Deep Learning from Sparse fNIRS Data to Augment High-Density Datasets Improves Decoding Performance

Mohammad Orabe[1,2], Theekshana Dissayanake[1,2], Shakiba Moradi[1,2], Alexander von Lühmann[1,2]
[1]BIFOLD — Berlin Institute for the Foundations of Learning and Data, Berlin, Germany
[2]Intelligent Biomedical Sensing (IBS) Lab, Technical University of Berlin, Berlin, Germany

m.orabe@icloud.com

Scan to view (PDF)

1 Motivation

High-density functional near-infrared spectroscopy (HD-fNIRS) and diffuse optical tomography (DOT) enable spatially resolved imaging of cortical activity [1]. However, **limited dataset sizes** and substantial variability across **subjects, measurement configurations, and signal representations** challenge the generalization of deep learning models. **Data augmentation** may address these limitations, but its effectiveness has not been systematically evaluated for **HD-fNIRS and DOT-based classification**.

We systematically evaluate **temporal, spectral, spatial, preprocessing, reconstruction, and channel-density augmentations** across multiple datasets and both **channel- and parcel-space representations**, aiming to identify strategies that improve **subject-independent classification**.

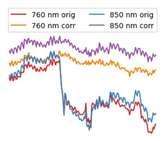
2 Datasets

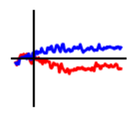
DATASET	MONTAGE	# CHANNELS	# SUBJECTS
Whole-head ball squeezing	HD	1220	13
Motor-cortex ball squeezing	HD	100	10
PFC Strop-Rest	HD	214	16
PFC Stroop-Rest	Sparse	52	16

3 Pipeline overview

Preprocessing

- Optical density conversion
- Motion correction (TDDR + wavelet)
- Channel Quality Assisment (SCI/PSP)
- Global-interference removal in OD
- Hemoglobin conversion to HbO / HbR [2, 3]

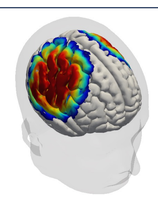


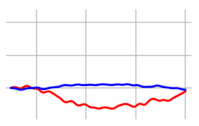


(Channel × Wavelength × Time)

DOT image reconstruction

- **Forward model:** Simulate light propagation through ICBM152 head + montage → build sensitivity matrix (Jacobian) [1, 3]
- **Inverse & reconstruction:** Invert sensitivity matrix (Tikhonov-regularized) → apply to optical density → parcel time courses





(Parcel × Chromo × Time)

Trial segmentation & baseline correction

Augmentation

Model training and evaluation

4 Augmentation Methods

1 / EEG-inspired augmentation

parcel space | parcel space

Mixes 5 sparse channel subsets into dense data with 0.5 prop.

Spatial domain
ch. symmetry | ch. dropout | ch. shuffle

Temporal domain
time mask | gaussian noise | time reverse | sign flip

Spectral domain
FT surrogate | frequency shift | bandstop filter

2 / Time-shift

parcel space | parcel space

Shift the trial window in time (Temporal offset).

3 / Channel-density sampling

parcel space

Mixes 5 sparse channel subsets into dense data with different ratios.

Sparse data source
same dataset | other dataset

Sparse subset ratio
20 % | 40 % | 60 % | 80 % | 100 %

5 Channel-density sampling

A closer look — one of the three augmentation methods from section 04, examined here in depth

Simulate lower-density montages by removing channels first, then recomputing parcel-space signals from the sparse configuration. Each dataset yields five sparse configurations, which are then mixed back into training at fixed ratios. Dense datasets are augmented with sparse ones, never the reverse.

How a sparse configuration is built

01 Start from raw recordings for each subject
02 Filter channels by source–detector distance
03 Keep only valid channels within the selected distance range
04 Project the remaining channels to parcel space using the forward model
05 Restrict the parcellation to the dataset-specific valid parcels

A / Same dataset

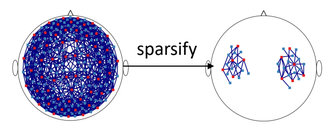
The dense data being trained on is augmented with sparse subsets drawn from that same dataset. Held-out subjects are excluded from all views.

B / different dataset

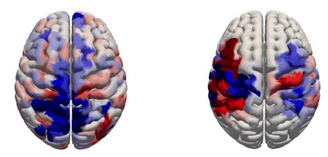
The dense data being trained on is augmented with sparse subsets drawn from a different dataset, used only for augmentation; evaluation always stays on the dense dataset.

Whole-head ball squeezing

5 sparse channel configs: 92 | 81 | 71 | 61 | 51



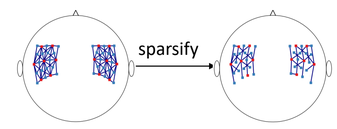
1220 channels → 51 channels



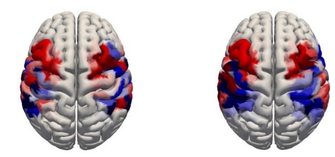
600 parcels → 600 parcels

Motor-cortex ball squeezing

5 sparse channel configs: 91 | 79 | 68 | 55 | 43



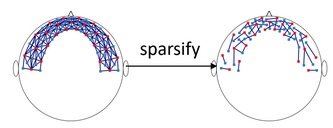
100 channels → 43 channels



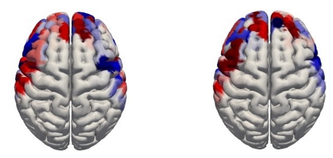
600 parcels → 600 parcels

PFC Stroop-Rest

5 sparse channel configs: 90 | 80 | 70 | 60 | 50

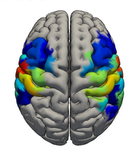


214 channels → 50 channels



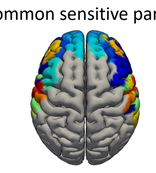
600 parcels → 600 parcels

Motor task
Restrict to common sensitive parcels



108 parcels

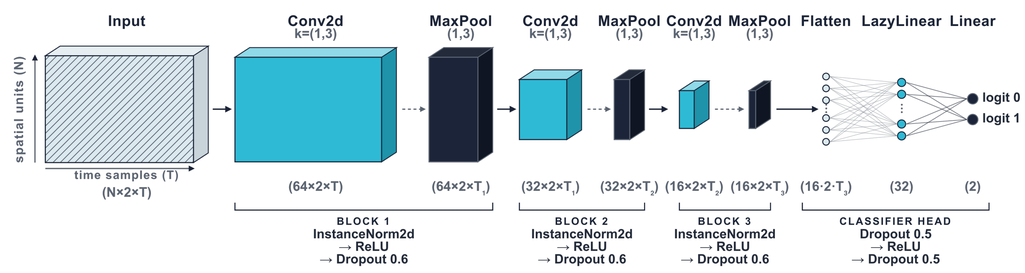
Cognitive task
Restrict to common sensitive parcels



135 parcels

6 Model: Convolutional Neural Network

A 2D CNN over fNIRS trial segments [4], binary classification. Spatial units N enter as the Conv2d channel dimension. Kernels and pools are (1,3): they act on time only, and the signal axis stays at 2 throughout; 2 wavelengths in channel space, 2 chromophores in parcel space.



Training & Evaluation protocol

EVALUATION
Strict leave-one-subject-out: one model per subject, tested only on that held-out subject.

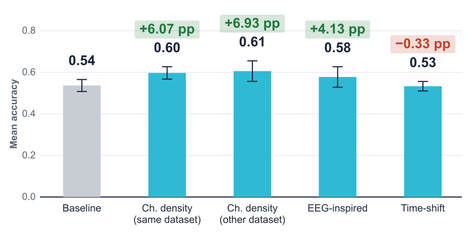
AUGMENTATION
Train folds only — the held-out subject's segments stay original and un-augmented.

OPTIMIZATION
AdamW, 200 fixed epochs per fold, cross-entropy loss function, weight decay as regularization.

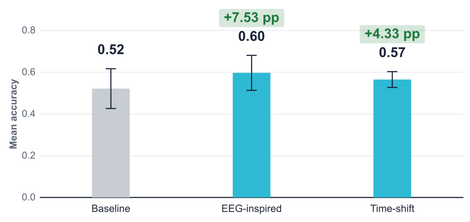
7 Results

Does data augmentation help the subjects who need it most?
Results for the worst 50% of subjects based on baseline accuracy.

Parcel space (image-reconstructed)



Channel space (sensor-level)



Mean accuracy of the three datasets, each at its best ratio or method; deltas are absolute percentage points (pp).

		Baseline (no aug.)	Ch. density same dataset	Ch. density other dataset	EEG-inspired	Time-shift
parcel space	Whole-head ball squeezing	0.52	0.57 (+5.1)	0.55 (+3.2)	0.52 (+0.8)	0.52 (+0.8)
	Motor-cortex ball squeezing	0.57	0.63 (+5.4)	0.64 (+6.5)	0.61 (+4.0)	0.56 (−1.4)
	PFC Stroop-Rest	0.52	0.60 (+7.7)	0.63 (+11.1)	0.60 (+7.6)	0.52 (−0.4)
channel space	Whole-head ball squeezing	0.43	—	—	0.50 (+6.9)	0.54 (+10.3)
	Motor-cortex ball squeezing	0.62	—	—	0.65 (+2.3)	0.61 (−1.5)
	PFC Stroop-Rest	0.51	—	—	0.64 (+13.4)	0.55 (+4.2)

Mean test accuracy across the worst-performing 50% of subjects. Deltas indicate the **absolute change in percentage points** relative to each dataset's baseline.

8 Conclusion

01 / LIFT
Data augmentation particularly benefits the worst-performing subjects, raising accuracy in nearly every dataset/space combination tested — up to **+7.53 pp** (EEG-inspired, channel space) and **+6.93 pp** (channel-density, other dataset, parcel space).

02 / DIVERSITY
Channel-density mixing **with other datasets** outperforms same-dataset mixing in **2 of 3 datasets**, suggesting the gain comes from montage diversity, not just more training samples.

03 / CONSISTENCY
EEG-inspired and channel-density augmentation help consistently across every configuration tested; time-shift is the only strategy that sometimes hurts accuracy, despite one large win **(+10.30 pp).**

9 References

[1] Wheelock MD, Culver JP, Eggebrecht AT. High-density diffuse optical tomography for imaging human brain function. Rev Sci Instrum, 2019. doi:10.1063/1.5086809.

[2] Yücel MA, Lühmann AV, Scholkmann F, et al. Best practices for fNIRS publications. Neurophotonics, 2021. doi:10.1117/1.NPh.8.1.012101.

[3] Middell E, Carlton LB, Moradi S, et al. Cedalion tutorial. Neurophotonics, 2026. doi:10.1117/1.NPh.13.S3.S32602.

[4] Dissanayake T, Müller KR, von Lühmann A. Deep learning from diffuse optical oximetry time-series. IEEE Rev Biomed Eng, 2025. doi:10.1109/RBME.2025.3617858.

Datasets
Whole-head ball squeezing: https://doi.org/10.82901/nemar.on007420
Motor-cortex ball squeezing: https://doi:10.1117/1.NPh.10.2.025007
PFC Stroop-Rest: https://doi:10.1117/1.NPh.12.3.035010