Visualising Convolutional Neural Network Decisions in Automated Sleep Scoring *

Fernando Andreotti, Huy Phan, and Maarten De Vos

Institute of Biomedical Engineering, University of Oxford, Oxford, UK
fernando.andreotti@eng.ox.ac.uk

Abstract. Current sleep medicine relies on the supervised analysis of polysomnographic recordings, which comprise amongst others electroencephalogram (EEG), electromyogram (EMG), and electrooculogram (EOG) signals. Convolutional neural networks (CNN) provide an interesting framework for automated sleep classification, however, the lack of interpretability of its results has hampered CNN's further use in medicine. In this study, we train a CNN using as input Continuous Wavelet transformed EEG, EOG and EMG recordings from a publicly available dataset. The network achieved a 10-fold cross-validation Cohen's Kappa score of $\kappa = 0.71 \pm 0.01$. Further, we provide insights on how this network classifies individual epochs of sleep using Guided Gradient-weighted Class Activation Maps (Guided Grad-CAM). The proposed approach is able to produce fine-grained activation maps on time-frequency domain for each signal providing a useful tool for identifying relevant features in CNNs.

Keywords: Convolutional Neural Networks · Class Activation Maps · Guided Backpropagation · Polysomnography · Wavelet Transform.

1 Introduction

Sleep is a fundamental biological process critical for the maintenance of physical and mental health. Associations between sleep disruption and various morbidities have been often reported, with some parainsomnias preceding serious neural disorders by many years [24]. Therefore, sleep monitoring is a matter of utmost importance. Current clinical praxis heavily relies on the analysis of polysomnographic (PSG) recordings, which include electroencephalogram (EEG), electromyogram (EMG), and electrooculogram (EOG) amongst other physiological signals. These signals are then interpreted based on clinical guidelines, such as R&K [14] and the AASM [2], which divide sleep into a few stages according to specific spectral content and characteristic waveform patterns (see Table 1). Manual scoring following these rules is the gold-standard in sleep medicine.

* This research was supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC) and the Engineering and Physical Sciences Research Council (EPSRC – grant EP/N024966/1). We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

Table 1: Brief summary of AASM rules [2] for sleep staging [23].

Stage	Description
Wake	Presence of EEG alpha rhythm (8-12 Hz); EMG with high-amplitude due to movements; EOG presents eye blinking artifacts, also visible on EEG around $0.5 - 2$ Hz.
N1	Attenuated alpha rhythm, presence of theta signal (4-7 Hz) on EEG; Muscle tone and slow eye movements decrease on EMG and EOG.
N2	EEG presents K-complexes, i.e. large low frequency negative peaks in the range < 1.5 Hz, and sleep spindles, i.e. bursts of oscillations in the sigma band (12-15 Hz).
N3	Slow wave activity for EEG (0.5-3 Hz), EMG tone is low and eye movements unusual.
REM	Rapid eye movements (REM) clearly visible in EOG; low-amplitude and mixed-frequency activity in EEG; muscle atonia on EMG.

In contrast to manual scoring, automated approaches provide objective means of classifying sleep stages. Traditional approaches make use of numerous hand-engineered features from the physiological signals in combination with classical machine learning methods, e.g. support vector machines or hidden Markov models. Throughout the past decade, Convolutional Neural Networks (CNNs) have been widely applied in fields such as computer vision and audio processing due to their ability to operate on raw data, not requiring the explicit definition of features. In the last couple of years, some early works applying CNNs in the field of sleep analysis began to emerge [22, 19, 3, 1, 12]. Despite the competitive performance achieved, due to the high degree of abstraction present in CNN hidden layers, interpretability is an issue that has hindered its further use in medicine.

Vilamala *et al.* [23] proposed to visualise CNN decisions in sleep analysis using sensitivity maps [13, 16]. For this purpose, the authors converted each sleep epoch of single-lead EEG signals into spectrograms by using the Multitaper Spectral Estimation. In order to convert spectra into RGB images (i.e. containing three colour channels), the authors artificially mapped the spectrogram intensities using an arbitrary colourmap. This was performed so that pre-trained image detection CNNs could be applied. The network of choice in [23] was the well-known VGGNet [17], which contains 13 convolutional layers followed by 3 fully connected (FC) layers and 138 million parameters. VGGNet was fine-tuned to the task of sleep scoring using the Physionet Sleep-EDF Database [8]. Despite presenting an interesting framework, [23] has a few limitations. First, the number and diversity of subjects available in the Sleep-EDF is limited to 20 young healthy subjects. Second, using a single EEG lead is restrictive as experts classify sleep stages based on multichannel and multimodal settings. Third, despite the abundance and availability of pre-trained image models, these networks were trained for a very different task than the one at hand. As shown in [25] the transferability of features decreases as the distance between the base task

and target task increases. Moreover, these networks usually assume three colour input channels. The workaround applied in [23] to obtain RGB channels out of a single spectrum provides no additional information to the network.

In this work, we proposed a simple CNN architecture that is trained from scratch using a large publicly available database. As input to the CNN we provide EEG, EOG and EMG signals, which are standard in sleep analysis. For visualising this network's weighs we apply the Guided Gradient-weighted Class Activation Maps [15]. This allows a detailed class-specific view of the network for each channel used as input, which may provide somnologists with additional interpretation tools for sleep analysis.

2 Data Material

Data was obtained from the Montreal Archive of Sleep Studies database (MASS-DB) [11], a large publicly available dataset comprising single night PSG recordings of 200 healthy participants ageing between 18 and 76 years, including 98 males aged 42.7±19.4 years and 102 females aged 38.1±18.9 years. The database is divided into 5 cohorts all of which were used in this study. To avoid a saturation on the number of needed channels in a recording setup we restrict our setup to three channels commonly used in the specific literature: a single central EEG lead (C4-A1 or C3-A2, where available), a differential EOG (ROC-LOC) and/or EMG (CHIN1-CHIN2).

As the MASS-DB comprises different study protocols, annotations using R&K were converted into AASM guidelines by assigning $S3$ and $S4$ stages to $N3$, while $\{S0, S1, S2\}$ were relabelled as $\{W, N1, N2\}$, respectively. Three of the MASS-DB cohorts contained 20s-epochs and were converted into 30s by including 5s of signal before and after each segment. A total of 228,870 epochs are available from the MASS-DB, being 13.6% W, 17.6% REM, 8.5% N1, 47.2% N2, 13.3% N3. To avoid biasing the network to either state, we undersampled each subject recording to the minority class, which results into 59,848 samples.

3 Methods

Various CNN models have been proposed for the task of sleep stage classification, most of which operate on raw single-channels (EEG or EOG). For instance, [19] proposed a network with two branches of 4 convolutional layers with distinct receptive fields aiming to generate feature maps with low and high frequency content. In [22], the authors propose a two-layer CNN model where after the first 1-dimensional convolution, filters are reshaped (stacked) and processed by a 2D convolution. In [3] a spatial filtering technique is applied to multiple EEG, EOG and EMG channels. Each group of signals is treated in separate CNN pipelines as images by applying two-layers of 2D convolutions. The study also confirms that multiple channels and sensors provide an increase in detection accuracy, similar results were obtained in [1].

Sleep stages are largely defined based on the spectral content of its signals (see AASM definitions in Table 1). Therefore, it is reasonable to make use of time-frequency transforms as in [23]. Similarly to [23], the Short-Time Fourier Transform (STFT) coupled with a compact CNN network was successfully applied in [12]. Time-frequency representations are also beneficial for visualisation purposes, as one can observe time events at multiple frequency scales. Therefore, in this study we aim at generating time-frequency transforms for each epoch and modality of signal (EEG, EOG and EMG) as described in the following Section 3.1. The proposed network is further explained in Section 3.2.

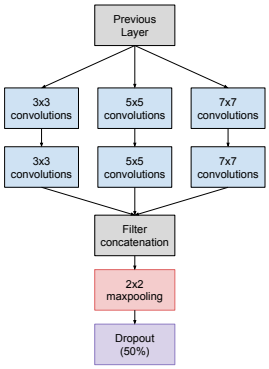
3.1 Preprocessing

All recordings were sampled at $f_s = 100$ Hz and divided in 30 s epochs following the AASM standard [2] for sleep scoring. Recordings were high-pass filtered using zero-phase 100^{th} order FIR filters with 0.1 Hz cutoff frequency for EEG/EOG signals and 10 Hz for EMG.

As presented in [21], Wavelet transforms are particularly suitable for analysing non-stationary signals (such as EEG), whereas STFT assumes periodicity and is signicantly affected its window size, which makes it fundamentally impossible to correctly determine the onset of rapid events such as sleep spindles. Therefore, based on [21], we apply the Continuous Wavelet Transform (CWT) with Morlet basis function as time-frequency transform to each 30 s epoch using Matlab's $cwt()$ function with default parameters. The resulting scalograms (dimensions $101\times(30 * f_s)$) were further reduced by using bilinear interpolation to 30×300 maps. The resulting time-frequency representations were then normalised to the range $[0, 1]$, which serve as input to the proposed network.

3.2 Network Architecture

The proposed network is detailed in Fig. 1. As mentioned, the network assumes as input scalograms for each channel, i.e. input dimensions $30\times300\times3$. The first two convolutional layers use valid padding and convolutional strides of (1,2) to reduce the input dimensions, thus decreasing the total number of parameters in the following layers. The following block performs convolutions on 3×3, 5×5 and 7×7 windows in a similar fashion as the naïve Inception block [20] and in [12] (shown in Fig. 1). The main idea of this block is to learn features at multiple temporal and frequency resolutions. At last, global average pooling is performed to reduce the spatial dimensions of each channel. Global pooling is directly followed by the softmax layer which outputs classes scores. The network contains a total of 9.0M parameters.



(a)

Layer Type	#Filters	Operation Size	Output Dim.
Input	-	-	(30,300,3)
Convolution 2D	64	(10×10)	(21,146,64)
Batch Normalisation	-	-	(21,146,64)
Activation	-	ReLU	(21,146,64)
Dropout	-	50%	(21,146,64)
Convolution 2D	128	(10×10)	(12,69,128)
Batch Normalisation	-	-	(12,69,128)
Activation	-	ReLU	(12,69,128)
Dropout	-	50%	(12,69,128)
Inception-like Block	256	(3×3),(5×5),(7×7)	3×(12,69,256)
Concatenate	-	-	(12,69,768)
MaxPool 2D	-	(2×2)	(6,34,768)
Dropout	-	50%	(6,34,768)
Global Average Pool	-	-	(768)
Dense	5	Softmax	5

(b)

Fig. 1: (a) Naïve Inception-like block used in this work to perform convolutions at multiple time-frequency scales, based on [20, 12]. Within the block, convolutional operations use zero-padding, are followed by Batch Normalisation [7] and Rectified Linear Units (ReLU) activation; (b) proposed CNN architecture for this study.

3.3 Training and Cross-Validation

The proposed model was developed in Keras/Tensorflow. Training was performed in 100 epochs, with batch size 64 and categorical cross entropy as loss function. Adam[9] optimiser was used with default parameters (learning rate $lr = 10^{-3}$, $\beta_1 = 0.9$ and $\beta_2 = 0.999$). L2 norm regularisation was applied to the last layer with value $\lambda = 10^{-5}$. To assess the performance of this network, a 10-fold cross-validation procedure was carried out on the balanced MASS-DB. Results are evaluated in terms of sensitivity (SE), positive predictive value (PPV), accuracy (ACC), F1-score (F1) and Cohen's Kappa score (κ).

3.4 Visualising CNNs

Deep visualisation is an active field of research. Current methods for visualising CNN weights with 2-dimensional data input (e.g. images or time-frequency representations) can be divided into *gradient-based* and *activation maximisation*.

Gradient-based methods aim at highlighting pixels of an input image I that have the most impact on the prediction score S_c for a given class c. In order to assess the rate of change in S_c with respect to small changes in image intensity, [16] proposed backpropagating the partial derivatives $\partial S_c/\partial I$ across the network. Therefore positive values on the gradient indicate an increase on the output score for that class. As the class score function is highly non-linear function of image I, first-order Taylor expansion is used to linearize this function [16]. Extensions of the backpropagation method focus on modifying to the original gradient function and result in qualitative improvements to visualisation [15]. Example of such methods include the Deconvolutional Networks [27], that modifies the backward

pass of ReLU by clipping negative gradients, and Guided Backpropagation [10, 18], on which activations are masked during both deconvolution and forward pass. This group of techniques can be used for visualising the last layer [16], or on each of the hidden neurons [4, 26]. These approaches are attractive due to their simplicity, however despite producing fine-grained visualisations, these methods are not class-discriminative as optimisation process tends to produce images that are hardly recognisable [15]. In order to produce more natural-looking images, some studies suggested biasing optimisation with natural image priors [10] or using regularisation techniques [26].

Activation maximisation methods focus on directly visualising the activation of some specific layer of a network given an input image. One approach is to estimate the importance of input pixels by visualising the probability of the correct class being chosen as a function of a mask occluding parts of the image, the so-called occlusion/perturbation sensitivity [27, 5, 23]. Another approach is to focus on the activation of the last layer before any FC layer, where higher-level visual concepts are captured. An example of such approach is are Class Activation Maps (CAMs) [28]. In order to retain the tensorial shape any flattening operation is substituted by global average pooling followed directly by the softmax (i.e. disregarding any FC layers). Based on the class scores S_c, w_k^c corresponding weights to class c for unit k, and $A_{xy}^k = \sum_{x,y} f_k(x, y)$ activation map of unit k in the last layer a class-specific heatmap L^c can be achieved for the image such as [28]:

$$L_{CAM}^c(x, y) = \sum_k w_k^c A_{xy}^k \tag{1}$$

Selvaraju *et al.* [15] proposed a generalisation of CAM by introducing the main concept of backpropagation from gradient-based approaches downstream from any A^k, i.e.

$$\alpha_k^c = \frac{1}{Z} \sum_x \sum_y \frac{\partial S_c}{\partial A_{xy}^k} \tag{2}$$

where α_k^c is the partial linearization of the network from feature map A^k and $\frac{1}{Z} \sum_x \sum_y$ represents the global average pooling operation perform over these feature maps. The Gradient-weighted CAM (Grad-CAM) then performs a weighted combination of forward activation maps followed by a ReLU:

$$L_{Grad-CAM}^c(x, y) = ReLU\Big(\sum_k \alpha_k^c A_{xy}^k \Big) \tag{3}$$

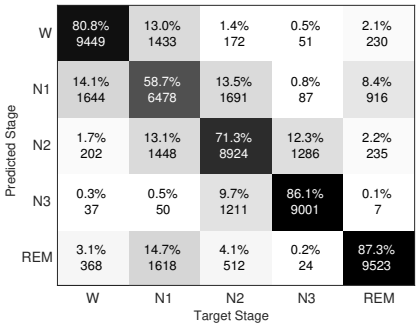
Grad-CAM enjoys the benefits from CAMs and produce class-discriminative localisation of image regions, while imposing fewer restrictions on the network's architecture. In fact, it can be used in any CNN-based architecture. However, Grad-CAM only produces an averaged heatmap for the image so that the channel information is lost. In order to show fine-grained gradient visualisations, [15] further proposed combining Guided Backpropagation and Grad-CAM by

point-wise multiplication of both bi-linear interpolated Grad-CAM and Guided Backpropagation heatmaps.

In this contribution, we aim at better understanding how the network classifies sleep stages regarding each signal modality (i.e. EEG, EOG and EMG). For this purpose, after pre-training the proposed CNN presented in Section 3.2, Guided Grad-CAM is applied to randomly selected segments of the MASS-DB. We qualitatively evaluate these Wavelet scalograms and heatmaps regarding their physiological meaning.

4 Results and Discussion

The results for the 10-fold cross-validation procedure using the proposed model and balanced MASS-DB set are described in Fig. 2. Despite having less than 10% the number of parameters of VGGNet, the proposed CNN is able of performing automated sleep staging when trained from scratch, resulting Kappa score of $\kappa = 0.71 \pm 0.01$. As usual in the literature, N1 stage classification performed worst as the state shares characteristics with wakefulness state. Moreover, the transition from wake to N1 is described as challenging and has been reported as a major source of inter-rater variability [6]. Our analysis differ from the one presented in [23] in that the database used in this study is larger and the different classes have been balanced. Similar results for the imbalanced MASS-DB were obtained in [19].



(a)

Stage	SE	PPV	ACC	F1
W	0.81	0.83	0.93	0.82
N1	0.59	0.60	0.84	0.59
N2	0.71	0.74	0.88	0.73
N3	0.86	0.87	0.95	0.87
REM	0.87	0.79	0.93	0.83

(b)

Fig. 2: Confusion matrix (a) and per class classification metrics (b) for 10-fold cross validation on balanced MASS-DB.

Figures 3 and 4 demonstrate the application of Guided Backpropagation and Guided Grad-CAM using the proposed trained model to segments of Wake, REM and N2 stages respectively. These stages were chosen for illustrating the visualisation method, due to their more distinguishable characteristics (e.g. presence of spindles/K-complexes on N2). A total of 100 randomly selected epochs from

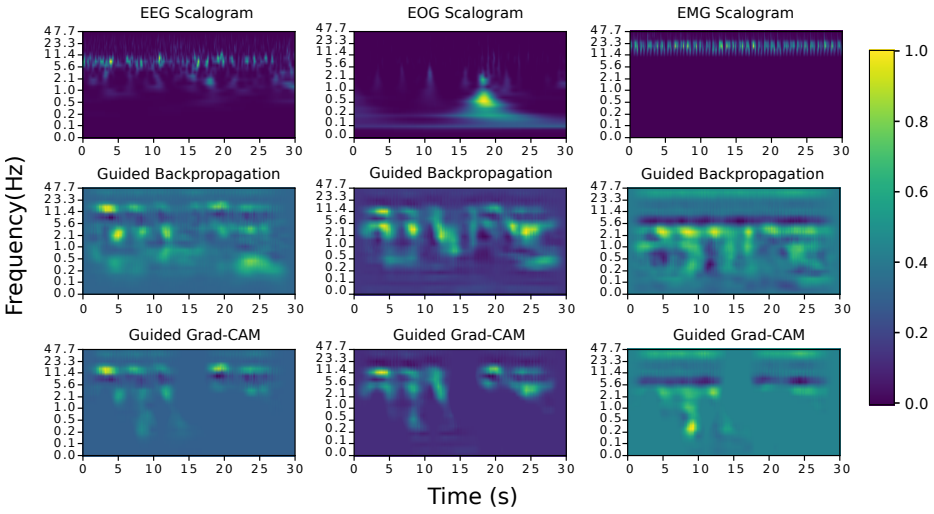
all cohorts and subjects were visually inspected from which Figs. 3 and 4 were selected. Overall Guided Grad-CAM produced cleaner heatmaps than Guided Backpropagation as described in [15]. More importantly, Guided Grad-CAM produces channel-specific maps, which are relevant when analysing multimodal data such as PSG recordings.

The presented results show that relevant CNN feature maps (i.e. weights) often correspond to regions of interest for particular sleep stages. For instance Fig. 3(a) presents high sensitivity around EEG alpha bands (i.e. 8-12 Hz), which is characteristic for both Wake stage. During the wake stage, the patterns for EOG and EMG seem erratic spreading along various frequency bands (Fig. 3(a)) whereas during REM (Fig. 3(b)) weighs lower frequent EOG bands. Usually the N2 stage is the mostly recognisable one, comprising K-complexes (negative peaks followed by positive peaks with duration > 0.5 s) and sleep spindles (bursts of oscillatory waves in the sigma band, i.e. 11-15 Hz). Comparing Fig. 4(a) and (b) we notice that the Guided Grad-CAM weighs heavily such patterns. These conclusions are similar to the ones found in [23], except we notice coherent physiological information across the different input channels used in this study.

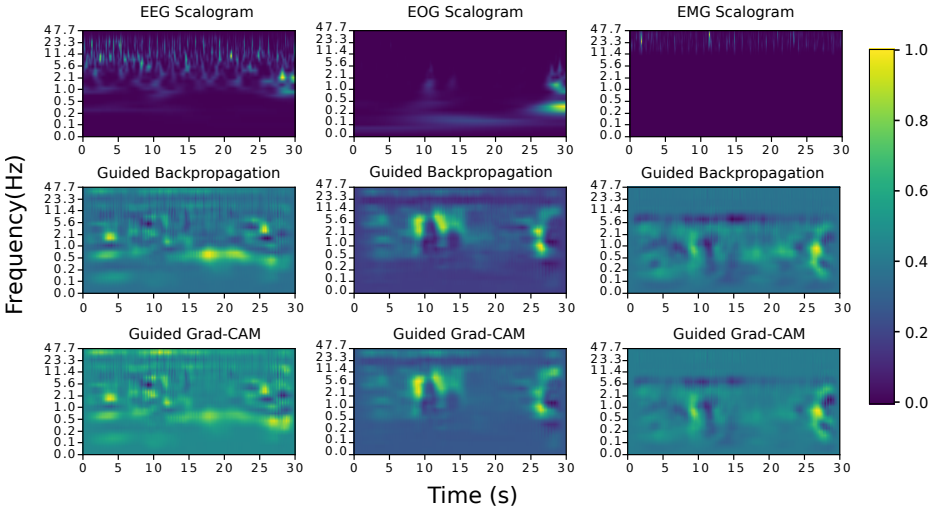
These exemplary epochs demonstrate the potential of such visualisation approaches for further interpreting how CNN weighs different sleep patterns and could greatly enhance the interaction of domain experts [23]. Nevertheless, conclusions must be carefully drawn as the method only provide a partial understanding of the network (e.g. not including FC layers). Moreover, as more aesthetically pleasing notions of image saliency are sought, different backpropagation heuristics (e.g. regularisation) are applied [5]. Further limitations of this method include i) epoch-specificity i.e. output depends on input representation/epoch; ii) the visualisation tool is not model-agnostic as each network will inevitably produce different outcomes. In this study we applied data from healthy subjects, future works should focus on comparing these results with feature maps generated from pathological data that may provide clinically relevant insights. Additionally, PSG analysis often takes into consideration surrounding epochs. The proposed CNN model treats each epoch independently, however, it is beneficial to consider the temporal/transition information contained in this signal. This can achieved by using Recurrent Neural Networks or soft-attention techniques.

5 Conclusion

In this contribution we shed a light onto how CNNs are able to distinguish sleep stages. For this purpose we trained a small CNN network from scratch which takes as input CWT scalograms from 3 different sensors (EEG, EOG and EMG). Further, we visualise regions of interest for the trained network by applying the Guided Grad-CAM method. The proposed approach is able to produce fine-grained activation maps on time-frequency representations of each individual signal providing a useful tool for identifying relevant features in CNNs.

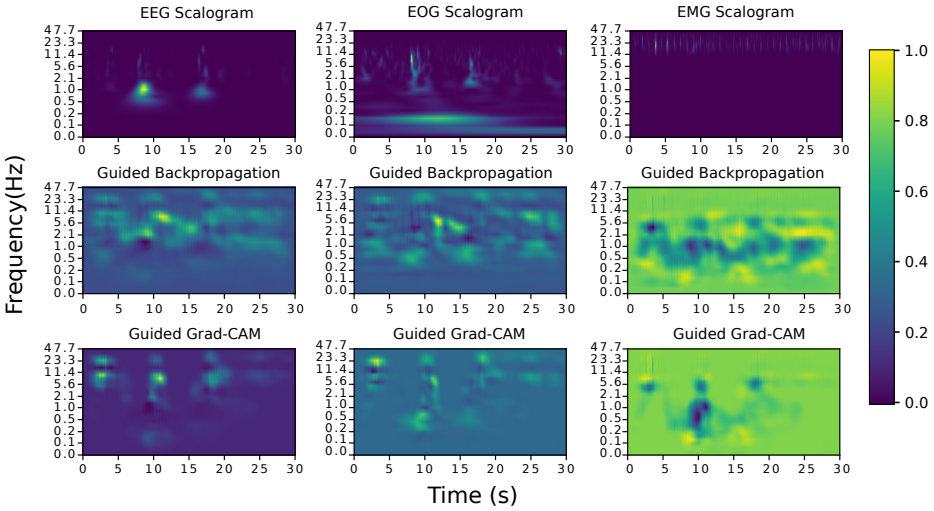


(a) Wake segment (subject SS4-12, epoch 1235).

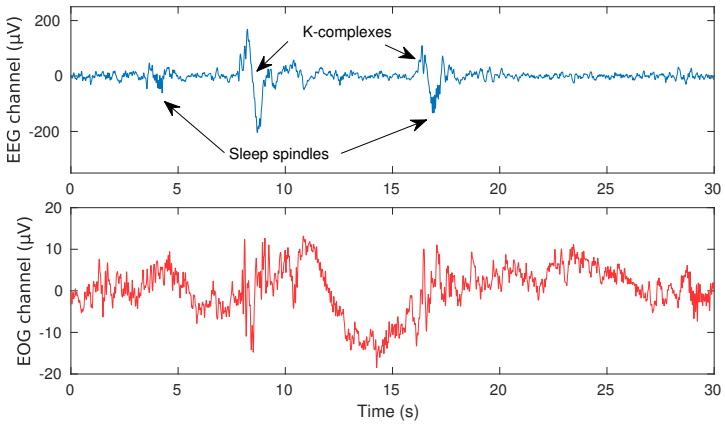


(b) REM segment (subject SS1-9, epoch 684).

Fig. 3: Randomly selected segments of Wake (a) and REM (b). Notice relevant EEG alpha rhythm (8-12 Hz) and EOG erratic content in (a). In (b) EEG contains mixed-frequencies, low frequency EOG is considered relevant but not EMG.



(a) N2 segment (subject SS4-8, epoch 717).



(b) Respective time signal with characteristic spindles and K-complexes.

Fig. 4: Randomly selected segment of N2 stage (a) and respective time signals (b). In (b) K-complexes and sleep spindles are present, characteristic of N2 stage. In (a) the network seems to detect both K-complexes (low frequency content) and sleep spindels (11-13 Hz).

References

1. Andreotti, F., Phan, H., Cooray, N., Lo, C., Hu, M.T.M., De Vos, M.: Multichannel Sleep Stage Classification and Transfer Learning using Convolutional Neural Networks. In: IEEE Engineering in Medicine and Biology Conference (EMBC) (2018)
2. Berry, R., Brooks, R., Gamaldo, C., Harding, S., Lloyd, R., Marcus, C., Vaughn, B.: The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. American Academy of Sleep Medicine (2015)
3. Chambon, S., Galtier, M., Arnal, P., Wainrib, G., Gramfort, A.: A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series. arXiv:1707.03321 pp. 1–14 (2017)
4. Erhan, D., Bengio, Y., Courville, A., Vincent, P.: Visualizing higher-layer features of a deep network. Tech. Rep. 1341, University of Montreal (2009)
5. Fong, R.C., Vedaldi, A.: Interpretable Explanations of Black Boxes by Meaningful Perturbation. In: IEEE International Conference on Computer Vision (ICCV). pp. 3449–3457 (2017)
6. Heidi, D., Peter, A., Josef, Z., Marion, B., Hans, D., Georg, G., Esther, H., Erna, L., Doris, M., Silvia, P., Bernd, S., Andrea, S., Georg, D.: Interrater reliability for sleep scoring according to the rechtschaffen & kales and the new aasm standard. J. Sleep Res. **18**(1), 74–84 (2009)
7. Ioffe, S., Szegedy, C.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv:1502.03167 (2015)
8. Kemp, B., Zwinderman, A.H., Tuk, B., Kamphuisen, H.A.C., Oberyé, J.J.L.: Analysis of a sleep-dependent neuronal feedback loop: The slow-wave microcontinuity of the EEG. IEEE Trans. Biomed. Eng. **47**(9), 1185–1194 (2000)
9. Kingma, D.P., Ba, J.L.: Adam: a Method for Stochastic Optimization. International Conference on Learning Representations 2015 pp. 1–15 (2015)
10. Mahendran, A., Vedaldi, A.: Understanding Deep Image Representations by Inverting Them. arXiv:1412.0035 p. 2014 (2014)
11. O'Reilly, C., Gosselin, N., Carrier, J., Nielsen, T.: Montreal Archive of Sleep Studies: an open-access resource for instrument benchmarking and exploratory research. J. Sleep Res. **23**(6), 628–635 (dec 2014)
12. Phan, H., Andreotti, F., Cooray, N., Chen, O.Y., De Vos, M.: DNN Filter Bank Improves 1-Max Pooling CNN for Automatic Sleep Stage Classification. In: IEEE Engineering in Medicine and Biology Conference (EMBC) (2018)
13. Rasmussen, P.M., Madsen, K.H., Lund, T.E., Hansen, L.K.: Visualization of nonlinear kernel models in neuroimaging by sensitivity maps. NeuroImage **55**(3), 1120–1131 (2011)
14. Rechtschaffen, A., Kales, A.: A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Tech. rep., National Institutes of Health publication, ; no. 204 (1968)
15. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In: IEEE International Conference on Computer Vision (ICCV). pp. 618–626 (2017)
16. Simonyan, K., Vedaldi, A., Zisserman, A.: Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. arXiv:1312.6034 pp. 1–8 (2013)

17. Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556 (2014)
18. Springenberg, J.T., Dosovitskiy, A., Brox, T., Riedmiller, M.: Striving for Simplicity: The All Convolutional Net. In: International Conference on Learning Representations (ICLR) (2015)
19. Supratak, A., Dong, H., Wu, C., Guo, Y.: DeepSleepNet: A Model for Automatic Sleep Stage Scoring Based on Raw Single-Channel EEG. IEEE Trans. Neural Syst. Rehabil. Eng. **25**(11), 1998–2008 (nov 2017)
20. Szegedy, C., Wei Liu, Yangqing Jia, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015)
21. Tsanas, A., Clifford, G.D.: Stage-independent, single lead EEG sleep spindle detection using the continuous wavelet transform and local weighted smoothing. Front. Hum. Neurosci. **9**(April), 1–15 (2015)
22. Tsinalis, O., Matthews, P.M., Guo, Y., Zafeiriou, S.: Automatic Sleep Stage Scoring with Single-Channel EEG Using Convolutional Neural Networks. arXiv:1610.1683 p. 12 (oct 2016)
23. Vilamala, A., Madsen, K.H., Hansen, L.K.: Deep Convolutional Neural Networks for Interpretable Analysis of EEG Sleep Stage Scoring. arXiv **1710.00633** (2017)
24. Wulff, K., Gatti, S., Wettstein, J.G., Foster, R.G.: Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease. Nat. Rev. Neurosci. **11**(8), 589–599 (2010)
25. Yosinski, J., Clune, J., Bengio, Y., Lipson, H.: How transferable are features in deep neural networks? In: NIPS (2014)
26. Yosinski, J., Clune, J., Nguyen, A., Fuchs, T., Lipson, H.: Understanding Neural Networks Through Deep Visualization. arXiv:1506.06579 (2015)
27. Zeiler, M.D., Fergus, R.: Visualizing and Understanding Convolutional Networks. In: European Conference on Computer Vision (ECCV). pp. 818–833 (2014)
28. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A.: Learning Deep Features for Discriminative Localization. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2921–2929 (jun 2016)