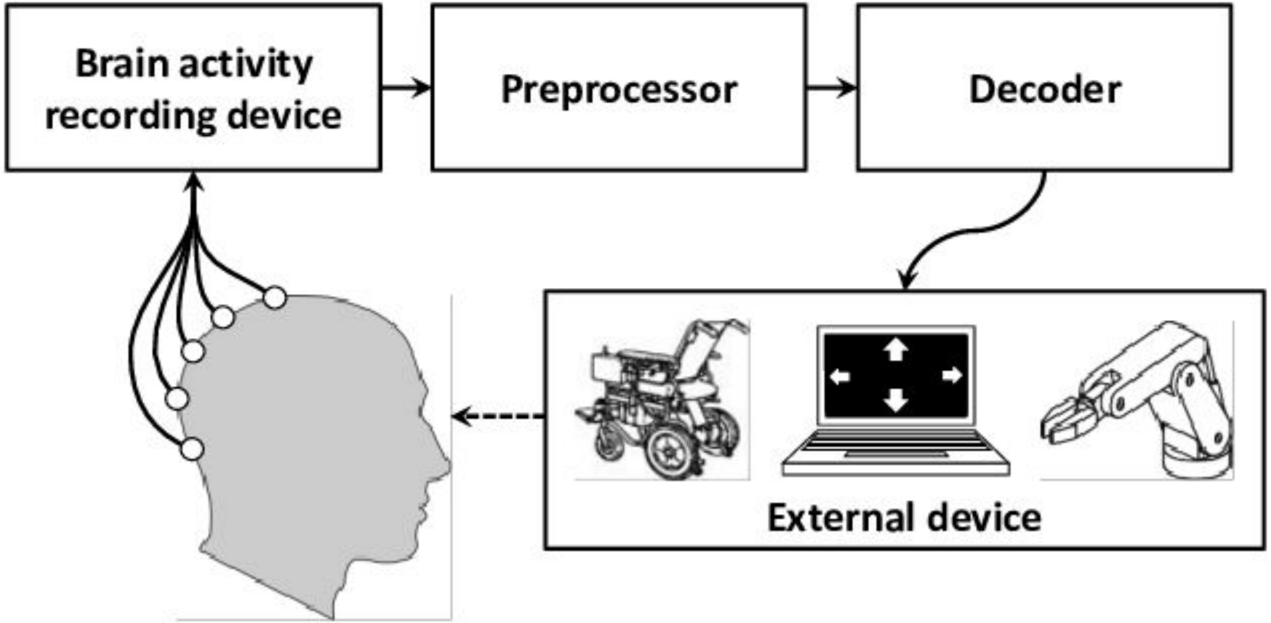
DIPARTIMENTO
MATEMATICA
DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

BCI for everyone

A subject-independent approach to EEG data

Introduction

What is BCI?



Brain activity recording device
Preprocessor
Decoder
External device

Why EEG data?

Compared to other approaches, EEG devices are:

- **Non invasive**
- Easier to use
- More affordable
- Of high portability

Related work

Related work

A study on current Deep Learning techniques for EEG classification in MI-based BCI offered some insights:

1. **Signal preprocessing** should be minimal
2. **Input formulation** should be raw signals
3. **CNNs** are the most popular architectural choice

Altaheri, H. *et al.* (2023) Deep learning techniques for classification of electroencephalogram (EEG) motor imagery (MI) signals: a review.
https://doi.org/10.1007/s00521-021-06352-5

Related work

Another study investigated data augmentation methods on EEG data, including:

1. **ChannelsDropout**

2. **BandstopFilter**

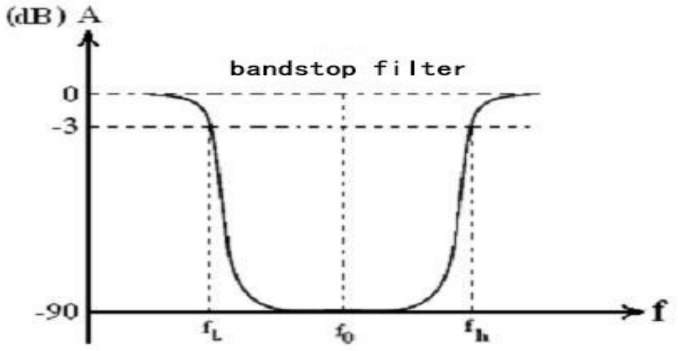


Fig. 1 Frequency response of a bandstop filter

Rommel, C. *et al.* (2022). Data augmentation for learning predictive models on EEG: a systematic comparison.
https://doi.org/10.48550/arXiv.2206.14483

Signals and features

Dataset

Motor imagery dataset for BCI applications

60 hours of EEG recordings:

- 13 participants
- 75 recording sessions

Four different MI paradigms

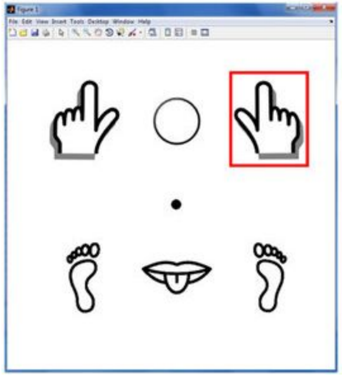


Fig. 2 HaLT graphical user interface

Kaya, M. *et al.* (2018) A large electroencephalographic motor imagery dataset for electroencephalographic brain computer interfaces.
https://doi.org/10.1038/sdata.2018.211

Experimental Design

Signals: recorded at 200 Hz and filtered by a 0.53-70 Hz bandpass filter

Trial: action signals (1 s) + off-time (1.5 to 2 s)

Trials repeated 900 times per recording session

Channels: Last channel used for data synchronization

Kaya, M. *et al.* (2018) A large electroencephalographic motor imagery dataset for electroencephalographic brain computer interfaces.
https://doi.org/10.1038/sdata.2018.211

Signal preprocessing

1. Low bandpass filtering with 38 Hz cut-off

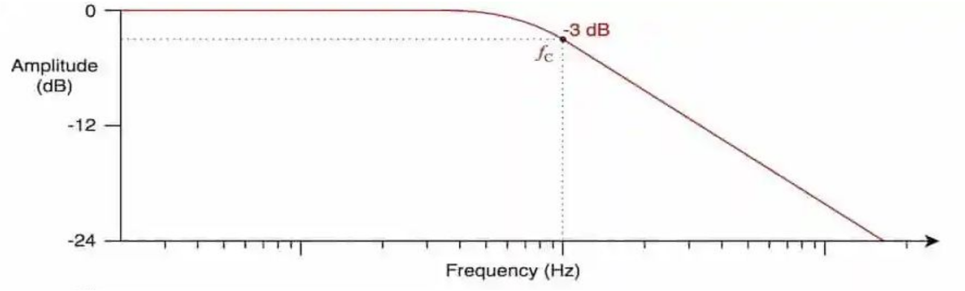
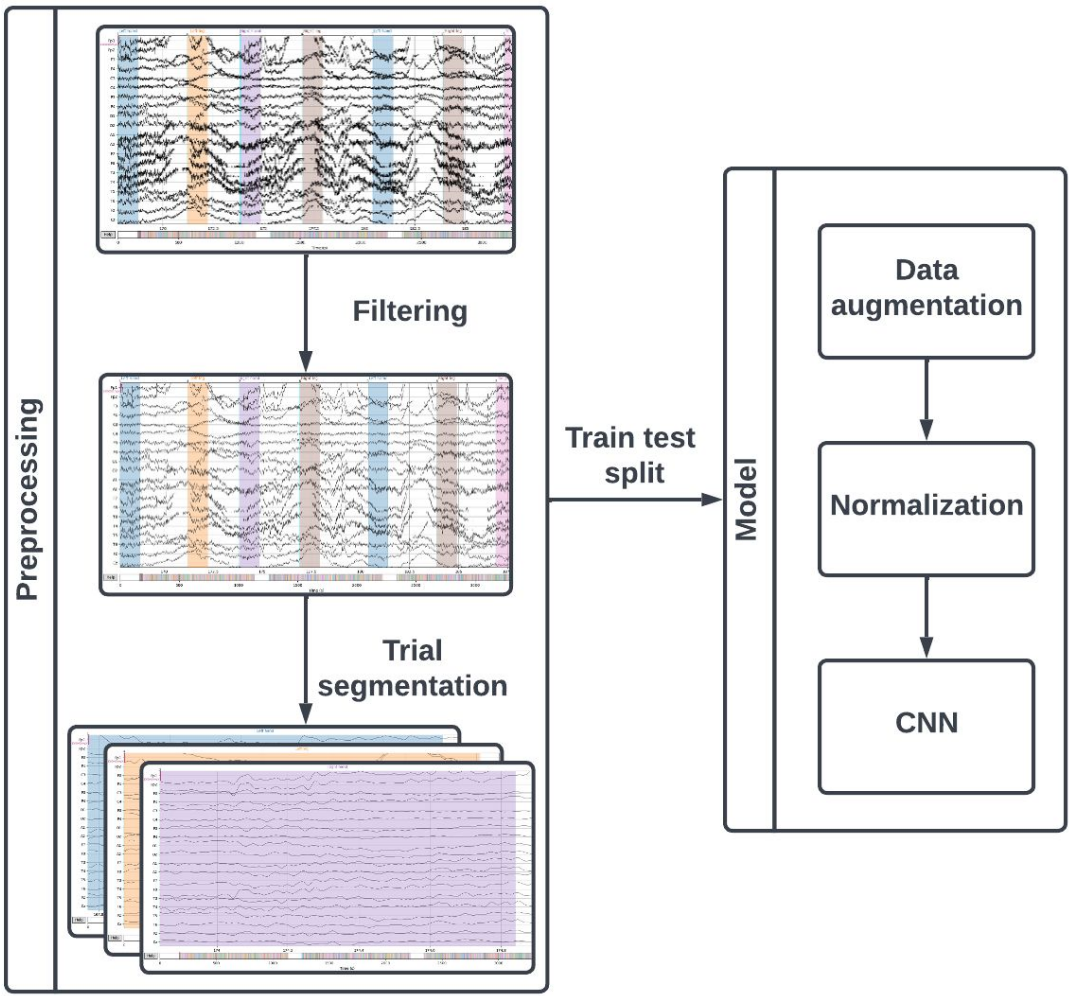


Fig. 3 Frequency response of a low bandpass filter

2. Segmentation of trials (210 samples)

Processing pipeline

Processing pipeline



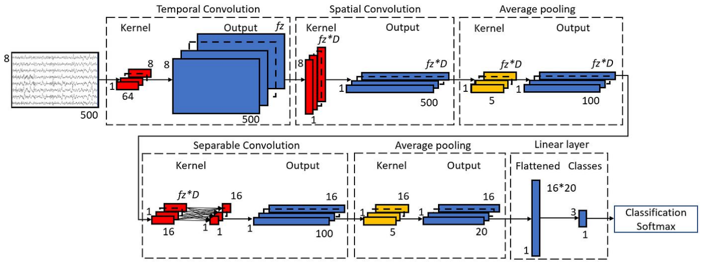
Learning Framework

Learning Framework

Goal: Evaluate subject-independent performance through a LOSO CV technique

- We considered the following architectures:

1. EEGNet (2016, Lawhern)
2. EEG-Inception (2021, Zhang)
3. Shallow ConvNet (2017, Schirrmeister)
4. Deep ConvNet (2017, Schirrmeister)



Learning Framework

Goal: Evaluate subject-independent performance through a LOSO CV technique

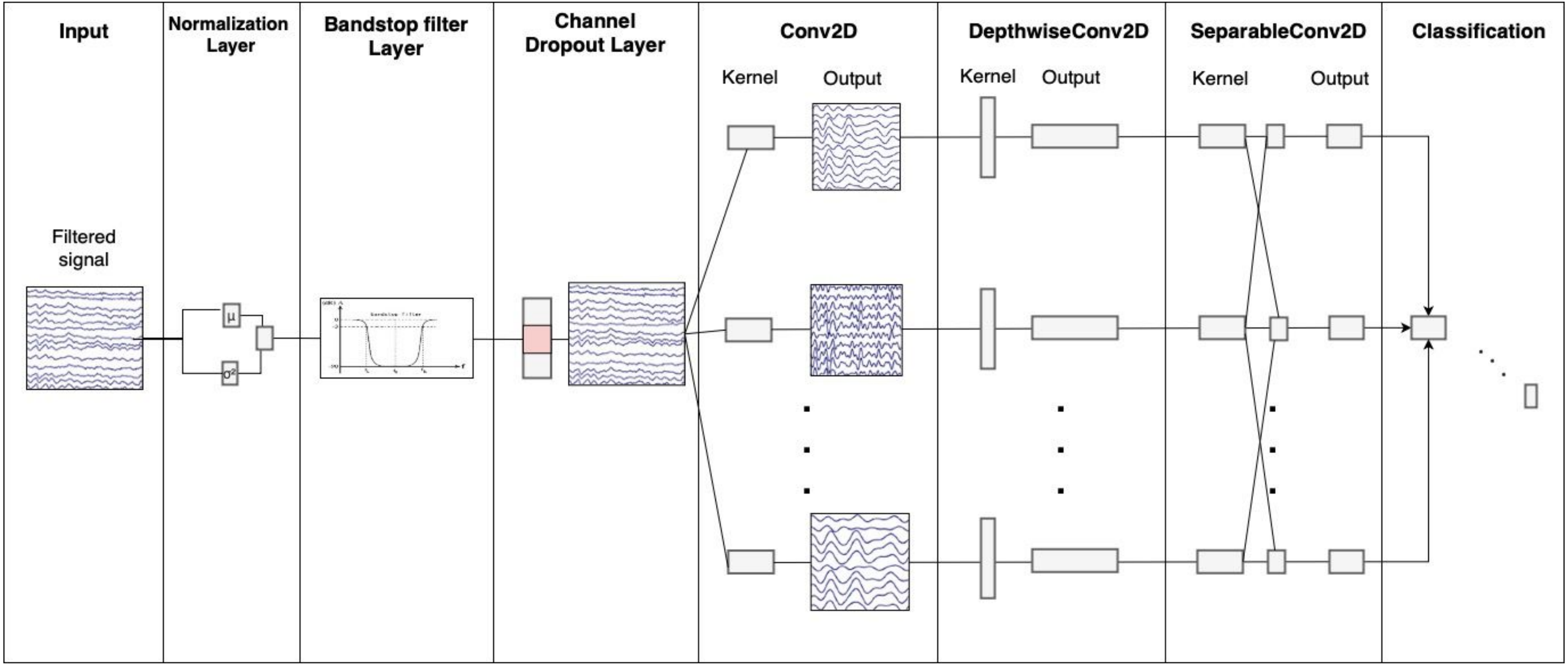
1. We chose the EEGNet architecture as a startpoint for our experimentations.
2. This decision was solely based on the lightweight nature of this architecture, which meant:
    - **Fewer parameters** to train and store.
    - **Preferable memory usage** and time complexity values.

Proposed architecture

Total params: 58,854 (229.90 KB)

Trainable params: 57,638 (225.15 KB)

Non-trainable params: 1,216 (4.75 KB)



Experiments setup

LOSO Cross Validation

	Training set									Validation set
Fold 1	B	C	F	G	H	I	K	L	...	A
Fold 2	A	C	F	G	H	I	K	L	...	B
Fold 3	A	B	F	G	H	B	K	L	...	C
Fold 4	A	B	C	G	H	B	K	C	...	F
Fold 5	A	B	C	F	H	I	K	L	...	G
Fold 6	A	B	C	F	G	I	K	L	...	H
Fold 7	A	B	C	F	G	H	K	L	...	I
Fold 8	A	B	C	F	G	H	I	L	...	K
Fold 9	A	B	C	F	G	H	I	K	...	L

Test set

E	M	J

Hyperparameters selection

Hyperparameters for our models

Hyperparameter	**Value**
Kernel length	75
Temporal filters (F1)	32
Pointwise filters (F2)	64
Spatial filters (D)	16
Dropout rate	0.25

Model candidates

		Model
Base model	EEGNet-based model	(I)
Bandstop Filter	BM + BF (width 0.5) BM + BF (width 1.0)	(II) (III)
Channel Dropout	BM + CD	(IV)
Mixed approach	BM + CD + BF (width 0.5) BM + CD + BF (width 1.0)	(V) (VI)

Results

Results

Model	**Details**	**CV Mean Accuracy**
I	EEGNet-based model	0.4967
II	BM + BF (width **0.5**)	0.5069
III	BM + BF (width **1.0**)	**0.5254 (Best result)**
IV	BM + **CD**	0.5120
V	BM + CD + BF (width **0.5**)	0.4977
VI	BM + CD + BF (width **1.0**)	0.5013

Results

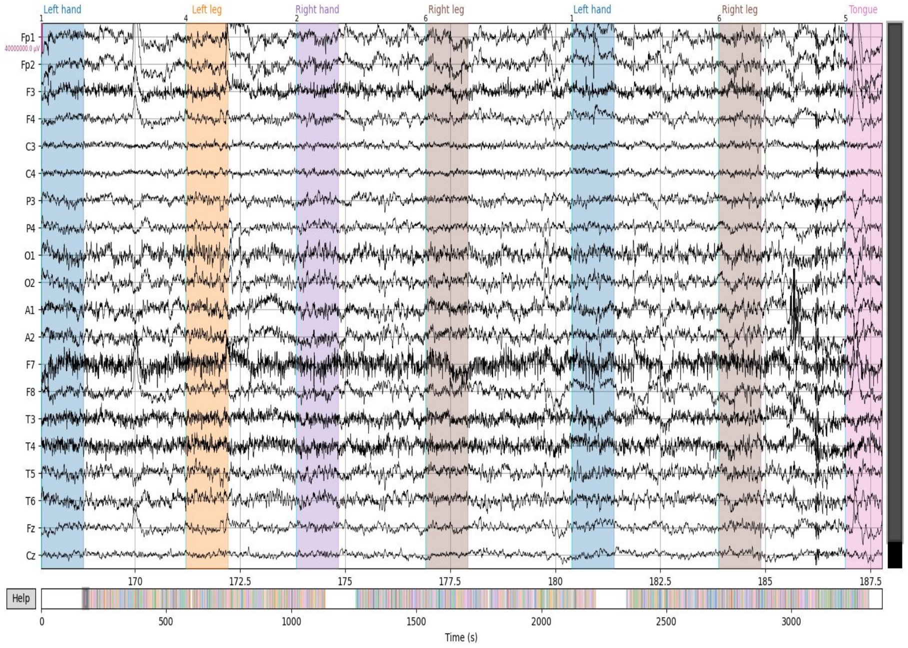
- Prediction of 6 classes and subject-independent model factors contributed to lower overall accuracies.
- No data augmentation resulted in the lowest accuracy.
- Best accuracy achieved with a Bandstop Filter of width 1.
- Significant variability in validation results among subjects.

Results

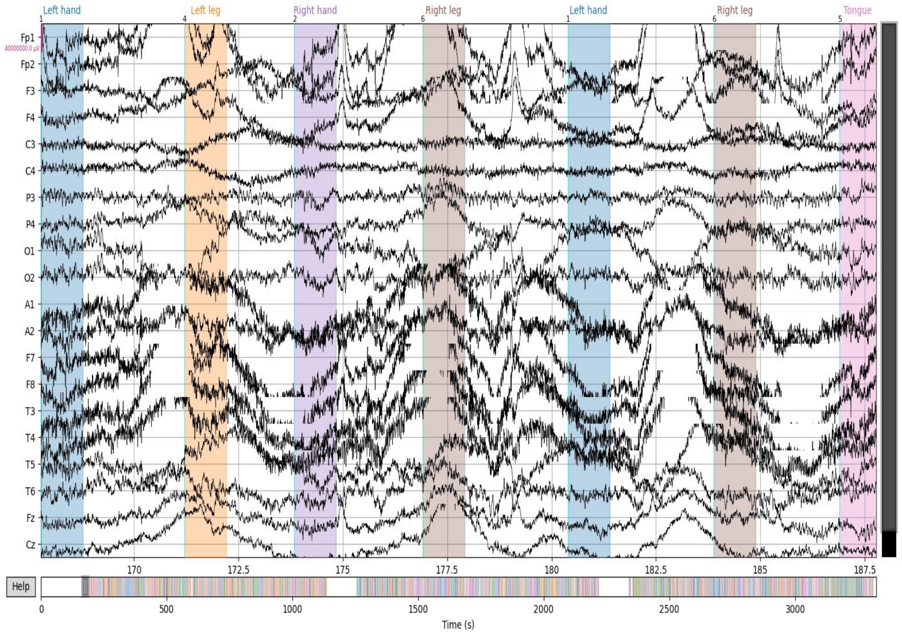
- Prediction of 6 classes and subject-independent model factors contributed to lower overall accuracies.
- No data augmentation resulted in the lowest accuracy.
- Best accuracy achieved with a Bandstop Filter of width 1.
- **Significant variability in validation results among subjects.**

Results analysis

Raw signal from Subject ‘L’

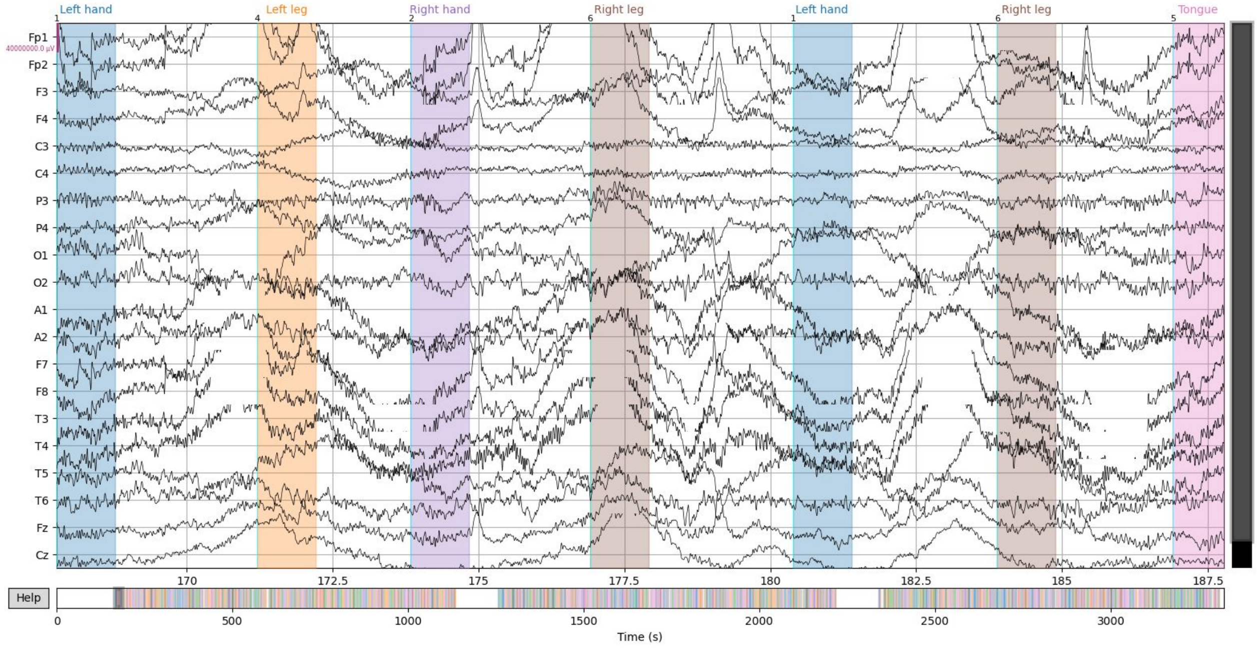


Raw signal from Subject ‘H’



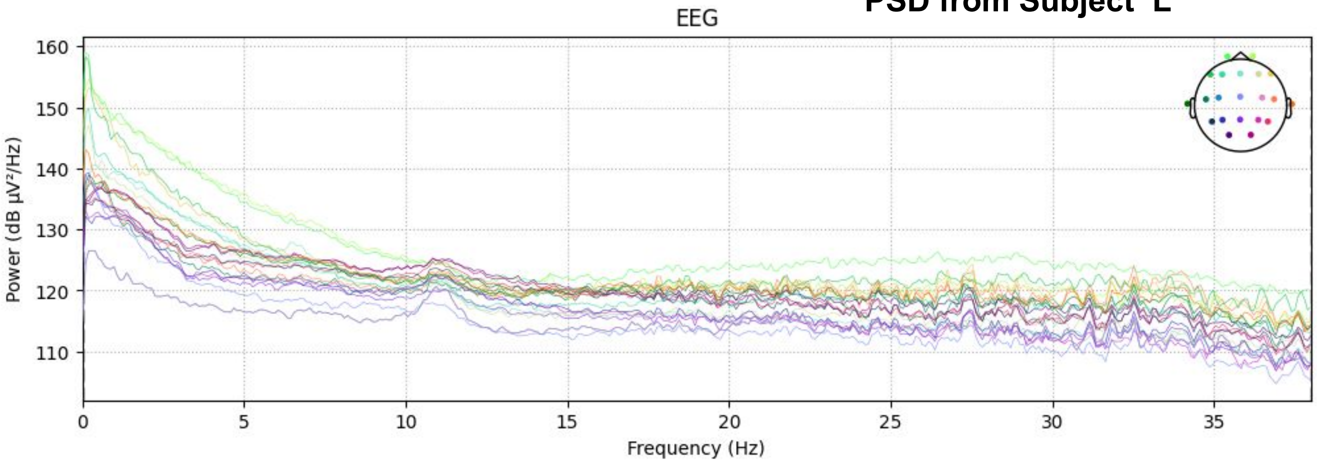
Results analysis

Filtered signal from Subject 'H'

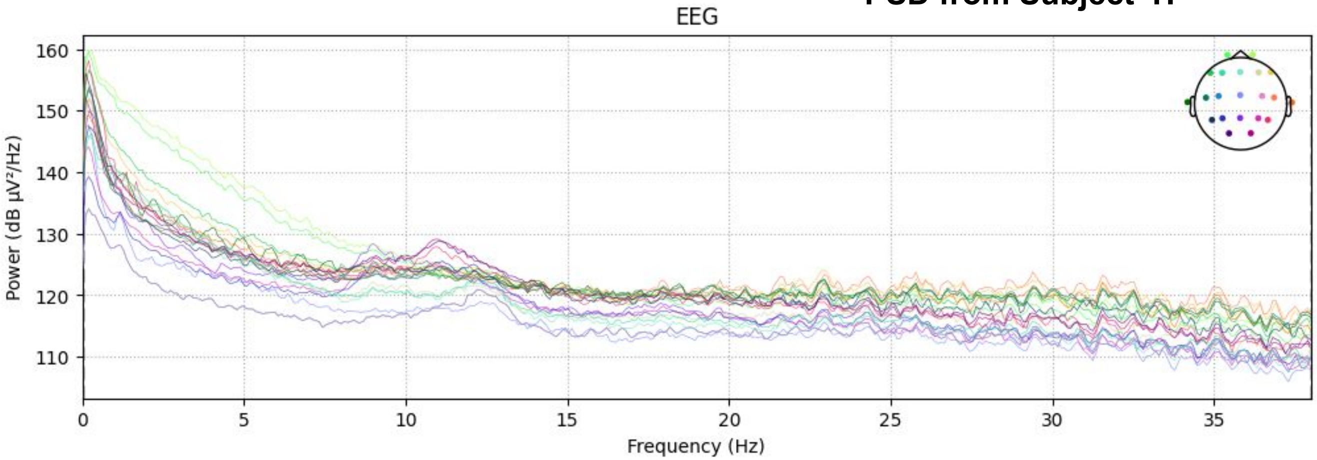


Results analysis

PSD from Subject ‘L’



PSD from Subject ‘H’



Results

Best model

Base Model + Bandstop Filter (width 1.0)

Performance metrics

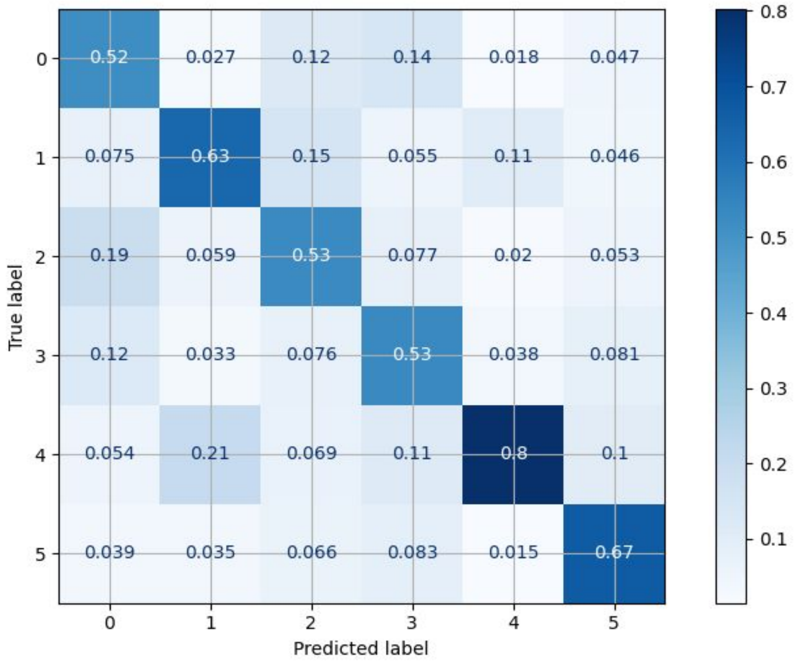
	Subjects	Accuracy
Training set	A, B, C, F, G, H, I, K, L	0.6650
Test set	E, M , J	0.5943

Results

Confusion Matrix

Best model

Base Model + Bandstop Filter (width 1.0)



- Prone to correctly predict the **'tongue'** and **'right leg'** classes , compared to other classes. While the class with the least accuracy is **"left hand"**

Concluding Remarks

- The obtained results indicate that for a subject-independent classification task, incorporating additional layers that enhance regularization improves performance.

- An appealing direction of improvement would be performing data augmentation with GANs networks, which has showed a significant increase in performance in EEG-related tasks.

F. Fahimi et al , "Generative adversarial networks-based data augmentation for brain computer interface" https://ieeexplore.ieee.org/documexnt/9177281/

Concluding Remarks

- Another interesting approach for a future work is using a Multilevel Weighted Feature Fusion architecture on the layers of the CNN. This approach has yielded competitive results in subject-independent tasks.

- Subject-independent EEG analysis remains a very challenging task. Extensive work is needed to account for the significant variability in regards to session, recording method, environment conditions, but particularly to subjects since it would make the adoption of this technology easier and more interpretable for different individuals.

S. U. Amin et a   "Multilevel Weighted Feature Fusion Using Convolutional Neural Networks for EEG Motor Imagery Classification" https://ieeexplore.ieee.org/document/8629079