# End-to-end machine learning on raw EEG signals for sleep stage classification


E. Gunnlaugsson[2], H. Ragnarsdóttir[1], H.M. Þráinsson[1], E. Finnsson[1], S.Æ. Jónsson[1], H. Helgadóttir[1], J.S. Ágústsson[1], P. Herman[2]

1. Nox Research, Nox Medical, Reykjavík, Iceland,
2. Kungliga Tekniska Högskolan, Division of Computational Science and Technology, Stockholm, Sweden



Presented at World Sleep 2019, Vancouver

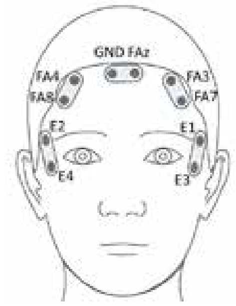


halla@noxmedical.com


## Introduction

We propose a method for automatic sleep stage classification by feeding raw EEG data into a Convolutional Neural Network (CNN), bypassing all feature extraction.

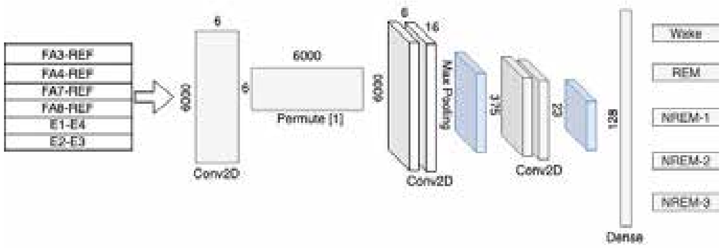
Most previous machine learning based methods have consisted of an EEG feature analysis step before machine learning based classification. Having this EEG feature analysis step before significantly slows down the procedure and can impose unsought side effects to the task, especially by constraining the information available in the classification task. It could therefore be very advantageous to bypass this EEG feature analysis step and use the raw signals directly for classification.




## Methods

This work explored recent advances in machine learning that aim to bypass pre-calculated features and use EEG/EOG data directly [1, 2, 3]. Various different neural network architectures were developed and validated on more than 400 anonymized nocturnal sleep studies recorded using the novel Self Applied Somnography (SAS) setup.

The architecture that gave the best results consisted of 5 convolutional layers, utilizing batch normalization and dropout layers for regularization and max-pooling layers to reduce the number of parameters inside the network.




The first convolutional layer implements *spatial filtering* introduced by Chambon et. al [1], the rest of the network consists of the following:

- Convolutional blocks, using ReLU activation function
- Max-pooling layer after each convolutional block
- Dense layer with 128 nodes, using ReLU activation function
- Dense output layer with 5 nodes, one for each sleep stage, using Softmax activation function

## Results

Model testing was done by using 290 recordings for training, 59 for validation and 60 for testing. Same test set was used for both methods

A baseline F1-score of 0.77 was achieved on the test data by a feed forward neural network (FFNN) trained using ~200 pre-calculated EEG and EOG features developed by Nox Research. Tests using the proposed method result in a F1-score of 0.73 not far from the baseline score.

Computational time tests for the baseline model resulted in a run time of ~185 seconds for a single recording, including both feature calculation and model prediction. The proposed method classified the same recordings in ~0.25 seconds on average, or 0.1% of the time required by the baseline model. All available recordings were used for these calculations

Normalized confusion matrix

| | Wake | REM | NREM-1 | NREM-2 | NREM-3 |
|---|---|---|---|---|---|
| Wake | 0.78 | 0.04 | 0.11 | 0.06 | 0.01 |
| REM | 0.06 | 0.69 | 0.09 | 0.16 | 0.0 |
| NREM-1 | 0.17 | 0.12 | 0.38 | 0.33 | 0.0 |
| NREM-2 | 0.03 | 0.03 | 0.05 | 0.86 | 0.03 |
| NREM-3 | 0.01 | 0.0 | 0.0 | 0.46 | 0.53 |

| | Baseline | Proposed |
|---|---|---|
| Classification time | 185±39s | 0.25±0.12s |
| Test set F1-score | 0.77 | 0.73 |

## Training

Training was done on batches of 64 epochs randomly chosen from the training set. The *Adam* optimizer was used for weight tuning. Early stopping was used to stop the training process when performance stopped improving on a hold-out validation set.

The network then restored to the weights that gave best performance on the hold-out validation set.

## Conclusion

Results show that models using raw EEG data are a realistic replacement to methods relying on pre-calculated hand-crafted features when classifying sleep stages using the SAS setup.

Further architectures are yet to be tested and methods like model ensembling and training-set data cleaning could be utilized to bring the accuracy of the method closer to state-of-the-art methods.

Additionally, these results encourage the development of similar methods for different EEG classification tasks, like arousal detection and spindle/K-complex detection.


## Acknowledgement

This project was developed as a master's thesis, supervised by Pawel Herman associate professor at Kungliga Tekniska Högskolan, Stockholm. The project was done in collaboration with Reykjavík University and Nox Medical, and supported by the Icelandic Centre for Research RANNÍS (175256-0611) and the European Union's Horizon 2020 SME Instrument (733461)

European Commission

Horizon 2020
European Union funding
for Research & Innovation

rannís


## References


1. Chambon, Stanislas, et al. "A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series." IEEE Transactions on Neural Systems and Rehabilitation Engineering 26.4 (2018): 758-769.
2. Bresch, Erik, Ulf Grossekathofer, and Gary Garcia-Molina. "Recurrent deep neural networks for real-time sleep stage classification from single channel EEG." Frontiers in computational neuroscience 12 (2018): 85.
3. Zhang, Junming, and Yan Wu. "A new method for automatic sleep stage classification." IEEE transactions on biomedical circuits and systems 11.5 (2017): 1097-1110.