Deep Learning Analysis for Estimating Sleep Syndrome Detection Utilizing the Twin Convolutional Model FTC2

Tim Cvetko and Tinkara Robek

Gimnazija Brešice, Brežice, Slovenia
E-mail: cvetko.tim@gmail.com; tinkara.robek@gmail.com

Abstract. Manual sleep stage scoring is frequently performed by sleep specialists by visually evaluating the patient’s neurophysiological signals acquired in sleep laboratories. This is a difficult, time-consuming, & laborious process. Because of the limits of human sleep stage scoring, there is a greater need for creating Automatic Sleep Stage Classification (ASSC) systems. Sleep stage categorization is the process of distinguishing the distinct stages of sleep & is an important step in assisting physicians in the diagnosis & treatment of associated sleep disorders. In this research, we offer a unique method & a practical strategy to predicting early onsets of sleep disorders, such as restless leg syndrome & insomnia, using the Twin Convolutional Model FTC2, based on an algorithm composed of two modules. To provide localized time-frequency information, 30 second long epochs of EEG recordings are subjected to a Fast Fourier Transform, & a deep convolutional LSTM neural network is trained for sleep stage categorization. Automating sleep stages detection from EEG data offers a great potential to tackling sleep irregularities on a daily basis. Thereby, a novel approach for sleep stage classification is proposed which combines the best of signal processing & statistics. In this study, we used the PhysioNet Sleep European Data Format (EDF) Database. The code evaluation showed impressive results, reaching accuracy of 90.43, precision of 77.76, recall of 93.32, F1-score of 89.12 with the final mean false error loss 0.09. All the source code is available at https://github.com/timothy102/eeg.

Keywords: Sleep stage classification, algorithm design, onset detection, convolutional neural networks, signal processing, recurrent neural networks, fourier transform, CEEMDAN, preprocessing.

INTRODUCTION

Sleep stage classification is crucial for our day-to-day functioning & disease prevention, such as obesity, type 2 diabetes, & different psychiatric disorders that may develop in association with the shortage of sleep. Maintaining a healthy sleep schedule is significant for obtaining our natural circadian rhythm & the normal course of sleep. This must include REM (rapid-eye movement) & non-REM phases. NREM is specified by slow-wave activity (delta power). REM phase is characterized by the eye movements which we track with EOG (electrooculogram), low muscle tone which we track with EMG (electromyogram), low amplitude waves, & generally mixed frequency EEG (electroencephalogram). It is also a stage where dreams occur. The brain is as active during REM as during wakefulness, only in different parts (correspondingly, the prefrontal cortex, responsible for judgment & decision making is deactivated during both phases). The REM state is initialized by the neurons located in the brainstem. The process of recording sleep in the laboratories for the purposes of examining its structure (mainly in terms of slow-wave activity) is known as polysomnography. Polysomnograms can show the diagnostics of various disorders during sleep. They are commonly used for identifying disorders such as obstructive sleep apnea, excessive daytime sleepiness, & insomnia. Polysomnograms record numerous types of information: EEG (electroencephalogram; for recording brain waves), EOG (electrooculogram; for recording eye movements), EMG (electromyogram; recording of the electrical activity of muscle surfaces), EKG (electrocardiogram). Combining the gathered data allows us to identify the individual’s stage of sleep. For the purposes of recognizing sleep apnea, it is also required to record snoring, nasal & oral airflow, movement of the abdomen & chest, & pulse oximetry. There are multiple types of waves measured in EEG which tell whether the person is asleep or awake & what stage they are in. Alpha activity (frequency:
8 to 13 Hz) is present in the occipital cortex when a person is resting with their eyes closed & goes away when they wake up or fall asleep. Theta activity (4 to 7 Hz) is the most common frequency during sleep. Delta activity is known as slow-wave activity (0.5 to 2/4 Hz) with higher amplitudes & is mostly present in the frontal areas of the cortex. Sleep spindles (11 to 16 Hz) normally indicate stage 2 of NREM sleep, with a duration of 2–3 seconds. K complexes are most common in stage 2 of NREM Their amplitude is not specified, while the time span must be at least 0.5 seconds. Sleep stages are regulated by individual neurotransmitters.

Wakefulness is a heterogeneous state, maintained by neurotransmitters acetylcholine, serotonin, norepinephrine, histamine, dopamine, hypocretin, & glutamate. NREM is initialized by GABA & adenosine, while acetylcholine also sends signals for REMonset. Interactions between neurotransmitters determine behavioral states at any given time. During sleep, they define muscle tone, eye movements & EEG activity. The same neurotransmitter can have opposite effects during wakefulness & sleep & in different brain regions. The preoptic area of the brain (anterior hypothalamus) is known to act as a sleep center, while the posterior hypothalamus acts as a wakefulness center. Electrical stimulation of the preoptic area, therefore, has the ability to initiate sleep. Apart from the objective reasons for not sleeping long enough (the recommended 8 hours, for obtaining full sleep cycles), there are also multiple sleeping disorders present in today’s society. Sleeping disorder is usually a composition of multiple neurological modifications that may be caused by external or internal factors. These factors include our lifestyle, genetics, associated diseases, psychiatric disorders, & so on. Disordered sleeping is alarmingly common in modern society. In the field of various disorders, the most common persist to stay the following: insomnia, sleep-disordered breathing (obstructive sleep apnea, hyperventilation), central disorders of hypersomnolence (narcolepsy, idiopathic hypersomnia, Kleine-Levin Syndrome), circadian rhythm disorders, parasomnias (sleepwalking, terrors, sleep eating disorders, sleep enuresis) & sleep-related movement disorders (restless leg syndrome, periodic limb movement disorder). Insomnia & other disorders that result in trouble sleeping are often associated with multiple chronic health conditions. They have an impact on blood pressure & blood sugar levels, therefore they put us at risk of coronary diseases & metabolic imbalances. Middle-aged / older parts of the population & obese individuals are at greater risk of developing these issues. In addition, sex also plays a role in developing some sleeping disorders males are more likely to develop obstructive sleep apnea, while women are predisposed to insomnia. The first step for identifying problems with sleep is by clinical approach with ESS (Epworth Sleepiness Scale) [1]. Insomnia can arise as chronic insomnia or in a short-term form. Its typical symptoms are problems initiating or maintaining sleep, non-restorative sleep, & waking too early which result in daily distress. These issues occur due to the brain’s arousal systems which are not turning off. It is diagnosed in a laboratory with polysomnography. It can be a cause for endothelial dysfunction, oxidative stress, inflammation, hypertension, & even stroke in severe cases. Normally, it is treated with CPAP (continuous positive airway pressure).

We intend to use these findings along with the FTC2 model for sleep staging predictions to predict onset of sleep syndromes. We firmly believe these methods strongly correlate to the modern sleep issues thus enhancing the medical procedure & diagnosis.

**METHODOLOGY**

For the purpose of this paper, we have used the following methods: Ceemdan, Fast Fourier Transform, Convolutional Neural Networks, LSTM (Long Short-Term Networks), Savitzky-Golay filter, Welch Power Spectrum.

**Event Detection During Sleep**

Although detecting periods of raised or lowered activity during sleep is an appealing task, it is rather interesting to detect transient, short signals. Furthermore, we can identify events such as QRS complexes in the ECG or EEG events such as slow waves (single waves of around 0.5–2 Hz with a high amplitude of about 75 uV), sleep spindles, & other sleep-related events during sleep (transient waxing & waning events of about 0.5–2 seconds duration & 15–50 uV maximal amplitude). For the sake of this study & simplicity, we do not distinguish between slow waves, K-complexes, & slow oscillations, but instead group them all together.

**Detailed Description of the Methodology**

**CEEMDAN preprocessing**

We used the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) as a noise aided EMD approach for the preprocessing. In recent years, a number of signal processing approaches have been created. The Fast Fourier Transform is the foundation of standard signal processing methods. However, this approach cannot concurrently yield time-domain & frequency-domain evaluated findings. Many time-frequency analysis approaches, such as the Wigner-Ville distribution (WWD) introduced in, have been used to diagnose vibration signals [2]. This approach computes the analytic signal corresponding to the input signal, constructs a weighted kernel function, then analyses the kernel using a Discrete Fourier Transform (DFT). According to [3], when evaluating the analytic signal needed by the algorithm, the time domain definition implemented as a finite impulse response (FIR) filter is more practical & efficient than the frequency domain definition. The wavelet transform ([4, 5]) utilizes the Fast
Fourier Transform to extract frequency based wavelets. One of the most potent signal processing techniques for defect diagnostics is empirical mode composition. The intrinsic mode function is derived from a signal’s local characteristic timings & can deconstruct the signal into a series of full & nearly orthogonal components (IMFs). The IMFs represent the signal’s inherent oscillation mode & act as basis functions, which are defined by the signal rather than by predetermined kernels. As a result, it is a self-adaptive signal processing approach ideal for nonlinear & nonstationary processes, as well as for defect feature extraction in spiral-bevel gears. According to [5], CEEMDAN has been demonstrated to be a successful approach for evaluating nonstationary signals that are followed by high noise. The raw vibration signal is first decomposed using the CEEMDAN technique to get a set of intrinsic modal functions (IMFs). Based on the correlation coefficient, the IMFs that incorporated the most defect information were chosen as the ideal IMFs. Following that, the optimum IMFs’ permutation entropy values are computed. The overlapping parameter approach is used to optimise the two main parameters, embedding dimension & delay time, in order to achieve correct permutation entropy values. The support vector machine (SVM) is employed as the classifier for failure mode detection in order to check the sensibility of the permutation entropy features, & the diagnostic accuracy can validate its sensibility.

FTC2 model architecture

The standard practice for image recognition, computer vision & more are deep convolutional neural networks. Models like, LeNet [7], AlexNet [8], ImageNet [9] & more work exclusively on 2D data. They are also referred to as 2D CNNs. With the same idea in mind, 1D CNN have been developed. Despite the obvious fact, that the 1D CNNs apply a 1D kernel to data, some of the advantages are: 1D CNNs don’t require heavy matrix operations (nor at the forward pass or the backpropagation). Potentially, what this means is that they are, compared to their 2D counterparts, computation less complex & can run on CPU hardware.

The FTC2 is passed through a Leaky Rectified Linear Unit for nonlinearity purposes. The input signal then passes a max-pooling layer with the stride of 2, following that is batch normalization layer & finally the Bidirectional layer. Each twin structure contains three such blocks. The outputs of the convolutional-LSTM blocks are then concatenated across axis 1 in order to be finally passed through a fully connected 64 unit dense layer. We use the softmax activation to receive the final scores for each sleep stage. The CNN consists of three different-sized filters with the idea in mind being that small filters extract temporal information & large filters to extract frequency information. The idea behind using variable filter widths originates from the signal processing field to have a trade-off between extracting time domain & frequency domain characteristics [6]. This allows you to benefit from both temporal & frequency domain data in the classification process. The feature map generated from the 1D concatenated output from the twin convolutional blocks are intertwined with long short-term memory (LSTM) units to capture the complex context dependencies between the inputs & the targets.

Utilizing the Conv-1D & the Bidirectional LSTM

The Conv-LSTM network combines the best of both worlds, one-dimensional convolutional layers handles spatial data using a kernel to train a set of filters that best represent the hyperplane & the LSTM which encodes temporal inputs using cell state memory. To put it differently, Conv-LSTM predicts the future state of a cell given vectors on a spatial grid & past states of neighbours. Convolutional-LSTM layer works by passing the 1D kernel over a set of training points & its result a.k.a feature map is passed onto the LSTM cell. In this project, Instead of using the typical LSTM cell, we used a bidirectional recurrent neural network. Put simply, this was enabled due to the limitations of LSTM having the limitation of being restricted from processing previous input state. Given the inputs $x = x_0, x_N$, which are single dimensional data of a 30-second clip, which should assure stability & little overlap, these inputs are passed into the twin architecture.

Loss Calculation

Sleep stage classification encounters the class imbalance issue which we have addressed accordingly. Such imbalance will have had severe tendency to bias the better represented part of the dataset. To alleviate this we have taken both the extended mean false error (MFE) & the mean squared false error (MSFE) for multiclass classification.

$$MFE_L = \frac{q}{N} \sum_{i=1}^{N} C(x_i - \hat{x})^2$$

Unsupervised Method for Detecting Onsets of Sleep Syndromes

Syndromes in question are restless leg syndrome (RLS), rapid Eye Movement (REM) sleep behaviour disorder, insomnia, sleep apnea & narcolepsy. In attempt to detect these complex syndromes through EEG data, we have come up with the following: Each of the syndromes will be followed with an explainable paragraph, respectively.

Restless leg syndrome

Lisuride, a dopaminergic medication, showed a rise in alpha, a decrease in beta, & delayed actions on brain function as evaluated by computerised EEG. It was
proposed that reverse EEG abnormalities might play a role in the pathophysiology of RLS. Changes in alpha activity generate long-lasting alpha arousal responses throughout the transition from alertness to sleep stage 1, & they continue to increase during sleep stage 2. This malfunction is most likely caused by a hereditary susceptibility of the EEG alpha rhythm & disinhibits the diencephalospinal dopamine pathway, mostly during sleep but also during alertness. Disinhibition creates a foundation for PLM activation, as well as unsettling feelings in the brainstem & a need to move, as well as motor restlessness in the cerebral cortex, most notably in the legs. All of these factors contribute to severe insomnia. Forced deviations from alpha to theta or beta activity are undesirable in RLS patients, & resting EEGs indicate a dopamine receptor-specific ‘individual sensitivity.’ This vulnerability is reduced following lisuride with appropriate CEEG adjustments. We have designed an unsupervised statistical inference model for approximating the restless leg syndrome in EEG. First, the Savitzky–Golay filter was applied to each 30-second window with the goal of smoothing the data, that is, increasing the accuracy of the data without altering the signal trend. Convolution is performed by fitting consecutive subsets of neighbouring data points with a low-degree polynomial using the linear least squares approach. The data was then subjected to an orthogonal discrete Fourier Transform to decompose functions based on space or time into functions based on spatial or temporal frequency. At last, the final RLS score is computed by applying the sigmoid activation function to the fourier-transformed data.

Sleep apnea

Carbon dioxide may accumulate in the circulation during sleep owing to pauses in breathing. Chemoreceptors detect it when it reaches a crucial threshold. These receptors provide a signal to the brain, causing the sleeping individual to wake up & breathe in air. As a result, a change in sleep phases occurs, causing fluctuations in the activity level of distinct frequency bands of the EEG signal. As a result, as compared to a full-b & EEG signal, more identifiable characteristics can be maintained in frequency...
band-limited signals for apnea diagnosis. As a result, instead of analysing full-b& EEG data for apnea event identification, features of band-limited signals are utilised. EEG signal is partitioned into five frequency bands [10] including delta (0.25 to 4 Hz), theta (4 to 8 Hz), alpha (8 to 12 Hz), sigma (12 to 16 Hz) & beta (16 to 40 Hz). Spectral filtering is done in fast Fourier transform domain to achieve this division.

In order to attain a relative score for approximating sleep apnea possibility, we have adopted the method proposed in [7]. This suggested technique introduces a subject-specific classification framework based on feature for the categorization of apnea & non-apnea occurrences. The feature is considered to be the inter-b& energy ratio of a band-limited EEG signal, which is believed to have discriminating characteristics for apnea & non-apnea episodes. Energy contents in various frequency bands alter dramatically during sleep apnea compared to non-apnea occurrences.

**REM behaviour sleep disorder**

Despite what was presented in [9], we have come about a rather simple, but clever way of detecting the REM instability score. A relative score was calculated by using a greedy algorithm for longest continuing sequence [10] to obtain a value that should correspond to the disturbance of REM sleep in an individual.

**Narcolepsy**

Patients with narcolepsy had lower alpha power, greater delta & theta power during alertness, & higher alpha & beta power during rapid eye movement (REM) sleep, according to [11]. The former two groups also had reduced sleep efficiency & a greater percentage of positive for REM-related symptoms than the other two groups. In light of this, we integrated the REM instability score & the acquired frequency range with wavelet transform on the initial & final 10% of sleep, suggesting worse sleep-to-wake transitions or vice versa.

**Insomnia**

Insomnia, often known as insomnia, is a sleep condition characterised by an inability to fall or stay asleep for as long as desired. Following the previous work done by Mohd Maroof Siddiqui, Geetika Srivastava, & Syed Hasan Saeed in [11]. Because the majority of EEG signals are constrained within the range of 25 Hz, each clipped signal is now preprocessed & then run through the Hanning window low pass filter to remove the high frequency components that eventually suggest noise. Hence, the filter based in FIR filter design of order 200 with cut off frequency of 25 Hz with shape of hanning window. At this stage, we adopt a method for extracting frequency length windows via the Welch Power Spectrum [12]. This score, obtained with softmax & the REM instability score combine for the insomnia score.

**Sleep depth**

Given a great sleep stage classifier, we concluded to not only move in the direction of detecting abnormalities, but to also quantify sleep depth. First, the raw records are pushed through a Savgol filter to noise filtering. Furthermore, a Fast Fourier Transform has given us a distribution of frequencies. At this stage, we evaluated the FFT distribution to obtain the sleep depth score along with the REM instability score.

<table>
<thead>
<tr>
<th>Syndromes</th>
<th>RLS disorder</th>
<th>Apnea</th>
<th>Narcolepsy</th>
<th>Insomnia</th>
<th>Sleep depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>177983</td>
<td>6674</td>
<td>8899</td>
<td>4449</td>
<td>48945</td>
</tr>
<tr>
<td>Percent</td>
<td>7.80</td>
<td>0.29</td>
<td>0.41</td>
<td>3.0</td>
<td>22</td>
</tr>
<tr>
<td>NHCS</td>
<td>7–10</td>
<td>0.5–1</td>
<td>0.5</td>
<td>3–7</td>
<td>10–30</td>
</tr>
<tr>
<td>Percent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**EXPERIMENTAL RESULTS**

**Data**

The Physionet Sleep-EDF dataset was utilised for the objectives of this study. The sleep-edf database comprises 197 PolySomnoGraphic sleep recordings from the whole night, including EEG, EOG, chin EMG, & event markers. Some recordings additionally include information about breathing & body temperature. The Sleep-EDF dataset comprises two studies: one on the effects of age on sleep in healthy people (SC = Sleep Cassette) & another on the effects of temazepam on sleep (ST = Sleep Telemetry). The dataset contains sleep recordings of whole-night polysomnograms (PSGs) with a sampling rate of 100 Hz. EEG (from the Fpz-Cz & Pz-Oz electrode sites), EOG, chin electromyography (EMG), & event markers are all included in each
record. Fewer records frequently include oro-nasal respiration & rectal body temperature. The hypnograms (sleep phases; 30-second epochs) were identified manually by well-trained personnel in accordance with the Rechtschaffen & Kales standard. Each level was assigned to a distinct class (stage). The classes include W, REM, N1, N2, N3, N4, M. Data syndrome comparison was done according to [13].

<table>
<thead>
<tr>
<th>Dataset</th>
<th>W</th>
<th>N1</th>
<th>N2</th>
<th>N3-N4</th>
<th>REM</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep-EDF-13</td>
<td>8285</td>
<td>2804</td>
<td>17799</td>
<td>5703</td>
<td>7717</td>
<td>42308</td>
</tr>
<tr>
<td>Sleep-EDF-18</td>
<td>65951</td>
<td>21522</td>
<td>96132</td>
<td>13039</td>
<td>25835</td>
<td>222479</td>
</tr>
</tbody>
</table>

It is worth noting that the sleep stages in the Physionet’s Sleep-EDF database are not normally distributed. There is a significantly higher number of W & N2 stages than others which is causing class imbalance. We have discussed this at Loss calculation.

The model was trained for 20 epochs. We used the Adam optimizer to minimize the MFE loss with a batch size of 16 & a learning rate alpha = 0.001. We also added the L2 regularization penalty with beta being = 0.001 to mitigate the overfitting. For this goal, we used Python as the programming language & TensorFlow as the framework.

By employing the Conv-1D & the Bidirectional LSTM twin network, we have successfully taken advantage of the temporal & spatial component of the signal processing nature. To reduce the class imbalances problem, we adopted the weighted MFE loss.

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
<th>Final MFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>90.43</td>
<td>77.76</td>
<td>93.32</td>
<td>89.12</td>
<td>0.09</td>
</tr>
</tbody>
</table>

CONCLUSION & FUTURE WORK

In this paper, we have successfully proposed a new deep learning approach to the challenging problem of sleep stage classification that is essential for further analysis of the human brain activity. We utilized a two-part convolutional – LSTM block along with a multi-layer perceptron to classify sleep stages from EEG data with the accuracy of. By incorporating a vastly important CEEM-DAN pre-processing method, we build a more stable & noise-free variant of our data, which enabled building an end-to-end trainable model for not only classifying sleep stages, but also detect early onset of frequently occurring sleep syndromes. Our algorithm takes advantage of both the spatial & temporal dimension of signal processing, as well as statistical analysis to come up with a score that represents the sleep stage the most. As a limitation, we would like to add that no data is known to be rare diluted patients. For future work, given proper data, we would like to do a follow-up. In the future, we intend to extend this using multimodal polysomnography to enhance the model’s performance & usability. Following our work, we will investigate how these syndrome predictions/scores align with actual scores & will hopefully serve as a professional service for automatic sleep stage & syndrome onset detection.

REFERENCES