Fast calibration with Deep Learning

Workshop 6: Learning from small datasets

Pierre Guetschel

Collaborative with Théodore Papadopoulo and Michael Tangermann
Radboud University, Donders Institute, Nijmegen, Netherlands

June 7, 2023 BCI Meeting, Brussels

Introduction

Deep learning is desirable for BCI.

Obstacles:

1. Long training time of the models;
2. Large amounts of training data required.

Introduction

Deep learning is desirable for BCI.

Obstacles:

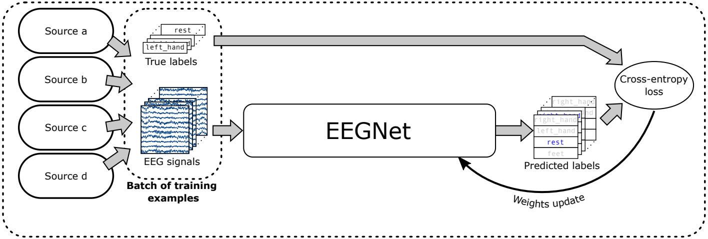
1. Long training time of the models;
2. Large amounts of training data required.

A promising direction: **transfer learning**,
i.e. re-use pre-trained models.

Plan:

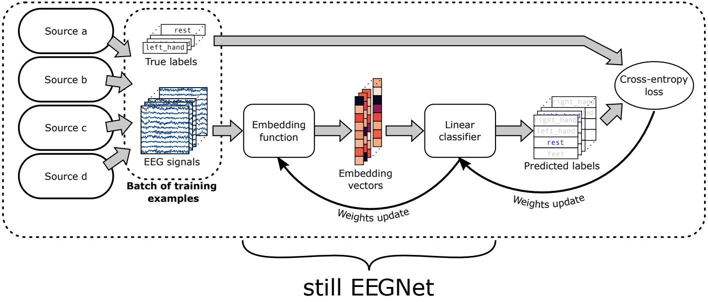
1. A simple strategy for cross-dataset transfer with calibration;
2. Tutorial for re-using pre-trained models, technically.

A simple and fast calibration strategy



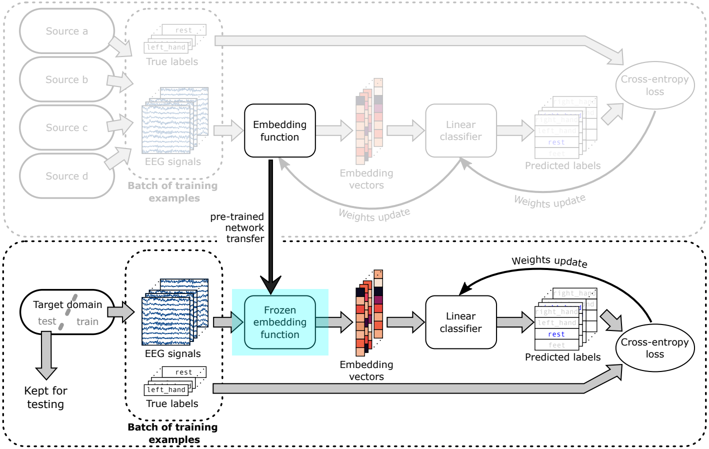
- **Phase**: 1/2 pre-training;
- **Example domains**: sessions, subjects (c.f. Analysis 1), datasets (c.f. Analysis 2), paradigms.

A simple and fast calibration strategy



- **Phase**: 1/2 pre-training;
- **Example domains**: sessions, subjects (c.f. Analysis 1), datasets (c.f. Analysis 2), paradigms.

A simple and fast calibration strategy



- **Phase**: 2/2 calibration;
- **Example domains**: sessions, subjects (c.f. Analysis 1), datasets (c.f. Analysis 2), paradigms.

Analysis 1: Cross-subject

Analysis 1: Cross-subject

[Guetschel et. al, *2022 IEEE MetroXRAINE*]

- High Gamma Dataset [Schirrmeister et. al (2017) *Hum. Brain Mapp.*];
- 14 subjects, 4 classes, 120-260 trials;
- **Pre-training**: leave-one-subject-out cross-validation;
- **Calibration**: within-session cross-validation (variable #trials).

Analysis 1: Cross-subject | Example embedding space

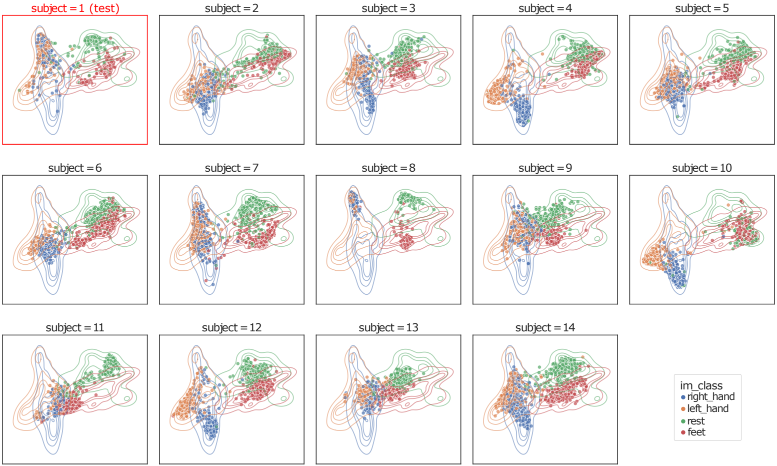


Figure: Embeddings projected by UMAP

[Guetschel et. al, *2022 IEEE MetroXRAINE*]

Analysis 1: Cross-subject | Classification accuracy

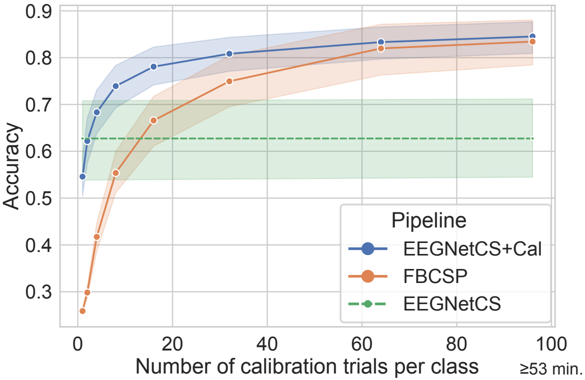


Figure: Classification accuracy averaged across subjects
[Guetschel et. al, *2022 IEEE MetroXRAINE*]

Analysis 2: Cross-dataset

Analysis 2: Cross-dataset

- 12 mental imagery datasets of MOABB;
- **Pre-training**: One whole dataset;
- **Calibration**: within-session cross-validation (variable #trials);
- **Calibration configurations**: `f-rh`, `lh-rh`, and `all`.

Dataset	#subjects	#trials	classes
AlexMI	8	20	f,r,rh
BNCI2014001	10	288	f,lh,rh,t
BNCI2014004	10	1800	lh,rh
BNCI2015001	13	400	f,rh
BNCI2015004	10	160	f,n,rh,s,wa
Cho2017	53	100	lh,rh
Lee2019_MI	55	200	lh,rh
Ofner2017	15	60	r,ree,ref, rhc,rho,rp,rs
PhysionetMI	109	23	bh,f,lh,r,rh
Schirrmeister2017	14	120	f,lh,r,rh
Weibo2014	10	80	bh,f,lh,lhrf, r,rh,rhlf
Zhou2016	4	480	f,lh,rh

- **bh**: hands
- **f**: feet
- **lh**: left hand
- **lhrf**: left hand right foot
- **n**: navigation
- **r**: rest
- **ree**: right elbow extension
- **ref**: right elbow flexion
- **rh**: right hand
- **rhc**: right hand close
- **rhlf**: right hand left foot
- **rho**: right hand open
- **rp**: right pronation
- **rs**: right supination
- **s**: subtraction
- **t**: tongue
- **wa**: word ass

Analysis 2: Cross-dataset | Results table

Test classes

Feet vs right hand

Pretraining dataset \ Test dataset	AlexMI	BNCI2014001	BNCI2015001	BNCI2015004	PhysionetMI(I)	Schirrmeister2017	Weibo2014	Zhou2016
AlexMI	60	62	56	51	56	77	52	58
BNCI2014001	70	**82**	68	52	64	90	67	87
BNCI2014004	70	66	65	**54**	60	84	66	81
BNCI2015001	70	74	**76**	53	64	89	**77**	87
BNCI2015004	50	61	55	50	54	72	51	57
Cho2017	52	58	55	50	59	66	51	62
Lee2019_MI	71	71	68	51	66	84	74	88
Ofner2017(I)	50	50	50	50	50	50	50	50
PhysionetMI(I,E)	**73**	78	70	53	**69**	90	72	85
Schirrmeister2017	72	75	71	50	64	**94**	73	88
Weibo2014	71	73	65	**54**	64	86	65	81
Zhou2016	68	71	71	52	64	86	73	**94**

Left hand vs right hand

Pretraining dataset \ Test dataset	BNCI2014001	BNCI2014004	Cho2017	Lee2019_MI	PhysionetMI(I)	Schirrmeister2017	Weibo2014	Zhou2016
AlexMI	57	62	61	55	61	60	50	54
BNCI2014001	**78**	70	64	64	68	61	59	80
BNCI2014004	68	**85**	62	67	63	65	65	84
BNCI2015001	68	70	63	63	62	59	61	72
BNCI2015004	56	62	60	53	60	61	48	55
Cho2017	63	69	**71**	58	67	59	52	62
Lee2019_MI	77	79	66	**74**	**73**	**68**	**73**	90
Ofner2017(I)	50	50	50	50	50	50	50	50
PhysionetMI(I,E)	68	67	62	65	70	60	60	74
Schirrmeister2017	69	68	62	66	64	64	62	78
Weibo2014	65	66	64	60	68	61	54	61
Zhou2016	72	74	65	69	66	66	68	**91**

All classes

Pretraining dataset \ Test dataset	AlexMI	BNCI2014001	BNCI2015004	Ofner2017(I)	PhysionetMI(I)	Schirrmeister2017	Weibo2014	Zhou2016
AlexMI	39	33	22	15	30	41	20	38
BNCI2014001	47	**52**	**24**	**17**	34	54	28	63
BNCI2014004	46	37	23	15	28	49	26	64
BNCI2015001	48	43	23	**17**	31	49	27	59
BNCI2015004	33	31	22	15	29	38	19	38
Cho2017	35	32	20	15	28	34	18	43
Lee2019_MI	47	43	21	**17**	30	49	28	70
Ofner2017(I)	33	24	20	16	12	25	14	33
PhysionetMI(I,E)	**51**	47	**24**	16	**38**	56	**30**	62
Schirrmeister2017	50	46	22	16	31	**62**	28	66
Weibo2014	45	41	22	**17**	33	49	26	53
Zhou2016	47	43	23	**17**	31	50	28	**77**

- Number of calibration trials per class: **8**;
- **Blue**: AUC score (colour range: 50-100%);
- **Green**: Accuracy score (colour range: 0-100%);
- **Bold**: best of the column.

Analysis 2: Cross-dataset | Results table

Test classes

Feet vs right hand

Pretraining dataset \ Test dataset	AlexMI	BNCI2014001	BNCI2015001	BNCI2015004	PhysionetMI(I)	Schirrmeister2017	Weibo2014	Zhou2016
AlexMI	63	66	57	51	60	81	54	60
BNCI2014001	73	**85**	71	53	67	92	72	89
BNCI2014004	72	70	67	55	63	86	71	85
BNCI2015001	70	77	**79**	53	67	90	**80**	89
BNCI2015004	52	65	57	52	56	76	51	60
Cho2017	56	61	57	49	61	70	51	65
Lee2019_MI	74	75	71	51	70	88	78	90
Ofner2017(I)	50	50	50	50	50	50	50	50
PhysionetMI(I,E)	**78**	82	73	**56**	**71**	92	77	88
Schirrmeister2017	74	80	75	51	67	**96**	77	91
Weibo2014	71	78	69	54	67	89	72	83
Zhou2016	71	75	74	53	67	88	78	**95**

Left hand vs right hand

Pretraining dataset \ Test dataset	BNCI2014001	BNCI2014004	Cho2017	Lee2019_MI	PhysionetMI(I)	Schirrmeister2017	Weibo2014	Zhou2016
AlexMI	61	64	62	56	65	64	51	56
BNCI2014001	**81**	75	66	69	72	65	64	83
BNCI2014004	71	**87**	64	70	66	68	69	86
BNCI2015001	72	73	64	66	66	63	64	78
BNCI2015004	58	65	62	55	63	64	50	57
Cho2017	66	72	**74**	60	70	62	52	65
Lee2019_MI	79	82	68	**77**	**75**	**72**	**77**	91
Ofner2017(I)	50	50	55	50	50	50	50	50
PhysionetMI(I,E)	72	70	65	68	74	65	64	80
Schirrmeister2017	74	73	64	70	68	69	66	84
Weibo2014	68	69	65	63	71	64	57	66
Zhou2016	75	78	66	72	69	70	70	**93**

All classes

Pretraining dataset \ Test dataset	AlexMI	BNCI2014001	BNCI2015004	Ofner2017(I)	PhysionetMI(I)	Schirrmeister2017	Weibo2014	Zhou2016
AlexMI	43	37	23	17	34	46	23	41
BNCI2014001	49	**56**	**25**	**18**	38	58	32	68
BNCI2014004	45	41	23	15	31	53	29	68
BNCI2015001	48	47	23	**18**	34	53	30	63
BNCI2015004	34	34	23	16	32	43	22	41
Cho2017	36	34	21	16	31	37	20	46
Lee2019_MI	48	47	23	17	33	53	32	74
Ofner2017(I)	33	24	20	17	12	25	14	33
PhysionetMI(I,E)	**56**	51	**25**	17	**43**	60	**36**	67
Schirrmeister2017	54	50	23	17	35	**66**	33	69
Weibo2014	46	46	23	**18**	38	54	31	55
Zhou2016	49	46	24	17	34	54	32	**80**

- Number of calibration trials per class: **16**;
- **Blue**: AUC score (colour range: 50-100%);
- **Green**: Accuracy score (colour range: 0-100%);
- **Bold**: best of the column.

Analysis 2: Cross-dataset | Learning curve 1/2

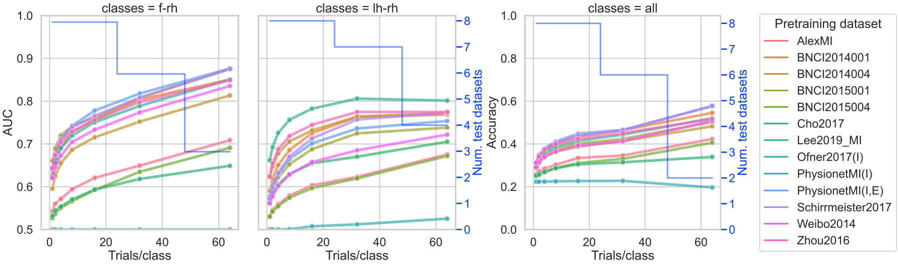


Figure: Scores wrt. num. of calibration trials per class (from test dataset); averaged over test datasets.

Best donors (pre-training datasets):

- **f-rh**: No clear winner, still BNCI2015001 best if few trials and PhysionetMI best if more trials;
- **lh-rh**: Lee2019 by a large margin (many subjects: 55).

Analysis 2: Cross-dataset | Learning curve 2/2

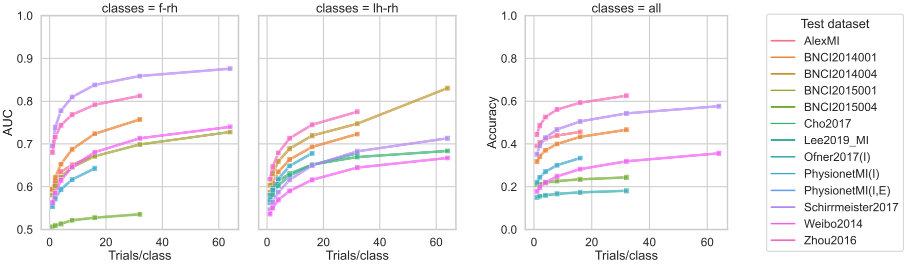


Figure: Scores wrt. num. of calibration trials per class (from test dataset); averaged over pre-training datasets.

Best receivers (test datasets):

- **f-rh**: Schirrmeister2017 (probably simpler because executed movements);
- **lh-rh**: Zhou2016 if few trials, BNCI2014004 if more trials.

(Re)Usability / Reproducibility

(Re)Usability / Reproducibility

Cross BCI dataset transfer works well with deep learning.

- C.f. previous results;
- Top 3 at BEETL competition (2021) all used DL.

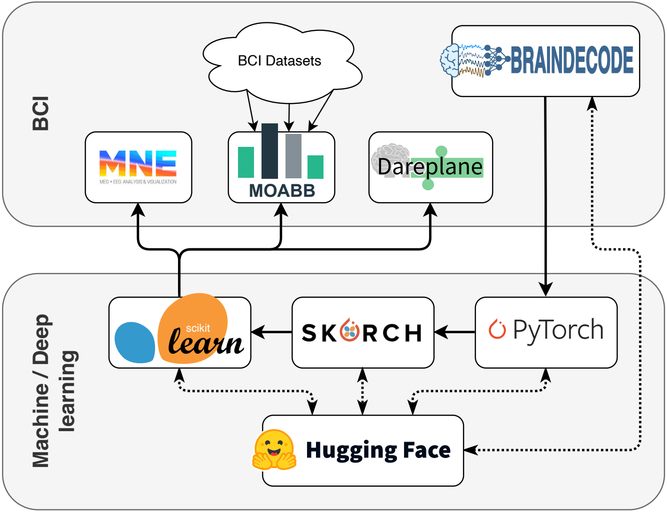
Why is cross-dataset transfer relevant?

- To reuse pre-trained models.

How can we re-use/share models?

- It can appear challenging technically;
- Tutorial $\rightarrow$

Libraries - open-source constellation / ecosystem



If you are interested, come and talk to me at the end of the session or poster 2-F-48.

But concretely... the code

Load a pre-trained model in 3 lines!

```
import torch
from huggingface_hub import hf_hub_download
from braindecode.models import EEGNetv4

path = hf_hub_download(
    repo_id='PierreGtch/EEGNetv4',
    filename='EEGNetv4_Lee2019_MI/model-params.pkl')
net = EEGNetv4(3, 2, 385).eval()
net.load_state_dict(torch.load(path, map_location='cpu'))
```

Full Notebook:

Explains how to:

- Download pre-trained models;
- Integrate into classification pipelines;
- Share your own models.

Take home message

- Cross-mental-imagery-dataset transfer works well using deep learning, even with simple pipelines;
- The infrastructures are there to *effortlessly* share and reuse models;
- We should make it a standard practice (within the BCI community) to share our pre-trained models;
  - Using pre-trained models allows for fast prototyping;
  - Using pre-trained models lowers the threshold for new studies (less data needed);
- Both our method and workflow could be used on other paradigms than mental imagery and other types of transfer than cross-dataset.

Full Notebook:

Poster

Come and chat tomorrow (session 2), spot 2-F-48.