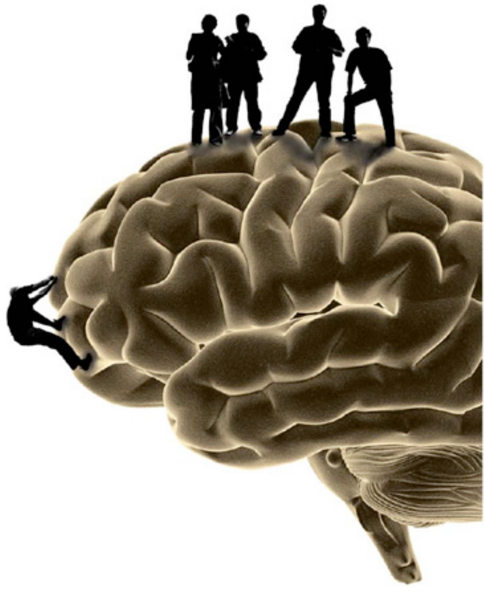
Data augmentation for learning predictive models on EEG

Alexandre Gramfort
http://alexandre.gramfort.net

work done while I was at:



June 2023

Outline

Cédric Rommel, Joseph Paillard, Thomas Moreau, Alexandre Gramfort (2022), Data augmentation for learning predictive models on EEG: a systematic comparison, Journal of Neural Engineering

Cédric Rommel

Joseph Paillard

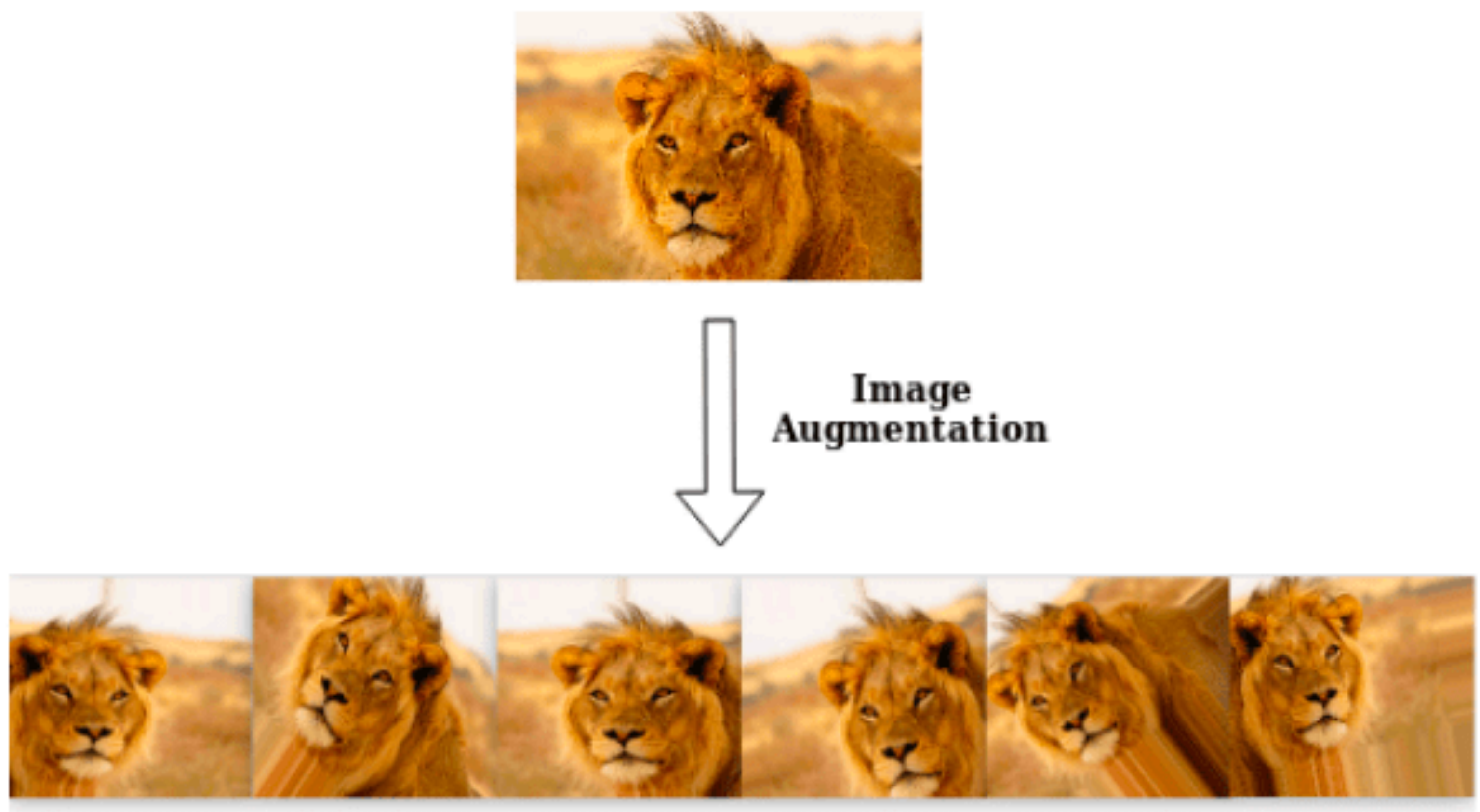
Thomas Moreau

https://github.com/eeg-augmentation-benchmark/eeg-augmentation-benchmark-2022

Data Augmentation

<u>Practical definition:</u> *Adding new synthetic examples to the training set by transforming existing ones in a target-preserving way*

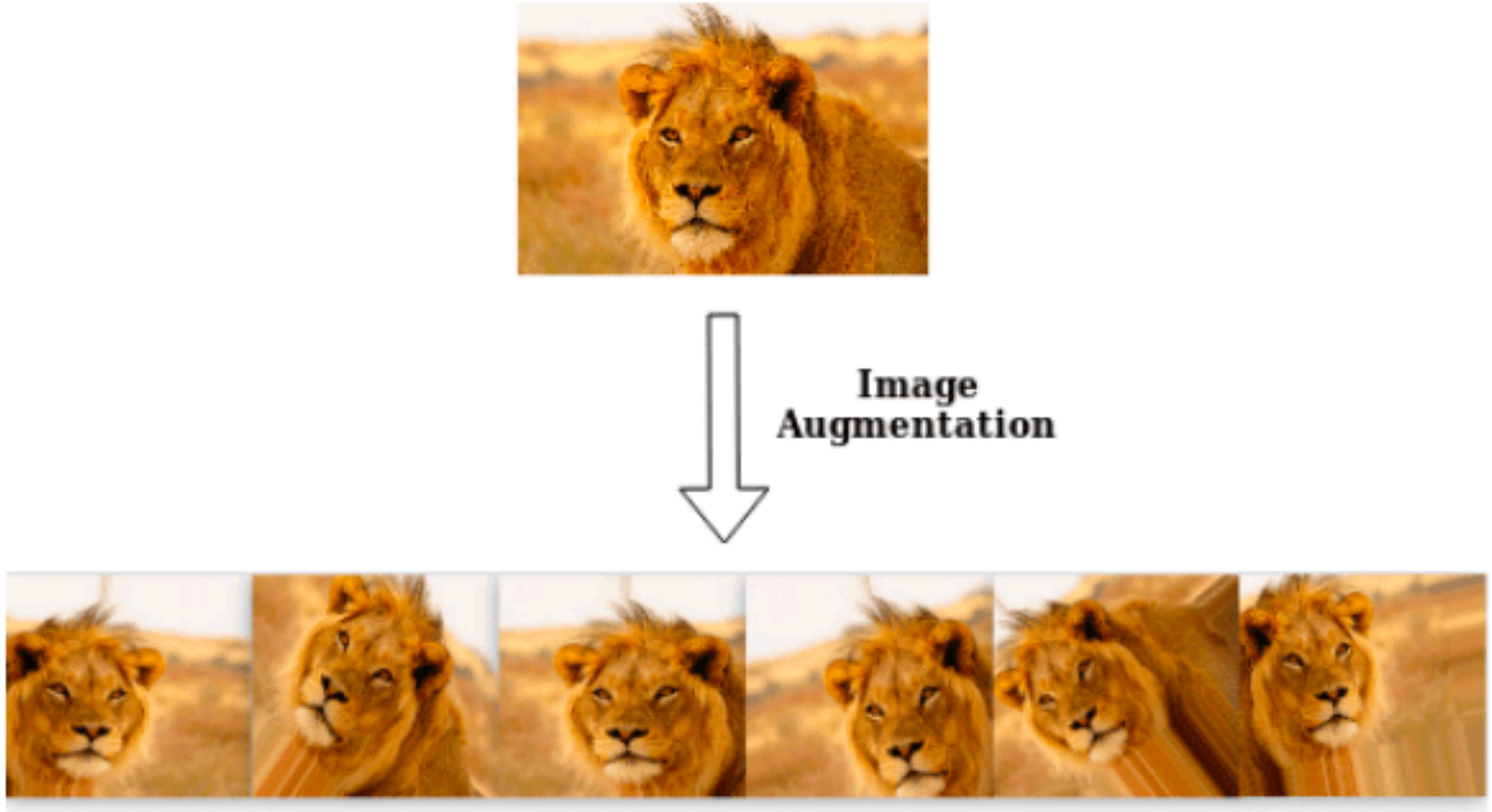
Augmentations encode invariances of the decision function and data known *a priori*



Data Augmentation

<u>Practical definition:</u> *Adding new synthetic examples to the training set by transforming existing ones in a target-preserving way*

Augmentations encode invariances of the decision function and data known *a priori*



Idea: distill the known invariances into the model (regularization/inductive bias)

EEG signals augmentation

- Neuroscience is a natural field of application for DA
- Brain data is **scarce**:
  1. Complex/costly to acquire
  2. Expert knowledge required to label it
- Yet, invariances are **not as intuitive** as for images

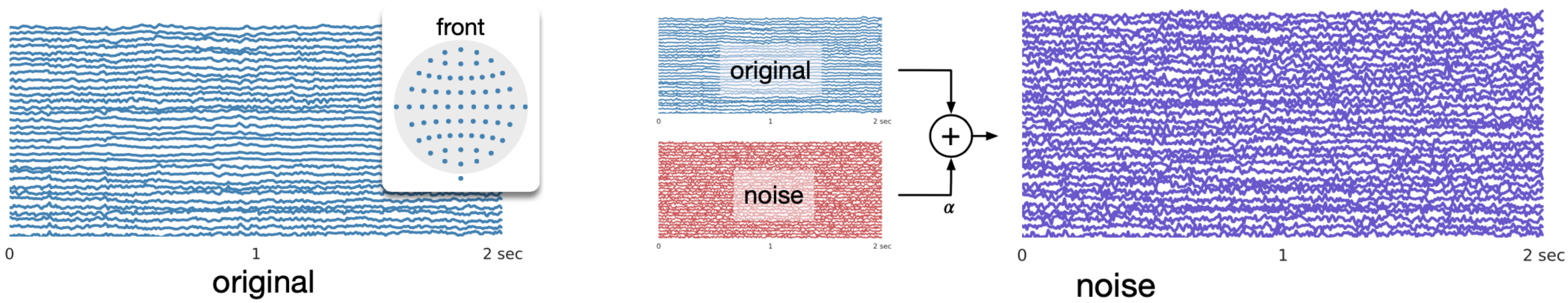
“Ok, if I flip the picture of a cat it still represents a cat…”

“But what transformations preserve sleep stage information in EEG signals?”

Existing transformations can be categorized:

noise, time, frequency, sensors

EEG DA: noise addition

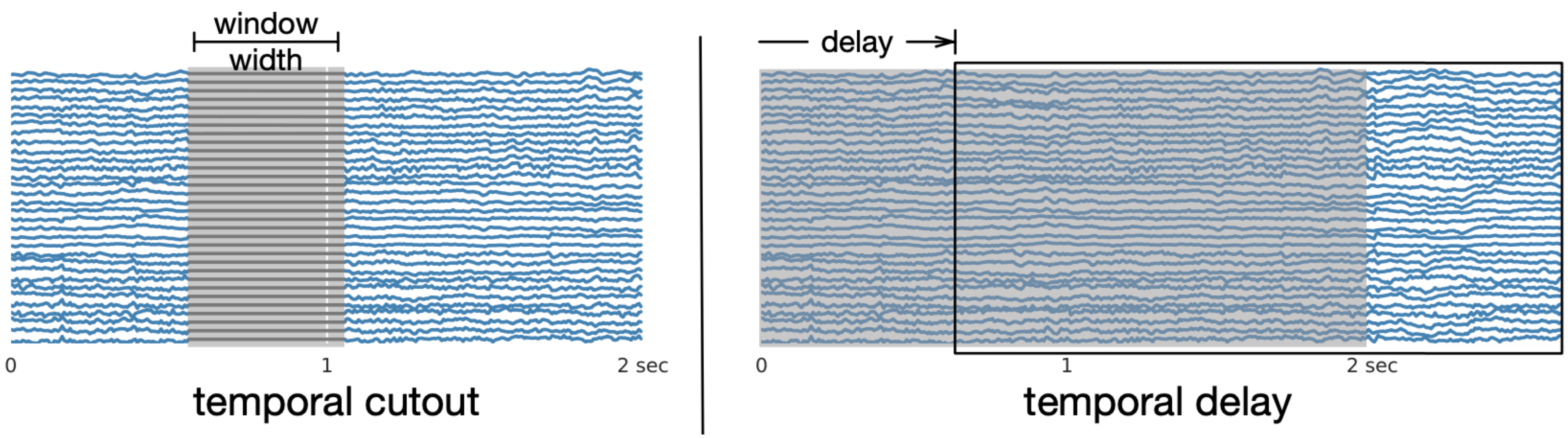


Picture from Cheng et al. (2020), Subject-Aware Contrastive Learning for Biosignals.

Add (Gaussian) **white noise** to the signal [1] or to features derived from it [2]

[1] Wang et al. (2018) Data Augmentation for EEG-Based Emotion Recognition with Deep Convolutional Neural Networks. In MultiMedia Modeling.
[2] Yin and Zhang (2017). Cross-Session Classification of Mental Workload Levels Using EEG and an Adaptive Deep Learning Model. In Biomedical Signal Processing and Control.

Time domain augmentations

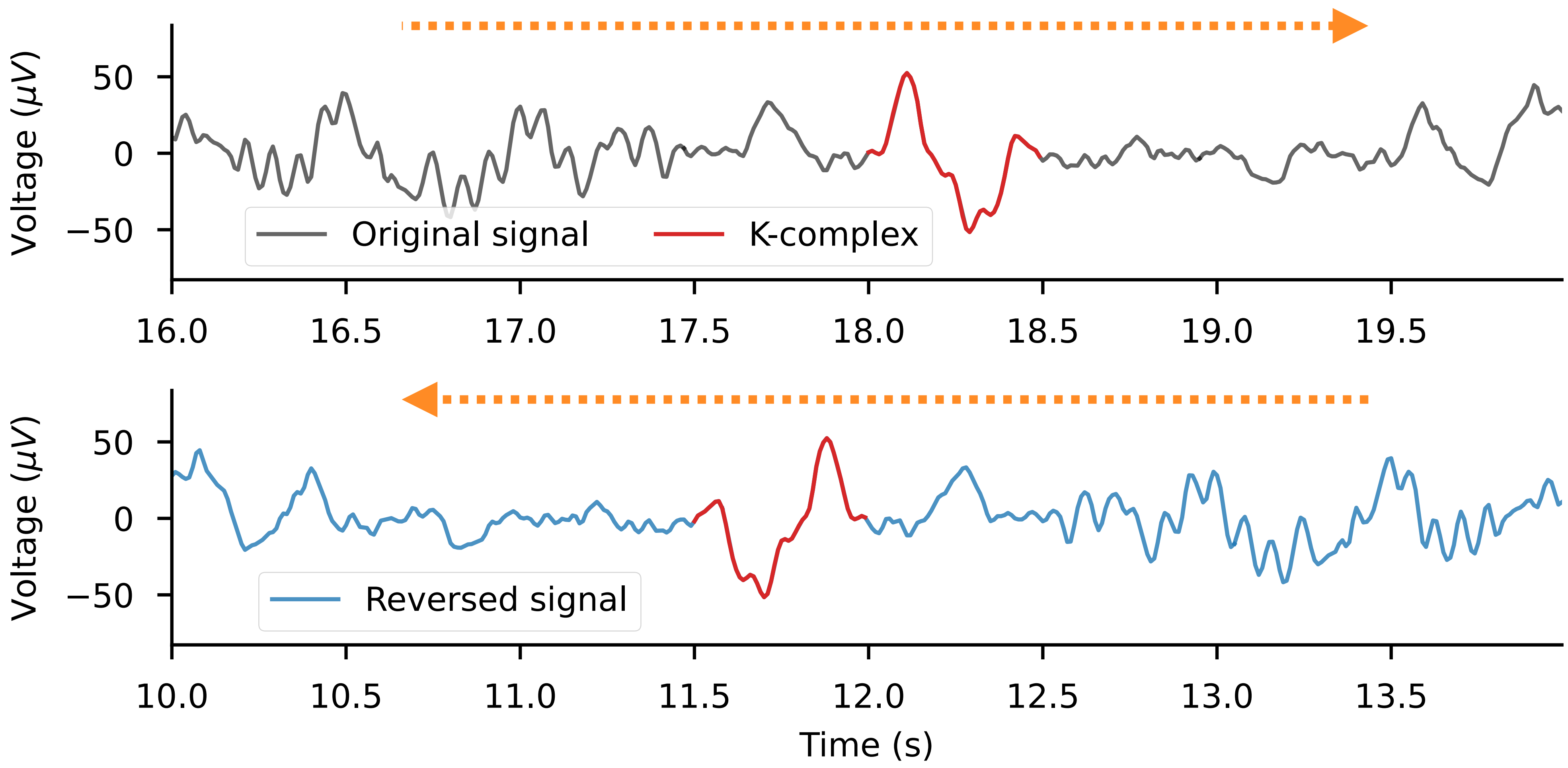


Picture from Cheng et al. (2020), Subject-Aware Contrastive Learning for Biosignals.

Temporal cutout/masking: *teach that predictive information is encoded in overall signal*

Temporal delay: *teach that human-annotations naturally have small misalignments*

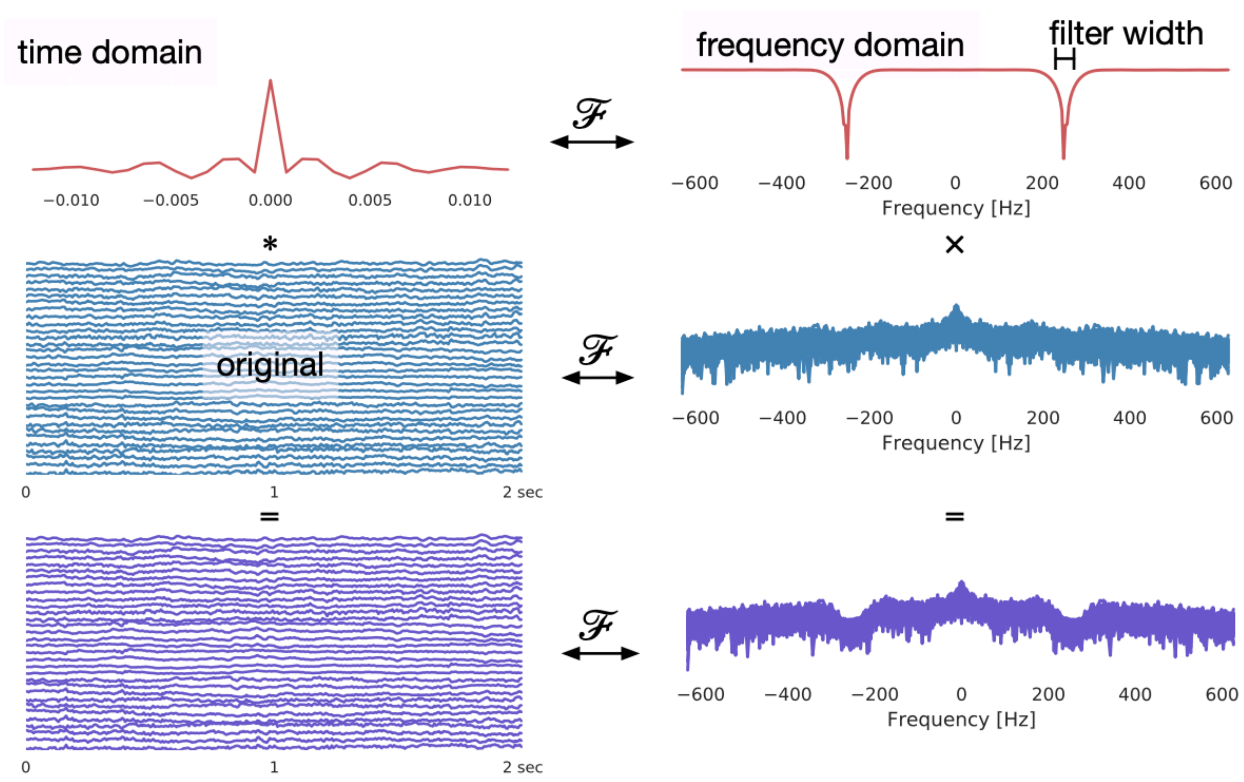
Time reverse



If we reverse time (flip signals w.r.t. time axis), relevant informations (e.g. relative proportions of frequencies) should be preserved

Frequency-based augmentation

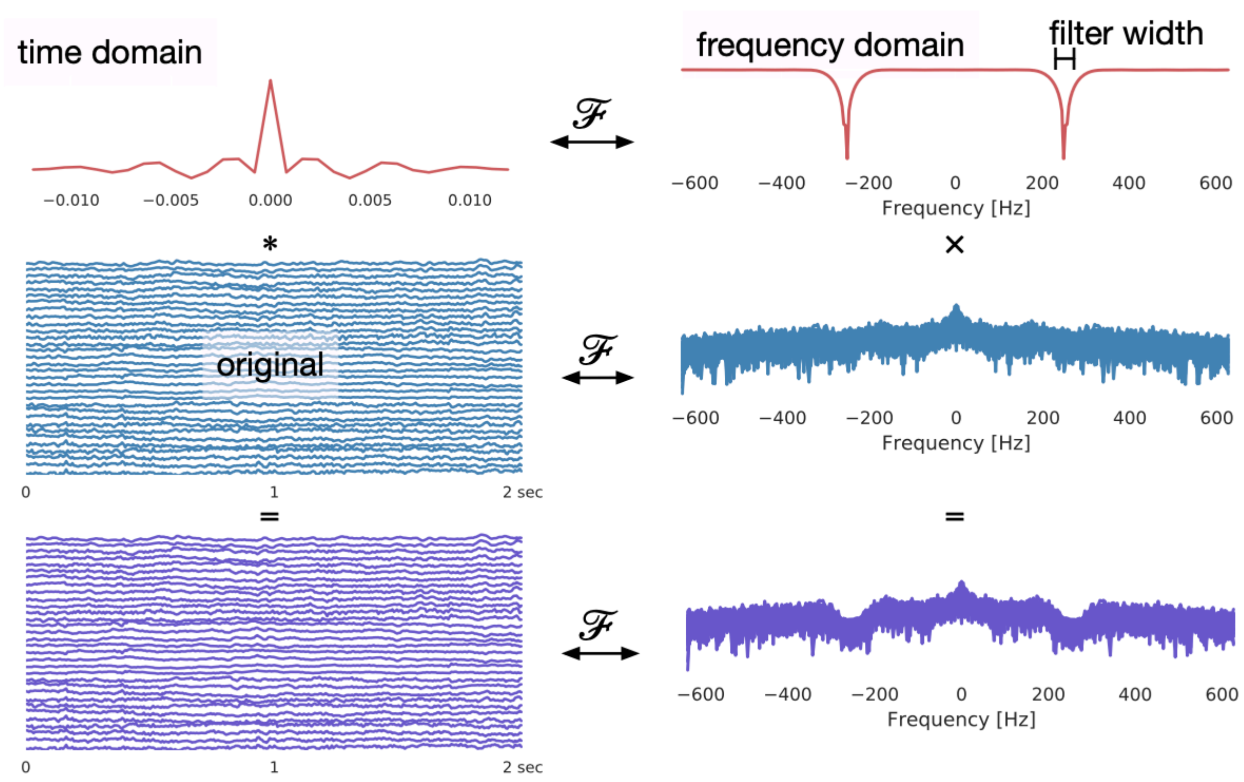
<u>Bandstop filtering</u>: *similar to Temporal cutout, but in the frequency domain*



Picture from Cheng et al. (2020), Subject-Aware Contrastive Learning for Biosignals.

Frequency-based augmentation

__Bandstop filtering__: *similar to Temporal cutout, but in the frequency domain*



Picture from Cheng et al. (2020), Subject-Aware Contrastive Learning for Biosignals.

__FT Surrogate__:

Idea: *Assume portions of EEG signals are stationary $\Rightarrow$ uniquely defined by their Fourier amplitudes*

[1] propose to randomize signals phases in the Fourier domain

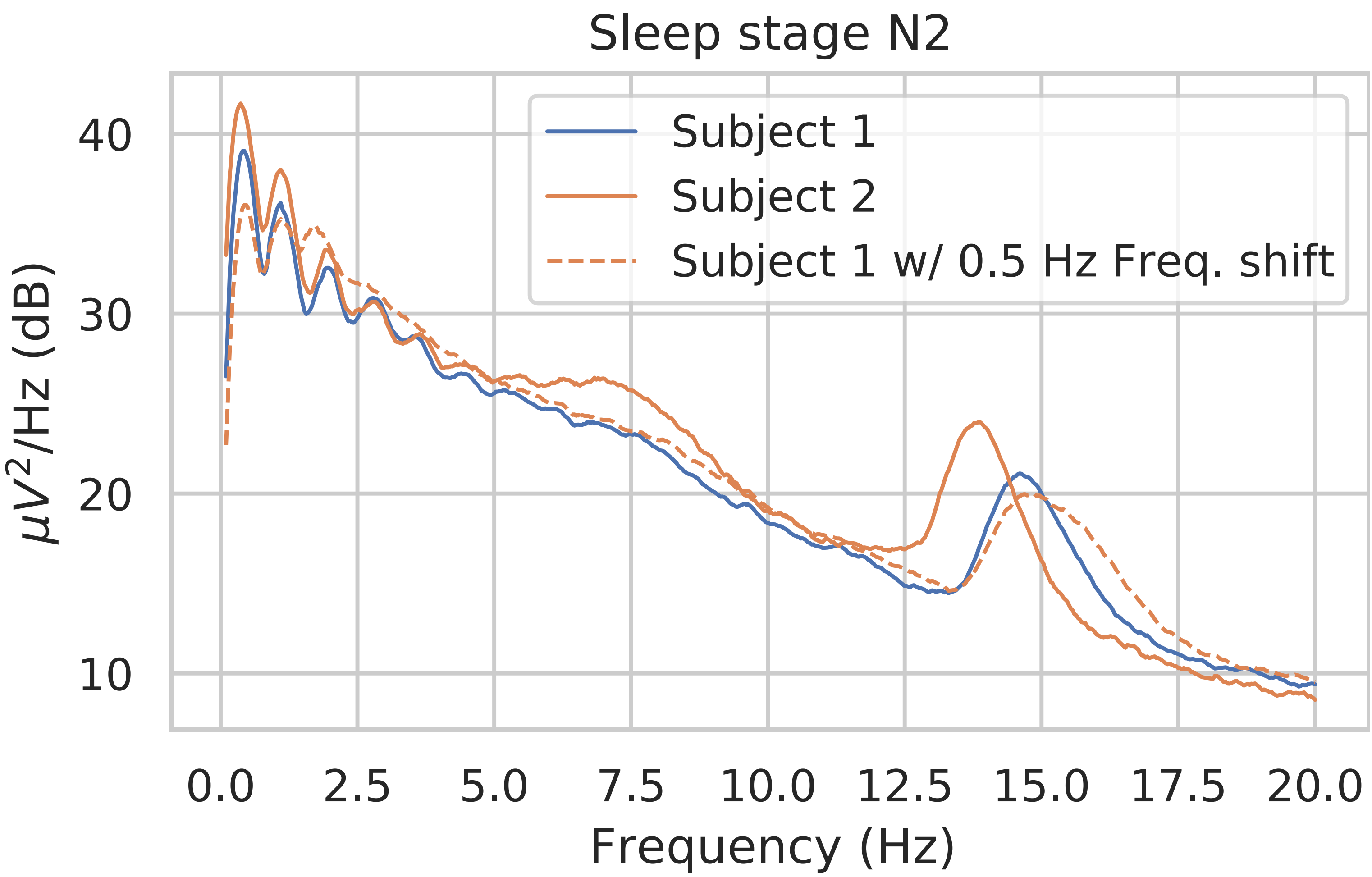
$$\mathscr{F}\left[\mathit{FTSurrogate}(x)\right](f) := \mathscr{F}[x](f) \cdot \exp(2\pi i \Delta\varphi),$$

where $\Delta\varphi$ is uniformly sampled from $[0, 2\pi)$

[1] Schwabedal et al. (2019). Addressing Class Imbalance in Classification Problems of Noisy Signals by using Fourier Transform Surrogates.

Frequency shift

- **Observation:** *different subjects have slightly shifted frequency peak for a given sleep stage*

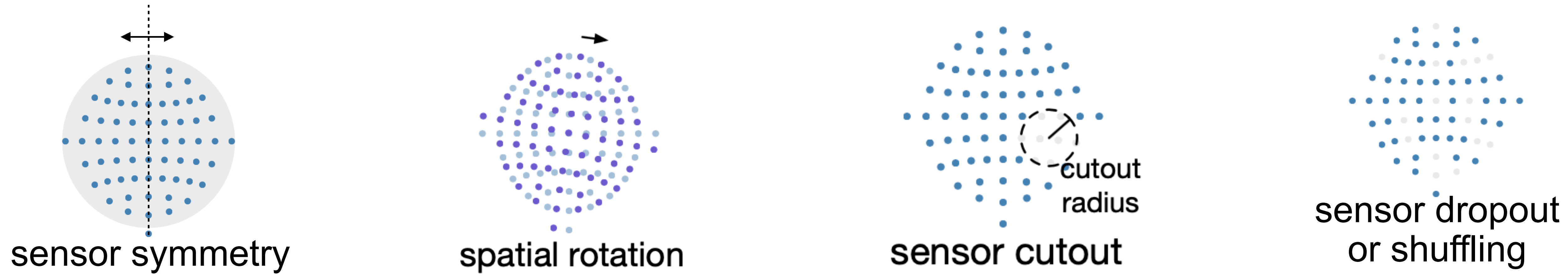


We propose to simulate such shifts using frequency modulation:

FrequencyShift$[x](t) :=$ *Re* $\left(H[x](t) \cdot \exp(2\pi i \Delta f t)\right),$ *where H denotes the analytic signal transform and* Δf *is the shift.*

Sensors domain

- **Sensors symmetry:** *exploit bilateral symmetry of the brain [1]*
- **Sensors rotation:** *simulate small perturbations in the cap placement over the head [2]*
- **Sensors cutout, dropout and shuffling:** *encourage robustness to recording inconsistencies [3,4]*

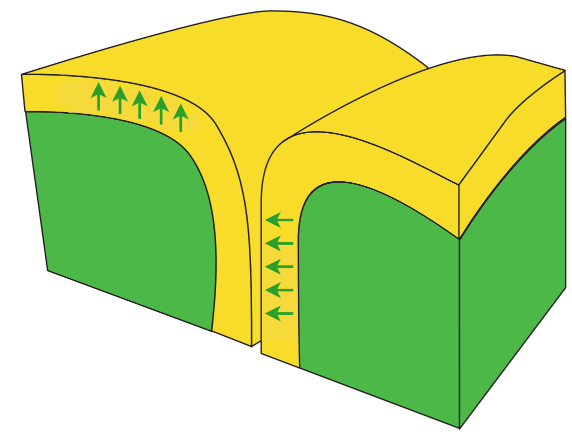


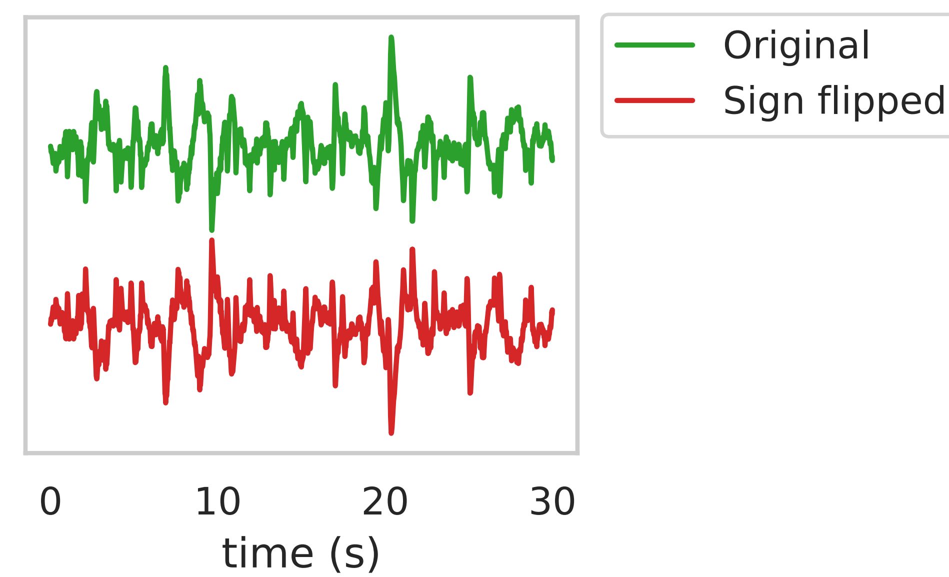
Picture from [3]

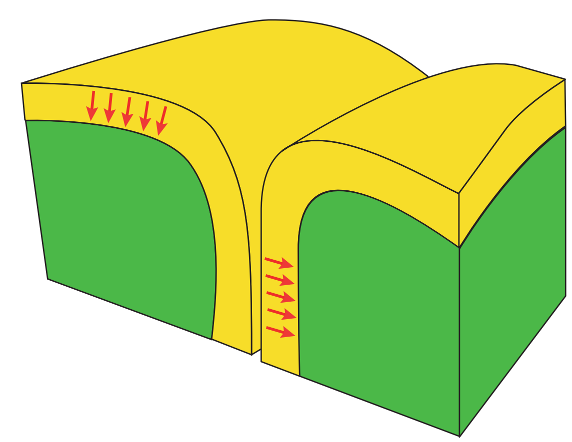
[1] Deiss et al. (2018). HAMLET: interpretable human and machine co-learning technique.
[2] Krell & Kim (2017). Rotational data augmentation for electroencephalographic data. In *EMBC*.
[3] Cheng et al. (2020), Subject-Aware Contrastive Learning for Biosignals.
[4] Saeed et al. (2021). Learning from heterogeneous eeg signals with differentiable channel reordering. In *ICASSP*.

Sign Flip

- Intuition:
  - *information encoded in the electrical potentials are likely to be invariant to the polarity of the electric field*
  - *both charges moving towards deeper and more superficial layers of the cortex are likely to happen*
- We propose to randomly flip the EEG signals sign







Augmentation	Type	Reference	Short description
`FTSurrogate`	F	Schwabedal et al. [42]	Randomize Fourier phases of all channels.
`BandstopFilter`	F	Mohsenvand et al. [27], Cheng et al. [10]	Randomly filter a small frequency band of all channels.
`FrequencyShift`	F	Rommel et al. [36]	Randomly translate all channels PSD by small shift.
`GaussianNoise`	T	Wang et al. [44]	Add Gaussian white noise to the signals.
`SmoothTimeMask`	T	Mohsenvand et al. [27]	Randomly pick a portion of the signal and set it to zero.
`SignFlip`	T	Rommel et al. [36]	Randomly flip the sign of all channels.
`TimeReverse`	T	Rommel et al. [36]	Randomly reverse the axis of time in all channels.
`ChannelsSymmetry`	S	Deiss et al. [12]	Randomly swap signals from right hemisphere to left hemisphere and *vice-versa*.
`ChannelsDropout`	S	Saeed et al. [40]	Randomly pick a given number of channels and set their signals to zero.
`ChannelsShuffle`	S	Saeed et al. [40]	Randomly pick a given number of channels and permute their signals.
`SensorsRotation`	S	Krell and Kim [22]	Interpolate channels signals on randomly rotated positions.

Table 1: Data augmentation methods studied in this work. Types stand for Frequency (F), Time (T) and Spatial (S) transformations.

[Rommel et al. JNE 2023]

Augmentation	Type	Reference	Short description
FTSurrogate	F	Schwabedal et al. [42]	Randomize Fourier phases of all channels.
BandstopFilter	F	Mohsenvand et al. [27], Cheng et al. [10]	Randomly filter a small frequency band of all channels.
FrequencyShift	F	Rommel et al. [36]	Randomly translate all channels PSD by small shift.
GaussianNoise	T	Wang et al. [44]	Add Gaussian white noise to the signals.
SmoothTimeMask	T	Mohse	
SignFlip	T	Romm	
TimeReverse	T	Romm	
ChannelsSymmetry	S	Deiss	
ChannelsDropout	S	Saeed	
ChannelsShuffle	S	Saeed	
SensorsRotation	S	Krell	rotated positions.

- All of the augmentations can be <u>more or less applied</u> during training
- Can be the <u>probability</u> of using it for a given data point or the <u>magnitude</u> of the transformation
- One can test the impact of <u>progressively</u> activating these transformations.

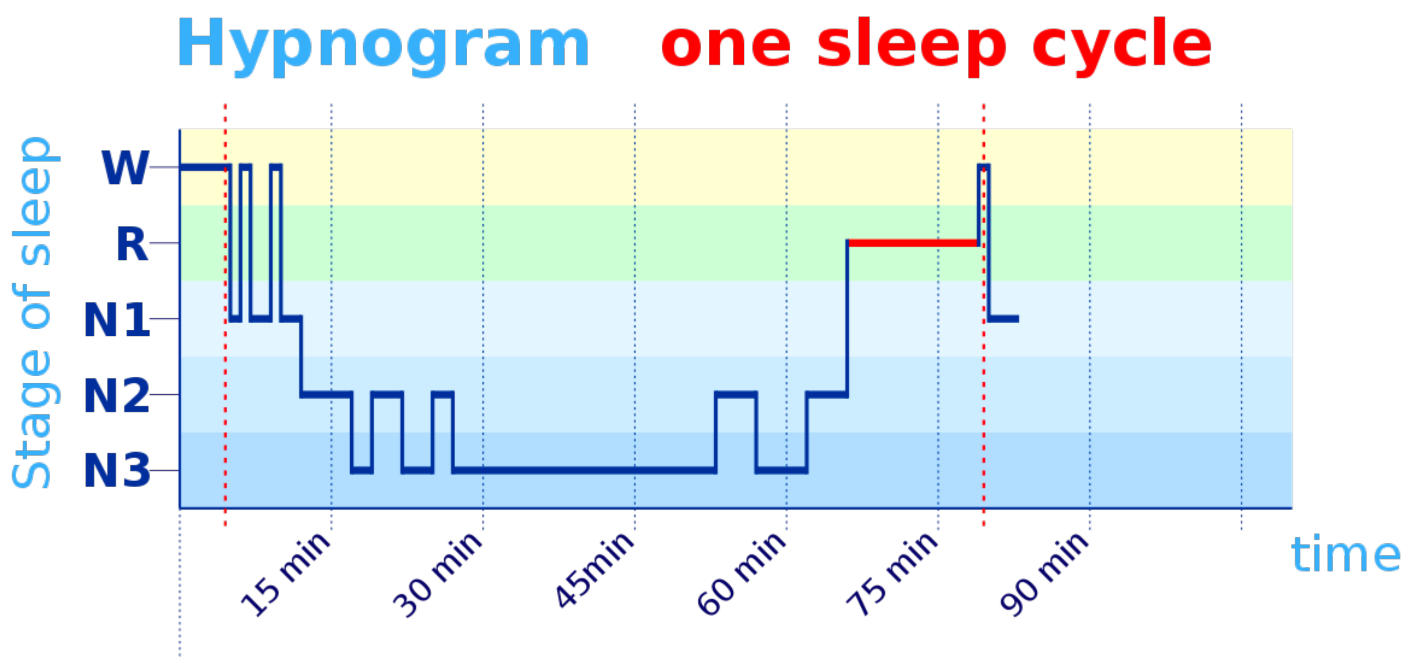
Table 1: Data augmentation methods studied in this work. Types stand for Frequency (F), Time (T) and Spatial (S) transformations.

[Rommel et al. JNE 2023]

Evaluation on sleep staging

<u>Sleep staging</u>: *assigning to windows of 30 seconds of signals a label among:*

- *Wake (W)*
- *Rapid Eye Movement (REM)*
- *Non-REM of depth 1 (N1)*
- *Non-REM of depth 2(N2)*
- *Non-REM of depth 3 (N3)*



For all following results, we used the Sleep Stager model from [1].

[1] Chambon et al. (2018), A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series.

Evaluation on motor imagery

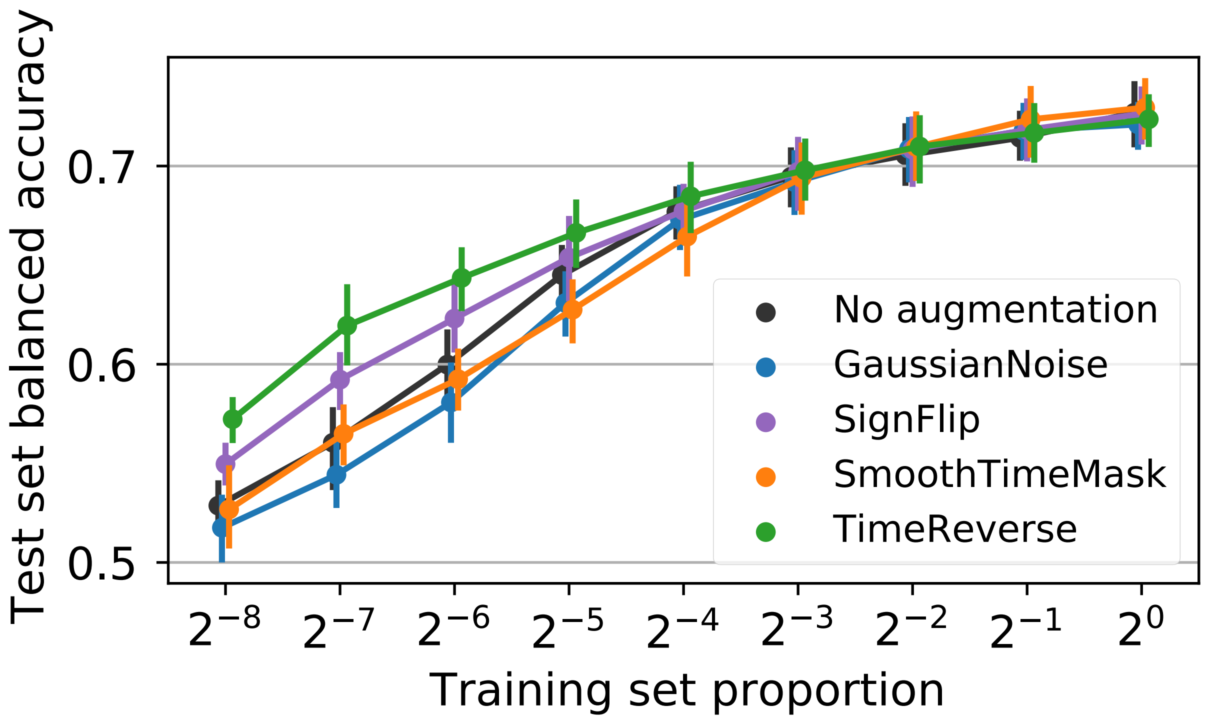
<u>Motor imagery (dataset BCI IV 2a)</u>: *assigning to windows of about 4 seconds of signals a label among:*

- *Foot*
- *Left-hand*
- *Right-hand*
- *Tongue*

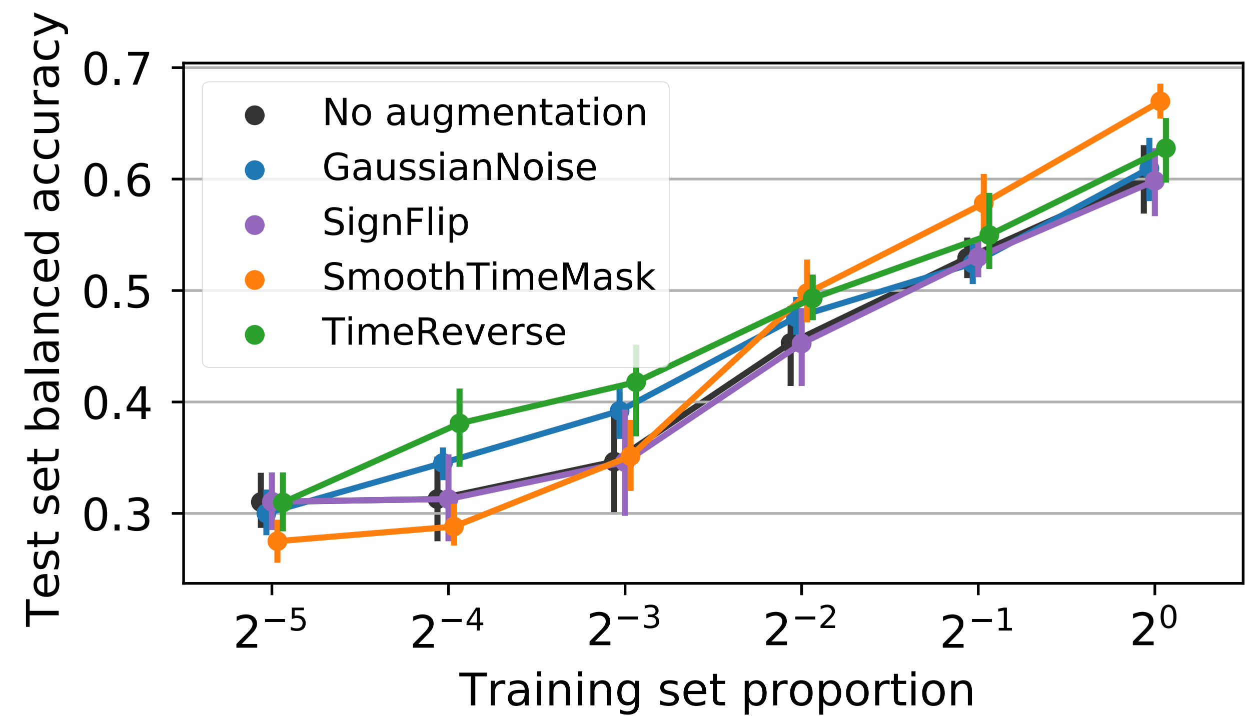
For all following results, we used the CNN FilterBank-CSP model from [1].

[1] Robin Tibor Schirrmeister, Jost Tobias Springenberg, Lukas Dominique Josef Fiederer, Martin Glasstetter, Katharina Eggensperger, Michael Tangermann, Frank Hutter, Wolfram Burgard, and Tonio Ball. Deep learning with convolutional neural networks for EEG decoding and visualization. Human Brain Mapping, 38 (11):5391–5420, 2017.

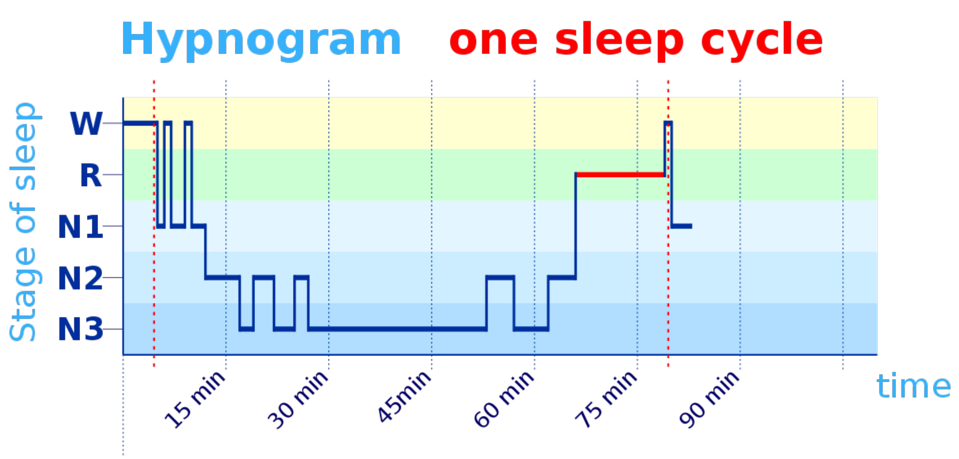
Results time domain: Learning curves



(a) *SleepPhysionet*

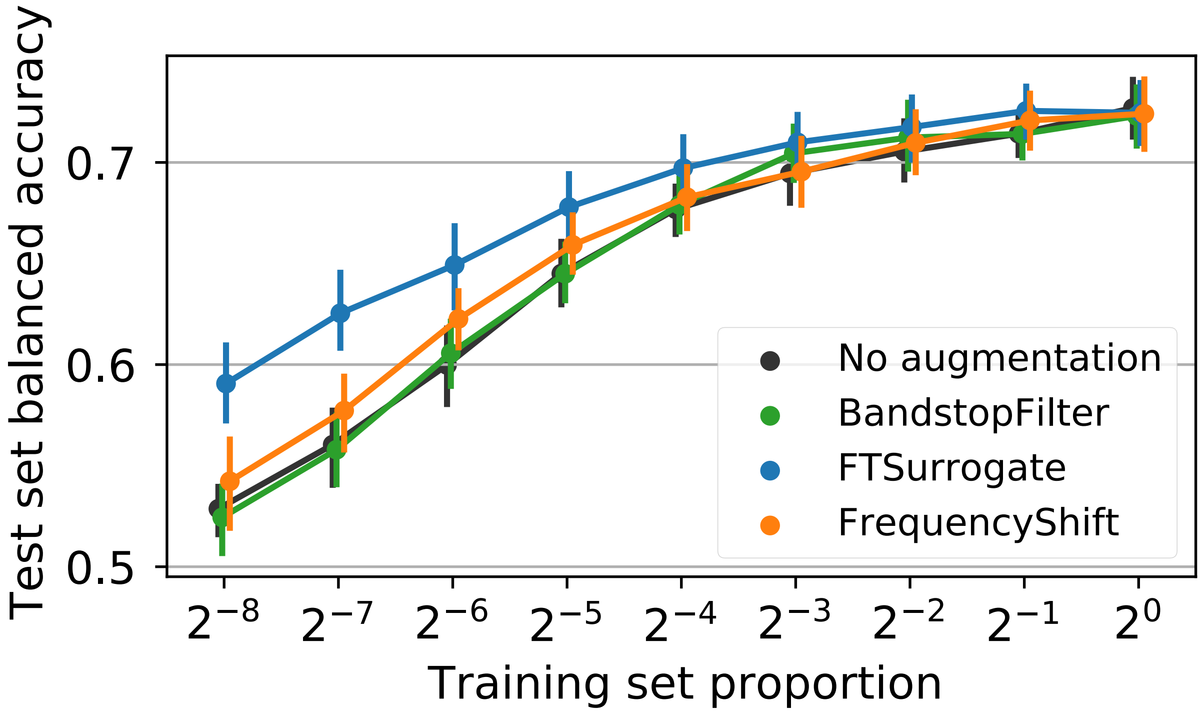


(b) *BCI IV 2a*

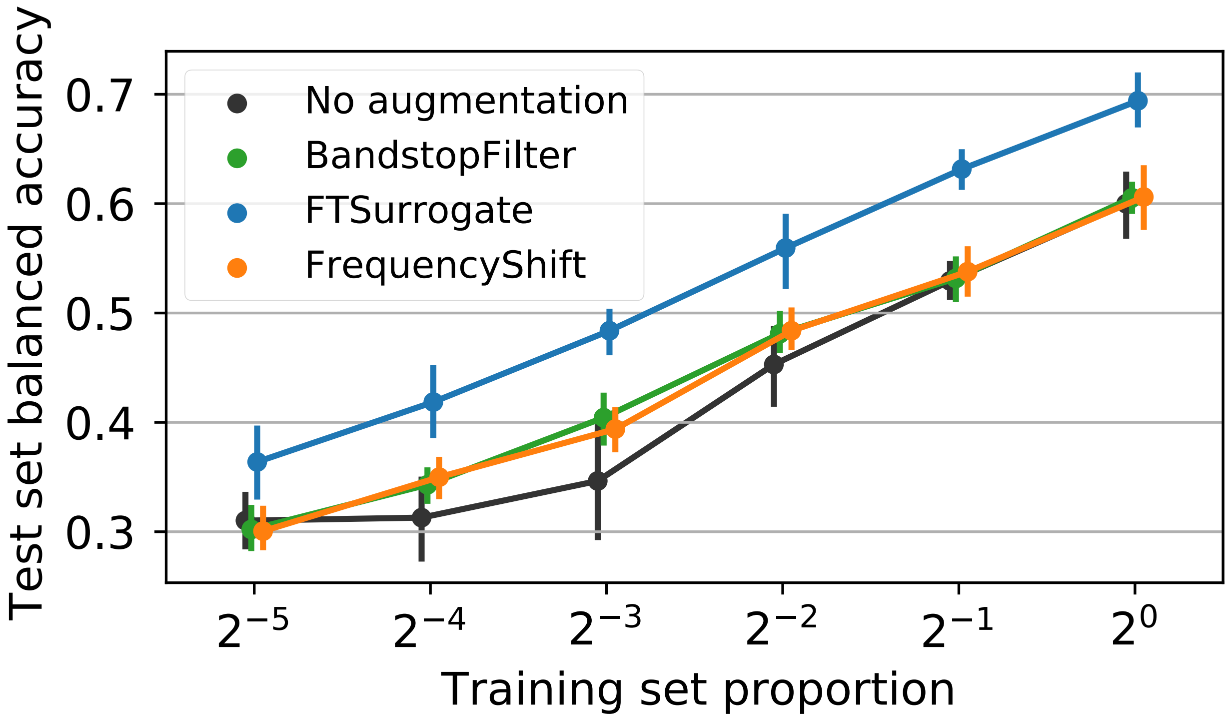


- The more data the better
- The strongest gain by augmentation is for low sample regime

Results frequency & channel domain

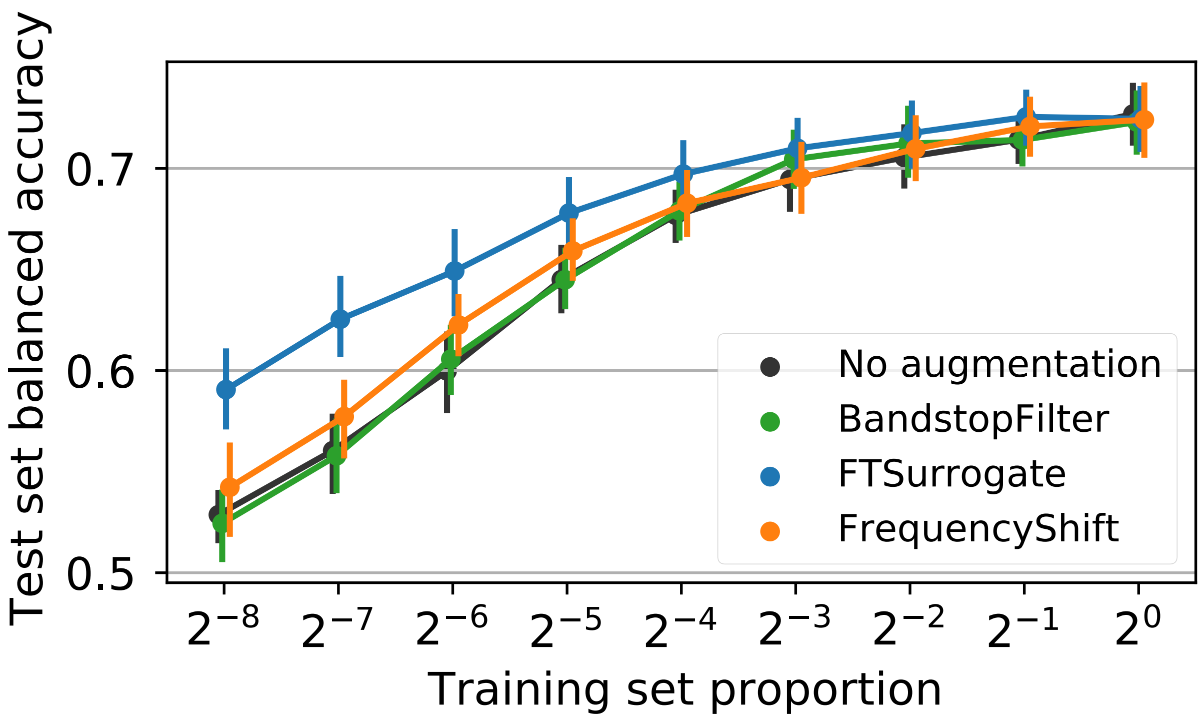


(a) *SleepPhysionet*

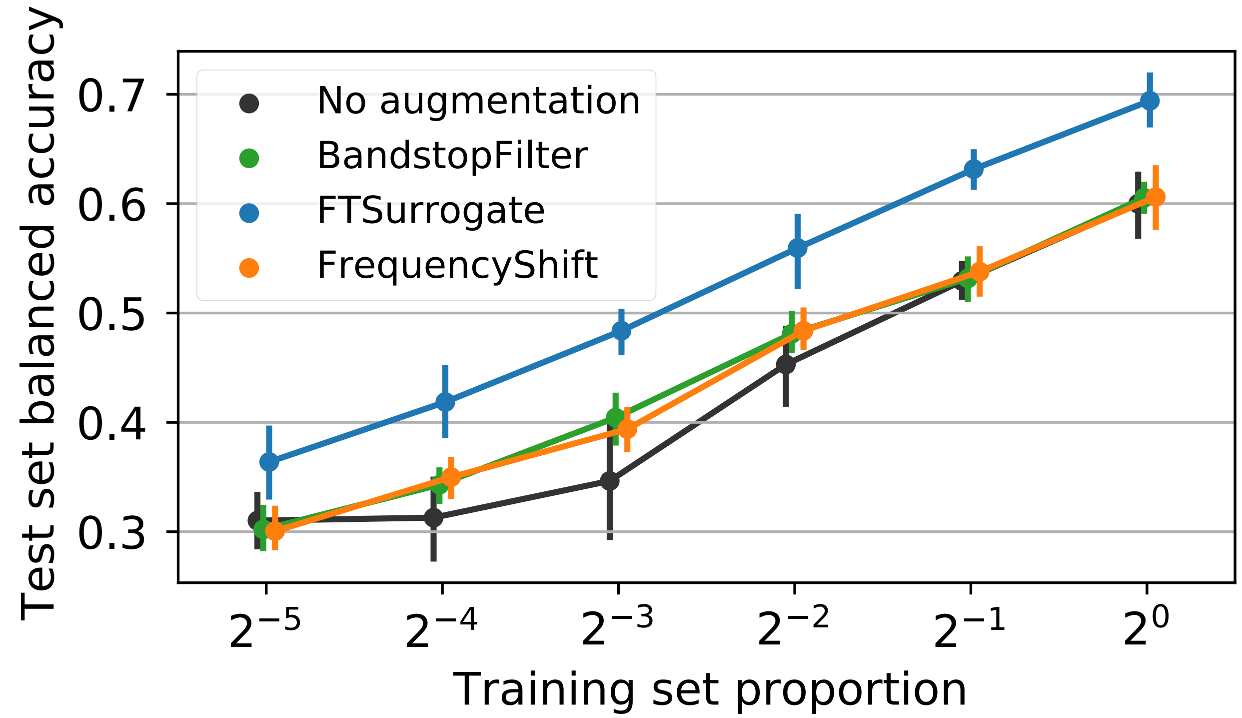


(b) *BCI IV 2a*

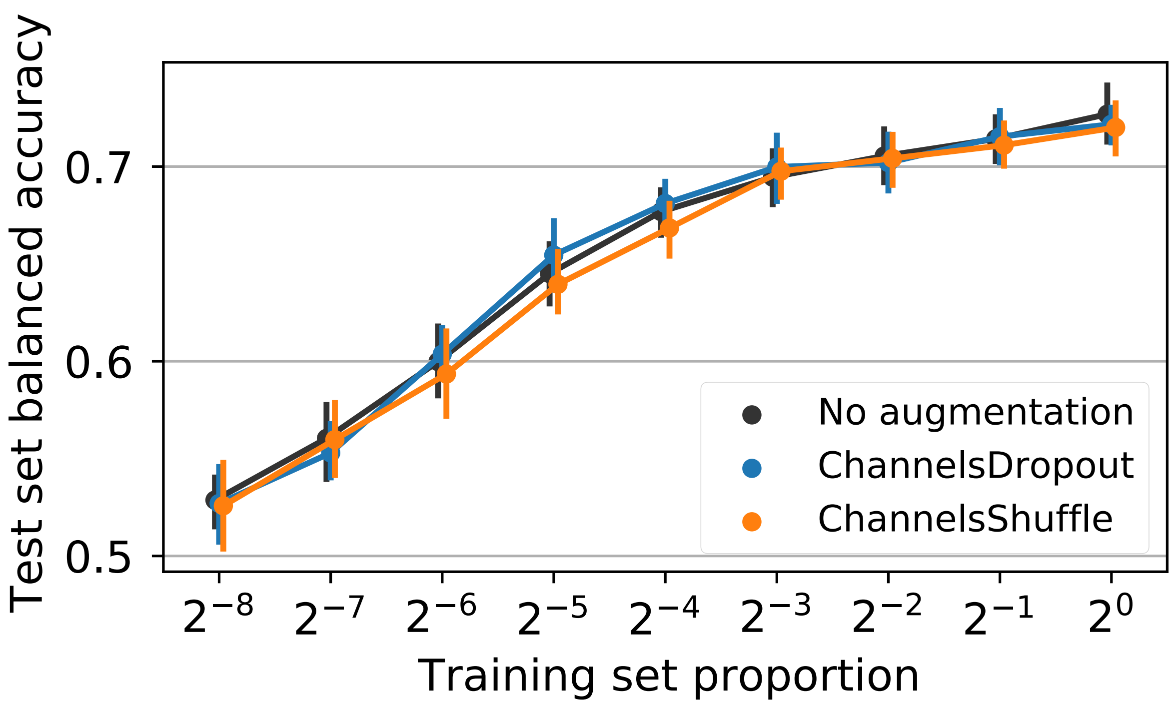
Results frequency & channel domain



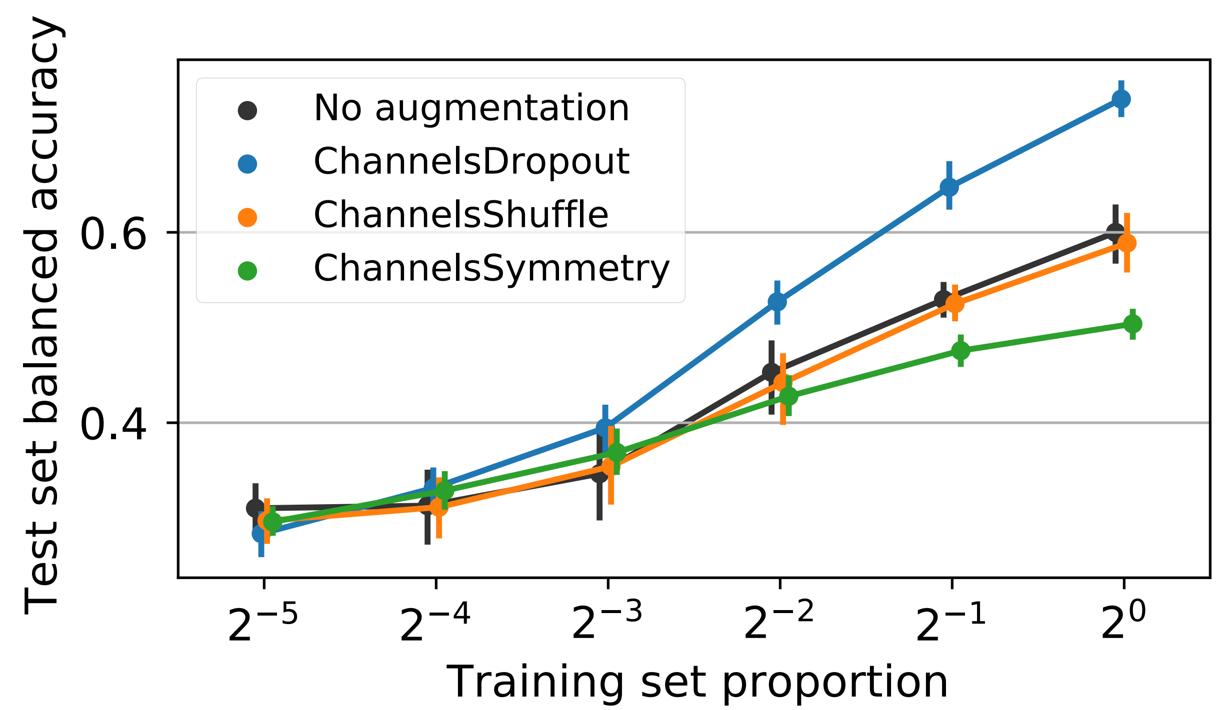
(a) *SleepPhysionet*



(b) *BCI IV 2a*



(a) *SleepPhysionet*

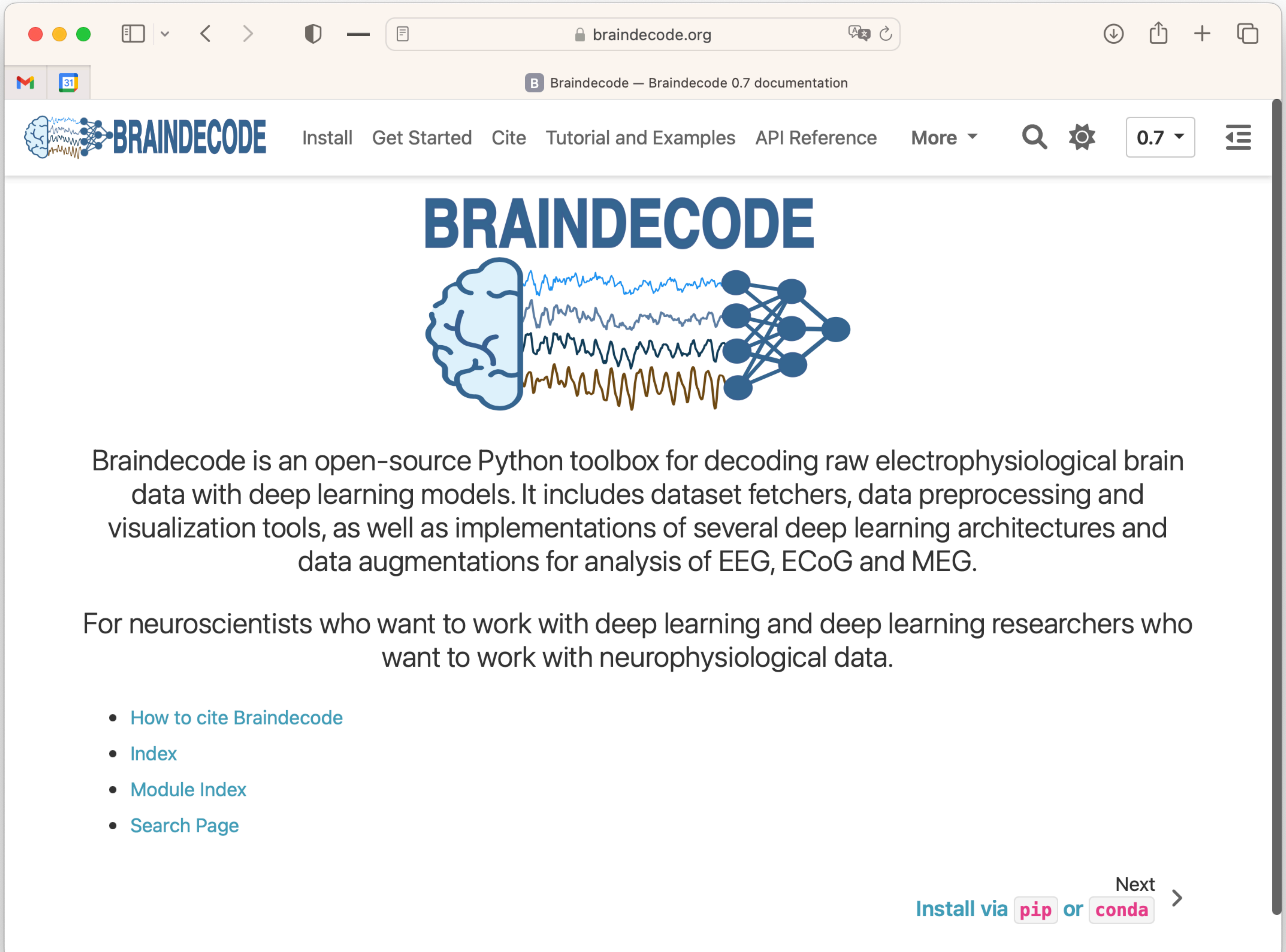


(b) *BCI IV 2a*

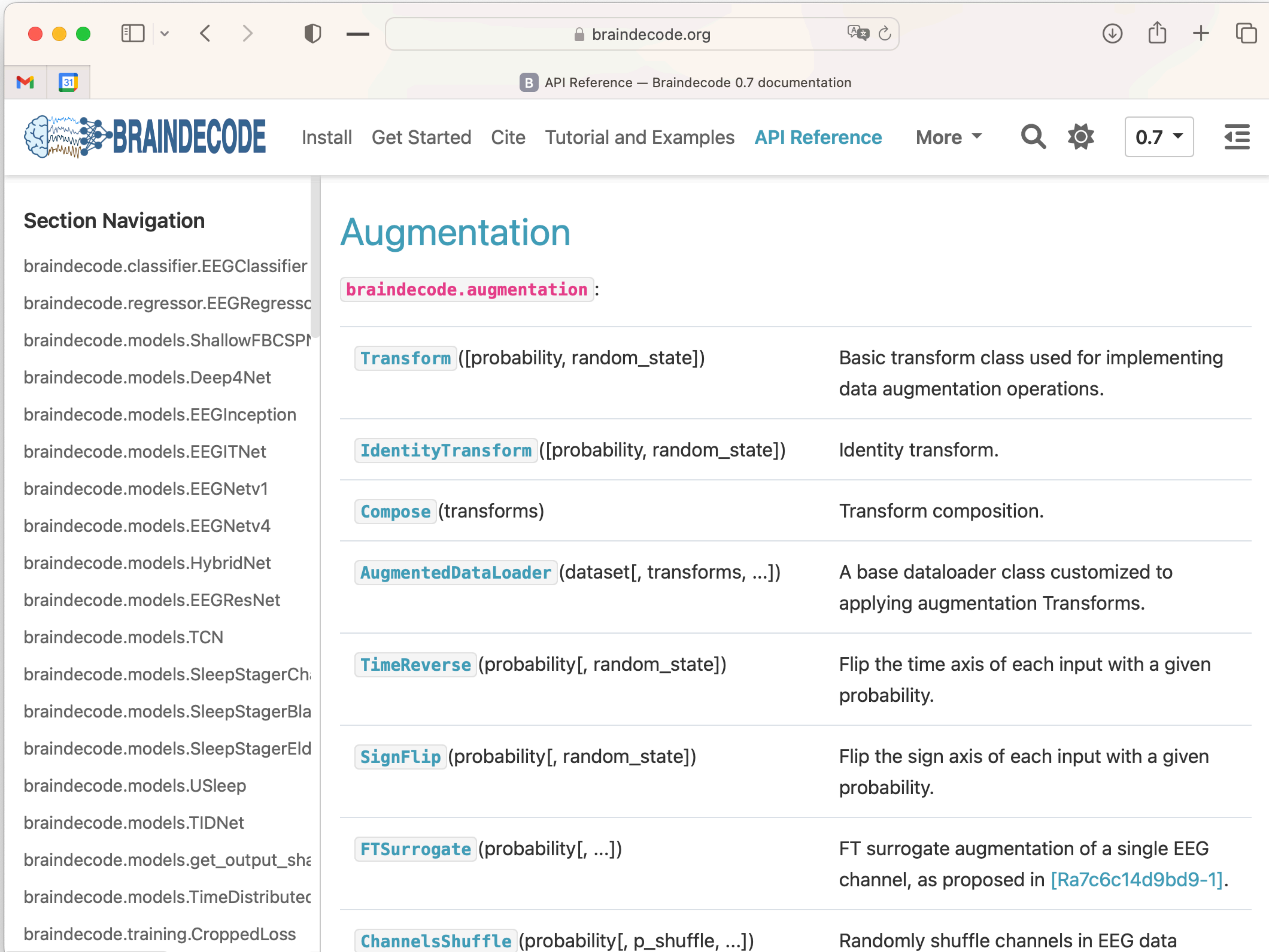
Take home messages

- Train time data augmentation for EEG signals can boost predictive performance
- Performance is the strongest in the low data regime.
- FT Surrogate and time reverse efficiency show that phase is not very informative
- Smooth time mask and channel dropout are competitive on the BCI task.

https://braindecode.org



https://braindecode.org



braindecode.org

API Reference — Braindecode 0.7 documentation

BRAINDECODE Install Get Started Cite Tutorial and Examples API Reference More 0.7

Section Navigation

- braindecode.classifier.EEGClassifier
- braindecode.regressor.EEGRegressc
- braindecode.models.ShallowFBCSPI
- braindecode.models.Deep4Net
- braindecode.models.EEGInception
- braindecode.models.EEGITNet
- braindecode.models.EEGNetv1
- braindecode.models.EEGNetv4
- braindecode.models.HybridNet
- braindecode.models.EEGResNet
- braindecode.models.TCN
- braindecode.models.SleepStagerCh
- braindecode.models.SleepStagerBla
- braindecode.models.SleepStagerEld
- braindecode.models.USleep
- braindecode.models.TIDNet
- braindecode.models.get_output_sha
- braindecode.models.TimeDistributec
- braindecode.training.CroppedLoss

Augmentation

`braindecode.augmentation`:

`Transform`([probability, random_state])	Basic transform class used for implementing data augmentation operations.
`IdentityTransform`([probability, random_state])	Identity transform.
`Compose`(transforms)	Transform composition.
`AugmentedDataLoader`(dataset[, transforms, ...])	A base dataloader class customized to applying augmentation Transforms.
`TimeReverse`(probability[, random_state])	Flip the time axis of each input with a given probability.
`SignFlip`(probability[, random_state])	Flip the sign axis of each input with a given probability.
`FTSurrogate`(probability[, ...])	FT surrogate augmentation of a single EEG channel, as proposed in [Ra7c6c14d9bd9-1].
`ChannelsShuffle`(probability[, p_shuffle, ...])	Randomly shuffle channels in EEG data

To go beyond

Learning to augment EEG data

Rommel C., Moreau T., Paillard J., Gramfort, A. (2018), "CADDA: Class-wise Automatic Differentiable Data Augmentation for EEG Signals", arXiv:2106.13695 (2021), Proc. ICLR 2022

Rommel C., Moreau T., Gramfort, A. (2018), "Deep invariant networks with differentiable augmentation layers", arXiv:2202.02142 (2022), Proc. NeurIPS 2022

Thanks !

Cedric Rommel (Post-doc Inria 2020-2022)
Thomas Moreau (PI, Inria)
Joseph Paillard (Intern 2021)

Contact

http://alexandre.gramfort.net

GitHub : @agramfort *Twitter : @agramfort*

Support

European Research Council (ERC SLAB-StG-676943)
ANR Chaire IA BrAIN

anr erc Inria
université PARIS-SACLAY