U-Time: A Fully Convolutional Network for Time Series Segmentation Applied to Sleep Staging

Mathias Perslev[1], Michael Hejselbak Jensen[1], Sune Darkner[1], Poul Jørgen Jennum[2] & Christian Igel[1]

[1] Department of Computer Science, University of Copenhagen, Denmark [2] Danish Center for Sleep Medicine, Rigshospitalet, Denmark

Overview

- Most deep learning systems for physiological time-series analysis combine convolutional and recurrent layers. However, they are difficult to tune and optimize. They often require task-specific modifications.
- We suggest *U-Time*; a fully feed-forward model for time-series segmentation based on the U-Net image model[6, 7]. Evaluated on the problem of *sleep staging*, we find that U-Time has the following benefits:
  - Higher segmentation accuracy compared to CNN-LSTM type models
  - High robustness in its hyperparameters across datasets
  - Very fast inference (full patient scoring in seconds on laptop CPUs)
  - Ability to output segmentations of higher-than-labels temporal resolution through in-built multiple instance learning mechanism.

Method

- *U-Time* is a fully convolutional encoder-decoder network. It maps 1D time-series to contiguous segments of labels on a chosen temporal scale.
- Let $\mathbf{x} \in \mathbb{R}^{\tau S \times C}$ be a signal of C channels sampled at rate S for τ seconds. Let e be the frequency at which we want to segment $\mathbf{x}$. Our goal is to map $\mathbf{x}$ to $\lfloor \tau e \rfloor$ labels where each label represents $i = S/e$ connected samples.
- The input x to *U-Time* are T fixed-length connected segments of signal each of length i. The model $f(x;\theta) : \mathbb{R}^{T \times i \times C} \to \mathbb{R}^{T \times K}$ with parameters θ maps x to K class confidence scores for all T segments. It consists of three logical submodules:
  - The ***encoder*** and ***decoder*** map x to an intermediate, high frequency segmentation in $\mathbb{R}^{T \times i \times K}$. The ***segment classifier*** performs average pooling over segments of length i to project the segmentation down in frequency to match the label-space in $\mathbb{R}^{T \times K}$.
  - The ***segment classifier*** serves as a trainable link to the intermediate, high frequency segmentation.
  - The segmentation frequency, e, may be increased at inference time.

Experiments

- We evaluated *U-Time* for 'sleep-stage' segmentation of raw EEG data.
- *U-Time* was trained to segment an EEG into sleep stages in the label set $\{\mathrm{wake}, \mathrm{N1}, \mathrm{N2}, \mathrm{N3}, \mathrm{REM}\}$ with frequency $e = 1/30$ Hz.
- We evaluated U-Time across 7 different datasets without hyperparameter modifications. The datasets span healthy and diseased populations.
- We compared to single- and multi-channel methods from literature as well as our single-channel re-implementation of the state-of-the-art sleep staging CNN-LSTM described in [1] trained on the U-Time pipeline.
- We conducted a significant number of baseline permutation studies on the S-EDF-39 and DCSM datasets.

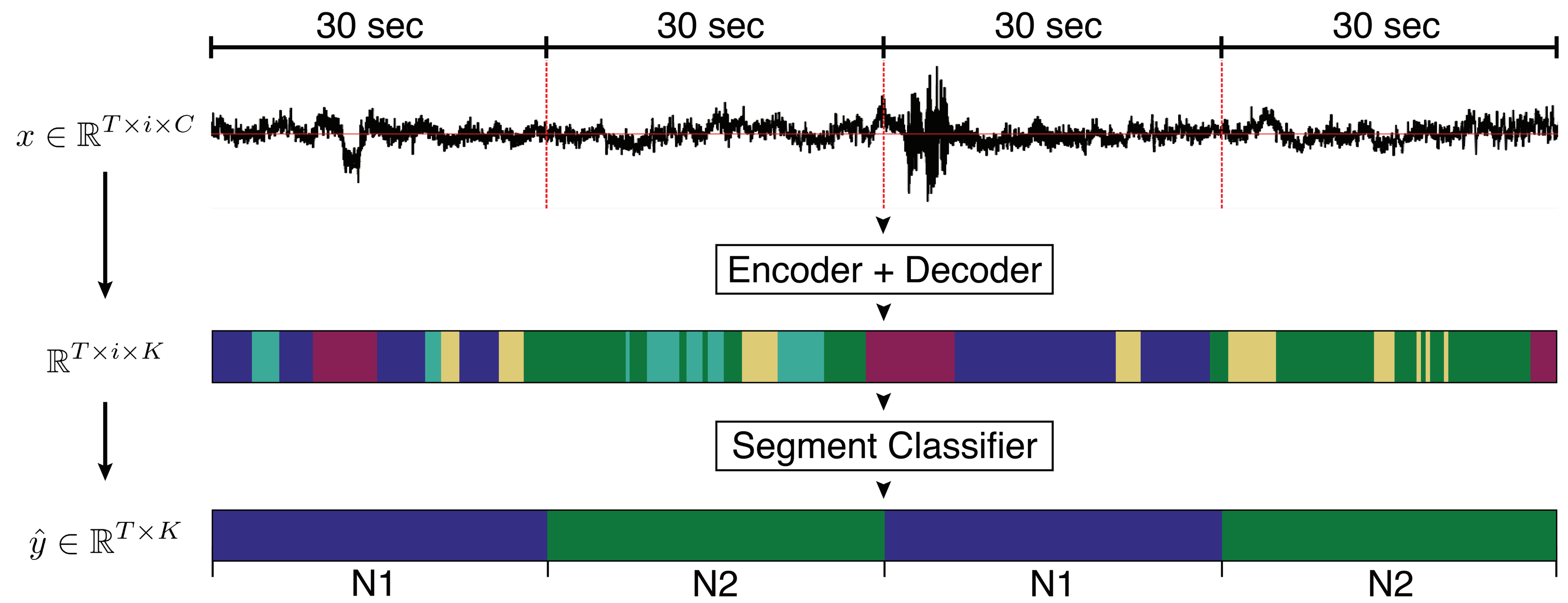

Figure 1: Illustration of how *U-Time* maps potentially very long input sequence to segmentations at a chosen temporal scale (here 1/30 Hz) by first segmenting the signal at every data-point and then aggregating these scores to form final predictions.

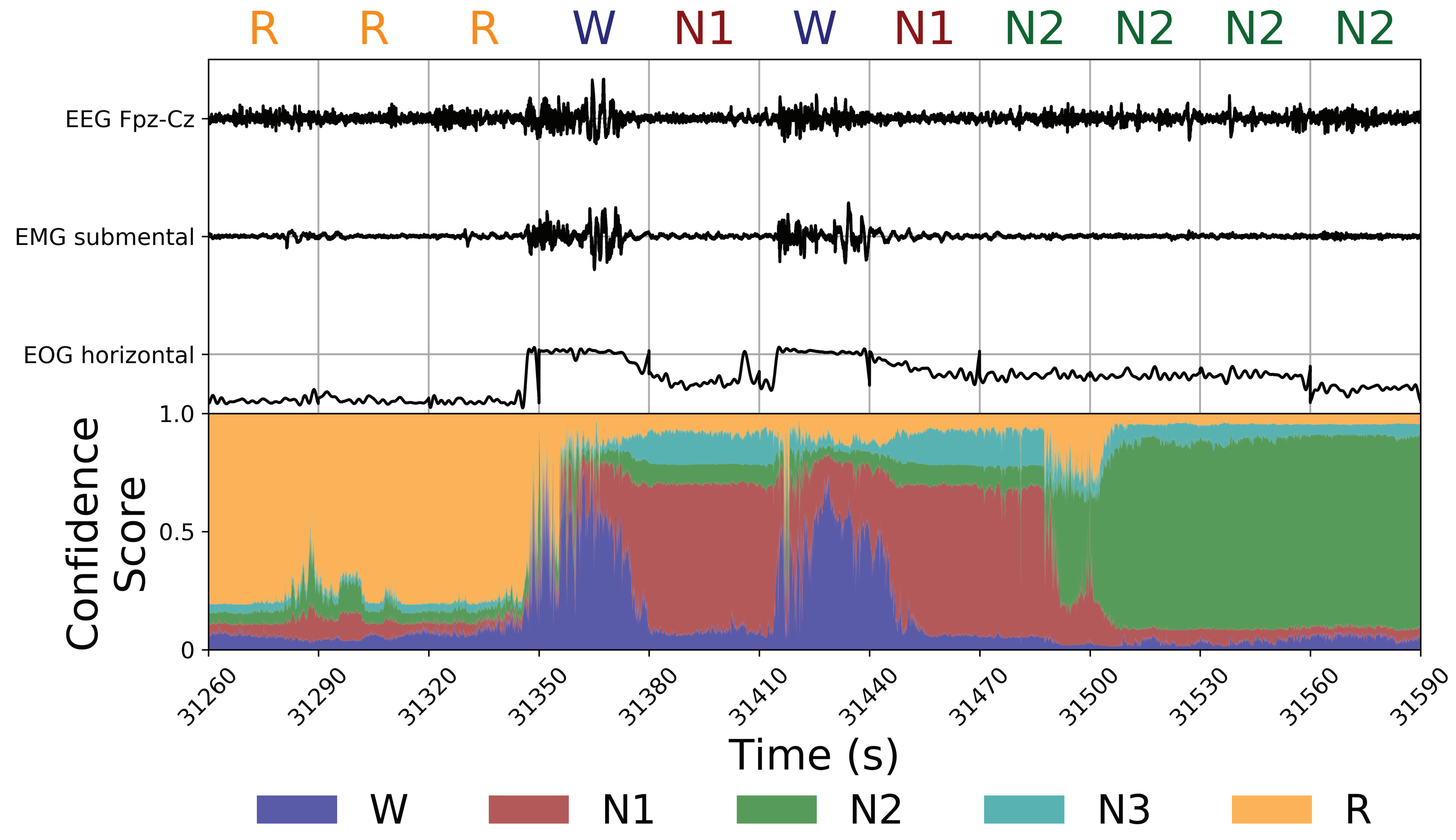

Figure 2: Visualization of the class confidence scores of *U-Time* on three input channels of a test subject from the Sleep-EDF-153 dataset. The segmentation frequency is set to match the input signal frequency of 100 Hz. I.e. *U-Time* outputs 100 sleep stage scorings per second.

Dataset	Size	Sample Rate	Channel	Scoring	Disorders
S-EDF-39	39	100	Fpz-Cz	R&K	None
S-EDF-153	153	100	Fpz-Cz	R&K	None
Physio-2018	994	200	C3-A2	AASM	Non-specific sleep disorders
DCSM	255	256	C3-A2	AASM	Non-specific sleep disorders
ISRUC	99	200	C3-A2	AASM	Non-specific sleep disorders
CAP	101	100-512	C4-A1/C3-A2	R&K	7 types of sleep disorders
SVUH-UCD	25	128	C3-A2	R&K	Sleep apnea, primary snoring

Table 1: Datasets overview. The 'Scoring' column reports the annotation protocol, 'Sample Rate' lists the original rate in Hz, and 'Size' gives the number of subjects included in our study after exclusions.

Results

		Eval		Global F1 scores					
Dataset	Model	Records	CV	W	N1	N2	N3	REM	mean
S-EDF-39	*U-Time*	39	20	0.87	0.52	0.86	0.84	0.84	0.79
	CNN-LSTM[1]	39	20	0.85	0.47	0.86	0.85	0.82	0.77
	VGGNet[2]	39	20	0.81	0.47	0.85	0.83	0.82	0.76
	CNN[3]	39	20	0.77	0.41	0.87	0.86	0.82	0.75
	Autoenc.[4]	39	20	0.72	0.47	0.85	0.84	0.81	0.74
S-EDF-153	*U-Time*	153	10	0.92	0.51	0.84	0.75	0.80	0.76
	CNN-LSTM	153	10	0.91	0.47	0.81	0.69	0.79	0.73
Physio-18	*U-Time*	994	5	0.83	0.59	0.83	0.79	0.84	0.77
	CNN-LSTM	994	5	0.82	0.58	0.83	0.78	0.85	0.77
DCSM	*U-Time*	255	5	0.97	0.49	0.84	0.83	0.82	0.79
	CNN-LSTM	255	5	0.96	0.39	0.82	0.80	0.82	0.76
ISRUC	*U-Time*	99	10	0.87	0.55	0.79	0.87	0.78	0.77
	CNN-LSTM	99	10	0.84	0.46	0.70	0.83	0.72	0.71
	Human obs.	99	-	0.92	0.54	0.80	0.85	0.90	0.80
CAP	*U-Time*	101	5	0.78	0.29	0.76	0.80	0.76	0.68
	CNN[5]	104	5	0.77	0.35	0.76	0.78	0.76	0.68
	CNN-LSTM	101	5	0.77	0.28	0.69	0.77	0.75	0.65
SVUH-UCD	*U-Time*	25	25	0.75	0.51	0.79	0.86	0.73	0.73

Table 2: U-Time results across 7 datasets. U-Time and our CNN-LSTM baseline process single-channel EEG data. Referenced models process either single- or multi-channel EEG data. F1 scores are computed across all subjects. Please refer to the supplementary material for per-subject summary metrics.

- Without hyperparameter tuning, U-Time performs at least at the level of methods from literature as well as the CNN-LSTM baseline.
- (Supplementary) We were not able to improve the CNN-LSTM baseline on neither S-EDF-153 nor DCSM across 13 architectural modifications.
- (Supplementary) U-Time might benefit from accepting multiple input channels. Expectedly, including an EOG (eye-movement) channel improves performance on the REM sleep stage.

References

[1] A. Supratak, H. Dong, C. Wu, and Y. Guo. DeepSleepNet: A model for automatic sleep stage scoring based on raw single-channel EEG. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25(11):1998–2008, 2017.
[2] A. Vilamala, K. H. Madsen, and L. K. Hansen. Deep convolutional neural networks for interpretable analysis of EEG sleep stage scoring. CoRR, abs/1710.00633, 2017.
[3] H. Phan, F. Andreotti, N. Cooray, O. Y. Chén, and M. D. Vos. Joint classification and prediction CNN framework for automatic sleep stage classification. CoRR, abs/1805.06546, 2018.
[4] O. Tsinalis, P. M. Matthews, and Y. Guo. Automatic sleep stage scoring using time-frequency analysis and stacked sparse autoencoders. Annals of Biomedical Engineering, 44(5):1587–1597, 2016.
[5] F. Andreotti, H. Phan, N. Cooray, C. Lo, M. T. M. Hu, and M. De Vos. Multichannel sleep stage classification and transfer learning using convolutional neural networks. In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pages 171–174, 2018.
[6] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention (MICCAI), volume 9351 of LNCS, pages 234–241. Springer, 2015.
[7] M. Perslev, E. B. Dam, A. Pai, and C. Igel. One network to segment them all: A general, lightweight system for accurate 3D medical image segmentation. In Medical Image Computing and Computer Assisted Intervention (MICCAI), volume 11765 of LNCS, pages 30–38. Springer, 2019.

UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE