Training Neural Networks with Domain Pattern-Aware Auxiliary Task for Sleep Staging

Taeheon Lee*, Jeonghwan Hwang*, Honggu Lee

Looxid Labs, Seoul, South Korea

LOOXID LABS

1 Introduction

Background

- Sleep staging is central to sleep medicine where changes in sleep stages are inferred from physiological recordings like EEG and EOG.
- Sleep staging is labor-intensive since trained technicians should manually inspect every epoch (30s segments) and annotate sleep stages for eight-hour long whole-night EEG recordings.
- Several machine learning approaches have been proposed to automate sleep staging. Among them, deep learning models have shown promising results.

Challenges and Motivation

- Due to small dataset sizes, neural networks often failed to acquire generalizable representations.
- Neural networks are often uninterpretable whereas clinical domains require explanations for decisions.
- Previous studies have shown that guiding neural networks with domain-knowledge features can improve performances and interpretability [1].
- Existence of certain EEG patterns provides an invaluable cue for scoring a stage from EEG recordings [2].

Objectives

- We introduce an auxiliary task to guide neural networks to learn significant EEG patterns for sleep staging.
- Specifically, our objective is to make neural networks to learn the existences of significant EEG patterns for sleep staging.

2 Main Contribution

Performance Improvements

- The auxiliary task significantly improves prediction accuracy.

Model Interpretability

- Convolutional filters learned important EEG patterns for identifying sleep stages. Important parts in data can be located from models.

References

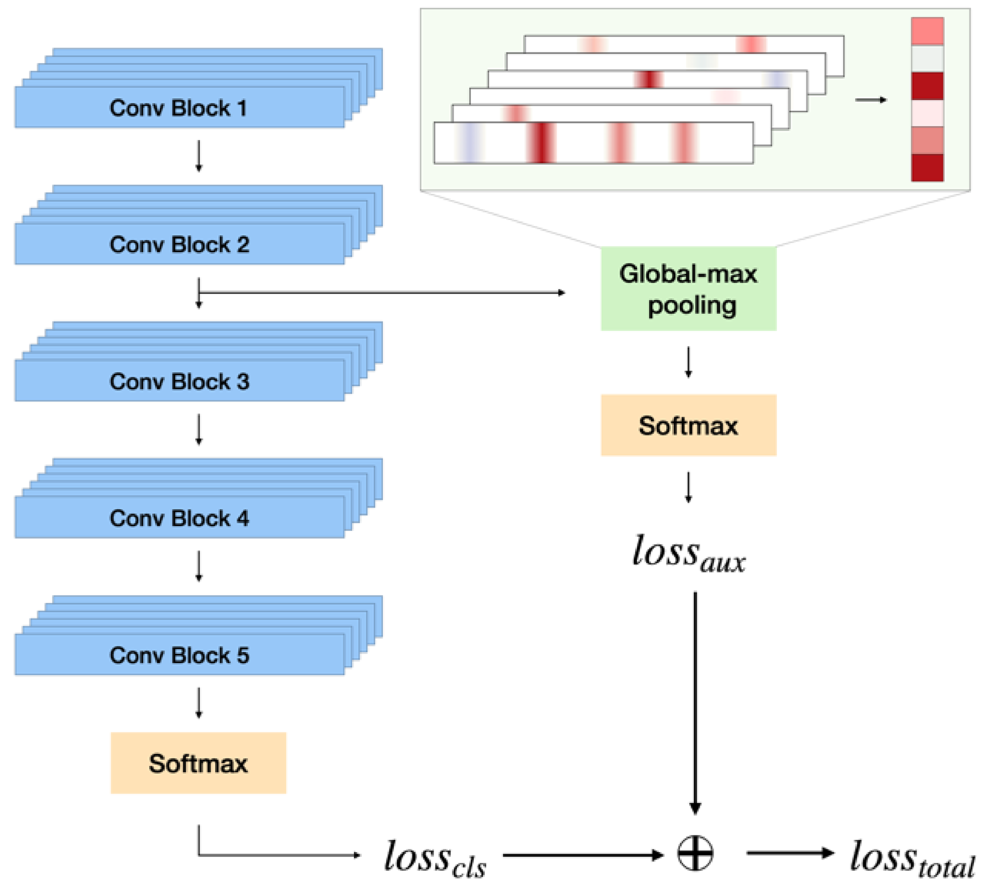
[1] Li, Xiaoyu, et al. "Domain knowledge guided deep atrial fibrillation classification and its visual interpretation." Conference on Information and Knowledge Management (2019).

[2] Berry, Richard B., et al. "The AASM manual for the scoring of sleep and associated events." Rules, Terminology and Technical Specifications, Darien, Illinois, American Academy of Sleep Medicine (2012).

[3] Dai, Wei, et al. "Very deep convolutional neural networks for raw waveforms." International Conference on Acoustics, Speech, & Signal Processing (2017).

[4] Sors, Arnaud, et al. "A convolutional neural network for sleep stage scoring from raw single-channel EEG." Biomedical Signal Processing and Control 42 (2018).

3 Methods



An Auxiliary task is proposed on the base loss function to model the existence of significant EEG patterns with the intermediate layers of convolutional neural networks.

Baseline Loss

From SoftMax probabilities $\hat{y}_i^{cls}$ for each class i, loss values are generally calculated with true label y_i.

$$loss_{cls} = -\sum_i y_i \log \hat{y}_i^{cls}$$

Auxiliary Loss

On intermediate layers of CNN, global max pooling operator can be applied to activation values of intermediate layers, $h^l \in \mathcal{R}^{1\times C\times L}$. $o^l \in \mathcal{R}^{1\times C}$, an output from global max pooling, represents whether features corresponding to each filter exist in the input.

On the representation vector o^l, auxiliary linear classifier (fc_{aux}) is attached and trained to classify stages from global max pooled activation values from the intermediate layer. Loss from this task can be calculated with predicted stages $\hat{y}_i^{aux}$ from this classifier.

$$loss_{aux} = -\sum_i y_i \log \hat{y}_i^{aux}$$

Classifier on global max pooling operator guides model to learn characteristic features whose existence in data can provide powerful information for identifying each class.

Total Loss

Total loss during training can be stated as a weighted sum between baseline loss and auxiliary loss with weight parameter λ

$$loss_{total} = loss_{cls} + \lambda \cdot loss_{aux}$$

4 Experimental Setup

Neural Networks

- **M-11, M-18**: CNN classifier for timeseries data [2]
- **SORS**: CNN for sleep staging from EEG signals [3]

Dataset

- **EDF**: 39 healthy subjects
- **ISRUC**: 100 recordings from subjects with sleep disorders

5 Classification Results

F1 scores estimated for each sleep stage and Macro averaged F1 scores

EDF dataset (F1 scores)

	Method	W	N1	N2	N3	REM	MF1
M-11	Base	0.85	0.25	0.83	0.81	0.70	0.688
	Ours-1	0.85	**0.30**	**0.86**	**0.84**	**0.76**	**0.721**
	Ours-2	0.85	**0.33**	**0.86**	**0.83**	**0.75.**	**0.725**
	Ours-3	0.85	**0.32**	**0.86**	**0.84**	**0.76.**	**0.726**
M-18	Base	0.82	0.27	0.82	0.80	0.67	0.676
	Ours-1	**0.86**	**0.31**	**0.86**	**0.84**	**0.76**	**0.726**
	Ours-2	**0.85**	**0.32**	**0.86**	**0.83**	**0.74**	**0.720**
	Ours-3	**0.84**	0.28	**0.85**	0.82	**0.73**	**0.704**
SORS	Base	0.87	0.32	0.85	0.81	0.73	0.717
	Ours-1	0.87	**0.44**	**0.86**	0.82	**0.79**	**0.755**
	Ours-2	0.86	**0.40**	**0.86**	0.82	**0.80**	**0.748**
	Ours-3	0.87	**0.43**	**0.87**	0.82	**0.81**	**0.758**

ISRUC dataset (F1 scores)

	Method	W	N1	N2	N3	REM	MF1
M-11	Base	0.84	0.41	0.75	0.84	0.66	0.698
	Ours-1	**0.86**	**0.46**	**0.77**	**0.85**	**0.69**	**0.727**
	Ours-2	**0.86**	**0.46**	**0.77**	**0.85**	**0.69**	**0.724**
	Ours-3	**0.86**	**0.45**	**0.76**	**0.86**	**0.69**	**0.727**
M-18	Base	0.84	0.38	0.74	0.83	0.66	0.690
	Ours-1	**0.86**	**0.45**	**0.77**	**0.86**	**0.70**	**0.727**
	Ours-2	**0.86**	**0.46**	**0.77**	**0.85**	**0.69**	**0.726**
	Ours-3	**0.86**	**0.45**	**0.76**	**0.85**	**0.68**	**0.720**
SORS	Base	0.84	0.45	0.77	0.86	0.72	0.727
	Ours-1	**0.86**	**0.52**	**0.78**	0.86	0.72	**0.749**
	Ours-2	**0.87**	**0.53**	**0.78**	0.86	**0.76**	**0.762**
	Ours-3	**0.87**	**0.53**	**0.78**	0.87	**0.74**	**0.758**

*** Bold letters indicate significant improvement (*P*< .05) compared to Baseline. (Significances were estimated with Wilcoxon Signed-Rank test)**

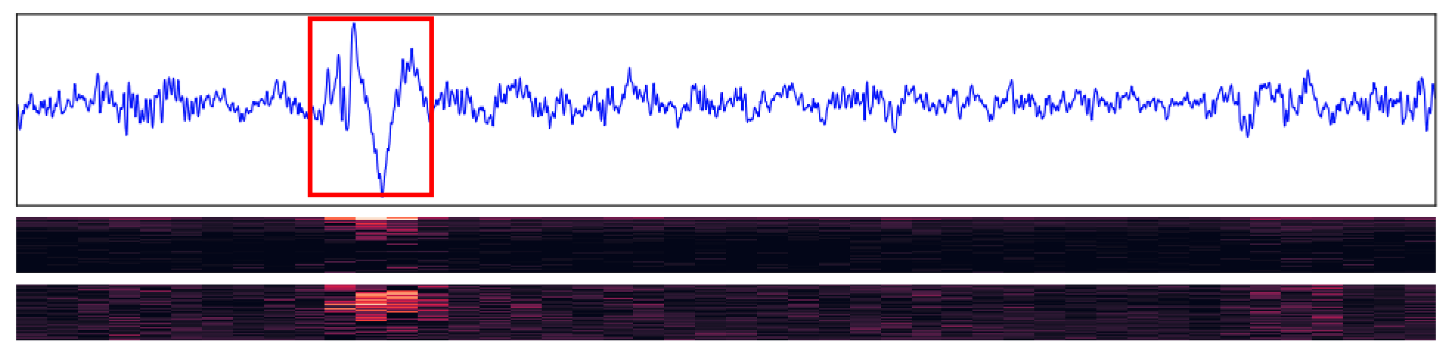
6 Investigation on Model Components

Input EEG Data and Activation Values from the Trained Convolutional Layer

EEG signals with k-complexes

Proposed Method

Baseline

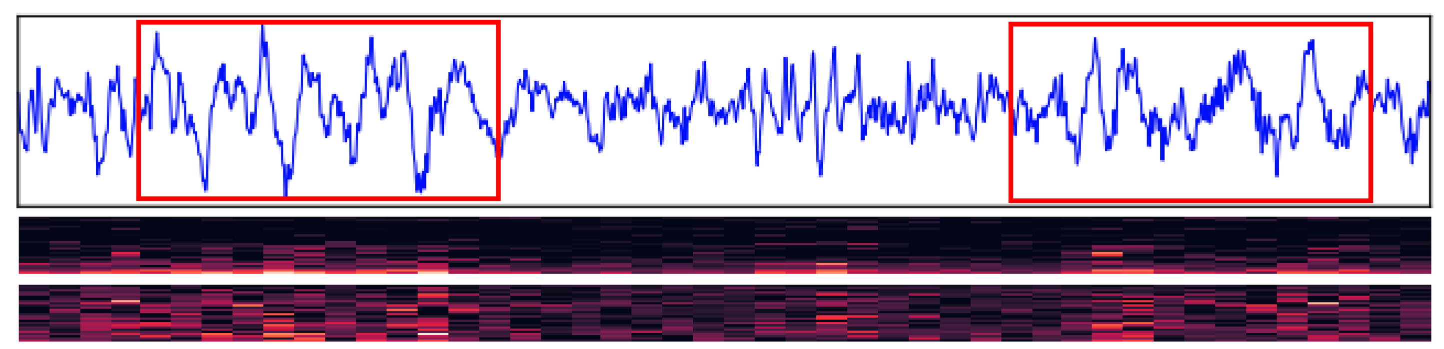


*** Existence of k-complex, which consists of negative and positive peaks in EEG amplitudes, is closely related to N2 stage.**

EEG signals with delta waves

Proposed Method

Baseline



*** Delta waves, slow fluctuations of EEG signals, are closely related to N3 stages.**