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, and 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 and is an important step in assisting physicians in the diagnosis and treatment of associated sleep disorders. In this research, we offer a unique method and a practical strategy to predict early onset of sleep disorders, such as restless leg syndrome and 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 electroencephalogram (EEG) recordings are subjected to a fast Fourier transform, and a deep convolutional long short-term networks neural network is trained for sleep stage categorization. Automating sleep stage detection from EEG data offers a great potential to tackle sleep irregularities on a daily basis. Thereby, a novel approach for sleep stage classification is proposed, which combines the best of signal processing and statistics. In this study, we used the PhysioNet Sleep European Data Format (EDF) database. The code evaluation showed impressive results, reaching an accuracy of 90.43, precision of 77.76, recall of 93,32, F1 score of 89.12, and the final mean false error loss of 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 and prevention of diseases, such as obesity, type 2 diabetes, and 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 and the normal course of sleep. This must include rapid eye movement (REM) and non-REM phases. NREM is specified by slow-wave activity (delta power). REM phase is characterized by the eye movements that we track with electrooculogram (EOG), low muscle tone that we track with electromyogram (EMG), low amplitude waves, and generally mixed frequency electroencephalogram (EEG). It is also a stage where dreams occur. The brain is as active during REM as during wakefulness, only in different parts (correspondingly, the pre-frontal cortex, which is responsible

for judgment and 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 laboratories to examine 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, and insomnia. Polysomnograms record numerous types of information: EEG (for recording brain waves), EOG (for recording eye movements), EMG (for recording the electrical activity of muscle surfaces), and EKG (electrocardiogram). Combining the gathered data allows us to identify the individual's stage of sleep. To recognize sleep apnea, it is also required to record snoring, nasal and oral airflow, movement of the abdomen and chest, and pulse oximetry. There are multiple types of waves measured in EEG which tell whether the person is

asleep or awake and 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 and 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 and 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, and 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 such as acetylcholine, serotonin, nore-pinephrine, histamine, dopamine, hypocretin, and gluta-mate. NREM is initialized by GABA and adenosine, while acetylcholine sends signals for REM onset. Interactions between neurotransmitters determine the behavioral states at any given time. During sleep, they define muscle tone, eye movements, and EEG activity. The same neurotrans-mitter can have opposite effects during wakefulness and sleep and 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 rec-ommended 8 hours, for obtaining full sleep cycles), there are also multiple sleeping disorders present in today's soci-ety. Sleeping disorder is usually a composition of multiple neurological modifications that may be caused by exter-nal or internal factors. These factors include our lifestyle, genetics, associated diseases, and psychiatric disorders. 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, hyper-ventilation), cen-tral disorders of hypersomnolence (narcolepsy, idiopathic hypersomnia, Kleine–Levin syndrome), circadian rhythm disorders, parasomnias (sleepwalking, terrors, sleep eating disorders, sleep enuresis), and sleep-related movement disorders (restless leg syndrome, periodic limb movement disorder). Insomnia and other disorders that result in trouble sleeping are often associated with multiple chronic health conditions. They have an impact on blood pressure and blood sugar levels; therefore, they put us at risk of coronary diseases and metabolic imbalances. Middle-aged/older population and 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 Epworth Sleepiness Scale (ESS) [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, and waking too early, which result in daily distress. These issues occur due to the brain's arousal systems that are not turning off. It is diagnosed in a laboratory with polysomnography. It can be a cause for endothelial dysfunction, oxidative stress, inflammation, hypertension, and even stroke in severe cases. Normally, it is treated with continuous positive airway pressure (CPAP).

We intend to use these findings along with the FTC2 model for sleep staging predictions to predict the onset of sleep syndromes. We firmly believe that these methods strongly correlate with modern sleep issues, thus enhanc-ing the medical procedure and diagnosis.

METHODOLOGY

For the purpose of this study, we have used the follow-ing methods: CEEMDAN, fast Fourier transform (FFT), convolutional neural networks, long short-term networks (LSTM), Savitzky–Golay filter, and 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 μV), sleep spindles, and other sleep-related events during sleep (transient wax-ing and waning events of about 0.5–2 seconds duration and 15–50 μV maximal amplitude). For the sake of this study and simplicity, we do not distinguish between slow waves, K-complexes, and slow oscillations, but instead group them all together.

Detailed Description of the Methodology

CEEMDAN preprocessing

We used the complete ensemble empirical mode decompo-sition 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 FFT is the foundation of standard signal process-ing methods. However, this approach cannot concurrently yield time domain and frequency domain evaluated find-ings. Many time-frequency analysis approaches, such as the Wigner-Ville distribution (WWD), have been used to diagnose vibration signals [2]. This approach computes the analytic signal corresponding to the input signal, con-structs a weighted kernel function, then analyses the kernel using a discrete Fourier transform (DFT). According to Boashash and Black [3], when evaluating the analytic sig-nal needed by the algorithm, the time domain definition

implemented as a finite impulse response (FIR) filter is more practical and efficient than the frequency domain definition. The wavelet transform [4, 5] utilizes the FFT 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 and can deconstruct the signal into a series of full and nearly orthogonal components [intrinsic modal functions (IMFs)]. The IMFs represent the signal's inherent oscillation mode and act as basic 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 and nonstationary processes, as well as for defect feature extraction in spiral-bevel gears. According to Xujing Yao and Ji Han [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 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 optimize the two main parameters, embedding dimension and 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, and the diagnostic accuracy can validate its sensibility.

FTC2 model architecture

The standard practice for image recognition, computer vision, and more is deep convolutional neural networks. Models like LeNet [7], AlexNet [8], ImageNet [9], and more work exclusively on 2D data. They are also referred to as 2D CNNs. With the same idea in mind, 1D CNNs have been developed. Despite the obvious fact that the 1D CNNs apply a 1D kernel to data, they also have some advantages. 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, computationally less complex and can run on CPU hardware.

   The FTC2 is passed through a leaky rectified linear unit for nonlinearity purposes (Figure 1). The input signal then passes a max-pooling layer with the stride of 2, following that is the batch normalization layer, and 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 that small filters extract temporal information and large filters 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 and frequency domain characteristics [6]. This allows you to benefit from both temporal domain and frequency domain data in the classification process. The feature map generated from the 1D concatenated output from the twin convolutional blocks is intertwined with LSTM units to capture the complex context dependencies between the inputs and the targets.

Utilizing the Conv-1D and the bidirectional LSTM

The Conv-LSTM network combines the best of both worlds. One-dimensional convolutional layers handle spatial data using a kernel to train a set of filters that best represent the hyperplane and 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 and past states of neighbors. Convolutional-LSTM layer works by passing the 1D kernel over a set of training points, and 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 = x0, xN, which are one-dimensional data of a 30-second clip, which should assure stability and little overlap, these inputs are passed into the twin architecture.

Loss calculation

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

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

Unsupervised Method for Detecting Onsets of Sleep Syndromes

Syndromes in question are restless leg syndrome (RLS), rapid eye movement (REM), sleep behavior disorder, insomnia, sleep apnea, and narcolepsy. In an attempt to detect these complex syndromes through EEG data, we have come up with the following sections.

Restless leg syndrome

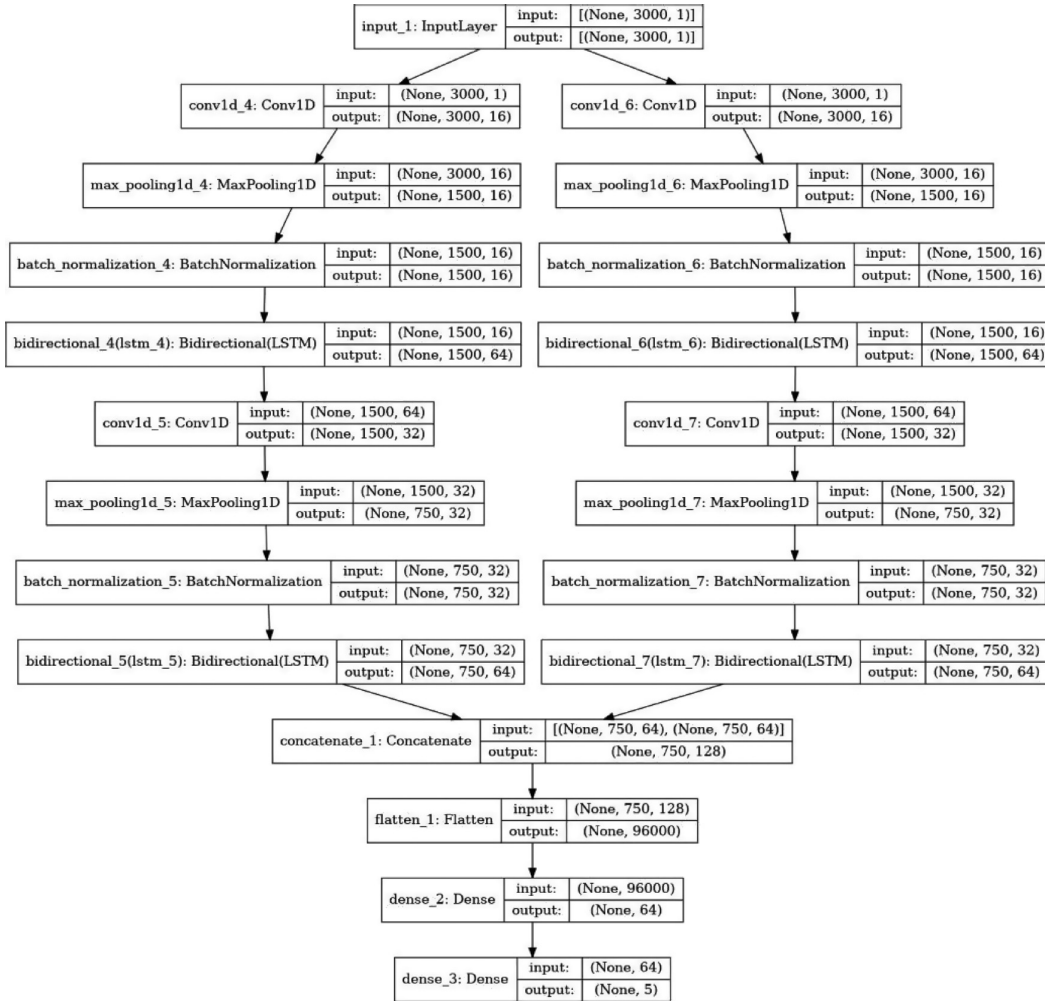Lisuride, a dopaminergic medication, showed a rise in alpha, a decrease in beta, and delayed actions on brain

Figure 1. A schema of the FTC2 architecture.

function as evaluated by computerized 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, and they continue to increase during sleep stage 2. This malfunction is most likely caused by a hereditary susceptibility of the EEG alpha rhythm and 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, and 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, and 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 neighboring data points with a low-degree polynomial using the linear least squares approach. The data were 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 and 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 and EEG signal, more identifiable characteristics can be maintained in frequency band-limited signals for apnea diagnosis. As a result,

instead of analyzing full-b and EEG data for apnea event identification, features of band-limited signals are utilized. EEG signal is partitioned into five frequency bands [10], namely, delta (0.25 to 4 Hz), theta (4 to 8 Hz), alpha (8 to 12 Hz), sigma (12 to 16 Hz), and beta (16 to 40 Hz). Spectral filtering is done in the FFT domain to achieve this division.

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

REM behavior sleep disorder

Despite what was presented in the study by Krizhevsky, Sutskever and Hinton [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 the 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 and theta power during alertness, and higher alpha and beta power during REM sleep, according to Deng et al. [11]. The former two groups also had reduced sleep efficiency and a greater percentage of positivity for REM-related symptoms than the other two groups. In light of this, we integrated the REM instability score and the acquired frequency range with wavelet transform on the initial and final 10% of sleep, suggesting worse sleep-to-wake transitions or vice versa.

Insomnia

Insomnia is a sleep condition characterized by an inability to fall or stay asleep for as long as desired.

Following the previous work, a picture was generated with the MNE Python library by Mohd Maroof Siddiqui, Geetika Srivastava, and Syed Hasan Saeed in their study [11] (Figure 2). Because the majority of EEG signals are constrained within the range of 25 Hz, each clipped signal is now preprocessed and then run through the Hanning window low-pass filter to remove the high-frequency components that eventually suggest noise. Hence, the filter is based on the FIR filter design of order 200 with a cutoff frequency of 25 Hz with the shape of the Hanning window. At this stage, we adopt a method for extracting frequency
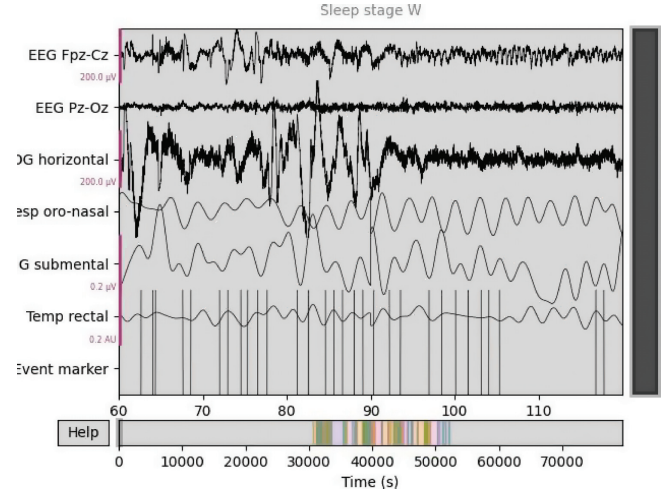

Figure 2. Polysomnography plot for a single patient.

length windows via the Welch power spectrum [12]. This score, obtained with softmax and the REM instability score, was combined 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 also to quantify sleep depth. First, the raw records are pushed through a Savgol filter to noise filtering. Furthermore, a FFT 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.

Syndromes	RLS	REM disorder	Apnea	Narcolepsy	Insomnia	Sleep depth
Total	177983	6674	8899	4449	48945	164738
Percent	7.80	0.29	0.41	3.0	22	74.0
NHCS Percent	7–10	0.5–1	0.5	3–7	10–30	66

EXPERIMENTAL RESULTS

Data

The PhysioNet Sleep-EDF dataset was utilized for the objectives of this study. The sleep-edf database comprises 197 PolySomnoGraphic sleep recordings from the whole night, including EEG, EOG, chin EMG, and event markers. Some recordings additionally include information about breathing and body temperature. The Sleep-EDF dataset comprises two studies, namely, one on the effects of age on sleep in healthy people (SC = Sleep Cassette) and the other 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 and Pz-Oz electrode sites), EOG, chin EMG, and event markers are all included in each record. Fewer records frequently include oro-nasal respiration and rectal body temperature. The hypnograms (sleep

phases; 30-second epochs) were identified manually by well-trained personnel in accordance with the Rechtschaffen an Kales standard. Each level was assigned to a distinct class (stage). The classes include W, REM, N1, N2, N3, N4, and M. Data syndrome comparison was done according to Schensted [13].

Dataset	W	N1	N2	N3–N4	REM	Total
Sleep-EDF-13	8285	2804	17799	5703	7717	42308
Sleep-EDF-18	65951	21522	96132	13039	25835	222479

It is worth noting that the sleep stages in PhysioNet's Sleep-EDF database are not normally distributed. There is a significantly higher number of W and N2 stages than others that cause class imbalance. We have discussed this in 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 and a learning rate alpha = 0.001. We also added the L2 regularization penalty with beta = 0.001 to mitigate the overfitting. For this goal, we used Python as the programming language and TensorFlow as the framework.

By employing the Conv-1D and the bidirectional LSTM twin network, we have successfully taken advantage of the temporal and spatial components of the signal processing nature. To reduce the class imbalance problem, we adopted the weighted MFE loss.

Accuracy	Precision	Recall	F1-score	Final MFE
90.43	77.76	93.32	89.12	0.09

CONCLUSION AND FUTURE WORK

In this study, 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 multilayer perceptron to classify sleep stages from EEG data with the accuracy of . By incorporating a vastly important CEEM-DAN preprocessing method, we build a more stable and noise-free variant of our data, which enabled building an end-to-end trainable model for not only classifying sleep stages but also detecting the early onset of frequently occurring sleep syndromes. Our algorithm takes advantage of both the spatial and temporal dimensions of signal processing, as well as the 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 are known to bare 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 and usability. Following our work, we will investigate how these syndrome predictions/scores align with actual scores and will hopefully serve as a professional service for automatic sleep stage and syndrome onset detection.

REFERENCES

[1] https://epworthsleepinessscale.com/about-the-ess/

[2] Shie Qian and Dapang Chen, "Decomposition of the Wigner-Ville distribution and time-frequency distribution series," in IEEE Transactions on Signal Processing, vol. 42, no. 10, pp. 2836–2842, Oct. 1994, doi: 10.1109/78.324750.

[3] B. Boashash and P. Black, "An efficient real-time implementation of the Wigner-Ville distribution," in IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 35, no. 11, pp. 1611–1618, November 1987, doi: 10.1109/TASSP.1987.1165070.

[4] Katsuhiro Inoue, Akihiko Takajo, Makoto Maeda and Shigeaki Matsuoka, "Tuning method of modified wavelet transform in human sleep EEG analysis," 2007 International Conference on Control, Automation and Systems, 2007, pp. 2784–2789, doi: 10.1109/ICCAS.2007.4406842.

[5] Xujing Yao, Ji Han, COVID-19 Detection via Wavelet Entropy and Biogeography-Based Optimization, COVID-19: Prediction, Decision-Making, and its Impacts, doi: 10.1007/978-981-15-9682-7-8, (69–76), (2021).

[6] https://www.hindawi.com/journals/sv/2019/7806015/.

[7] Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review – Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Architecture-of-LeNet-5-LeCun-et-al-1998-fig2-317496930 [accessed 3 May, 2021].

[8] M. X. Cohen, Analyzing neural time series data: theory and practice. MIT press, 2014.

[9] Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C.J.C. Burges, L. Bottou and K.Q. Weinberger (ed.), Advances in Neural Information Processing Systems 25 (pp. 1097–1105). Curran Associates, Inc.

[10] Saha S, Bhattacharjee A, Fattah SA. Automatic detection of sleep apnea events based on inter-band energy ratio obtained from multi-band EEG signal. Healthc Technol Lett. 2019;6(3):82–86. Published 2019 Jun 3. doi: 10.1049/htl.2018.5101.

[11] J. Deng, W. Dong, R. Socher, L. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248–255, doi: 10.1109/CVPR.2009.5206848.

[12] Ferri R, Rundo F, Silvani A, Zucconi M, Bruni O, Ferini-Strambi L, Plazzi G, Manconi M. REM Sleep EEG Instability in REM Sleep Behavior Disorder and Clonazepam Effects. Sleep. 2017 Aug 1;40(8). doi: 10.1093/sleep/zsx080. PMID: 28482056.

[13] Schensted, C. (1961). Longest Increasing and Decreasing Subsequences. Canadian Journal of Mathematics, 13, 179–191. doi: 10.4153/CJM-1961-015-3.

[14] Sasai-Sakuma T, Inoue Y. Differences in electroencephalographic findings among categories of narcolepsy-spectrum disorders. Sleep Med. 2015 Aug;16(8):999–1005. doi: 10.1016/j.sleep.2015.01.022. Epub 2015 Apr 1. PMID: 26026626.

[15] Siddiqui MM, Srivastava G, Saeed SH. Diagnosis of insomnia sleep disorder using short time frequency analysis of PSD approach applied on EEG signal using channel ROC-LOC. Sleep Sci. 2016;9(3):186–191. doi: 10.1016/j.slsci.2016.07.002.

[16] Zhao, Li and He, Yang. (2013). Power Spectrum Estimation of the Welch Method Based on Imagery EEG. Applied Mechanics and Materials. 278–280. 1260–1264. doi: 10.4028/www.scientific.net/AMM.278-280.1260.

[17] https://www.cdc.gov/nchs/nvss/index.html [Accessed 4.5. 2021].