Carnet de Bord

Bachelor Thesis
Supervised learning on EEG signals for clinical diagnosis

Vocabulary

Vocabulary

- EEG signals = differences of potentials in between various areas of the brain, measured in Volts
- MEG signals = flux of magnetic fields / strength of a magnetic field at a given location in a given direction, measured in Tesla or Tesla / m
- Splitting the dataset: train (fit the parameters i.e. weights), validation (tune the hyperparameters i.e. architecture → avoid overfitting) & test sets
- k-fold cross validation: split the data into k folds / sets where each fold is used as a testing set at some point

Time series analysis - Vocabulary

- SNR = Signal to Noise Ratio → should be as high as possible, MEG has a better SNR than EEG
- Multimodal signal = multiple modalities i.e. EEG, EOG, EMG etc.
- Power Spectrum Density (PSD) / periodogram: répartition fréquentielle de la puissance d’un signal suivant les fréquences qui le composent
  - Different sleep stages have different signatures (by visualising the PSD) ???
- Spectrogram = frequency content over time

Important commands

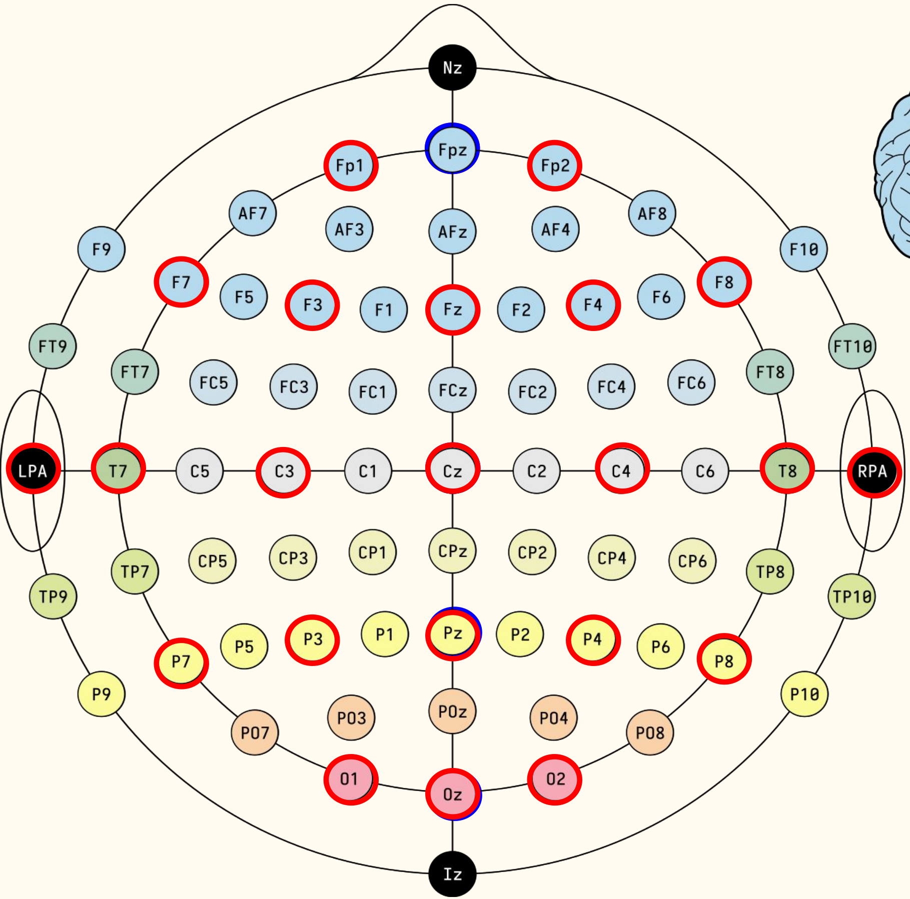
`nice -n 5 python myscript.py` to share processing resources

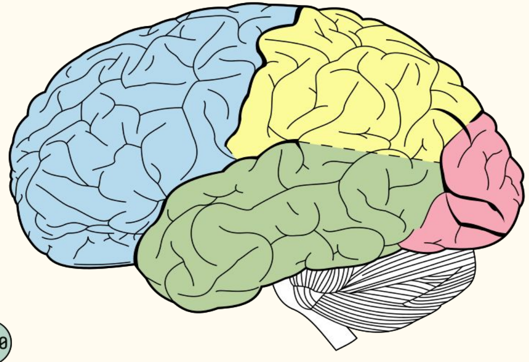
To install Braindecode: `pip install -U https://github.com/braindecode/braindecode/archive/master.zip`

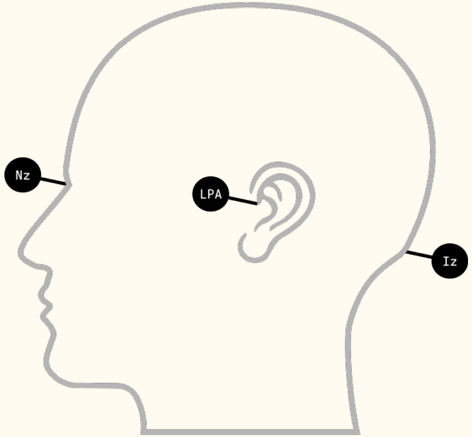
```
eval "$(/home/parietal/msolal/miniconda3/bin/conda shell.bash hook)"
```

Datasets

Combined

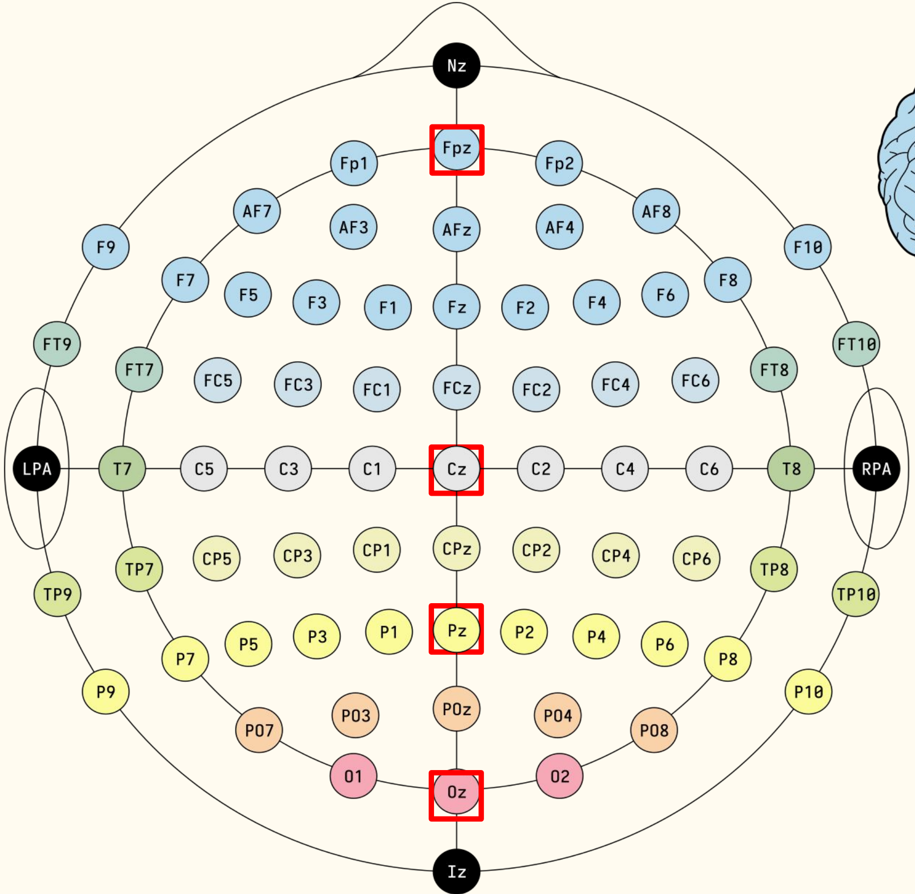


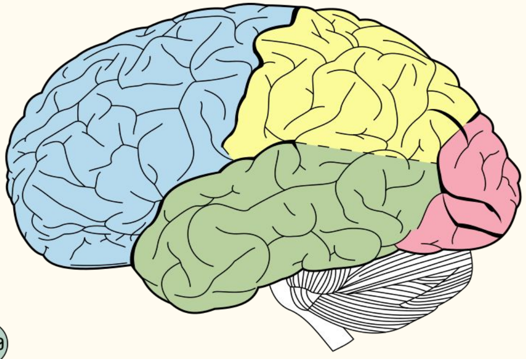


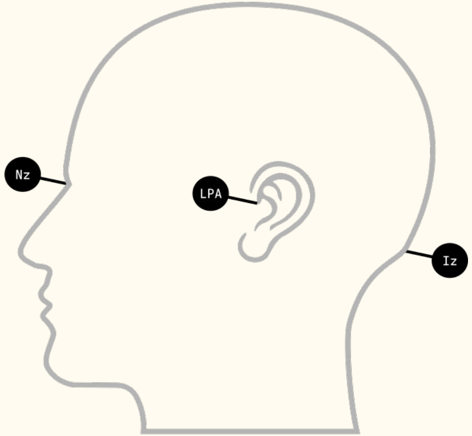


Blue: SleepPhysionet
Red: MASS
Green: Clinical

SleepPhysionet



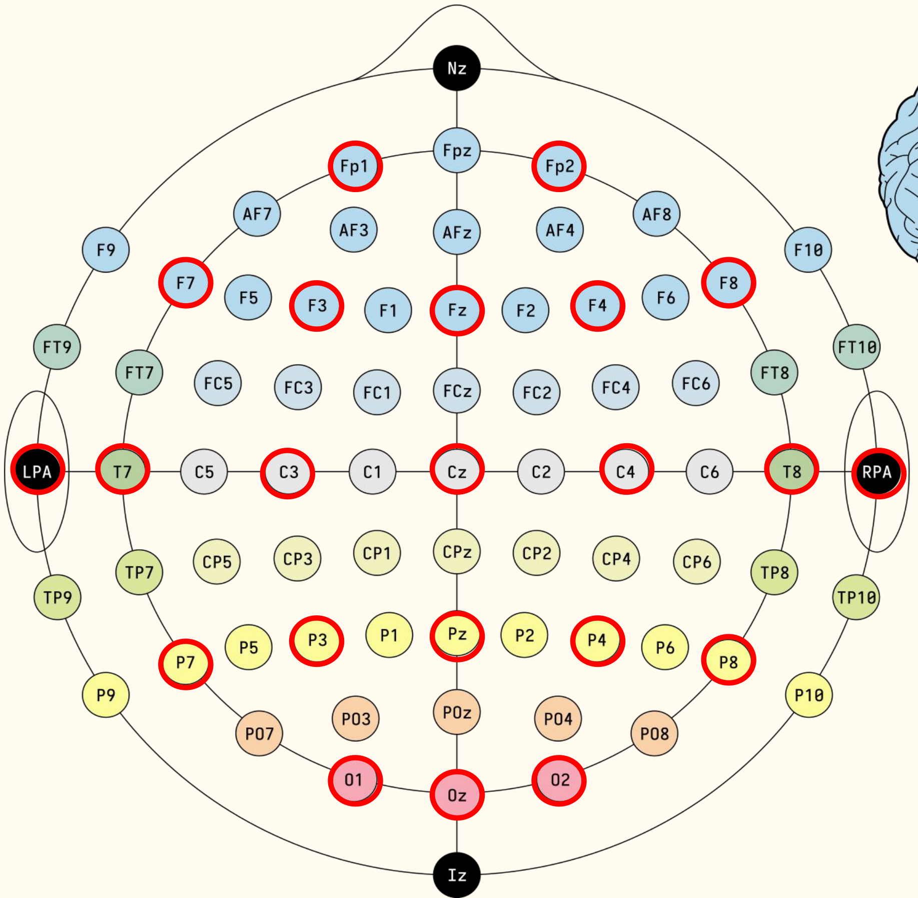


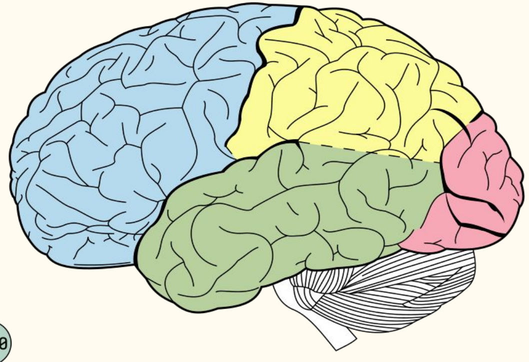


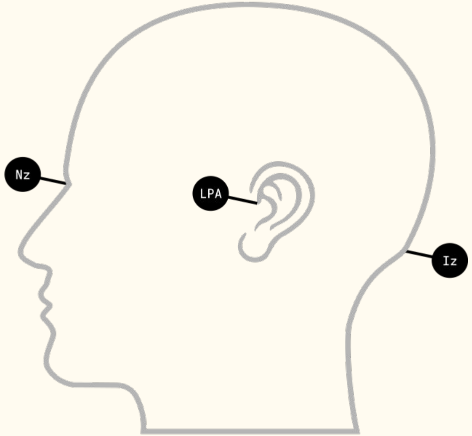
- 1 EOG channel: horizontal
- 1 EMG channel: submental

https://physionet.org/content/sleep-edfx/1.0.0/

MASS







- 2 EOG channels: left and right
- 3 bipolar EMG channels: Chin1, Chin2, Chin3

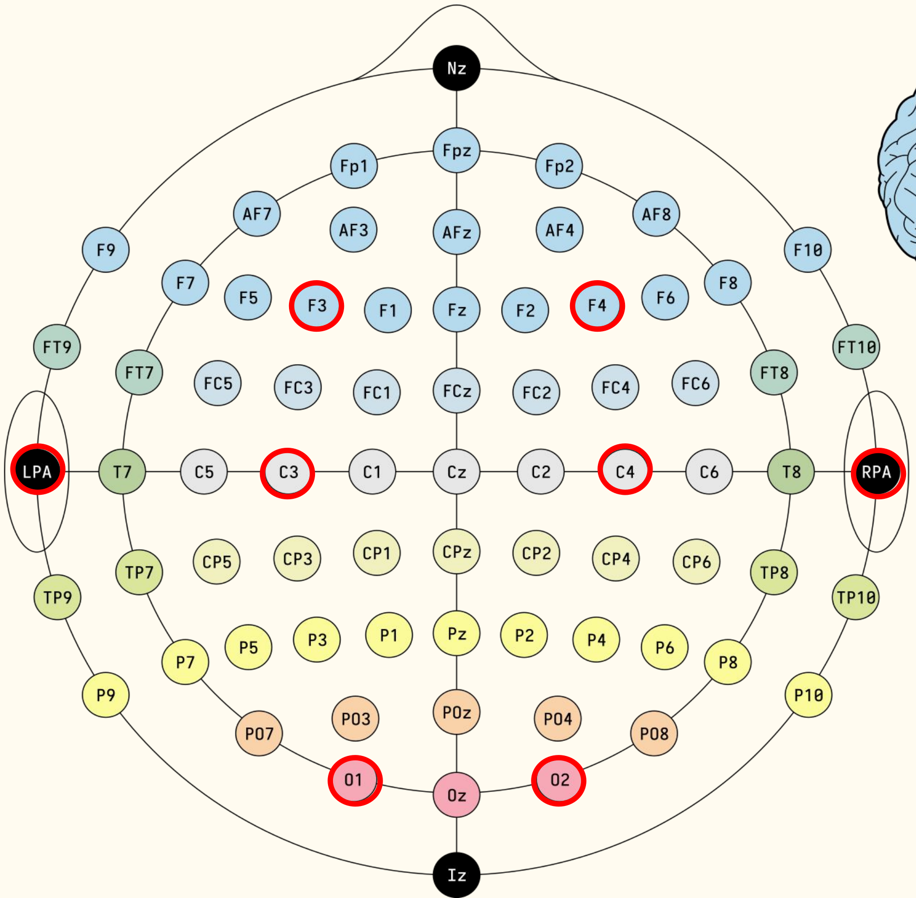
https://www.readcube.com/articles/10.1111/jsr.12169

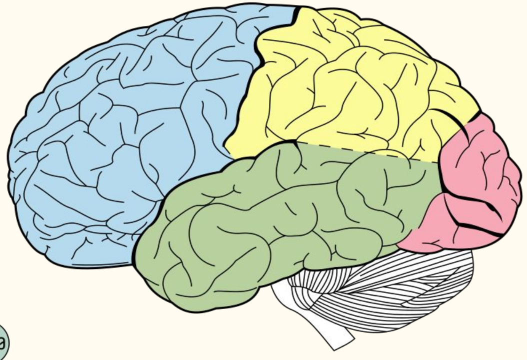
- `CC-SS-XXXX`: `CC` is the cohort number, `SS` is a subset number and `XXXX` is a record number
- 62 subjects including 28 male and 34 female
- Age: mean, standard deviation, min, max:

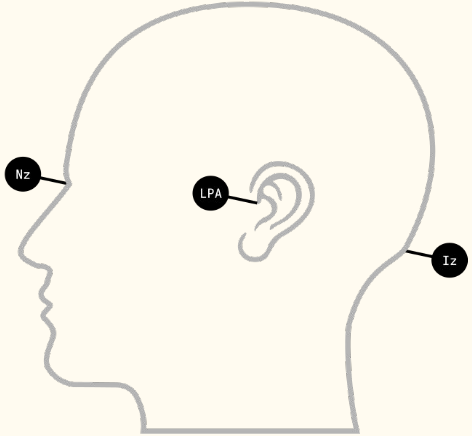
C1/SS3	Total	42.5	18.9	20	69
	Men	40.4	19.4	20	69
	Women	44.2	18.6	20	69

- Channels for records `01-03-0001-39` (except `0036`) + `01-03-0041-44` (A):
  - `['ECG I', 'EEG A2-CLE', 'EEG C3-CLE', 'EEG C4-CLE', 'EEG Cz-CLE', 'EEG F3-CLE', 'EEG F4-CLE', 'EEG F7-CLE', 'EEG F8-CLE', 'EEG Fp1-CLE', 'EEG Fp2-CLE', 'EEG Fz-CLE', 'EEG O1-CLE', 'EEG O2-CLE', 'EEG Oz-CLE', 'EEG P3-CLE', 'EEG P4-CLE', 'EEG Pz-CLE', 'EEG T3-CLE', 'EEG T4-CLE', 'EEG T5-CLE', 'EEG T6-CLE', 'EMG Chin1', 'EMG Chin2', 'EMG Chin3', 'EOG Left Horiz', 'EOG Right Horiz']`
- Channels for records `01-03-0040` + `01-03-0047-64` (B):
  - `['ECG I', 'EEG C3-LER', 'EEG C4-LER', 'EEG Cz-LER', 'EEG F3-LER', 'EEG F4-LER', 'EEG F7-LER', 'EEG F8-LER', 'EEG Fp1-LER', 'EEG Fp2-LER', 'EEG Fz-LER', 'EEG O1-LER', 'EEG O2-LER', 'EEG Oz-LER', 'EEG P3-LER', 'EEG P4-LER', 'EEG Pz-LER', 'EEG T3-LER', 'EEG T4-LER', 'EEG T5-LER', 'EEG T6-LER', 'EMG Chin1', 'EMG Chin2', 'EMG Chin3', 'EOG Left Horiz','EOG Right Horiz', 'Resp Belt Abdo', 'Resp Belt Thor']`
- LER = linked-ear reference with a 10k\Omega resistance; CLE = computed linked-ear

Clinical


Nz
Fpz
Fp1
Fp2
AF7
AF3
AFz
AF4
AF8
F9
F7
F5
F3
F1
Fz
F2
F4
F6
F8
F10
FT9
FT7
FC5
FC3
FC1
FCz
FC2
FC4
FC6
FT8
FT10
LPA
T7
C5
C3
C1
Cz
C2
C4
C6
T8
RPA
TP7
CP5
CP3
CP1
CPz
CP2
CP4
CP6
TP8
TP9
TP10
P7
P5
P3
P1
Pz
P2
P4
P6
P8
P9
P10
PO7
PO3
POz
PO4
PO8
O1
Oz
O2
Iz




Nz
LPA
Iz

- MASS dataset:
  - T5, T6 became P7, P8
  - T3, T4 became T7, T8
- A1, A2 are also called M1, M2 - near the ears

Weekly logbook

Week 1.

11th - 15th January

- Cours AI4Health
  - Regarder les vidéos du cours
  - Faire les notebooks
- Articles
  - [1] A Deep Learning Architecture for temporal sleep stage classification using multivariate and multimodal time series, Chambon et al.
  - [2] A Deep Learning Architecture to detect events in EEG signals during sleep, Chambon et al.
- Get familiar with Google Colab
- DATAIA Seminars.
  - Can current neural networks provide a satisfactory model of the human brain, S. Dehaene

Sleep Stage Classification [1]

- 5 sleep stages: Wake, REM (Rapid Eye Movement), N1, N2, N3 (Non Rapid Eye Movement)
  - have distinct time and frequency patterns and differ in proportions over a night
- Approaches.
  - App1. Extract features using expert knowledge about the signal
  - App2. Learn feature representation from the raw signal directly
- Challenge. The data is imbalanced (more data for certain classes)
  - Sol1. Reweight the loss function so that the cost of making an error on a rare sample is larger
  - Sol2. BALANCED SAMPLING, feed the network with batches of data which contain as many data points from each class (each batch has 20% of samples of each class since there are 5 classes) → requires using balanced accuracy metrics

Event Detection [2]

- SleepStagerChambon2018 cf notebook 5 from AI4Health

Typical pre-processing techniques

- Get rid of the artifacts (eye blinks, high frequency / environmental noise, heart beats, drift i.e. sensors moving)
- Low-pass filtering: 30Hz cutoff frequency
- Downsampling to a sampling rate around 100Hz / 128Hz
  - Speed-up computations
  - I’m not sure I understand what this means
- Normalizing each 30s sample (zero mean and unit variance)
- Sometimes convert from volt to microvolt

Week 2.

18th - 22nd January

- Articles
  - [3] Uncovering the structure of clinical EEG signals with self-supervised learning, Banville et al. + notebook
  - BT Agathe
- HSS/BIO 361 Sciences cognitives
- Get familiar with PyTorch https://pytorch.org/tutorials/beginner/deep_learning_60min_blitz.html
- Sleep Staging w. MASS dataset

Self Supervised Learning

- Self Supervised Learning (SSL)
    - Unsupervised learning approach that learns representations from unlabeled data, exploiting the structure of the data to provide supervision
    - Reframe an unsupervised problem as a supervised one
    - Comprises a PRETEXT and a DOWNSTREAM task
        - Downstream task: what we are interested in but limited / no annotations
        - Pretext task: must be sufficiently related to the ownstream task such that similar representations should be employed to carry it out + must be possible to generate annotations for this task using the data alone

SSL to study the structure of clinical EEG signals [3]

- Relative Positioning (RP)
  - how to discriminate pairs of time windows based on their relative position
- Temporal Shuffling (TS)
- Contrastive Predictive Coding (CDC)
  - moins bien compris
- Downstream tasks: sleep staging & pathology detection
- Datasets: PC18 (Physionet Challenge) + TUHab (TUH abnormal)

Week 3.

25th - 29th January

- Articles
  - Montreal Archive of Sleep Studies: an open-access resource for instrument benchmarking and exploratory research, O'Reilly et al.
  - MEG/EEG group study with MNE: recommendations, quality assessments and best practices, Jas et al.
- Sleep staging w. MASS dataset
- Created GitHub repository
- Sleep staging cleanup //Braindecode
- Olivier's data

Montreal Archive of Sleep Studies, Session 3 (MASS SS3)

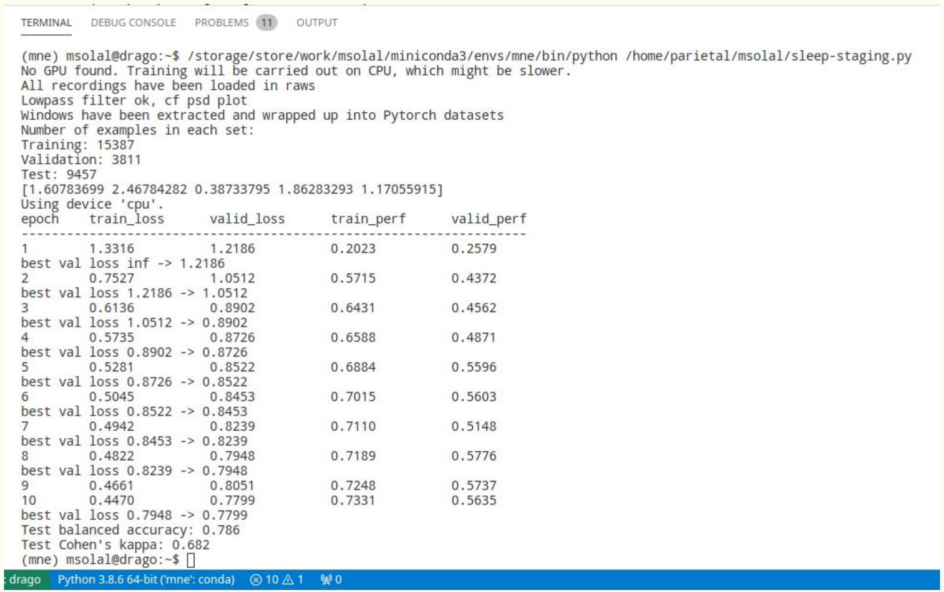
- `CC-SS-XXXX`: `CC` is the cohort number, `SS` is a subset number and `XXXX` is a record number
- 62 subjects including 28 male and 34 female
- Age: mean, standard deviation, min, max:

C1/SS3	Total	42.5	18.9	20	69
	Men	40.4	19.4	20	69
	Women	44.2	18.6	20	69

- Channels for records `01-03-0001-39` (except `0036`) + `01-03-0041-44` (A):
  - `['ECG I', 'EEG A2-CLE', 'EEG C3-CLE', 'EEG C4-CLE', 'EEG Cz-CLE', 'EEG F3-CLE', 'EEG F4-CLE', 'EEG F7-CLE', 'EEG F8-CLE', 'EEG Fp1-CLE', 'EEG Fp2-CLE', 'EEG Fz-CLE', 'EEG O1-CLE', 'EEG O2-CLE', 'EEG Oz-CLE', 'EEG P3-CLE', 'EEG P4-CLE', 'EEG Pz-CLE', 'EEG T3-CLE', 'EEG T4-CLE', 'EEG T5-CLE', 'EEG T6-CLE', 'EMG Chin1', 'EMG Chin2', 'EMG Chin3', 'EOG Left Horiz', 'EOG Right Horiz']`
- Channels for records `01-03-0040` + `01-03-0047-64` (B):
  - `['ECG I', 'EEG C3-LER', 'EEG C4-LER', 'EEG Cz-LER', 'EEG F3-LER', 'EEG F4-LER', 'EEG F7-LER', 'EEG F8-LER', 'EEG Fp1-LER', 'EEG Fp2-LER', 'EEG Fz-LER', 'EEG O1-LER', 'EEG O2-LER', 'EEG Oz-LER', 'EEG P3-LER', 'EEG P4-LER', 'EEG Pz-LER', 'EEG T3-LER', 'EEG T4-LER', 'EEG T5-LER', 'EEG T6-LER', 'EMG Chin1', 'EMG Chin2', 'EMG Chin3', 'EOG Left Horiz','EOG Right Horiz', 'Resp Belt Abdo', 'Resp Belt Thor']`
- LER = linked-ear reference with a 10k\Omega resistance; CLE = computed linked-ear
- Record `01-03-0036`: `No matching events found for Sleep stage 3/4 (event id 4)`
- Why are we using SS3?

Sleep Staging with the MASS dataset

- MASS SS3 - Montreal Archive of Sleep Studies, session 3
- Had to wait a few days to get the SSH access
- Took code from AI4Health tuto 6 and adapted it to MASS dataset
- Got nice results with 30 subjects (incl. 10 testing)
- use only records A + ignore subject `01-03-0036`
- Split the code:
  - main, datasets, models, training, visualisation



```
TERMINAL    DEBUG CONSOLE    PROBLEMS 11    OUTPUT

(mne) msolal@drago:~$ /storage/store/work/msolal/miniconda3/envs/mne/bin/python /home/parietal/msolal/sleep-staging.py
No GPU found. Training will be carried out on CPU, which might be slower.
All recordings have been loaded in raws
Lowpass filter ok, cf psd plot
Windows have been extracted and wrapped up into Pytorch datasets
Number of examples in each set:
Training: 15387
Validation: 3811
Test: 9457
[1.60783699 2.46784282 0.38733795 1.86283293 1.17055915]
Using device 'cpu'.
epoch    train_loss      valid_loss      train_perf      valid_perf
-----------------------------------------------------------------------
1        1.3316          1.2186          0.2023          0.2579
best val loss inf -> 1.2186
2        0.7527          1.0512          0.5715          0.4372
best val loss 1.2186 -> 1.0512
3        0.6136          0.8902          0.6431          0.4562
best val loss 1.0512 -> 0.8902
4        0.5735          0.8726          0.6588          0.4871
best val loss 0.8902 -> 0.8726
5        0.5281          0.8522          0.6884          0.5596
best val loss 0.8726 -> 0.8522
6        0.5045          0.8453          0.7015          0.5603
best val loss 0.8522 -> 0.8453
7        0.4942          0.8239          0.7110          0.5148
best val loss 0.8453 -> 0.8239
8        0.4822          0.7948          0.7189          0.5776
best val loss 0.8239 -> 0.7948
9        0.4661          0.8051          0.7248          0.5737
10       0.4470          0.7799          0.7331          0.5635
best val loss 0.7948 -> 0.7799
Test balanced accuracy: 0.786
Test Cohen's kappa: 0.682
(mne) msolal@drago:~$
drago   Python 3.8.6 64-bit ('mne': conda)   ⊗ 10 ⚠ 1   0
```

30 subjects, incl 10 test

Olivier's data

- 466 records with annotations
- Check missing CSV / EDF files
- O2*: ABlo590819 ['EEG F3', 'EEG F4', 'EEG C3', 'EEG C4', 'EEG O1', 'EEG M1', 'EEG M2', 'EMG Ment', 'ECG', 'EMG Tib-L', 'E MG Tib-R', 'Resp Therm', 'Resp Thor', 'Resp Abd', 'SaO2 SaO2', 'SaO2 HR', 'SaO2 Pulse', 'BodyPos BodyPos', 'EEG O2*', 'EOG eog-r', 'EOG eog-l', 'Sound Ronf', 'Resp Flux', 'Resp Flw2', 'BodyPos Pos']
- M1#2, M2#2: AFan740527 ['EEG F3', 'EEG F4', 'EEG C3', 'EEG C4', 'EEG O1', 'EEG O2', 'EOG E2', 'EOG E1', 'EMG Chin1', 'EMG Tib-L' , 'EMG Tib-R', 'Resp Cann Raw', 'Resp Thor', 'Resp Abd', 'SaO2 SaO2', 'EEG M1#2', 'EEG M2#2', 'Sound Ronf', 'Resp Ther', 'ECG', 'Unspec Flux', 'Unspec PULS', 'BodyPos Pos', 'Unspec BP RAW', 'Unspec BP LVL', 'Unspec BEAT']
- no O2*, no M1#2, no M2#2: BOfr660513 ['EEG F3', 'EEG F4', 'EEG C3', 'EEG C4', 'EEG O1', 'EEG O2', 'EEG M1', 'EEG M2', 'EMG Chin1', 'ECG', 'EMG Tib-L', 'EMG Tib-R', 'Resp Therm', 'Resp Thor', 'Resp Abd', 'SaO2 SaO2', 'SaO2 HR', 'SaO2 Pulse', 'Sound Mic', 'BodyPos BodyPos', 'EOG E1#1', 'EOG E2#1', 'Resp Flux', 'Resp Flw2', 'BodyPos Pos']
-

Week 4.

1st - 5th February

- Code organisation: Braindecode style
- Experiments: MASS & Sleep Physionet Dataset
- Olivier's data: formatting of annotations + conversion to BIDS
- MASS dataset conversion to BIDS

Olivier’s data

- Annotations files: keep all validated annotations + add events missing in between validated events
- Combine raw and annotations files into BIDS format
  - We set `line_freq` to 50Hz
- Have a look at the different channels, rename them correctly and give them the right types
- Created ClinicalData class to deal with these files

Data formatting

- Convert all datasets to BIDS format?
- Converting MASS to the BIDS format
  - Get the correct channels names
  - Get the correct channels types
  - Get rid of EEG prefix in EEG channels /!\ also changes the unit
  - How to deal with artefacts annotations?

Week 5.

8st - 13th February

- BIDS conversion: MASS & clinical datasets
- Clinical annotations

BIDS Conversion - MASS & Clinical

- Rename channels without changing the units
- Respiratory channels should be of type ‘resp’ and not ‘misc’
- Reference channel isn’t always the same

Experiments

- Sleep scoring on MASS dataset
  - ERROR: 20 and 21 channels → should all have 20 channels
- Sleep scoring on clinical dataset
  - overlapping events
  - problems with annotations in Olivier's data → have a look into that

Clinical annotations

- Select validated events
- get_prev and get_next to get rid of NO events overlapping with YES events
- Merge identical events, i.e. events with same description
- Compute offset + check if offset and onset are ok

Week 6.

15th - 19th February

- Clinical annotations
- Experiments

Clinical annotations

- Finished writing code to get correct annotations
- Talk w. Olivier about overlapping NO events
- Still need to figure out what to do when idx is the last one of the data
  - `MEga720912`
  - `MOas940705`
  - `WAre750130`
  - `OZch790212`
  - `PRma490730`
- `COma390828` mouline dans le vide, what is happening?
- Should change events.tsv files, no need to restart entire BIDS conversion?

Experiments

- MASS-60-usual_split
- SleepPhysionet-60-usual_split
- What about crop_wake_mins?
  - adjusted tmax and tmin

Week 7.

22th - 27th February

- Clinical annotations
- Experiments
- Shuffle windows dataset before split
- 10 epochs
- Downsample to 100Hz to be able to compare SleepPhysionet & MASS
- Preprocessing? Comment gagner du temps? Save pre-processed dataset

Clinical dataset

- Removed recordings `['HEna690502annot.csv', 'DEmi471006annot.csv', 'PAch711105annot.csv', 'TRhe570716annot.csv', 'MOsy850124annot.csv', 'GRre720825annot.csv', 'CAad580107annot.csv', 'DAci800823annot.csv', 'LOel520325annot.csv', 'BApa990220annot.csv', 'MOre780128annot.csv']` because they had no validated data
- Removed recordings with `False` onset-offset

Lightning Talk

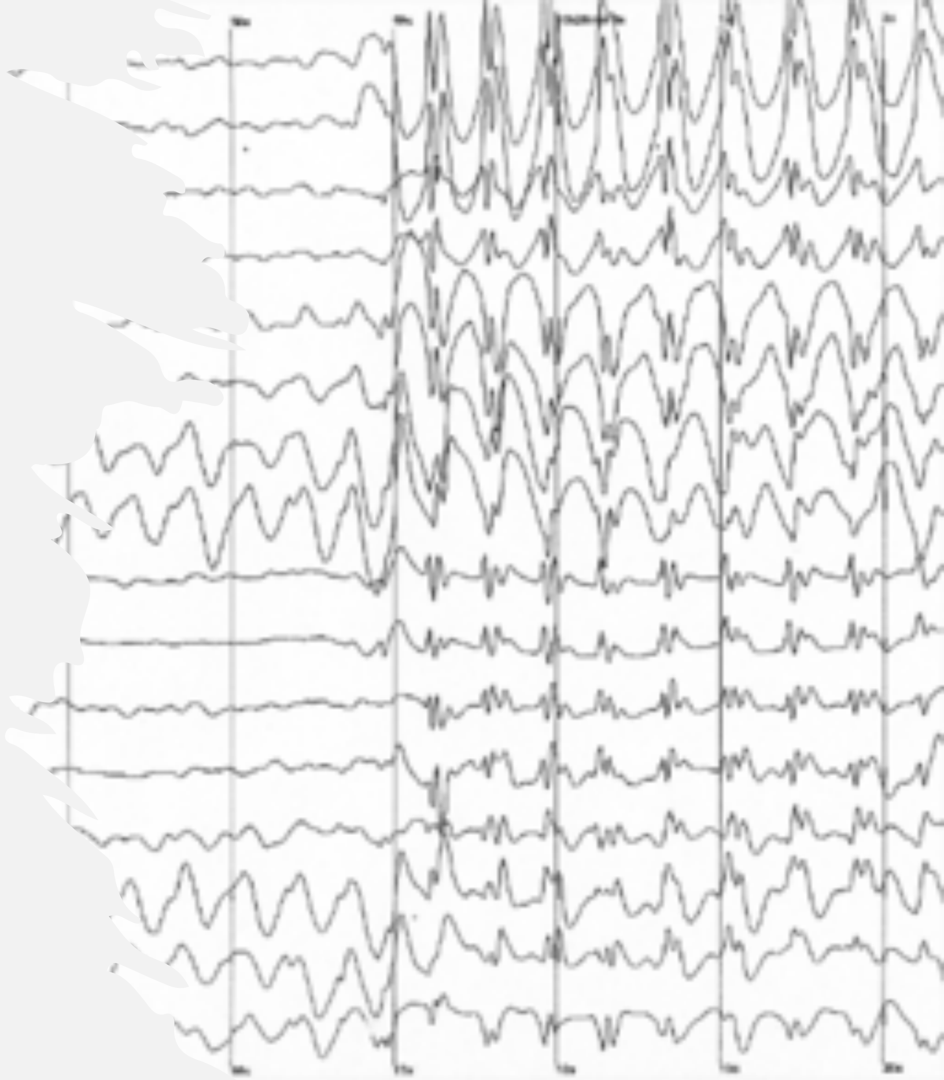
Lightning Talk

My Bachelor Thesis in 3 minutes

Sleep Stage Scoring

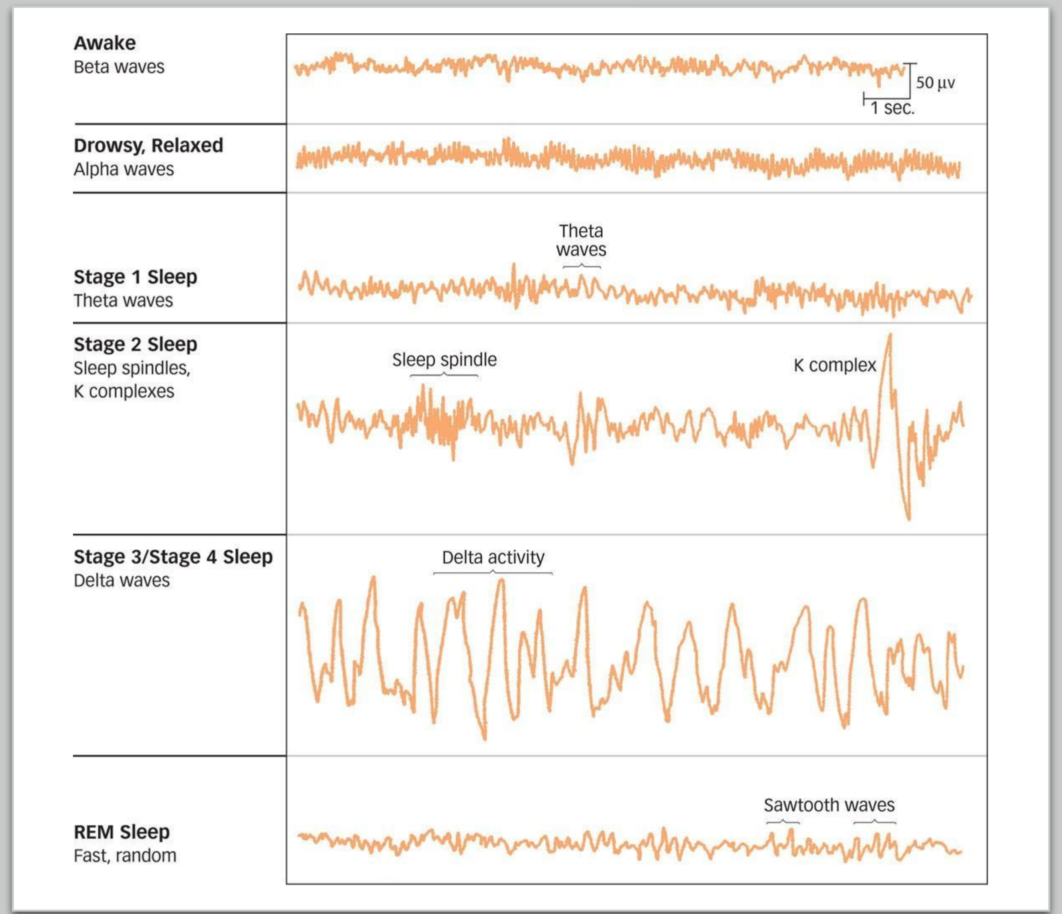
Maëlys Solal, BX2021

Supervised by Alexandre Gramfort and Dr Olivier Pallanca



Sleep

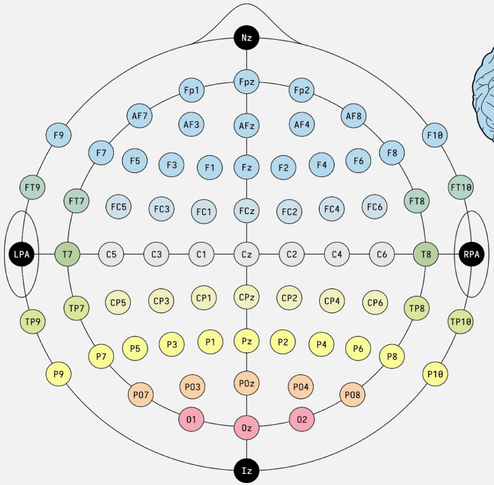
- **Sleep stages and sleep cycle**
  - Wake (W)
  - Rapid Eye Movement (REM)
  - Non REM1 (N1)
  - Non REM2 (N2)
  - Non REM3 (N3)
- **Polysomnography**
  - Includes EEG (brain's electrical activity), EOG (eyes), EMG (muscles), ECG (heart)
  - Used for clinical diagnosis (insomnia, sleep apnea etc.)
- **Sleep stage scoring**
  - According to a precise set of rules
  - Done by hand by sleep scorers
  - Quite tedious process

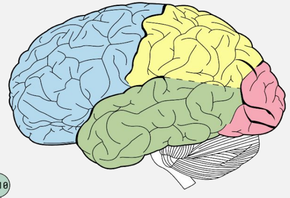


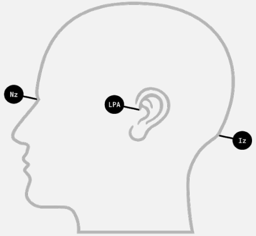
Motivations and main objectives

Data

- MASS (Montreal Archive of Sleep Study)
- SleepPhysionet
- Clinical dataset from Dr Olivier Pallanca







Model

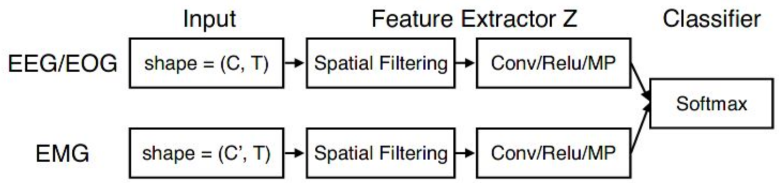


Fig. 1. Network general architecture: the network processes C EEG/EOG channels and C' EMG channels through separate pipelines. For each modality, it performs spatial filtering and applies convolutions, non linear operations and max pooling (MP) over the time axis. The outputs of the different pipelines are finally concatenated to feed a softmax classifier.

Ref. *Chambon, S., Galtier, M., Arnal, P., Wainrib, G. and Gramfort, A. (2018) A Deep Learning Architecture for Temporal Sleep Stage Classification Using Multivariate and Multimodal Time Series. IEEE Trans. on Neural Systems and Rehabilitation Engineering 26: (758-769)*

Results: Balanced Accuracy depending on training and testing dataset

	SleepPhysionet	MASS
SleepPhysionet	0.680	0.545
MASS	0.650	0.734

- EEG channels: FpzCz, PzOz

- Learning rate: 0.0005
- Batch size: 8
- Number of epochs: 10

Converting PSG data to BIDS

- Simple way to organise neuroimaging and behavioral data

mass-bids	08/03/2021 16:44	folder	
SS3	26/02/2021 10:25	folder	
derivatives	08/03/2021 23:15	folder	
sub-030001	10/02/2021 15:01	folder	
eeg	26/02/2021 11:40	folder	
sub-030001_channels.tsv	10/02/2021 16:32	TSV document	1,6 KiB
sub-030001_eeg.edf	10/02/2021 16:32	unknown	396,1 MiB
sub-030001_eeg.json	10/02/2021 16:32	JSON document	513 B
sub-030001_events.tsv	10/02/2021 16:32	TSV document	295,5 KiB
sub-030001_scans.tsv	10/02/2021 16:32	TSV document	71 B
sub-030002	10/02/2021 15:33	folder	

BIDS
BRAIN IMAGING DATA STRUCTURE

- Meticulous work : channels names, types, units, annotations, EEG reference
- Simplifies how we load the data
  - `write_raw_bids`
  - `read_raw_bids`
- Simplifies preprocessing + saving preprocessed data

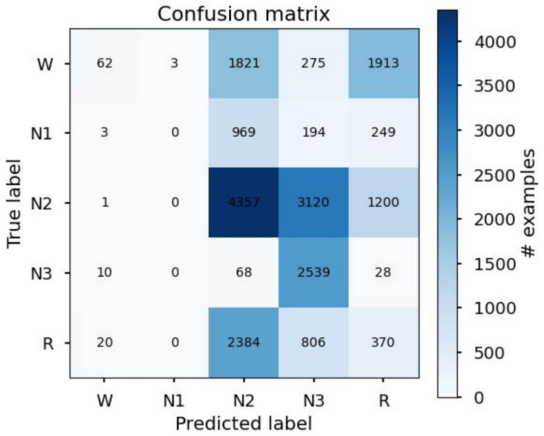
MNE-BIDS

MNE
MEG + EEG ANALYSIS & VISUALIZATION

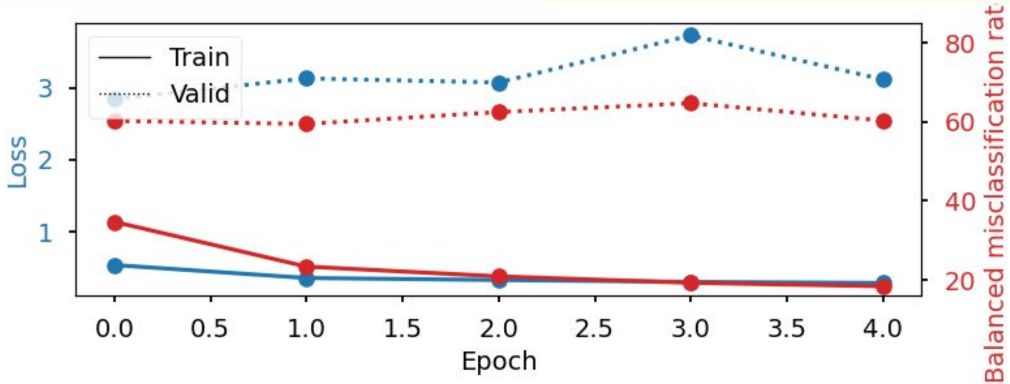
Gorgolewski, K.J., Auer, T., Calhoun, V.D., et al. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Scientific Data, 3 (160044). doi:10.1038/sdata.2016.44

Results

MASS-all-usual_split (62 subjects, 0.6_0,2_0,2)




```
Training: <Epochs |  699 events (all good), 0 - 29.9961 sec, baseline off, ~819.2 MB, data loaded, with metadata,
 'Sleep stage 1': 58
 'Sleep stage 2': 349
 'Sleep stage 3': 110
 'Sleep stage R': 143
 'Sleep stage W': 39>
Validation: <Epochs |  535 events (all good), 0 - 29.9961 sec, baseline off, ~627.0 MB, data loaded, with metadata,
 'Sleep stage 1': 57
 'Sleep stage 2': 265
 'Sleep stage 3': 58
 'Sleep stage R': 141
 'Sleep stage W': 14>
Test: <Epochs |  1754 events (all good), 0 - 14.998 sec, baseline off, ~2.01 GB, data loaded, with metadata,
 'Sleep stage 1': 75
 'Sleep stage 2': 749
 'Sleep stage 3': 338
 'Sleep stage R': 322
 'Sleep stage W': 270>
```

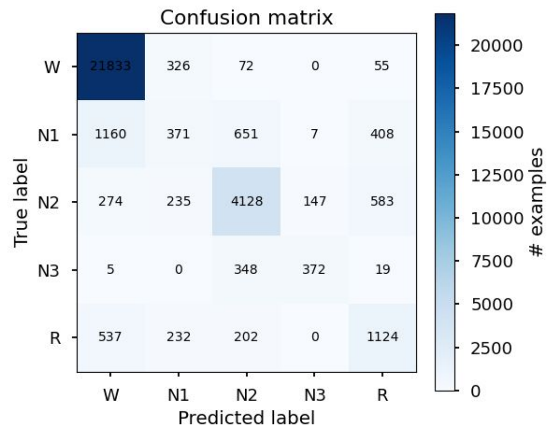


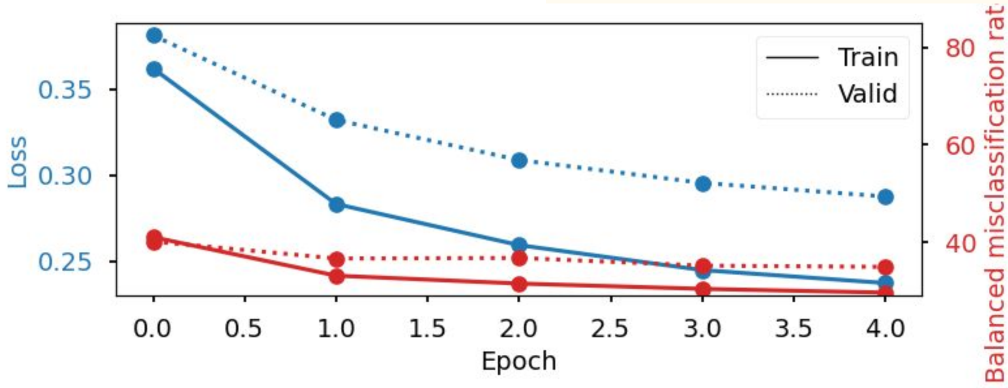
epoch	train_bal_acc	train_loss	valid_bal_acc	valid_loss	dur
1	0.6532	0.5442	0.3978	2.8428	2600.2484
2	0.7665	0.3652	0.4053	3.1271	2221.6011
3	0.7911	0.3369	0.3753	3.0704	2227.2182
4	0.8075	0.3123	0.3528	3.7219	2612.5081
5	0.8161	0.2959	0.3969	3.1127	2199.2252

Test balanced accuracy: 0.316
Test Cohen's kappa: 0.113

	precision	recall	f1-score	support
0	0.65	0.02	0.03	4074
1	0.00	0.00	0.00	1415
2	0.45	0.50	0.48	8678
3	0.37	0.96	0.53	2645
4	0.10	0.10	0.10	3580
accuracy			0.36	20392
macro avg	0.31	0.32	0.23	20392
weighted avg	0.39	0.36	0.30	20392

SleepPhysionet-60-usual_split (60 subjects, 0.6_0,2_0,2)



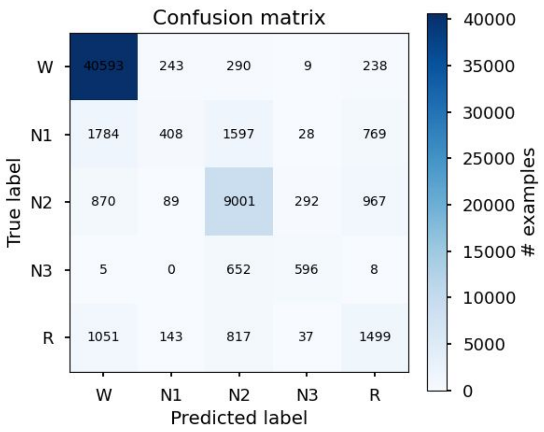


epoch	train_bal_acc	train_loss	valid_bal_acc	valid_loss	dur
1	0.5893	0.3621	0.5984	0.3812	708.2698
2	0.6684	0.2832	0.6332	0.3321	762.9625
3	0.6846	0.2592	0.6321	0.3087	750.1039
4	0.6955	0.2447	0.6478	0.2955	749.9859
5	0.7032	0.2372	0.6504	0.2875	730.4074

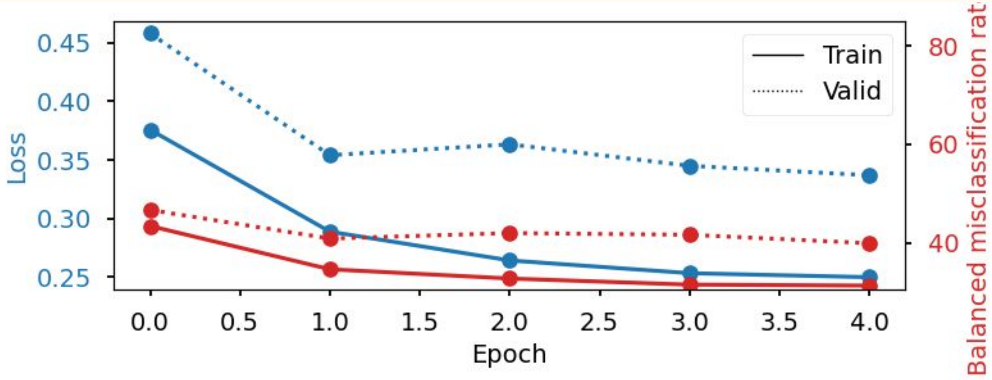
Test balanced accuracy: 0.586
Test Cohen's kappa: 0.670

	precision	recall	f1-score	support
0	0.92	0.98	0.95	22286
1	0.32	0.14	0.20	2597
2	0.76	0.77	0.77	5367
3	0.71	0.50	0.59	744
4	0.51	0.54	0.52	2095
accuracy			0.84	33089
macro avg	0.64	0.59	0.60	33089
weighted avg	0.82	0.84	0.82	33089

SleepPhysionet-all-05_02_03 (76 subjects, 0.5_0,2_0,3)




```
Training: <Epochs |  2650 events (all good), 0 - 29.99 sec, baseline off, ~121.3 MB, data loaded, with metadata,
 'Sleep stage 1': 58
 'Sleep stage 2': 250
 'Sleep stage 3': 220
 'Sleep stage 4': 220
 'Sleep stage R': 125
 'Sleep stage W': 1997>
Validation: <Epochs |  2630 events (all good), 0 - 29.99 sec, baseline off, ~120.4 MB, data loaded, with metadata,
 'Sleep stage 1': 163
 'Sleep stage 2': 377
 'Sleep stage 3': 4
 'Sleep stage R': 110
 'Sleep stage W': 1976>
Test: <Epochs |  2702 events (all good), 0 - 29.99 sec, baseline off, ~123.7 MB, data loaded, with metadata,
 'Sleep stage 1': 521
 'Sleep stage 2': 272
 'Sleep stage 3': 50
 'Sleep stage 4': 50
 'Sleep stage R': 113
 'Sleep stage W': 1746>
```

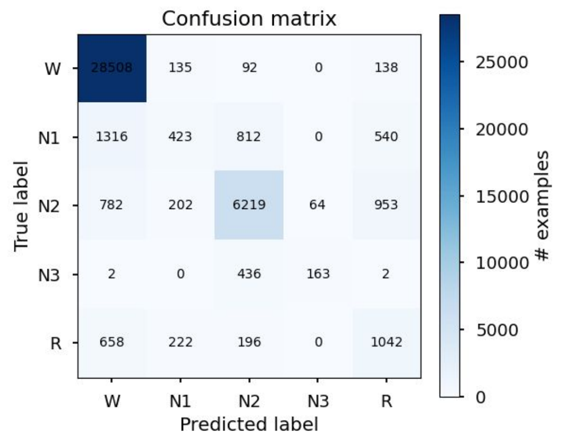


epoch	train_bal_acc	train_loss	valid_bal_acc	valid_loss	dur
1	0.5658	0.3758	0.5333	0.4581	524.8289
2	0.6538	0.2886	0.5913	0.3540	505.7235
3	0.6726	0.2640	0.5802	0.3633	457.0923
4	0.6850	0.2531	0.5837	0.3449	467.0661
5	0.6870	0.2497	0.6009	0.3369	461.1684

Test balanced accuracy: 0.554
Test Cohen's kappa: 0.669

	precision	recall	f1-score	support
0	0.92	0.98	0.95	41373
1	0.46	0.09	0.15	4586
2	0.73	0.80	0.76	11219
3	0.62	0.47	0.54	1261
4	0.43	0.42	0.43	3547
accuracy			0.84	61986
macro avg	0.63	0.55	0.56	61986
weighted avg	0.81	0.84	0.82	61986

SleepPhysionet-all-usual_split (76 subjects, 0.6_0,2_0,2)

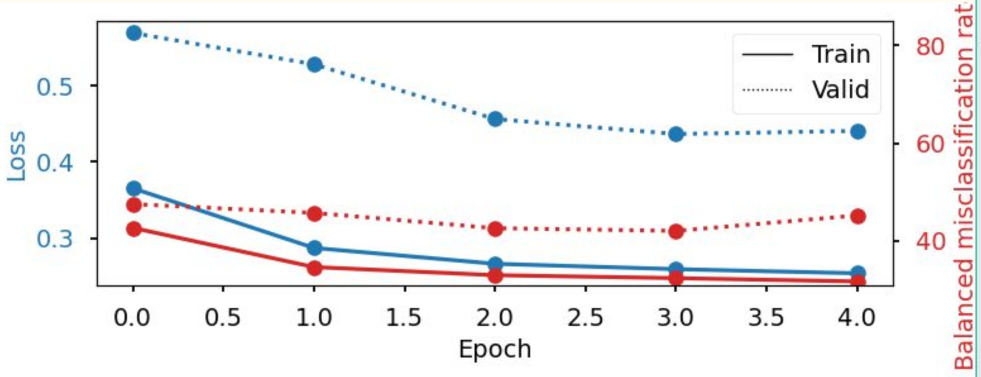



```
Training: <Epochs |  2650 events (all good), 0 - 29.99 sec, baseline off, ~121.3 MB, data loaded, with metadata,
 'Sleep stage 1': 58
 'Sleep stage 2': 250
 'Sleep stage 3': 220
 'Sleep stage 4': 220
 'Sleep stage R': 125
 'Sleep stage W': 1997>
Validation: <Epochs |  2739 events (all good), 0 - 29.99 sec, baseline off, ~125.4 MB, data loaded, with metadata,
 'Sleep stage 1': 78
 'Sleep stage 2': 349
 'Sleep stage 3': 275
 'Sleep stage 4': 275
 'Sleep stage R': 120
 'Sleep stage W': 1917>
Test: <Epochs |  2758 events (all good), 0 - 29.99 sec, baseline off, ~126.3 MB, data loaded, with metadata,
 'Sleep stage 1': 131
 'Sleep stage 2': 481
 'Sleep stage 3': 39
 'Sleep stage R': 120
 'Sleep stage W': 1987>
```

epoch	train_bal_acc	train_loss	valid_bal_acc	valid_loss	dur
1	0.5735	0.3652	0.5248	0.5695	422.9851
2	0.6537	0.2869	0.5428	0.5283	439.7777
3	0.6703	0.2660	0.5741	0.4564	511.0618
4	0.6763	0.2590	0.5796	0.4366	547.5684
5	0.6831	0.2535	0.5481	0.4407	531.8142

Test balanced accuracy: 0.529
Test Cohen's kappa: 0.675

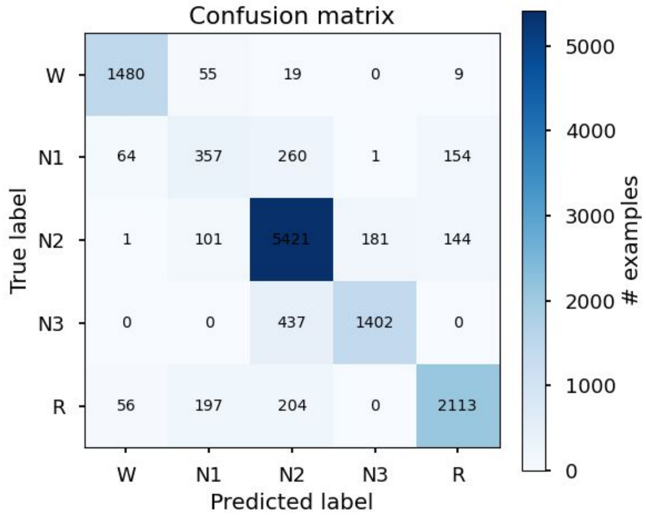
	precision	recall	f1-score	support
0	0.91	0.99	0.95	28873
1	0.43	0.14	0.21	3091
2	0.80	0.76	0.78	8220
3	0.72	0.27	0.39	603
4	0.39	0.49	0.43	2118
accuracy			0.85	42905
macro avg	0.65	0.53	0.55	42905
weighted avg	0.83	0.85	0.83	42905

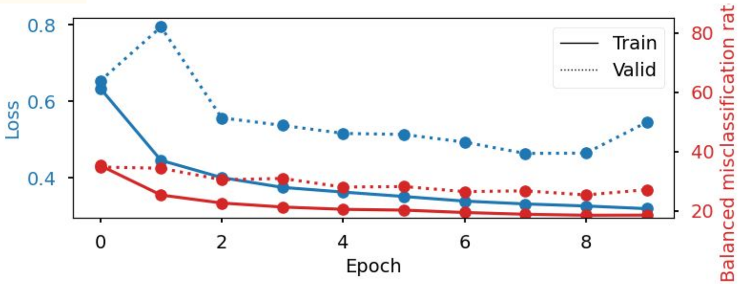


Changes

- 10 epochs instead of 5
- Shuffle windows before splitting into train, valid and test datasets

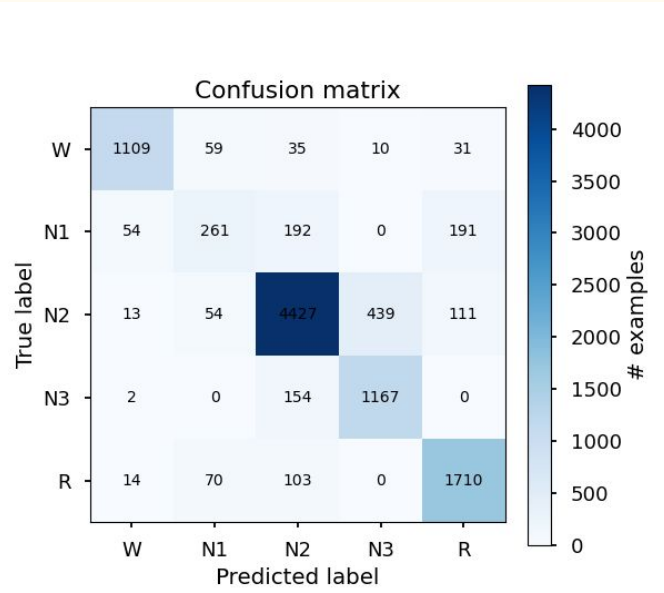
MASS_256-all-batch8_10epochs-shuffle-usual_split (sampling freq = 256)

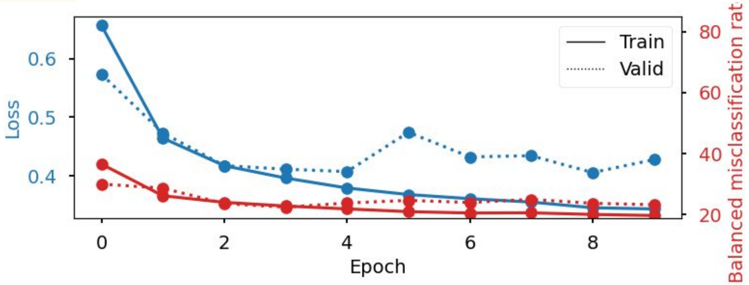





```
Training: <Epochs |  732 events (all good), 0 - 29.9961 sec, baseline off, ~857.8 MB, data loaded, with metadata,
 'Sleep stage 1': 42
 'Sleep stage 2': 324
 'Sleep stage 3': 162
 'Sleep stage R': 195
 'Sleep stage W': 9>
Validation: <Epochs |  773 events (all good), 0 - 29.9961 sec, baseline off, ~905.9 MB, data loaded, with metadata,
 'Sleep stage 1': 61
 'Sleep stage 2': 220
 'Sleep stage 3': 167
 'Sleep stage R': 107
 'Sleep stage W': 218>
Test: <Epochs |  611 events (all good), 0 - 29.9961 sec, baseline off, ~716.0 MB, data loaded, with metadata,
 'Sleep stage 1': 23
 'Sleep stage 2': 308
 'Sleep stage 3': 149
 'Sleep stage R': 108
 'Sleep stage W': 23>
Using device 'cuda'.
  epoch    train_bal_acc    train_loss    valid_bal_acc    valid_loss      dur
-------  ---------------  ------------  ---------------  ------------  -------
      1           0.6469        0.6339           0.6520        0.6534  51.4051
      2           0.7461        0.4459           0.6560        0.7940  55.2796
      3           0.7730        0.4011           0.6947        0.5568  54.6039
      4           0.7864        0.3755           0.6902        0.5373  54.8501
      5           0.7942        0.3633           0.7194        0.5163  54.1410
      6           0.7969        0.3517           0.7175        0.5136  54.8169
      7           0.8049        0.3398           0.7346        0.4935  55.6171
      8           0.8108        0.3323           0.7320        0.4642  55.5627
      9           0.8142        0.3270           0.7457        0.4649  55.9751
     10           0.8137        0.3200           0.7282        0.5453  55.3381
Test balanced accuracy: 0.777
 Test Cohen's kappa: 0.785

              precision    recall  f1-score   support

           0       0.92      0.95      0.94      1563
           1       0.50      0.43      0.46       836
           2       0.85      0.93      0.89      5848
           3       0.89      0.76      0.82      1839
           4       0.87      0.82      0.85      2570

    accuracy                           0.85     12656
   macro avg       0.81      0.78      0.79     12656
weighted avg       0.85      0.85      0.85     12656
```

MASS_100-all-batch8_10epochs-shuffle-usual_split (sampling_freq=100)






```
Training: <Epochs |  746 events (all good), 0 - 29.9961 sec, baseline off, ~874.2 MB, data loaded, with metadata,
 'Sleep stage 1': 39
 'Sleep stage 2': 286
 'Sleep stage 3': 83
 'Sleep stage R': 138
 'Sleep stage W': 200>
Validation: <Epochs |  1536 events (all good), 0 - 14.998 sec, baseline off, ~1.76 GB, data loaded, with metadata,
 'Sleep stage 1': 202
 'Sleep stage 2': 762
 'Sleep stage 3': 76
 'Sleep stage R': 276
 'Sleep stage W': 220>
Test: <Epochs |  651 events (all good), 0 - 29.9961 sec, baseline off, ~762.9 MB, data loaded, with metadata,
 'Sleep stage 1': 27
 'Sleep stage 2': 333
 'Sleep stage 3': 148
 'Sleep stage R': 125
 'Sleep stage W': 18>
Using device 'cuda'.
  epoch    train_bal_acc    train_loss    valid_bal_acc    valid_loss      dur
-------  ---------------  ------------  ---------------  ------------  -------
      1           0.6343        0.6564           0.6997        0.5735  54.3571
      2           0.7381        0.4644           0.7132        0.4733  54.4066
      3           0.7596        0.4173           0.7637        0.4175  54.8864
      4           0.7718        0.3963           0.7755        0.4114  55.6856
      5           0.7812        0.3793           0.7619        0.4073  57.0859
      6           0.7902        0.3681           0.7532        0.4744  56.2749
      7           0.7947        0.3611           0.7602        0.4324  55.8299
      8           0.7941        0.3553           0.7499        0.4344  56.8281
      9           0.7994        0.3456           0.7628        0.4056  57.3421
     10           0.8029        0.3436           0.7674        0.4278  56.2074
Test balanced accuracy: 0.785
 Test Cohen's kappa: 0.782

              precision    recall  f1-score   support

           0       0.93      0.89      0.91      1244
           1       0.59      0.37      0.46       698
           2       0.90      0.88      0.89      5044
           3       0.72      0.88      0.79      1323
           4       0.84      0.90      0.87      1897

    accuracy                           0.85     10206
   macro avg       0.80      0.79      0.78     10206
weighted avg       0.85      0.85      0.85     10206
```

References

BIDS: Gorgolewski, K.J., Auer, T., Calhoun, V.D., Craddock, R.C., Das, S., Duff, E.P., Flandin, G., Ghosh, S.S., Glatard, T., Halchenko, Y.O., Handwerker, D.A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B.N., Nichols, T.E., Pellman, J., Poline, J.-B., Rokem, A., Schaefer, G., Sochat, V., Triplett, W., Turner, J.A., Varoquaux, G., Poldrack, R.A. (2016). **The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments**. Scientific Data, 3 (160044).
doi:10.1038/sdata.2016.44

EEG-BIDS: Pernet, C. R., Appelhoff, S., Gorgolewski, K.J., Flandin, G., Phillips, C., Delorme, A., Oostenveld, R. (2019). **EEG-BIDS, an extension to the brain imaging data structure for electroencephalography**. Scientific data, 6 (103).
doi:10.1038/s41597-019-0104-8