DEVELOPING AN AUTOMATIC SLEEP STAGE SCORING MODEL BASED ON RAW SINGLE-CHANNEL EEG

Thanh Luu, Anh Le, Lua Ngo*
School of Biomedical Engineering, International University, Vietnam National University, HCMC
*Corresponding author: Ph.D. Ngo Thi Lua, email: ntlua@hcmiu.edu.vn

INTRODUCTION

- Sleep disturbances are a diverse group of disorders, varying from mild conditions such as sleep deprivation and insomnia to more severe cases such as sleep apnea or narcolepsy.
- Failing to address these sleep-related issues can have significant repercussions on both the quality and duration of sleep, potentially resulting in serious health conditions and increasing the risk of cardiovascular diseases and mental health difficulties.
- Given the growing prevalence of sleep disorders, the need for an accurate, complete diagnosis and monitoring of the response to treatment of sleep status becomes progressively crucial.
- Herein, we developed an automated system that employs machine learning and electroencephalogram (EEG) signals to accurately identify and classify different sleep stages.

DATA DESCRIPTION

The publicly Physionet Sleep-EDF 2018 dataset (Sleep Cassette subset)

Included **153 PSG recordings** belonging to **78 subjects** containing the following signals of interest

- two EEG (Fpz-Cz and Pz-Cz)
- one EOG (horizontal)
- one EMG (submental chin) signal.

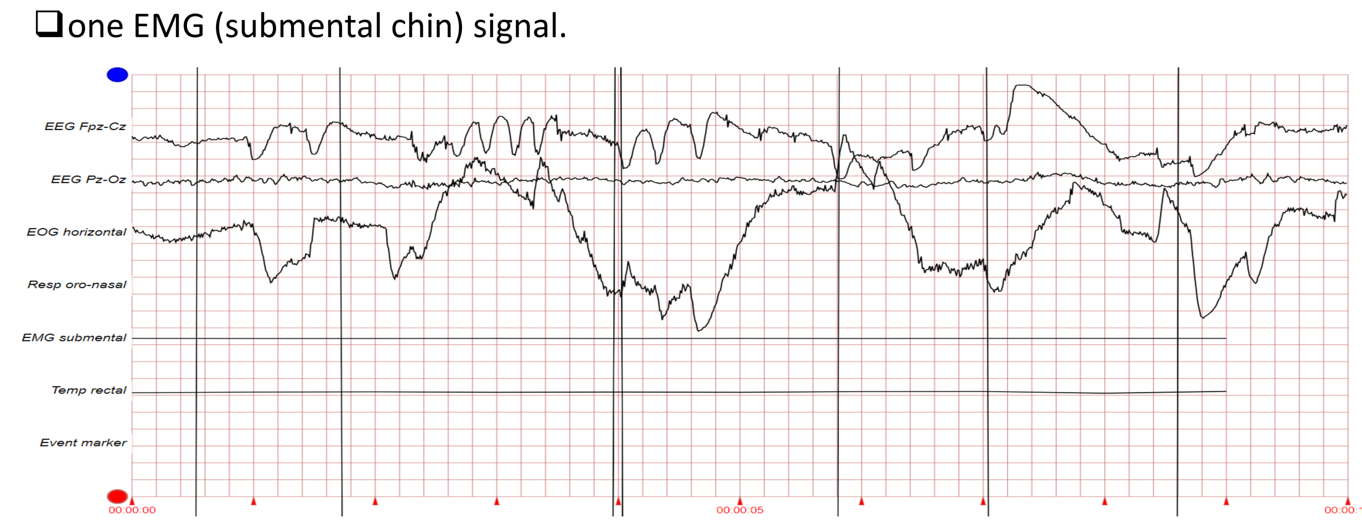

Figure 1. Visualize waveforms of a sample file "SC4001E0-PSG.edf"

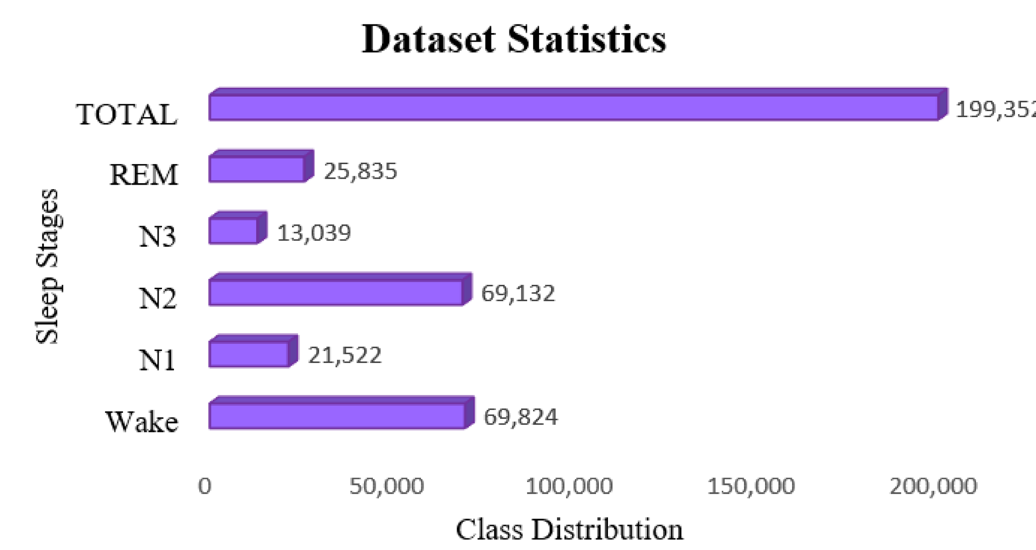

Figure 3. Details of number of 30-s epoch for each sleep stage

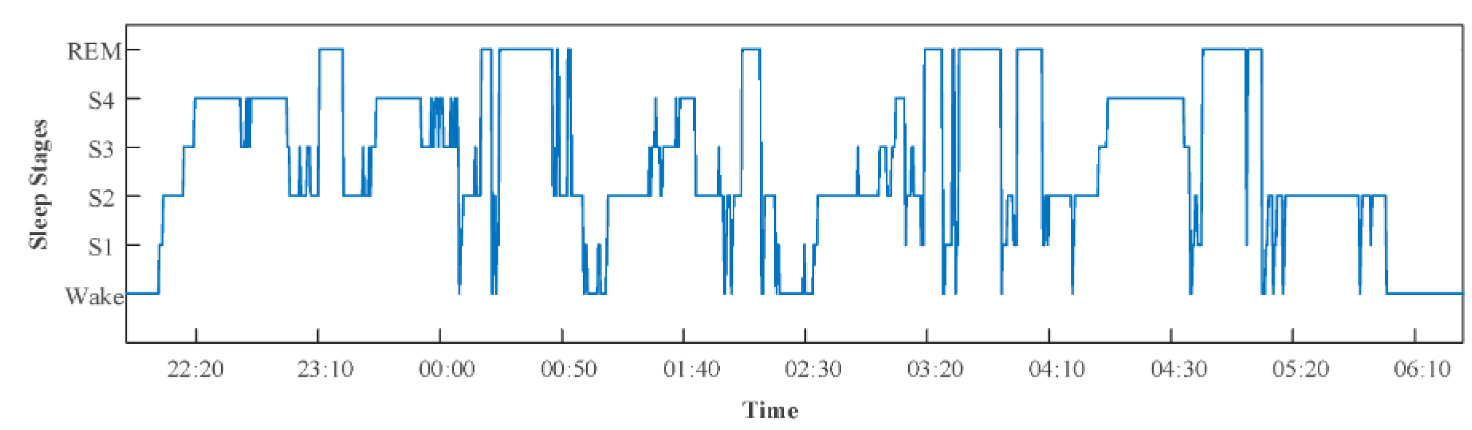

Figure 2. Hypnogram scoring of a sample file "SC4001E0-Hypnogram.edf"

- The label of the sleep stage for each 30-s EEG epoch is provided by the data distributors according to AASM rules.
- Five staging classes (WAKE, N1, N2, N3, and REM).
- Epoch duration: 30s
- Event marker: 1Hz
- The EEG, EOG, and EMG signals were recorded at 100Hz

METHODOLOGY

Model Workflow

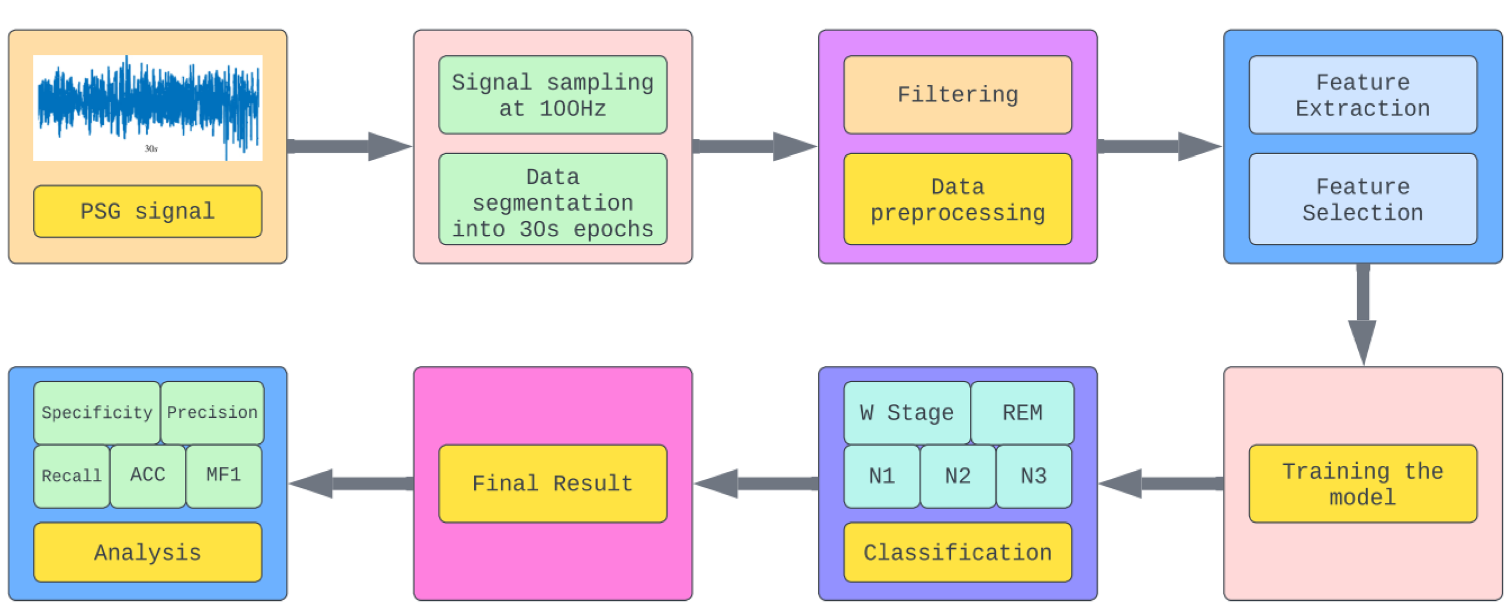

Figure 4. Model Workflow

Convolutional Neural Network (CNN) Architecture

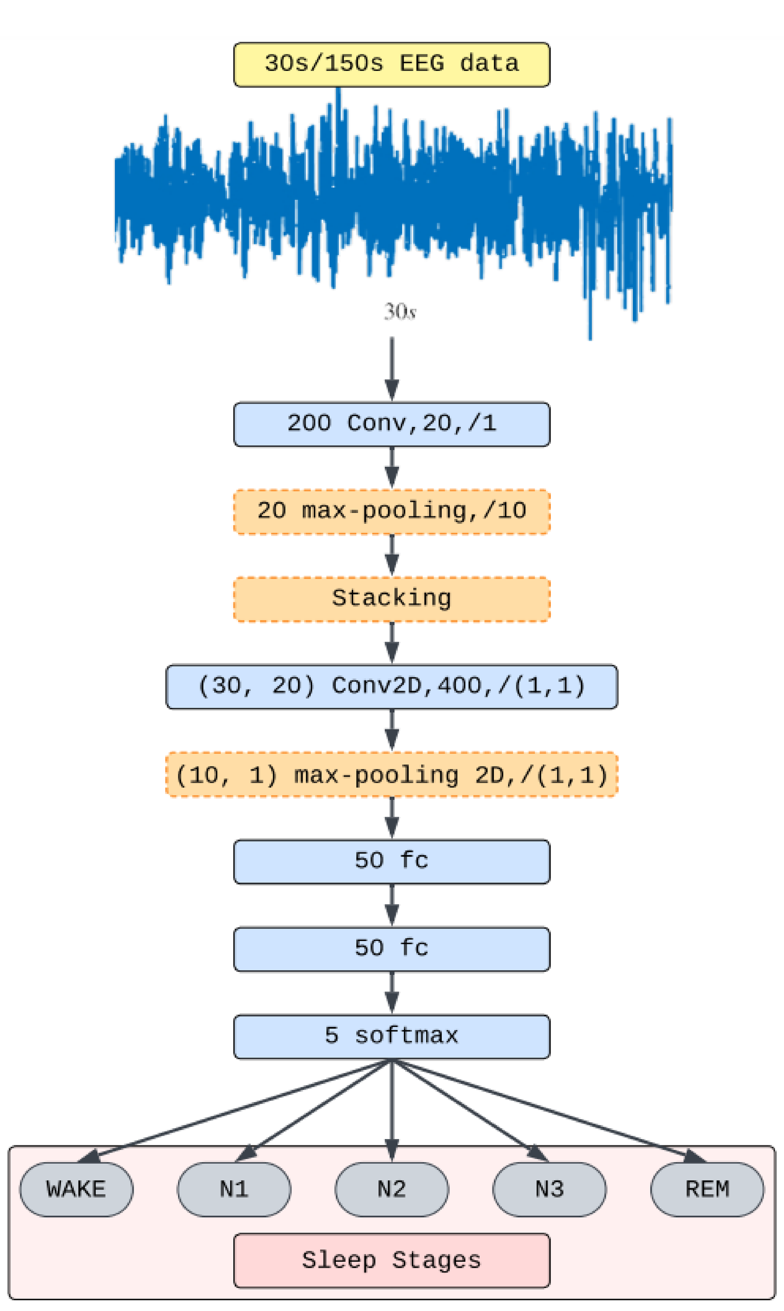

Figure 5. An overview of proposed model

Layer	Type	Unit	Activation	Size	Stride	Output Shape
Input						(B, 15000,1)
C1	Conv1D	20	ReLU	200		(B,14801,20)
P1	MaxPool1D			20		(B,1479,20)
R	Reshape					(B,1479,20,1)
C2	Conv2D	400	ReLU	(30,20)	(1,1)	(B, 1450, 1, 400)
P2	MaxPool2D			(10,1)	(2,1)	(B,721,1,400)
F	Flatten					(B,288400)
F1	Dense	50	ReLU			(B,50)
F2	Dense	50	ReLU			(B,50)
F3	Dense	5	Softmax			(B,5)

Table 1. CNN Architecture

Data Segmentation into 30-s epochs

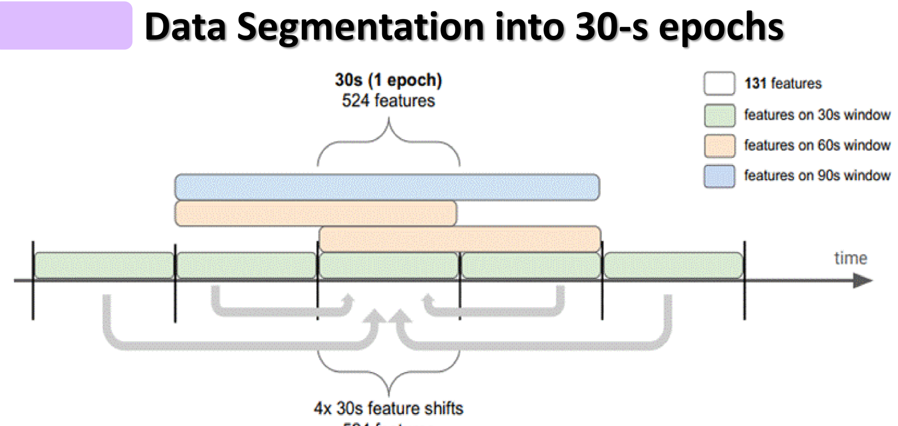

Figure 6. Visualize of 30-s input

Evaluation & Metrics

✓ **Overall accuracy (ACC)**

$$ACC = \frac{\sum_{c=1}^{C} TP_c}{N}$$

✓ **Macro F1-score (MF1)**

$$MF1 = \frac{\sum_{c=1}^{C} F1_c}{C}$$

- The 20-fold cross-validation was used to assess the generalizability of the methods
- TP_C: True positive of class c
- $F1_c$: Per-class F1-score of class c
- C: Number of sleep stages
- N: Total number of test samples

RESULTS & DISCUSSION

Model Results

		Wake	N1	N2	N3	REM
Actual	Wake	**17129.64**	1751.89	0	0	389.31
	N1	1362.59	**13820.51**	2335.86	0	2141.21
	N2	194.66	3893.10	**13236.54**	389.31	1751.89
	N3	0	0	583.97	**18881.54**	0
	REM	583.97	4282.41	778.62	0	**14015.16**
Precision		0.88	0.58	0.78	0.97	0.77
Recall		0.88	0.71	0.68	0.97	0.72
F1 Score		0.88	0.63	0.72	0.97	0.73
Accuracy: 0.79			WF1: 0.75		Kappa: 0.70	

Table 2. Confusion matrix and classification performance of proposed model

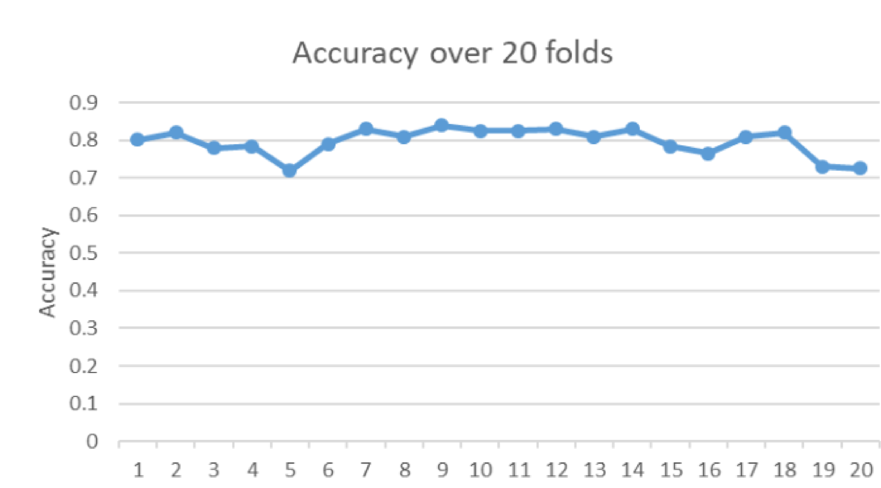

Figure 7. Accuracy of proposed models over 20 folds

	PREDICTED					
	Stage	W	N1	N2	N3	REM
ACTUAL	W	**88**	9	0	0	2
	N1	7	**71**	12	0	11
	N2	1	20	**68**	2	9
	N3	0	0	3	**97**	0
	REM	3	22	4	0	**72**

Figure 8. Model's confusion matrix in the percentage of actual class

Discussion

	Method	CV	EEG Channel	Overall Performance			Per-class Performance (F1)				
				ACC	MF1	κ	W	N1	N2	N3	REM
	Proposed model	20-fold	Fpz-Cz	79.0	75.0	0.70	88.0	**63.0**	72.0	**97.0**	73.0
Inter.	SleepEEG Net [18]	20-fold	Fpz-Cz	**84.3**	**79.7**	**0.79**	**89.2**	52.2	**86.8**	85.1	**85.0**
	Tsinalis et al. [29]	20-fold	Fpz-Cz	78.9	73.7	-	71.6	47.0	84.6	84.0	81.4
	Tsinalis et al. [17]	20-fold	Fpz-Cz	74.8	69.8	-	65.4	43.7	80.6	84.9	74.5

Table 3. Comparison of performance obtained by the study with other state-of-the-art algorithms

- The part in bold is the best result of this column
- The proposed model are competitive with the state-of-the-art results
- CNN models has the potential as valuable tools for the automatic sleep stage scoring in clinical practice

CONCLUSION

- The model presented in this study utilizes a combination of frequency-specific activities over time to predict the class for each sample.
- Despite its simplicity, the model gives a good result and is comparable with other state-of-the-art methods.
- The application of AI and DL in the automatic sleep stage scoring has significant economic, social, and global impacts ➔ Enhancing the efficiency and accuracy of sleep stage scoring through AI-based systems can lead to improved patient outcomes, alleviate the strain on healthcare systems.

Future Directions

- Train on the new dataset obtained from university's laboratory to develop an automatic sleep scoring tool specifically for Vietnamese community.
- Address the misclassifications in the CNN model & improve the overall performance by utilize the CNN model as the primary model while incorporate other models as auxiliary models such as the BiRNNs, Attention Decoder or the Bi-LSTMs.
- Extend this work using multimodal PSG signals including EEG, EOG and EMG to boost the performance of the sleep stage classification.

REFERENCES

1. Sleep-EDF database expanded. [cited 2023 June 18]; Available from: https://physionet.org/content/sleep-edfx/1.0.0/.
2. Mousavi, S., F. Afghah, and U.R. Acharya, SleepEEGNet: Automated sleep stage scoring with sequence to sequence deep learning approach. PloS one, 2019. 14(5): p. e0216456.
3. Tsinalis, O., et al., Automatic sleep stage scoring with single-channel EEG using convolutional neural networks. arXiv preprint arXiv:1610.01683, 2016.
4. Tsinalis, O., P.M. Matthews, and Y. Guo, Automatic sleep stage scoring using timefrequency analysis and stacked sparse autoencoders. Annals of biomedical engineering, 2016. 44: p. 1587-1597.
5. Supratak, A., et al., DeepSleepNet: A model for automatic sleep stage scoring based on raw single-channel EEG. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2017. 25(11): p. 1998-2008
6. Al-Hussaini, I., et al. SLEEPER: interpretable Sleep staging via Prototypes from Expert Rules. in Machine Learning for Healthcare Conference. 2019. PMLR.