vol. 271, pp. 107–118, Sep. 2016.

[12] H. Li, C. Peng, and D. Ye, "A study of sleep staging based on a sample entropy analysis of electroencephalogram," *Bio-Med. Mater. Eng.*, vol. 26, no. s1, pp. S1149–S1156, 2015.

[13] A. F. Farag, S. M. El-Metwally, and A. A. Morsy, "A sleep scoring system using EEG combined spectral and detrended fluctuation analysis features," *J. Biomed. Sci. Eng.*, vol. 7, no. 8, p. 584, 2014.

[14] Y. Zhang *et al.*, "Automatic sleep staging using multi-dimensional feature extraction and multi-kernel fuzzy support vector machine," *J. Healthcare Eng.*, vol. 5, no. 4, pp. 505–520, 2014.

[15] S. A. Imtiaz and E. Rodriguez-Villegas, "Automatic sleep staging using state machine-controlled decision trees," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 378–381.

[16] T. Sousa, A. Cruz, S. Khalighi, G. Pires, and U. Nunes, "A two-step automatic sleep stage classification method with dubious range detection," *Comput. Biol. Med.*, vol. 59, pp. 42–53, Apr. 2015.

[17] S.-F. Liang, C.-E. Kuo, F.-Z. Shaw, Y.-H. Chen, C.-H. Hsu, and J.-Y. Chen, "Combination of expert knowledge and a genetic fuzzy inference system for automatic sleep staging," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 10, pp. 2108–2118, Oct. 2016.

[18] S. A. Imtiaz, Z. Jiang, and E. Rodriguez-Villegas, "An ultralow power system on chip for automatic sleep staging," *IEEE J. Solid-State Circuits*, vol. 52, no. 3, pp. 822–833, Mar. 2017.

[19] E. Malaekah, C. R. Patti, and D. Cvetkovic, "Automatic sleep-wake detection using electrooculogram signals," in *Proc. IEEE Conf. Biomed. Eng. Sci. (IECBES)*, Dec. 2014, pp. 724–728.

[20] C. C. R. Sady, U. S. Freitas, A. Portmann, J.-F. Muir, C. Letellier, and L. A. Aguirre, "Automatic sleep staging from ventilator signals in non-invasive ventilation," *Comput. Biol. Med.*, vol. 43, no. 7, pp. 833–839, 2013.

[21] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *Proc. BMVC*, vol. 1, no. 3, 2015, p. 6.

[22] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. CVPR*, Jun. 2015, pp. 1–9.

[23] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 91–99.
[24] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 779–788.
[25] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and tell: A neural image caption generator," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3156–3164.
[26] K. Xu *et al.*, "Show, attend and tell: Neural image caption generation with visual attention," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 2048–2057.
[27] A. Supratak, H. Dong, C. Wu, and Y. Guo, "DeepSleepNet: A model for automatic sleep stage scoring based on raw single-channel EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 11, pp. 1998–2008, Nov. 2017.
[28] S. Chambon, M. N. Galtier, P. J. Arnal, G. Wainrib, and A. Gramfort, "A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 4, pp. 758–769, Apr. 2018.
[29] M. E. Raichle, "The brain's default mode network," *Annu. Rev. Neurosci.*, vol. 38, pp. 433–447, Jul. 2015.
[30] C. J. Stam, M. Breakspear, A.-M. V. C. van Walsum, and B. W. van Dijk, "Nonlinear synchronization in EEG and whole-head MEG recordings of healthy subjects," *Hum. Brain Mapping*, vol. 19, no. 2, pp. 63–78, 2003.
[31] O. A. Rosso *et al.*, "Wavelet entropy: A new tool for analysis of short duration brain electrical signals," *J. Neurosci. Methods*, vol. 105, no. 1, pp. 65–75, 2001.
[32] M. Mohanraj, S. Jayaraj, and C. Muraleedharan, "Applications of artificial neural networks for thermal analysis of heat exchangers—A review," *Int. J. Therm. Sci.*, vol. 90, pp. 150–172, Apr. 2015.
[33] D. Rastovic, "Targeting and synchronization at Tokamak with recurrent artificial neural networks," *Neural Comput. Appl.*, vol. 21, no. 5, pp. 1065–1069, Jan. 2011.
[34] D. Rastovic, "From non-Markovian processes to stochastic real time control for Tokamak plasma turbulence via artificial intelligence techniques," *J. Fusion Energy*, vol. 34, no. 2, pp. 207–215, Apr. 2014.
[35] P. Chriskos *et al.*, "Automatic sleep stage classification applying machine learning algorithms on EEG recordings," in *Proc. IEEE 30th Int. Symp. Comput.-Based Med. Syst. (CBMS)*, Jun. 2017, pp. 435–439.
[36] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Comput.*, vol. 7, no. 6, pp. 1129–1159, 1995.
[37] F. Tadel, S. Baillet, J. C. Mosher, D. Pantazis, and R. M. Leahy, "Brainstorm: A user-friendly application for MEG/EEG analysis," *Comput. Intell. Neurosci.*, vol. 2011, Jan. 2011, Art. no. 8. [Online]. Available: https://dl.acm.org/citation.cfm?id=1992539
[38] R. D. Pascual-Marqui, "Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details," *Methods Find Exp. Clin. Pharmacol.*, vol. 24, pp. 5–12, Jan. 2002.
[39] J. Theiler, "Spurious dimension from correlation algorithms applied to limited time-series data," *Phys. Rev. A, Gen. Phys.*, vol. 34, no. 3, p. 2427, 1986.
[40] C. A. Frantzidis, A. B. Vivas, A. Tsolaki, M. A. Klados, M. Tsolaki, and P. D. Bamidis, "Functional disorganization of small-world brain networks in mild Alzheimer's disease and amnestic mild cognitive impairment: An EEG study using relative wavelet entropy (RWE)," *Frontiers Aging Neurosci.*, vol. 6, p. 224, Mar. 2014.
[41] C. E. Shannon, "A note on the concept of entropy," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.
[42] L. Deuker *et al.*, "Reproducibility of graph metrics of human brain functional networks," *NeuroImage*, vol. 47, no. 4, pp. 1460–1468, 2009.
[43] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, p. 440, 1998.
[44] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, 2002.
[45] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
[46] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, 2010, pp. 807–814.
[47] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 815–823.
[48] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1701–1708.
[49] A. Kramer, J. Kümmel, E. Mulder, A. Gollhofer, P. Frings-Meuthen, and M. Gruber, "High-intensity jump training is tolerated during 60 days of bed rest and is very effective in preserving leg power and lean body mass: An overview of the cologne RSL study," *PLoS ONE*, vol. 12, no. 1, 2017, Art. no. e0169793.
[50] C. A. Frantzidis *et al.*, "Cortical connectivity analysis for assessing the impact of microgravity and the efficacy of reactive sledge jumps countermeasure to NREM 2 sleep," *Acta Astron.*, to be published. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0094576518311111
[51] C. A. Frantzidis, A.-K. I. Ladas, A. B. Vivas, M. Tsolaki, and P. D. Bamidis, "Cognitive and physical training for the elderly: Evaluating outcome efficacy by means of neurophysiological synchronization," *Int. J. Psychophysiol.*, vol. 93, no. 1, pp. 1–11, 2014.
[52] S. Haufe, V. V. Nikulin, K.-R. Müller, and G. Nolte, "A critical assessment of connectivity measures for EEG data: A simulation study," *NeuroImage*, vol. 64, pp. 120–133, Jan. 2013.
[53] M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon, "Functional connectivity in the resting brain: A network analysis of the default mode hypothesis," *Proc. Nat. Acad. Sci.*, vol. 100, no. 1, pp. 253–258, 2003.
[54] P. G. Sämann *et al.*, "Development of the brain's default mode network from wakefulness to slow wave sleep," *Cerebral Cortex*, vol. 21, no. 9, pp. 2082–2093, 2011.
[55] X. Nie *et al.*, "Functional connectivity of paired default mode network subregions in primary insomnia," *Neuropsychiatric Disease Treat.*, vol. 11, p. 3085, 2015.
[56] J.-M. Schoffelen and J. Gross, "Source connectivity analysis with MEG and EEG," *Hum. Brain Mapping*, vol. 30, no. 6, pp. 1857–1865, 2009.
[57] S. Schneider, V. Brümmer, T. Abel, C. D. Askew, and H. K. Strüder, "Changes in brain cortical activity measured by EEG are related to individual exercise preferences," *Physiol. Behav.*, vol. 98, no. 4, pp. 447–452, 2009.
[58] R. D. Pascual-Marqui *et al.*, "Low resolution brain electromagnetic tomography (LORETA) functional imaging in acute, neuroleptic-naive, first-episode, productive schizophrenia," *Psychiatry Res., Neuroimag.*, vol. 90, no. 3, pp. 169–179, 1999.

Panteleimon Chriskos received the B.Sc. degree in informatics from the Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece, in 2015, and the M.Sc. degree in informatics and communications, with a focus on digital media and computational intelligence, and medical informatics from AUTH, in 2017, where he is currently pursuing the Ph.D. degree with the Laboratory of Medical Physics, School of Medicine.

He has co-authored papers in international conferences and journals as well as two book chapters. His current research interests include digital signal and image processing, computational intelligence, and health informatics.

Christos A. Frantzidis received the Diploma degree in electrical and computer engineering, the M.Sc. degree in medical informatics, and the Ph.D. degree in neurophysiological evaluation of nonpharmaceutical cognitive and/or physical interventions aiming to promote healthy aging from the Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece, in 2006, 2008, and 2014, respectively.

He was a Visiting Researcher with the European Space Agency-funded bed-rest study held in the envihab premises of the German Aerospace Agency, Cologne, Germany. He is currently collaborating with the Greek AeroSpace Medical Aviation, Kalamaria, Greece, and the Laboratory of Medical Physics, School of Health Sciences, AUTH, toward brain-network and functional connectivity analysis. He is studying the effects of both physiological and pathological aging and the detrimental effects of microgravity on sleep physiology. He has authored the paper What are the symbols of Alzheimer? A Permutation Entropy-based symbolic analysis for the detection of early changes of the electroencephalographic complexity due to mild Alzheimer which received the third prize in the Best Student Paper Competition during the 12th IEEE Conference on BioInformatics and BioEngineering, Larnaca, Cyprus, in 2012.

Polyxeni T. Gkivogkli received the bachelor's degree in biology and the master's degree in neuroscience ethics from the Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece, in 2008 and 2011, respectively.

She is currently collaborating with the Laboratory of Medical Physics, Medical School, AUTH, where She is involved in sleep-related projects. She is also a Sleep Expert with emphasis on sleep macroarchitecture and grapho-elements (K-complexes and spindles) analysis.

Panagiotis D. Bamidis has founded and leads three working groups (Medical Education Informatics, Assistive Technologies and Silver Science, and Applied and Affective Neuroscience). He is currently an Assistant Professor with the Medical School, Aristotle University of Thessaloniki, Thessaloniki, Greece. He is also a Visiting Scientist with the Karolinska Institute, Solna, Sweden. Recently, he has coordinated four European Projects and led several national and international funded projects. His publication record exceeds 85 international refereed journal papers, 300 international peer reviewed conference papers, several book chapters/edited conference proceeding volumes, and more than 470 citations (h-index>12). Furthermore, he has been a referee in more than 25 journals and a guest editor in more than 13 journal special issues. His current research interests include affective and applied neuroscience/computing, multimodal interaction and HCI, technology enhanced learning in medical education, health information management, biomedical informatics, assistive technologies for active and healthy aging, special education/developmental disorders, and silvergaming/exergaming/silver science.

Mr. Bamidis is a member of the Administration Board of the Greek Federation of Alzheimer's Associations and Related Disorders, the Innovation Zone, Thessaloniki, and the Hellenic Biomedical Engineering Society. He has organized and/or chaired six international conferences (iSHIMR2001, iSHIMR2005, MEDICON2010, GASMA2010, SAN2011, and MEI2012) and was a Conference Producer of the Medical Education Informatics Conference and Spring School Series.

Chrysoula Kourtidou-Papadeli received the Ph.D. degree from the Department of Physiology, Aristotle University of Thessaloniki, Thessaloniki, Greece, in 1986, with a focus on conducting research on COPD patients.

She was the Head of pulmonary medicine (7 years) with the Eurodiagnosis Medical Center, Thessaloniki. She was a Research Collaborator (5 years) with the Sleep Laboratory, Department of Internal Medicine, Saint Paul General Hospital, Thessaloniki, where she conducting research on sleep disturbances. She was the Head of the Pulmonary Medicine Department, IASI Medical Center, in 2014, focusing on sleep and smoking cessation. She was the Director of the Aeromedical Center, Thessaloniki, in 2004, participating in smoking cessation interventions focused on pilots. She is currently a Specialist in pulmonary and aerospace medicine, Dayton, OH, USA, in 1993 and 1996, respectively. She participated in several research projects and has authored medline journal publications and chapters in several books.

Dr. Kourtidou-Papadeli is a member of the Association of Pulmonary Medicine of Northern Greece, the European Academy of Allergology and Pulmonary medicine, the International Sleep Association, the International Pilots Association Greece, ESA Astronaut Medical Selection Board in 2009, and the Greek Aerospace Medical Association, the President of Space Research, a member of the International Academy of Astronautics (Section III: Life Sciences), a fellow of the Aerospace Medical Association, the Chair of the International Committee, Space Medicine branch of AsMA, the International Academy of Astronautics (Academy Board of trustees), and the U.K. Association of Aviation Medical Examiners.