*sensors*

MDPI

*Article*

# Recognition of Patient Groups with Sleep Related Disorders using Bio-signal Processing and Deep Learning

**Delaram Jarchi** [1,2,*], **Javier Andreu-Perez** [1,2], **Mehrin Kiani** [1], **Oldrich Vysata** [3,4], **Jiri Kuchynka** [4], **Ales Prochazka** [3,5] **and Saeid Sanei** [6]

[1]   Smart Health Technologies Group, School of Computer Science and Electronic Engineering; University of Essex, Colchester CO4 3SQ, UK; javier.andreu@essex.ac.uk (J.A.P.); mehrin.kiani@essex.ac.uk (M.K.)
[2]   Embedded and Intelligent Systems Laboratory, School of Computer Science and Electronics, University of Essex, Colchester CO4 3SQ, UK
[3]   Department of Computing and Control Engineering, University of Chemistry and Technology in Prague, 166 28 Prague 6, Czech Republic; Oldrich.Vysata@fnhk.cz (O.V.); A.Prochazka@ieee.org (A.P.)
[4]   Department of Neurology, Faculty of Medicine in Hradec Králové, Charles University, 500 05 Hradec Králové, Czech Republic; Jiri.Kuchynka@fnhk.cz
[5]   Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University in Prague, 160 00 Prague 6, Czech Republic
[6]   School of Science and Technology, Nottingham Trent University, Nottingham NG11 8NS, UK; saeid.sanei@ntu.ac.uk
*   Correspondence: delaram.jarchi@essex.ac.uk

check for updates

**Abstract:** Accurately diagnosing sleep disorders is essential for clinical assessments and treatments. Polysomnography (PSG) has long been used for detection of various sleep disorders. In this research, electrocardiography (ECG) and electromayography (EMG) have been used for recognition of breathing and movement-related sleep disorders. Bio-signal processing has been performed by extracting EMG features exploiting entropy and statistical moments, in addition to developing an iterative pulse peak detection algorithm using synchrosqueezed wavelet transform (SSWT) for reliable extraction of heart rate and breathing-related features from ECG. A deep learning framework has been designed to incorporate EMG and ECG features. The framework has been used to classify four groups: healthy subjects, patients with obstructive sleep apnea (OSA), patients with restless leg syndrome (RLS) and patients with both OSA and RLS. The proposed deep learning framework produced a mean accuracy of 72% and weighted F1 score of 0.57 across subjects for our formulated four-class problem.

## 1. Introduction

Study of human sleep and the underlying physiological changes is crucial for detecting sleep disorders, improving the quality of sleep and avoiding daytime sleepiness. There are two main sleep disorders, including obstructive sleep apnea (OSA) [1,2] as a breathing-related sleep disorder, and restless leg syndrome (RLS) as a movement-related sleep disorder [3–5]. OSA is associated with full or partial occlusion of the upper airway during sleep which needs treatment; otherwise, it can affect the cardiovascular system and also lead to nocturnal death [6,7]. On the other hand, RLS is associated with uncontrollable sporadic or periodic leg movements during sleep.

Polysomnography (PSG) as a type of sleep study has been developed for noninvasive analysis of sleep. It includes a number of wearable sensors, such as an electrocardiogram (ECG) to record the cardiac activity from electrodes attached to chest; a pulse oximeter attached to a finger to record photoplethysmography (PPG) for blood oxygen saturation measurement; an electromyograph (EMG) that can be fitted to the chin and leg for movement analysis; an electroencephalograph (EEG) to record brain activity that is useful for sleep stage classification; an electrooculograph (EOG) to monitor eye movements; a nose sensor for monitoring nasal breathing; and an accelerometer to analyze chest movements. PSG has been used for detecting various sleep disorders and also measurement of sleep quality as the gold standard.

Many techniques have been developed to detect sleep stages and sleep disorders, such as OSA using PSG in research studies and clinical trials. EEG has been used to classify sleep stages using correlation graphs [8] and weighted undirected complex networks [9]. EEG has also been used to detect sleep apnea events using inter-band energy ratio [10]. In another study, EEG and EMG signals were used to detect sleep apnea exploiting EEG arousal [11]. A combination of EEG, ECG and EMG has been used in [12] to detect sleep apnea events in different sleep stages. A non-contact pressure-sensitive device was built [13] to detect OSA. Deriving breathing patterns from a single channel ECG is more straightforward than analysis of EEGs, which are prone to artifacts and drifts [14], and usually more than one EEG channel is required for apnea detection. The benefit of using ECG in our framework is to further motivate the development and adaptation of similar techniques in future studies for PPG analysis from wrist worn optical sensors to make unobtrusive and portable systems for identification of sleep disorders.

Reducing the number of wearable sensors in the PSG system for detection of sleep disorders alleviates the system complexity and focuses on the required information only. Single-lead ECG has been successfully used to identify OSA [15–17]. The key in analysis of OSA using ECG is to accurately extract the breathing signal, since OSA is directly related to respiration. All the techniques developed for ECG analysis can be simply adapted for PPG analysis which is a very non-intrusive way of measuring human physiological parameters.

EMG has been used for detection of neuromuscular diseases which can arise from sleep movement disorders [18]. EMG has also been shown to be particularly useful in the detection of RLS. There is evidence that RLS and OSA have been associated as there are more periodic leg movements during sleep apnea episodes [19]. However, concurrent and quantitative analysis of RLS and OSA have not been well studied.

The objective of this paper is to exploit strong signal processing and machine learning techniques for joint identification of OSA and RLS using ECG and EMG. This will open a new insight into identification and correlation analysis of respiratory and movement-related symptoms. More specifically, a deep-learning-based framework is proposed for automatic identification and classification of healthy subjects, patients with OSA, patients with RLS and patients with both OSA and RLS. This is also useful for group-specific sleep analysis of sleep stages and scoring [20,21].

The remainder of the paper is as follows. In Section 2, first the EMG analysis is introduced which includes description of features such as lower and higher-order statistics and a relatively new entropy-related measure called dispersion entropy (DE) to incorporate signal fluctuations [22]. This is followed by the ECG analysis subsection. In this subsection, first a strong signal processing technique for time-frequency analysis, i.e., synchrosqueezed wavelet transform (SSWT), is described. Then, an iterative pulse peak detection technique which uses SSWT as a core technique is introduced for reliable ECG pulse peak detection. Then, ECG derived features, including heart rate variability and respiratory-related measures are described. At the end of this section, a deep learning framework is introduced which exploits ECG and EMG features for classification of four groups, including healthy subjects, patients with OSA, patients with RLS and patients with both OSA and RLS. In Section 3, the results are provided. Finally, Section 4 concludes the paper, highlighting the major contributions, limitations and potential improvements of the proposed framework for future sleep analysis systems.

## 2. Materials and Methods

### *2.1. EMG Analysis*

In order to characterize the distribution of EMG data samples, quantitative measures should be used. In the following, lower and higher-order statistics and entropy-based measures which are useful to quantify changes in the EMG signals are described. In those cases where EMG is contaminated by ECG, it can be effectively restored using the method based on singular spectrum analysis proposed in [23].

#### 2.1.1. EMG Features from Moments

The first moment is simply the data mean and denoted as $E[X]$, where $X$ is considered as a real-valued random variable. The second moment is the variance; its square root can be calculated using $(E[(X - E[X])^2])^{\frac{1}{2}}$ as the standard deviation. These two moments are considered low-order statistics. Similarly, higher order statistics including skewness and kurtosis are derived using the following equations:

$$Skewness = \frac{E[(X - E[X])^3]}{(E[(X - E[X])^2])^{\frac{3}{2}}} \tag{1}$$

$$Kurtosis = \frac{E[(X - E[X])^4]}{(E[(X - E[X])^2])^2} \tag{2}$$

where $X$ is a real-valued random variable. Both lower and higher order statistics can be applied to EMG signals for feature extraction.

#### 2.1.2. EMG Features from Entropy Measurement

Entropy has been used to quantify the signal uncertainty and dynamical characteristics. Sample entropy [24] and permutation entropy [25] have been widely used to indicate the entropy measure. Here, we have selected a recently developed entropy-based method called DE which takes into account the signal fluctuations [22]. The algorithm to estimate the DE is summarized in four main steps below:

*Step One*: The samples $x_j (j = 1, 2, \ldots, N)$ of signal **x** are mapped into $c$ classes that are labeled from 1 to $c$ ($c$ is a user-defined parameter). As proposed in [22], normal cumulative distribution function (NCDF) is used to map $\mathbf{x} = \{x_1, x_2, \ldots, x_N\}$ into $\mathbf{y} = \{y_1, y_2, \ldots, y_N\}$, from 0 to 1, first. Then, a linear algorithm is applied to assign each $y_j$ value to an integer from 1 to $c$ ($z_j^c = round(c.y_j + 0.5)$ where $z_j^c$ is the $j$th member of the classified time-series).

*Step Two*: Embedding vectors with embedding dimension $m$ and time delay $d$ are created based on: $z_i^{m,c} = \{z_i^c, z_{i+d}^c, \ldots, z_{i+(m-1)d}^c\}, i = 1, 2, \ldots, N - (m - 1)d$. Then, each produced time series $z_i^{m,c}$ is mapped to a dispersion pattern $\pi_{v_0 v_1 \ldots v_{m-1}}$, where $z_i^c = v_0, z_{i+d}^c = v_1, \ldots, z_{i+(m-1)d}^c = v_{m-1}$. The number of potential dispersion patterns that can be assigned to each time series ($z_i^{m,c}$) is equal to $c^m$. This is due to the fact that the time-series contains $m$ members and each member can be assigned to an integer from 1 to $c$.

*Step Three*: For each possible dispersion pattern (there are $c^m$ potential patterns), a relative frequency is calculated as:

$$p(\pi_{v_0, v_1, \ldots, v_{m-1}}) = \frac{Number\{i | i \le N - (m-1)d, z_i^{m,c} \text{ has type } \pi_{v_0 v_1 \ldots v_{m-1}}\}}{N - (m-1)d} \tag{3}$$

Based on the above equation, $p(\pi_{v_0, v_1, \ldots, v_{m-1}})$ demonstrates the number of dispersion patterns $\pi_{v_0 v_1 \ldots v_{m-1}}$ which are assigned to $z_i^{m,c}$, divided by the total number of embedded time-series (embedding dimension is $m$).

*Step Four*: The DE value can be calculated using the Shannon's definition of entropy as [26]:

$$DE(\mathbf{x}, m, c, d) = -\sum_{\pi=1}^{c^m} p(\pi_{v_0, v_1, ..., v_{m-1}}) . ln(p(\pi_{v_0, v_1, ..., v_{m-1}})) \tag{4}$$

A summary of EMG features including moments and DE is provided in Table 1.

**Table 1.** A summary of features extracted from EMG and ECG signals is provided below.

| EMG Features | |
| --- | --- |
| **Feature** | **Feature type**[*] |
| Mean | Raw |
| Standard deviation | Raw |
| Skewness | Raw |
| Kurtosis | Raw |
| Dispersion Entropy | Raw |
| ECG Features | |
| **Feature** | **Feature type** |
| Maximum difference of pulse peaks | Time-domain |
| Minimum difference of pulse peaks | Time-domain |
| Mean difference of pulse peaks | Time-domain |
| Maximum amplitude of respiratory amplitude modulation | Time-domain |
| Minimum amplitude of respiratory amplitude modulation | Time-domain |
| Mean amplitude of respiratory amplitude modulation | Time-domain |
| Maximum of instantaneous frequencies of respiratory amplitude modulation | Frequency-domain |
| Minimum of instantaneous frequencies of respiratory amplitude modulation | Frequency-domain |
| Mean of instantaneous frequencies of respiratory amplitude modulation | Frequency-domain |
| Standard deviation of instantaneous frequencies of respiratory amplitude modulation | Frequency-domain |

[*] Raw feature type means the features are directly extracted form signal samples.

## 2.2. ECG Analysis

In order to extract features from ECG, the ECG pulse peaks (R peaks) should be detected. Although many peak detection algorithms have been developed for ECG peak detection, the reliability and generalisation of most of these algorithms have not been comprehensively evaluated. Here, we propose an iterative R peak detection algorithm which combines the temporal domain technique (simple pulse peak detection) with time-frequency spectral presentation of the signal. Such a hybrid technique that uses both temporal information of the ECG signal and time-frequency-based representation has significantly increased the reliability of pulse peak detection algorithm.

### 2.2.1. Synchrosqueezed Wavelet Transform

SSWT applies the time-frequency analysis method to produce the spectrum of the input signal in the first stage. In the second stage, a reassignement technique is applied to the resultant time-frequency spectrum [27]. The objective of the reassignment technique is to enhance the initial spectrum to produce a sharper spectrum with visible and narrower frequency ridges. SSWT was first targeted for auditory signal analysis following its development [27,28]. The SSWT algorithm is described in the following. *Stage 1*: The continuous wavelet transform (CWT) is applied to the input signal. Consider the CWT of the input signal **s** as:

$$W_s(a, b) = \int_{-\infty}^{\infty} s(t) a^{-1/2} \overline{\psi(\frac{t-b}{a})} dt \tag{5}$$

where $a$ is the wavelet scale, $t$ is the time index, $\psi$ is the selected mother wavelet ($\overline{\psi}$ represents its complex conjugate form) and $b$ is the position parameter. A purely harmonic signal has been selected to better explain the SSWT algorithm. Suppose that the mother wavelet is concentrated on the positive frequency axis such that $\hat{\psi}(\varepsilon) = 0$ for $\varepsilon < 0$. The selected harmonic presented as $s(t) = A\cos(\omega t)$ is

used in the following equations to demonstrate the operation of the SSWT. Based on the Plancherel's theorem [27], and using the harmonic signal as the input, CWT can be formulated as:

$$
\begin{aligned}
W_s(a,b) &= \int_{-\infty}^{\infty} s(t) a^{-1/2} \overline{\psi\left(\frac{t-b}{a}\right)} dt \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{s}(\varepsilon) a^{1/2} \overline{\hat{\psi}(a\varepsilon)} e^{ib\varepsilon} d\varepsilon \\
&= \frac{A}{4\pi} \int_{0}^{\infty} [\delta(\varepsilon - \omega) + \delta(\varepsilon + \omega)] a^{1/2} \overline{\hat{\psi}(a\varepsilon)} e^{ib\varepsilon} d\varepsilon
\end{aligned}
\tag{6}
$$

If $\hat{\psi}(\varepsilon)$ is concentrated around $\varepsilon = \omega_0$, then $W_s(a,b)$ is concentrated around $a = \omega_0/\omega$, and spreads out over a region of the horizontal line $a$:

$$
W_s(a,b) = \frac{A}{4\pi} a^{1/2} \overline{\hat{\psi}(a\omega)} e^{ib\omega}
\tag{7}
$$

If $\omega = \omega_0/a$ is similar but not necessarily exactly identical to the actual instantaneous frequency (IF) of the input signal, some non-zero energy for $W_s(a,b)$ will appear. This energy needs to be moved away from $\omega$, and this is the main objective of the synchrosqueezing concept. Reassigning of the frequency locations nearer to the actual IF is a key to SSWT which enhances time-frequency spectral representation. *Stage 2*: In the second stage, for any (a,b), for which $W_s(a,b) \neq 0$, the candidate IFs ($\omega_s(a,b)$) should be calculated:

$$
\omega_s(a,b) = -i(W_s(a,b))^{-1} \frac{\partial}{\partial b} W_s(a,b)
\tag{8}
$$

For the selected signal which is purely harmonic signal $s(t) = A\cos(\omega t)$, $\omega_s(a,b)$ are simply derived as $\omega$ (the frequency of the harmonic signal) [27]. The candidate IFs are used to recover actual frequencies. In the synchrosqueezing step, re-allocation technique is used to map the time domain into the time-frequency domain (using $(b,a) \Rightarrow (b, \omega_s(a,b))$).

The discrete-frequency synchrosqueezed transform of **s** is defined as $T_s(\omega_l, b)$:

$$
T_s(\omega_l, b) = (\Delta\omega)^{-1} \sum_{a_k : |\omega_s(a_k,b) - \omega_l| \leq \frac{\Delta\omega}{2}} W_s(a_k, b) a_k^{-3/2} (\Delta a)_k
\tag{9}
$$

where $\Delta\omega$ presents the width of those frequency bins $[\omega_l - \frac{1}{2}\Delta\omega, \omega_l + \frac{1}{2}\Delta\omega]$, $\Delta\omega = \omega_l - \omega_{l-1}$, $(\Delta a)_k = a_k - a_{k-1}$ and $T_s(\omega_l, b)$ is the synchrosqueezed transform at the centres $\omega_l$ of the sequential frequency bins. Using Equation (9), at each fixed time point $b$, the reassigned frequencies should be estimated for all the scales. This can be done by adding up all contributions incorporating $W_s(a,b)$ and considering the distance between the candidate IF $\omega_s(a,b)$ and $\omega_l$ (see Equation (9)). This distance must be within a specified frequency bin width ($\Delta\omega/2$).

### 2.2.2. Inverse Synchrosqueezed Wavelet Transform

To obtain the time-frequency spectrum of a desired input signal ($s(t)$) by SSWT, $T_s(\omega_l, b)$ must be calculated using Equation (9). As an important advantage of SSWT, as shown in [27], the original signal can be analytically reconstructed succeeding the synchronosqueezing step. As expected, $T_s(\omega_l, b)$ is concentrated more sharply around the actual IFs of the original signal (Equation (9)). This is due to the fact that the spectrum derived by the SSWT is more sparse than $W_s(a,b)$ which is obtained by wavelet transform (Equation (5)).

ISSWT is proposed to perform signal reconstruction. Its operation is based on inverting the CWT integrating over the frequencies associated with a desired component having a fully discretized version of Equation (9) presented as $\tilde{T}_s(w_l, t_m)$ and a set of fixed frequency ranges as the input to the ISSWT. These frequency ranges can be specified either by the user or by applying a standard least-squares

ridge extraction method [29]. Let us represent these frequencies as $l \in \mathscr{L}(t_m)$, where $m = 0, ..., n-1$, $t_m = t_0 + m\Delta t$, $a_j = 2^{j/n_v}\Delta t$, $j = 1, ..., Ln_v$ and $Ln_v$ is the number of scales. Given the frequencies of a desired component ($k^{th}$), the corresponding signal can be reconstructed using the following equation:

$$r_k(t_m) = 2R_\psi^{-1}\Re\left(\sum_{l\in\mathscr{L}(t_m)} \tilde{T}_s(w_l, t_m)\right) \tag{10}$$

where $R_\psi$ as a normalization constant has been defined in [28]. In this research, we used ISSWT to reconstruct the enhanced SSWT spectrum. To do this, first the original SSWT was applied to the ECG signal; then, the frequency ridge around HR frequency range was obtained. This ridgewas subjected to ISSWT and then SSWT was applied which is expected less noisy than the original ECG spectrum.

### 2.2.3. Iterative Pulse/R Peak Detection

One key to reliable ECG analysis is to do a robust peak detection. The proposed method here to extract R peaks of the ECG signals is based on an iterative algorithm. The iterative algorithm aims to combine the information from R peaks estimated in temporal domain and frequency ridge estimated from the time-frequency spectrum of the ECG. This has been done by comparing the estimated heart rate from the estimated ECG R peaks and frequency ridge with maximum energy around the heart rate frequency. One issue in comparing such HR frequencies from these two techniques (time-domain and time-frequency domain) is sampling intervals, and consequently the number of samples, which makes pairwise comparison difficult.

To illustrate the difference in sampling intervals, first consider that SSWT has been applied to the ECG signal, and that the frequency ridge (presenting HR frequency) with the highest energy in the range of valid heart rate has been derived. This will produce a new time-series corresponding to instantaneous frequencies. For this new time-series, the sampling interval is fixed as 1/fs where fs is the sampling frequency. The length of this time-series is equal to the number of ECG signal samples. On the other hand, the difference in consecutive R peaks is considered to derive a new time-series presenting HR frequency (or heart rate variability). The sampling interval is simply the difference in the sample numbers between R peaks divided by fs as diff(R_peaks_samples)/fs.The length of this time-series is equal to the number of R peaks minus one. Suppose there are 30 R peaks detected in a time window of 5000 samples. For the first time-series presenting HR derived from SSWT, the length of the time-series is 5000; for the second one, the length is 29 samples.

To compare these two time-series, a uniform sampling is required. For our experiment, the maximum frequency of 2 Hz for heart rate has been considered; therefore, the resampling of HR frequencies to 4 Hz signal satisfies the Nyquist sampling criterion. Once both time-series are resampled to a fixed rate, an absolute error of difference measure can be calculated. Such a measurement of error has been used to indicate the reliability of the R peak detection algorithm (temporal domain) compared with time-frequency spectral analysis. Calculated error might not be zero in the case of having a highly reliable R peak detection method, as the resampled signals cannot be overlapped thoroughly; however, the error has been used to tune the peak detection parameter (such as thresholds), and therefore, discard invalid detected R peaks to increase the reliability of and generalize the pulse peak of the detection algorithm.

The proposed iterative peak detection algorithm is summarized in Algorithm 1. The algorithm includes applying a simple peak detection algorithm with a threshold as input. For a set of different thresholds (**t**), the R peak detection has been applied and two time-series were derived to present HR frequency ($f_1$ from SSWT transform and $f_2$ as from detected R peaks). For various threshold levels, $f_2$ has been calculated and an absolute difference error of $f_1$ and $f_2$ has been obtained. Finally, a set of R peaks with minimum error have been selected as the final selected R peaks. If the agreement between two methods resulted in production of a much smaller number of R peaks or a sparse SSWT

spectrum, the ECG segment was marked as noisy with unreliable pulse peaks and not included in the analysis.

---

**Algorithm 1:** Iterative pulse peak detection.

---

- **Input ECG: s** $\leftarrow$ input ECG
- Set sampling frequency $fs$
- Set a vector of thresholds **t** for detecting the pulse peaks
- $[W]$ = WSST(**s**, $fs$) - SSWT transform
- $\mathbf{f}_1$ = TFRIDGE($W$) - frequency ridge in the range of $[0.5 - 2]$ Hz
- $\hat{\mathbf{f}}_1$ = INTERPRR($\mathbf{f}_1$...4...) - Resample frequencies into 4 Hz
- **for** *thr* **in t**
    $[p]$ = PEAKDET(**s**, *thr*)
    $\mathbf{f}_2$ = $fs$/DIFF(p) - HR frequency
    $\hat{\mathbf{f}}_2$ = INTERPRR($\mathbf{f}_2$...4...) - Resample frequencies into 4 Hz
    error(*thr*) = ABS($\hat{\mathbf{f}}_2$-$\hat{\mathbf{f}}_1$)
- Find set of pulse peaks associated with minimum error

---

To further illustrate the algorithm's performance, an example is shown in Figure 1. In this figure, nine different threshold levels have been shown in an increasing order from the top to bottom. In the right column of Figure 1, estimated HR frequency from R peaks and SSWT transform are overlaid after resampling into 4 Hz. In Figure 1b1–b5, overestimated R peaks have created large variations in derived HR frequency, while in Figure 1b7–b9, there exist underestimated R peaks which have also created large variations. The plot in Figure 1b6 corresponds to the minimum error between HR frequency derived from R peaks and SSWT and represents the most reliable detected R peaks on ECG signal. This plot has been shown in Figure 2 along with its estimated SSWT.
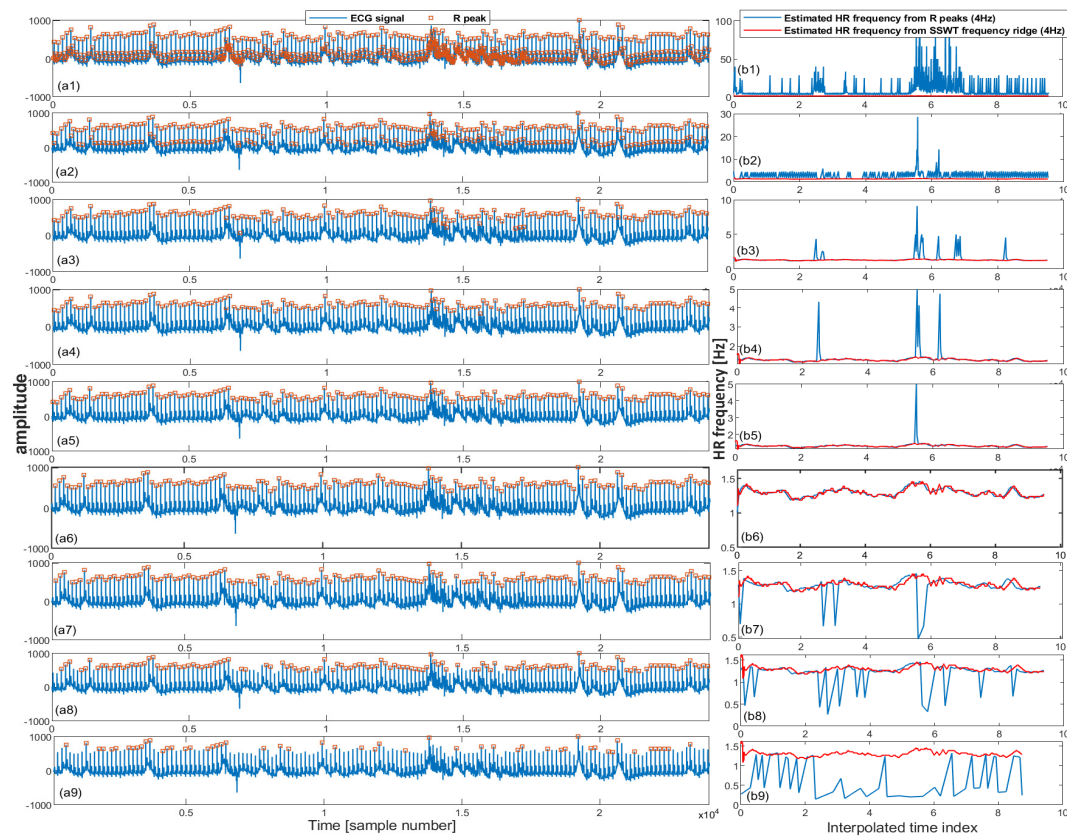
**Figure 1.** In each row, selected ECG segments and detected R peaks are shown on the left. In the right, the estimated HR frequency from the detected R peaks and synchrosqueezed wavelet transform (SSWT) are shown. The threshold to detect R peaks increases from top to bottom; (**b1**–**b5**) presents overestimated R peaks; (**b6**) presents the correctly detected R peaks hugely overlapped with the estimated HR frequencies; (**b7**–**b9**) presents under estimated R peaks.

Once R peaks are reliably detected, the ECG features can be extracted. Here, we have used two types of ECG features. The first set of features is directly derived from the detected R peaks. The second set of features is derived from respiratory modulation. The respiratory modulation is a time-series derived from R peak amplitudes. These two sets are features are explained in the following subsection.
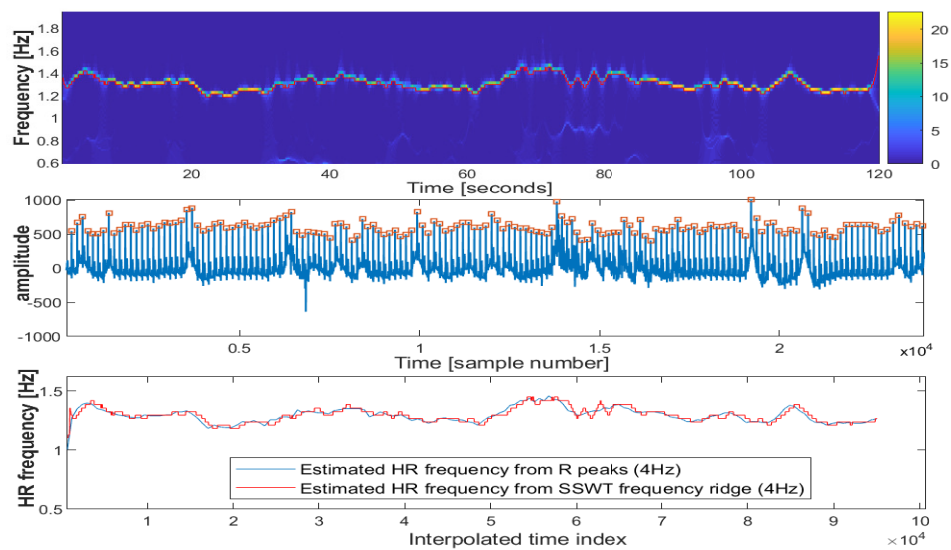
**Figure 2. Top**: estimated spectrum using SSWT. **Middle**: segmented ECG signal. **Bottom**: estimated HR frequency from R peaks and SSWT (resampled into 4 Hz) overlain.

### 2.2.4. ECG Features

ECG features extracted from pulse peaks include maximum, minimum and mean differences of consecutive R peaks. This set of features provides the timing of R peaks and is partly related to heart rate variability and one type of respiratory modulation (frequency modulation) [30]. The second set of ECG features is derived from the respiratory modulation. There are three respiratory modulations (frequency, amplitude and intensity) introduced in the literature to be derived from ECG R peaks to present respiratory [30,31]. Here, we calculated the amplitudes of R peaks to derive respiratory amplitude modulation. Then, maximum, minimum and mean amplitudes of derived time-series were calculated. Finally, SSWT was applied to the derived respiratory amplitude and the instantaneous frequencies related to the breathing range [0.05–1] Hz were estimated. Then, maximum, minimum, mean and standard deviation values of instantaneous frequencies were estimated. A summary of ECG features (along with EMG features) is shown in Table 1.

### 2.3. A Multimodal Deep Learning Neural Network for Sleep Disorder Detection

The EMG can be analyzed by directly extracting the raw data features from signal samples, while ECG should be analysed by applying the proposed techniques in Section 2.2 for reliable extraction of pulse peaks followed by derivation of respiratory modulation. Then, the features extracted from the EMG and ECG (Table 1) are both used to train a deep learning neural network (DNN). Using the proposed DNN, we obtained a performance of 72% accuracy, 57% F1 score, 53% precision and 62% recall. The accuracy was significantly better than that of the second best approach test. A concern when designing the architecture is that the deep neural network should consider that the ECG features summarize a full observational period of a trial (120 s), while for the EMG, each observation represents a sliding window of 2 s each over a trial 24 s. In terms of deep learning nomenclature, this means that the tensors are of different shapes (dimensions). An architecture that takes both temporal and structured data tensors and is able to integrate them into the neural network is a desirable option. Our proposed best performing model is a deep learning architecture wherein the temporal variations of the EMG features can be also accommodated. This is achieved by adding a merging layer that unites the input layer from the ECG with the output from the recurrent layers that takes the temporal EMG features as inputs, as shown in Figure 3. The layers of deep learning architecture are detailed as:

- **Input Layer EMG:** 5×3 tensor of the feature layer; the first dimension is the statistical parameters as presented in Table 1 and the second dimension is temporal window of the trial.
- **Input Layer ECG:** 1×10 tensor of statistical parameters of the ECG extracted from Table 1.
- **Two hidden RNN layers for EMG input:** Two hidden RNN layers with five units each.
- **Fully connected layers for ECG input (FC1):** A fully connected layer (FC) that performs a linear transformation with a weight matrix and bias; an input length of 10 and a number of hidden units of 20 with a rectifier linear unit (ReLU) activation function.
- **Merging Layer:** A layer that concatenates output tensors from the RNN layers and FC1 as a 1 dimensional tensor 1×25.
- **Drop-out layer:** A drop out layer with attention weights set to 0.4.
- **Fully connected layer (FC1):** A fully connected layer that applies a linear transformation (i.e., weight matrix and bias) with 25 inputs and 15 hidden units with ReLU activation.
- **Fully connected layer (FC2):** A fully connected layer that applies a linear transformation with 15 inputs and 15 hidden units with ReLU activation.
- **Fully connected layer (FC3):** A fully connected layer that applies a linear transformation with 15 inputs and 10 hidden units with ReLU activation.
- **Fully connected layer (FC4):** A fully connected layer that applies a linear transformation with 15 inputs and four hidden units.
- **Softmax layer (SF):** This layer re-scales between 0 and 1, with each output tensor element ($x_i$) from the fully connected layer FC4 as $x_i = exp(x_i) \backslash \sum_j exp(x_j)$.
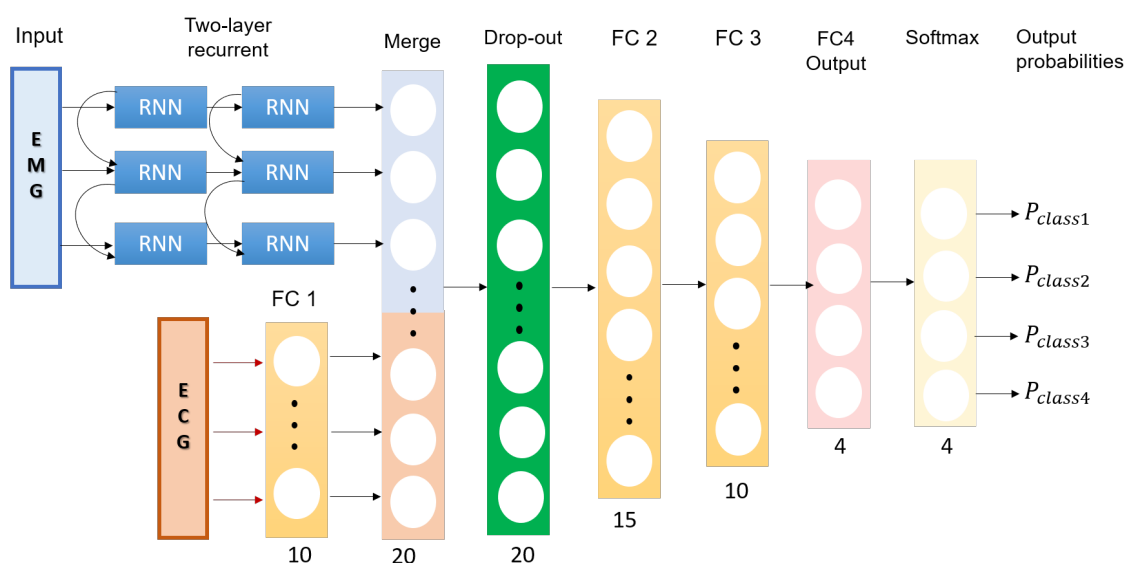


**Figure 3.** Architecture schema of the proposed multimodal EMG-ECG DNN.

The criterion used for this multi-class problem is the negative log likelihood loss. An adaptive learning rate optimizer (ADAM) is used to train the neural network. The initial layers of the DNN architecture were designed to incorporate the temporal information provided by EMG alongside the ECG information. That is, keeping as much information as possible from the extracted electrophysiology signal markers. Beyond the merging layer, the number of FC layers has been decided based on a simple hit-or-miss approach. That is, we gradually add layers until no further improvement can be achieved. For the selection of the hyper-parameters, including the number of hidden units for each FC layer, the drop out rate and the learning rate for ADAM, we employed a tree-structured Parzen estimator search algorithm (Bergstra, 2011). For the pruning and early termination of bad trials, we used a custom termination condition that checks that the loss reduction is larger than 10e-3 every 100 epochs during the training loop, and otherwise terminates it.

## 3. Results

A set of 76 patients including healthy subjects (33) and patients with OSA only (25), and both OSA and RLS (18), was used in [21] for sleep scoring. This dataset was extended to include patients with RLS only. Then, a subset of this extended dataset was selected using 10 subjects from each group to evaluate our developed framework. Therefore, in this research, a dataset of 40 subjects, including a set of healthy subjects (10) and patients with OSA only (10), RLS only (10) and both OSA and RLS (10) has been used.

One ECG channel and an EMG (attached to leg) were used to evaluate the proposed framework. The sampling frequencies of both ECG and EMG signals were fixed at 200 Hz. Both ECG and EMG signals are divided into 120 s data segments. Both signals were sliding every 60 s (50% overlap). The EMG signal aligned with ECG segment, is further divided into 20 s window size sliding every 2 s (90% overlap) to extract EMG features. This creates three values for each feature of the EMG signal (five features summarized in Table 1) and 15 features for each EMG window. However, the ECG signal is evaluated across 120 s to extract 10 features, summarized in Table 1, in each ECG window. Due to continuous recording of the signals, those signal segments corrupted with noise and artifacts were discarded. In total, the numbers of observations we obtained for the groups were: (4378~33%) for healthy subjects, (4590~34%) for patients with OSA only, (2225~17%) for RLS only and (2153~16%) for both OSA and RLS.

In this section we show the results from the recognition of the four subject groups by several state-of-the-art classifiers and the proposed method. The recognition was performed across subjects (i.e., cross-subject), wherein the training data did not belong to the same subject when the test was performed. This recognition scenario is rather complex because of the subject variability and symptomatic uncertainties; however, it is more desirable in clinical terms for real applications.

The methods used for benchmark purposes are a vanilla multi-layer perceptron (MLP) with 15 units in the hidden layer, in which parameters are trained using back-propagation with the Limited-memory BFGS solver and default settings of scikit-learn [32]; a support vector machine (SVM) using the libsvm package [33] with l2 penalty, hinge loss and multiclass one-vs-rest (OVR); a linear support vector machine (LSVM) using the liblinear library [34] with l2 penalty, hinge loss and OVR; random forest (RF) [35] made of 10 estimators and using gini as a split criterion and the default setup in scikit-learn; a k-nearest neighbor (KNN) [36] with number of neighbors set to 5 and default specifications [32]; XGBoost (XGB) using the dmlc XGBoost package [37] and its recommended settings; and AutoKeras (autok) [38] an AutoML approach based on neural architecture searching that tunes its architecture and hyper-parameters itself.

The patient-group recognition was evaluated in a cross-subject recognition scenario for all methods. All the tests were evaluated using 10-fold cross-validation. Examining the results of the cross-subject recognition scenario (Table 2 and Figure 4) we can see that the performance results are not extremely high but highly promising for such a limited sample size (10 samples for each group only). In a cross-subject recognition, our proposed multimodal DNN method performance is above most of the other classifiers and it outperforms XGBoost for most subjects (Figure 4). XGBoost is well known for beating many classifiers in machine learning competitions (such as Kaggle), including deep learning approaches [39]. Statistical differences in the performance statistics were found between the proposed DNN method and the rest. It is also worth noting that the AutoML approach tested, AutoKeras [38], that tries to automatically find an optimal DNN architecture, performs very poorly. An accuracy lower than 80% in this complex clinical scenario is expected because the symptomatic uncertainty comes into play. Some interesting performance patterns may begin to be appreciated in our study. In this cross-subject scenario our proposed method significantly outperforms all the others; in particular, the overall recall of our proposed method differs positively by a wide margin. The abilities of deep learning for transferring knowledge from one domain to another and integrating variable information are well known. Therefore, our intuition is that if the sample size increases, the performance of our proposed DNN method will increase with respect to other methods.

**Table 2.** Cross-subject results.

| Metric | Proposed | MLP | SVM | LSVM | RF | KNN | XGB | AUTOK |
|--------|----------|-----|-----|------|----|----|-----|-------|
| F1 Score | **0.57** ± 0.08 | 0.38 ± 0.07 | 0.36 ± 0.07 | 0.32 ± 0.09 | 0.37 ± 0.06 | 0.36 ± 0.06 | 0.44 ± 0.11 | 0.11 ± 0.07 |
| Accuracy | **0.72** ± 0.09 | 0.48 ± 0.09 | 0.48 ± 0.07 | 0.45 ± 0.1 | 0.46 ± 0.07 | 0.42 ± 0.05 | 0.5 ± 0.1 | 0.21 ± 0.1 |
| Precision | **0.53** ± 0.08 | 0.41 ± 0.07 | 0.43 ± 0.11 | 0.37 ± 0.11 | 0.43 ± 0.1 | 0.37 ± 0.06 | 0.46 ± 0.12 | 0.09 ± 0.07 |
| Recall | **0.62** ± 0.09 | 0.4 ± 0.08 | 0.39 ± 0.07 | 0.36 ± 0.08 | 0.38 ± 0.06 | 0.36 ± 0.05 | 0.45 ± 0.1 | 0.27 ± 0.07 |
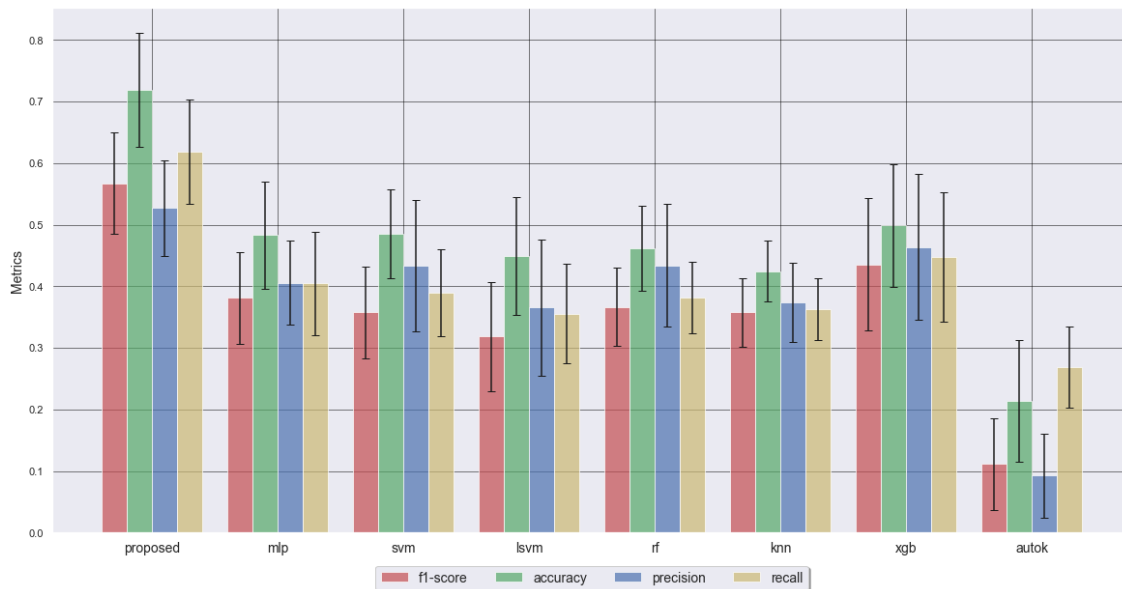


**Figure 4.** Results comparison of f-scores for the across-subject model.

## 4. Discussion and Conclusions

In this research, concurrent analysis of ECG and EMG signals have been used for recognition of breathing and movement-related disorders. Our research has proposed a set of markers from highly-noisy electrophysiological data that can be useful for the recognition of disease symptoms. Their adequacy was evaluated using a set of different machine learning classifiers. Furthermore, a DNN architecture that yields competitive and promising results by using and integrating this multimodal data was also presented. The proposed iterative peak detection algorithm can be applied for processing other signals in various applications, not just for musculoskeletal but for inflammatory disorders too [40]. The obtained accuracy in this study is very encouraging for continuing the research towards intelligent diagnoses and continuous monitoring of multiple sleep disorders. It should be noted that the accuracy is obtained by considering thousands of observations where some observations do not reflect sleep disorder symptoms, whereas they are included in determination of classification accuracy. In future work, a detailed analysis may be performed by accurate identification of sleep disorder symptoms only. However, this is very difficult, since it involves labeling hours of continuous recording of sensor data. Our method can be used for more detailed data analysis of patients diagnosed with sleep disorder symptoms by further subject-specific analysis of ECG and EMG patterns. For future works we aim at expanding the number of bio-signals employed in the recognition. The signal processing techniques developed for ECG analysis can be adapted for the analysis of PPG signals recorded in very unobtrusive ways.

**Author Contributions:** Conceptualization, D.J., J.A.P., A.P., and S.S.; Methodology, D.J., and J.A.P.; Software, D.J., J.A.P., and M.K.; Formal Analysis, D.J., J.A.P. and M.K.; Data Curation, O.V., and J.K.; Writing–Original Draft Preparation, D.J., J.A.P.; Writing–Review & Editing, D.J., J.A.P, M.K., O.V., J.K., A.P., and S.S.; Supervision, A.P. and S.S. All authors have read and agreed to the published version of the manuscript.

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

## References

1.  Fang, Y.; Jiang, Z.; Wang, H. A novel sleep respiratory rate detection method for obstructive sleep apnea based on characteristic moment waveform. *J. Healthc. Eng.* **2018**, *2018*, 1902176.

2.  Khandoker, A.H.; Palaniswami, M.; Karmakar, C.K. Support vector machines for automated recognition of obstructive sleep apnea syndrome from ECG recordings. *IEEE Trans. Inf. Technol. Biomed.* **2009**, *13*, 37–48.

3.  de Weerd, A.W.; Rijsman, R.M. Activity patterns of leg muscles in periodic limb movement disorder. *J. Neurol. Neurosurg. Psychiatry* **2004**, *75*, 317–319.

4.  Roux, F.J. Restless legs syndrome: Impact on sleep-related breathing disorders. *Respirology* **2013**, *18*, 238–245.

5.  Ferreri, F.; Rossini, P.M. Neurophysiological investigations in restless legs syndrome and other disorders of movement during sleep. *Sleep Med.* **2004**, *5*, 397–399.

6.  Hamilton, G.; Meredith, I.; Walker, A.; Solin, P. Obstructive sleep apnea leads to transient uncoupling of coronary blood flow and myocardial work in humans. *Sleep* **2009**, *32*, 263–270.

7.  Takama, N.; Kurabayashi, M. Influence of untreated sleep disordered breathing on the long-term prognosis of patients with cardiovascular disease. *Am. J. Cardiol.* **2009**, *103*, 730–734.

8.  Abdulla, S.; Diykh, M.; Luaibi Laft, L.; Saleh, K.; Deo, R.C. Sleep EEG signal analysis based on correlation graph similarity coupled with an ensemble extreme machine learning algorithm. *Expert Syst. Appl.* **2019**, *138*, 112790.

9.  Diykh, M.; Li, Y.; Abdulla, S. EEG sleep stages identification based on weighted undirected complex networks. *Comput. Methods Programs Biomed.* **2020**, *184*, 105116.

10. Saha, S.; Bhattacharjee, A.; Fattah, S.A. Automatic detection of sleep apnea events based on inter-band energy ratio obtained from multi-band EEG signal. *Healthc. Technol. Lett.* **2019**, *6*, 82–86.

11. Sugi, T.; Kawana, F.; Nakamura, M. Automatic EEG arousal detection for sleep apnea syndrome. *Biomed. Signal Process. Control* **2009**, *4*, 329–337.

12. Moridani, M.K.; Heydar, M.; Jabbari Behnam, S.S. A Reliable Algorithm Based on Combination of EMG, ECG and EEG Signals for Sleep Apnea Detection: (A Reliable Algorithm for Sleep Apnea Detection). In Proceedings of the 2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI), Tehran, Iran, 28 February–1 March 2019; pp. 256–262.

13. Mosquera-Lopez, C.; Leitschuh, J.; Condon, J.; Hagen, C.C.; Rajhbeharrysingh, U.; Hanks, C.; Jacobs, P.G. Design and Evaluation of a Non-Contact Bed-Mounted Sensing Device for Automated In-Home Detection of Obstructive Sleep Apnea: A Pilot Study. *Biosensors* **2019**, *9*, 90.

14. Andreu-Perez, J.; Cao, F.; Hagras, H.; Yang, G.Z. A self-adaptive online brain-machine interface of a humanoid robot through a general type-2 fuzzy inference system. *IEEE Trans. Fuzzy Syst.* **2016**, *26*, 101–116.

15. Bsoul, M.; Minn, H.; Tamil, L. Apnea MedAs-sist: Real-time sleep apnea monitor using single-lead ECG. *IEEE Trans. Inf. Technol. Biomed.* **2011**, *15*, 3, 416–427.

16. Jarchi, D.; Sanei, S.; Prochazka, A. Detection of sleep apnea/hypopnea events using synchrosqueezed wavelet transform. In Proceedings of the ICASSP 2019—2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; pp. 1199–1203.

17. Jarchi, D.; Sanei, S. Derivation of respiratory effort from photoplethysmography. In Proceedings of the 2019 27th European Signal Processing Conference (EUSIPCO), A Coruna, Spain, 2–6 September 2019; pp. 1–5.

18. Shokrollahi, M.; Krishnan, S. A Review of Sleep Disorder Diagnosis by Electromyogram Signal Analysis. *Crit. Rev. Biomed. Eng.* **2015**, *43*, 1–20.

19. Podlipnik, M.; Sarc, I.; Ziherl, K. Restless leg syndrome is common in patients with obstructive sleep apnoea. *ERJ Open Res.* **2017**, *3*, 20.

20. Prochazka, A.; Kuchynka, J.; Vysata, O.; Cejnar, P.; Valis, M.; Marik, V. Multi-class sleep stage analysis and adaptive pattern recognition. *Appl. Sci.* **2018**, *8*, 697.

21. Prochazka, A.; Kuchynka, J.; Vysata, O.; Yadollahi, M.; Sanei, S.; Marik, V.; Valis, M. Sleep scoring using polysomnography data features. *Signal Image Video Process.* **2018**, *12*, 1043–1051.

22. Rostaghi, M.; Azami, H. Dispersion Entropy: A measure for time-series analysis. *IEEE Signal Process. Lett.* **2016**, *23*, 610–614.

23. Sanei, S.; Lee, T.K.M.; Abolghasemi, V. A new adaptive line enhancer based on singular spectrum analysis. *IEEE Trans. Biomed. Eng.* **2012**, *59*, 428–434.

24. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Amer. J. Physiol. Heart Circ. Physiol.* **2000**, *278*, H2039–H2049.

25. Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*, 1–4.

26. Shannon, C.E. A mathematical theory of communication. *Bell Syst Tech. J.* **1948**, *27*, 623–656.

27. Daubechies, I.; Lu, J.; Wu, H. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Appl. Comput. Harmon. Anal.* **2011**, *30*, 243–261.

28. Thakur, G.; Brevdo, E.; Fuckar, N.S.; Wu, H.T. The Synchrosqueezing algorithm for time-varying spectral analysis: Robustness properties and new paleoclimate applications. *IEEE Trans. Signal Process.* **2011**, *93*, 1097–1094.

29. Carmona, R.A.; Wang, W.L.; Torresani, B. Characterization of signals by the ridges of their wavelet transforms. *IEEE Trans. Signal Process.* **1997**, *45*, 2586–2590.

30. Charlton, P.H.; Bonnici, T.; Tarassenko, L.; Clifton, D.A.; Beale, R.; Watkinson, P.J. An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram. *Physiol. Meas.* **2016**, *37*, 610–626.

31. Charlton, P.H.; Birrenkott, D.A.; Bonnici, T.; Pimentel, M.A.; Johnson, A.E.; Alastruey, J.; Tarassenko, L.; Watkinson, P.J.; Beale, R.; Clifton, D.A. Breathing rate estimation from the electrocardiogram and photoplethysmogram: A Review. *IEEE Rev. Biomed. Eng.* **2018**, *11*, 2–20.

32. Varoquaux, G.; Buitinck, L.; Louppe, G.; Grisel, O.; Pedregosa, F.; Mueller, A. Scikit-learn: Machine learning without learning the machinery. *GetMobile: Mob. Comput. Commun.* **2015**, *19*, 29–33.

33. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *6*, 1–27.

34. Fan, R.E.; Chang, K.W.; Hsieh, C.J.; Wang, X.R.; Lin, C.J. Liblinear: A library for large linear classification. *J. Mach. Learn. Res..* **2008**, *9*, 1871–1874.

35. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.

36. Short, R.; Fukunaga, K. The optimal distance measure for nearest neighbor classification. *IEEE Trans. Inf. Theory.* **1981**, *27*, 622–627.

37. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.

38. Jin, H.; Song, Q.; Hu, X. Auto-keras: An efficient neural architecture search system. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; pp. 1946–1956.

39. Nielsen, D. Tree Boosting with XGBoost-Why does Xgboost Win "Every" Machine Learning Competition? Master's Thesis, NTNU, Trondheim, Norway, 2016.

40. Andreu-Perez, J.; Garcia-Gancedo, L.; McKinnell, J.; Van der Drift, A.; Powell, A.; Hamy, V.; Keller, T.; Yang, G.Z. Developing fine-grained Actigraphies for rheumatoid arthritis patients from a single accelerometer using machine learning. *Sensors* **2017**, *17*, 2113.