Subject- and Task-Aware EEG Foundation Model

Sion An[1], Soopil Kim[2], and Sang Hyun Park[1,3⋆]

[1] Graduate School of Artificial Intelligence, Pohang Unversity of Science and Technology (POSTECH), Pohang, Republic of Korea
[2] AX Research Group for Semiconductors, Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu, Republic of Korea
[3] Department of Computer Science & Engineering, Pohang Unversity of Science and Technology (POSTECH), Pohang, Republic of Korea
{sionan624,sanghyunpark}@postech.ac.kr

Abstract. Recent advances in EEG foundation models have predominantly relied on self-supervised learning to extract generalizable representations from large-scale datasets. However, this paradigm overlooks the rich semantic information such as subject identities and task labels available in pre-training datasets. Therefore, it limits the model's ability to consider inter-subject variability and task heterogeneity in EEG datasets, often resulting in suboptimal generalization to unseen subjects and adaptation to various tasks. In this paper, we propose a novel **S**ubject- and **T**ask-Aware **E**EG Foundation **M**odel (STEM) that fully leverages available semantic information. This method disentangles feature representations into subject-aware and task-aware components by integrating a metric-based meta-learning with contrastive self-supervision. This maximizes the utility of semantic labels to guide the learning of robust features. Extensive experiments across three downstream benchmarks demonstrate that incorporating semantic supervision significantly enhances adaptation capability for various tasks and new subjects. Our findings suggest that actively utilizing subject and task information during pre-training is crucial for building robust EEG foundation models. The code is available at https://github.com/SionAn/MICCAI2026STEM.

Keywords: Electroencephalograph · Foundation model · Meta-learning.

1 Introduction

The foundation model for electroencephalography (EEG) analysis holds immense potential to innovate various applications such as brain-computer interfaces (BCI) [3,4], clinical diagnostics of neurological disorders [9,28], and the discovery of biomarkers [6,8]. Recent studies [13,30,37] have leveraged self-supervised learning (SSL) to pre-train models. They define pretext tasks such as masked signal reconstruction using EEG datasets to capture temporal and channel-wise features without manual annotations. For example, LaBraM [13] introduces a

⋆ Corresponding author

vector quantized spectral tokenizer that enables the model to learn frequency-aware representations, remaining robust despite the inherent noise of neural signals. To capture global dependencies, CBraMod [30] employs a two-directional attention that alternately encodes temporal and spatial information, efficiently modeling the complex interplay between channel and time dimensions. Furthermore, CSBrain [37] addresses the hierarchical nature of neural activity through a cross-scale modeling framework, integrating features across varying resolutions to capture the multi-scale dynamics of the brain. While SSL pre-training provides a strong feature extraction backbone, current paradigms overlook the rich semantic information such as subject identities and task labels that are often available in pre-training datasets. EEG signals have significant inter-subject variability and task heterogeneity arising from diverse neuroanatomical structures, electrode locations, and recording protocols. They cause EEG signals to differ drastically from person to person and dataset to dataset, even when performing the identical task. Consequently, by treating these heterogeneous signals merely as unlabeled data, existing models often struggle to learn a structured latent space capable of handling the large variability, leading to suboptimal adaptation to new subjects and diverse downstream tasks. Therefore, a key challenge lies in effectively leveraging the semantic richness embedded in existing datasets to learn representations that are robustly adaptable to new subjects and tasks.

In this paper, we introduce a novel **S**ubject- and **T**ask-Aware **E**EG Foundation **M**odel (STEM). Unlike previous methods that rely solely on SSL, STEM actively utilizes available semantic information to address both task and subject heterogeneity. Specifically, we integrate self-supervised contrastive learning with a supervised metric-based meta-learning strategy to learn a general representation space that is robust to variations across subjects and adaptable to various downstream tasks. STEM is explicitly trained to separate EEG representations into two distinct components: (1) a subject-aware feature that captures the unique neurophysiological fingerprint of each subject, and (2) a task-aware feature that contains the subject-invariant and task-relevant information required for downstream applications. This disentangled and meta-learned approach enables efficient adaptation, improving the model's generalization capabilities to new subjects and new tasks. The main contributions of this work are as follows:

1. We propose a novel EEG foundation model that moves beyond SSL pre-training by integrating a supervised metric-based meta-learning framework. This approach fully leverages the semantic information such as subject and task labels available in large-scale EEG datasets to structure the latent space.
2. We introduce a feature disentanglement with meta-learning to explicitly separate subject- and task-aware representations. This mitigates the variability in EEG signals, ensuring the model learns features that are both robust to individual neurophysiological differences and adaptable to new tasks.
3. We demonstrate through extensive experiments that our model achieves consistent improvements on three downstream benchmark datasets. The results validate that explicitly modeling subject and task semantics is essential for building general-purpose EEG foundation models.

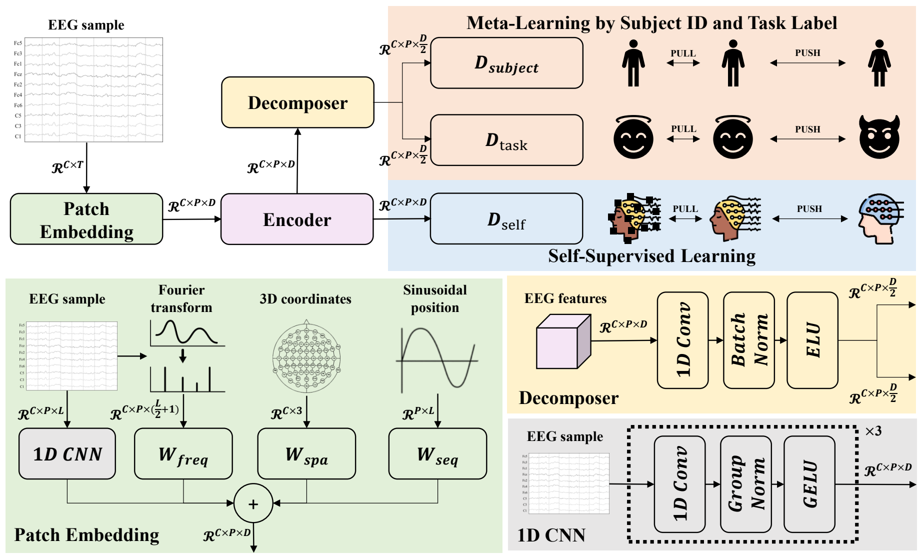

Fig. 1: The overview of STEM. The model comprises a shared encoder, a decomposer and three distinct decoders. EEG sample first is embedded with frequency, spatial and sequential embedding and then fed into the encoder. The encoded features are utilized in two pathways: (1) direct input to D_{self} for self-supervised contrastive learning with an augmented view, and (2) disentanglement via a decomposer into subject-aware and task-aware features, which are processed by $D_{subject}$ and D_{task}, respectively, for meta-learning.

2 Method

We present the **S**ubject- and **T**ask-Aware **E**EG Foundation **M**odel (STEM) designed to fully exploit the semantics embedded in the dataset *i.e.*, subject identities and task labels. The objective is to learn the generalizable feature space that can be effectively adapted to new subjects and new tasks. We employ the meta-learning strategy with disentanglement of the latent representation into subject-dependent and task-dependent features. As shown in Fig. 1, STEM comprises a shared encoder followed by three specific decoders that are jointly optimized metric-based meta-learning [1,2] and self-supervised contrastive learning.

For pre-training, we use the task-related dataset $\mathcal{D}^{TS} = \{(x_i, y_i^s, y_i^t)\}_{i=1}^{N^{TS}}$ and unlabeled dataset $\mathcal{D}^{RS} = \{(x_j, y_j^s)\}_{j=1}^{N^{RS}}$. Here, $x \in \mathbb{R}^{C\times T}$ represents the EEG signals with C channels and T time points. y^s denotes the subject ID of x and y^t denotes the task such as right-hand in motor imagery task corresponding to x. N^{TS} and N^{RS} are the number of EEG signals in $\mathcal{D}^{TS}$ and $\mathcal{D}^{RS}$, respectively.

Patch embedding. We adopt a patch-based tokenization strategy. Given x, we first segment into a sequence of non-overlapping patches along the time-axis.

x is reshaped into a sequence of patches $x_p \in \mathbb{R}^{C\times P\times L}$, where P is the number of patches and L is the length of each patches. These patches are then mapped into a latent space of dimension D using a 1D convolutional neural network ϕ consisting of three layers with a kernel size of 3. To encode the temporal order of patches, the spectral information of sequences, and the spatial configuration of electrodes, we employ three learnable embeddings: (1) a sequential embedding E_{seq} uses fixed sinusoidal position encoding [29]. It is then passed through a linear projection W_{seq}. (2) a frequency embedding E_{freq} first applies a Fourier transform to x. The magnitude of the complex spectrum is then computed, resulting in a spectral feature vector for each patch. It is then passed through a linear projection W_{freq} (3) a spatial embedding E_{spa} uses 3D coordinates of electrodes. It is then mapped to the latent space using a linear projection W_{spa}. The final patch embedding E is obtained by linear summation of them *i.e.*, $E = \phi(x_p) + E_{seq} + E_{freq} + E_{spa}$, encoding the spatiotemporal and spectral context of EEG signals.

The final embedding E is fed into the shared encoder. The output z is then distributed into two pathways. First, z is directly input to the self-supervised decoder D_{self}, resulting z^{self}. Simultaneously, z is processed by a decomposer that disentangles z into subject- and task-dependent features. These features are subsequently forwarded to the subject decoder D_{sub} and the task decoder D_{task}, resulting z^{sub} and z^{task}, respectively.

Subject-aware meta-learning. To consider the inter-subject variability and utilize semantic information regarding to subject identities, we employ the binary classification as meta-learning objective, operating under the assumption that identifying the subject constitutes a distinct meta-task. Specifically, we adopt an asymmetric sampling strategy where the positive is represented by samples $\mathcal{X}_+^{sub} = \{(x_{+,k}, y_{+,k}^s)\}_{k=1}^K$ sharing the same subject identity, while the negative is represented by a single sample (x_-, y_-^s) with a different subject identity. This stabilizes training by enforcing a direct comparison against a sample-level negative feature rather than an aggregated prototype.

For a query embedding z_q^{sub} with subject label y_q^s, we compute the positive prototype c_+^{sub} as the mean of representations from $\mathcal{X}_+^{sub}$, while the negative prototype c_-^{sub} corresponds to the extracted features from negative sample x_-. We define the probability distribution over the subjects for z_q^{sub} based on the cosine similarity *sim* to the prototypes as follows:

$$p(y_+^s = y_q^s | z_q^{sub}) = \frac{\exp(\text{sim}(z_q^{sub}, c_+^{sub})/\tau)}{\exp(\text{sim}(z_q^{sub}, c_+^{sub})/\tau) + \exp(\text{sim}(z_q^{sub}, c_-^{sub})/\tau)}, \quad (1)$$

where τ is a temperature hyperparameter. Finally, the subject-aware loss $\mathcal{L}_{sub}$ is calculated as the negative log-likelihood of the true class probability as follows: $\mathcal{L}_{sub} = -\log p(y_+^s = y_q^s | z_q^{sub})$.

Task-aware meta-learning. To facilitate effective adaptation to diverse tasks and leverage task semantics, we employ the 2-way K-shot metric-based meta-learning objective. Specifically, for a query z_q^{task} with task label y_q^t, we sample both a positive support set $\mathcal{X}_+^{task} = \{(x_{+,k}, y_{+,k}^t)\}_{k=1}^K$ and a negative support set $\mathcal{X}_-^{task} = \{(x_{-,k}, y_{-,k}^t)\}_{k=1}^K$, where $y_q^t = y_+^t \neq y_-^t$. We compute the task prototypes c_+^{task} and c_-^{task} by averaging the representations of $\mathcal{X}_+^{task}$ and $\mathcal{X}_-^{task}$, respectively. We define the probability distribution over the task for z_q^{task} based on the cosine similarity to the prototypes as follows:

$$p(y_+^t = y_q^t | z_q^{task}) = \frac{\exp(\mathrm{sim}(z_q^{task}, c_+^{task})/\tau)}{\exp(\mathrm{sim}(z_q^{task}, c_+^{task})/\tau) + \exp(\mathrm{sim}(z_q^{task}, c_-^{task})/\tau)}. \quad (2)$$

Finally, the task-aware loss $\mathcal{L}_{task}$ is calculated as the negative log-likelihood of the true class probability as follows: $\mathcal{L}_{task} = -\log p(y_+^t = y_q^t | z_q^{task})$.

Auxiliary self-supervised learning. We incorporate the self-supervised contrastive learning to leverage abundant unlabeled EEG signals such as resting-state EEG signals. Instead of masked signal reconstruction, we adopt the contrastive learning object to capture the global semantic representations. Specifically, we generate a positive augmented view x_{mask} for x by randomly masking. Both are processed by the shared encoder and the self-supervised decoder D_{self} to obtain the final representations z^{self} and z_{mask}^{self}. The model is trained to distinguish this positive pair from other samples in the batch by minimizing the contrastive loss $\mathcal{L}_{self}$ following [7] as follows:

$$\mathcal{L}_{self} = -\log \frac{\exp(\mathrm{sim}(z^{self}, z_{mask}^{self})/\tau)}{\sum_{j=1}^{2B} \mathbb{1}_{[j \neq i]} \exp(\mathrm{sim}(z^{self}, z_j^{self})/\tau)}, \quad (3)$$

where B is the batch size.

Overall objective. The training of STEM is conducted in two phases. In the pre-training phase, STEM is optimized using a linear combination of $\mathcal{L}_{sub}$, $\mathcal{L}_{task}$ and $\mathcal{L}_{self}$. This enforces the model to simultaneously learn a structured metric space that disentangles subject and task information while also capturing robust global features from the EEG signal. In the fine-tuning phase, the pre-trained STEM is adapted to specific downstream tasks. D_{self} is discarded and a classifier is appended to the output of D_{task}. The entire model is then fine-tuned end-to-end by minimizing a combination of $\mathcal{L}_{sub}$, $\mathcal{L}_{task}$, and the cross-entropy loss. This strategy ensures that STEM maintains the disentangled structure robust to inter-subject variability while optimizing for the downstream tasks.

3 Experimental Settings

Pretraining datasets. For the pre-training, we use 18 publicly available EEG datasets consisting of various tasks such as the motor imagery [5,11,15,22,23,24,33,36], emotion recognition [20,34,35], seizure detection [10], and others [16,17,19,25,26,27].

Table 1: The details of the downstream tasks.

Dataset	Task	Class	Channel	Hz	Train/Val/Test
PhysioNet-MI [21]	Motor Imagery	4	64	160	70/10/29
SHU-MI [18]	Motor Imagery	2	32	250	15/5/5
ISRUC [14]	Sleep staging	5	6	200	80/10/10

Following previous studies [30], the EEG signals are preprocessed using a 0.1–75 Hz bandpass filter and a 50/60 Hz notch filter to suppress powerline interference. Subsequently, the signals are resampled to 200 Hz. Each sample was segmented into 4-second windows, resulting in 20 patches *i.e.*, $P = 20$ and $L = 40$. We sort the channel order according to a standard montage provided by the MNE python library [12]. We scaled each sample to fall within the range of -3 to 3 μV.

Downstream tasks. For the evaluation, we use three datasets for two task-related EEG analysis [21,18] and a state-related EEG analysis [14] on cross-subject evaluation. The details are shown in Table 1. For task-related EEG analysis, we employ two motor imagery datasets such as PhysioNet-MI [21] and SHU-MI [18]. For PhysioNet-MI dataset, we first resample each sample to 200Hz. Subsequently we applied a 0.1 Hz highpass filter to remove low-frequency drift and a 60 Hz notch filter to remove power line interference. For SHU-MI dataset, we resample each sample to 200 Hz. For state-related EEG analysis, we employ a sleep staging dataset such as ISRUC [14]. We applied a 0.1–75 Hz bandpass filter and a 50 Hz notch filter to remove power line interference.

Implementation details. We employ the eight attention heads transformer-based architecture [30] for four layers encoder and two layers decoders. We set the feature dimension to 256. We sample three support samples *i.e.*, $K = 3$. The decomposer consists of a 1D convolutional layer with a kernel size of 3. It operates in a patch-wise manner, maintaining the same dimensionality as the input. We fully randomly mask the half of EEG patches for SSL. The model is trained using AdamW optimizer during 3M iterations with a cosine annealing learning rate scheduler gradually anneals the learning rate from 1e-4 to 1e-5, $\tau = 0.07$ and $B = 8$. For evaluation, we fine-tune the pre-trained model over 50 epochs using AdamW optimizer. We employ a sequential learning rate scheduler: a linear warm-up phase for the first 5 epochs to 5e-4, followed by cosine annealing to decay the learning rate to a minimum of 1e-6. The optimal model checkpoint is selected based on the highest Kappa or Precision-Recall score. We re-implement CbraMod [30], CSBrain [37] and EEG-DINO [31] by following their official codebases for a fair comparison.

4 Experimental Results

Performance comparison. Table 2 presents the performance on task-related EEG analysis using the PhysioNet-MI and SHU-MI datasets. STEM outperforms

Table 2: The comparison of task-related EEG analysis performance. The best score denoted in bold.

Method	Params (M)	PhysioNet-MI			SHU-MI		
		Acc	Kappa	F1	Acc	AUCPR	AUROC
BIOT [32]	3.2	0.6153	0.4875	0.6158	0.6179	**0.6770**	0.6609
LaBraM [13]	5.8	0.6173	0.4912	0.6177	0.6166	0.6761	0.6604
CBraMod [30]	4.0	0.6225	0.4966	0.6224	0.6184	0.6745	0.6754
CSBrain [37]	4.9	0.6117	0.4821	0.6111	0.6091	0.6720	**0.6818**
STEM	3.1	**0.6242**	**0.4989**	**0.6232**	**0.6226**	0.6753	0.6782

Table 3: The comparison of state-related EEG analysis performance. The best score denoted in bold.

Method	Params (M)	Acc	Kappa	F1
BIOT [32]	3.2	0.7527	0.7192	0.7790
LaBraM [13]	5.8	0.7633	0.7231	0.7810
CBraMod [30]	4.0	0.7673	0.7368	0.7947
CSBrain [37]	4.9	0.7633	0.7176	0.7792
EEG-DINO [31]	201	0.7725	0.7441	0.8004
STEM	3.1	**0.7863**	**0.7509**	**0.8036**

task-specific supervised learning baselines and foundation models, even with the minimum number of parameters. Specifically, STEM on PhysioNet-MI achieves the best performance on all metrics, demonstrating its superior capability in handling high inter-subject variability. Furthermore, STEM on SHU-MI attains the highest accuracy and maintains competitive performance in AUROC (2nd best) and AUCPR (3rd best) even if STEM has fewer parameters compared to other foundation models. These findings confirm that explicitly modeling subject and task semantics through our meta-learning framework yields more generalizable representations than methods relying solely on self-supervised learning.

We further evaluate STEM on state-related EEG analysis tasks using the ISRUC, as shown in Table 3. On the ISRUC dataset, STEM achieves state-of-the-art performance, outperforming the massive EEG-DINO model despite having only 3.1 M parameters. These results demonstrates that our subject- and task-aware meta-learning framework is significantly more efficient than simple model scaling, effectively capturing complex neurophysiological states. Across all datasets, the performance gaps of STEM over the baselines are statistically significant, with p-values smaller than 0.05.

Ablation study. To validate the effectiveness of our proposed components, we conducted an ablation study comparing STEM with three variants such as a reconstruction-based masked autoencoder (*w/ MAE*), models without subject-aware (*w/o Subject*), and task-aware (*w/o Task*) meta-learning. As shown in Table 4, STEM consistently achieves the highest accuracy across all three benchmarks, with statistically significant improvements over all variants at p-values

Table 4: The comparison of performance with varying training strategy. The best score denoted in bold.

Method	PhysioNet-MI			SHU-MI			ISRUC		
	Acc	Kappa	F1	Acc	AUCPR	AUROC	Acc	Kappa	F1
STEM	**0.6242**	**0.4989**	**0.6232**	**0.6226**	**0.6753**	0.6782	**0.7863**	**0.7509**	**0.8036**
w/ MAE	0.6146	0.4859	0.6134	0.6104	0.5231	0.5827	0.7707	0.7319	0.7918
w/o Subject	0.6203	0.4936	0.6170	0.6218	0.6661	**0.6936**	0.7825	0.7288	0.7917
w/o Task	0.6211	0.4967	0.6196	0.6177	0.6609	0.6844	0.7782	0.7315	0.7917

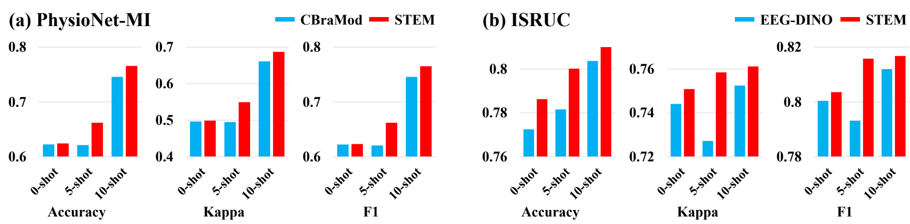

Fig. 2: Results on the PhysioNet-MI and ISRUC datasets with K' labeled EEG signals from target subject. The best baseline on each dataset *i.e.*, both CBraMod and EEG-DINO are marked as blue and STEM are marked as red.

smaller than 0.05. Specifically, STEM outperforms *w/ MAE*, demonstrating that our semantic contrastive learning yields more discriminative features than simple signal reconstruction. Furthermore, the performance drops observed in *w/o Subject* and *w/o Task* confirm that explicitly disentangling subject variability and leveraging task semantics are helpful in robust generalization. These results highlight the necessity of integrating both subject and task information to effectively handle the heterogeneity of EEG signals.

Few-shot adaptation. To evaluate the capability of STEM to personalize to new subject, we compare few-shot adaptation performance on unseen target subjects. Specifically, we first fine-tune the model on each downstream dataset using $\mathcal{L}_{sub}$ and $\mathcal{L}_{task}$ and additionally use K' support samples for each class following [2]. We predict the task label based on the cosine similarity between z_q^{task} and the prototypes obtained by averaging K' support samples for each class. As shown in Figure 2, While the best baseline on each dataset suffers from performance degradation in the 5-shot setting due to the overfitting, STEM consistently improves across all metrics as K' increases. This highlights the stability of STEM, demonstrating its reliability for personalized BCI applications.

5 Conclusion

In this paper, we propose a novel subject- and task-aware EEG foundation model that fully leverages available semantic information, overcoming the limitations of existing self-supervised pre-training paradigms. By employing a metric-based meta-learning framework, the proposed method explicitly disentangles EEG representations into subject-aware and task-aware components. This structured latent space effectively addresses the challenges of inter-subject variability and task heterogeneity in EEG datasets. Extensive experiments across diverse downstream tasks demonstrate that the proposed method achieves state-of-the-art performance and exhibits the efficiency compared to existing foundation models. Furthermore, our few-shot adaptation results highlight its reliability for personalization with minimal target data. Ultimately, STEM underscores the critical importance of actively utilizing both subject and task semantics to build EEG foundation models.

Acknowledgments. This study is supported by Smart Health Care Program funded by the Korean National Police Agency (No. RS-2022-PT000186), and the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.RS-2024-00439264, Development of High-Performance Machine Unlearning Technologies for Privacy Protection), and National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2025-00516124).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. An, S., et al.: Few-shot relation learning with attention for eeg-based motor imagery classification. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 10933–10938. IEEE (2020)
2. An, S., et al.: Dual attention relation network with fine-tuning for few-shot eeg motor imagery classification. IEEE Transactions on Neural Networks and Learning Systems **35**(11), 15479–15493 (2024)
3. An, S., et al.: Subject-adaptive transfer learning using resting state eeg signals for cross-subject eeg motor imagery classification. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 678–688. Springer (2024)
4. An, S., et al.: Subject-adaptive meta-learning for personalized bci: A fusion of resting-state eeg signal and task-specific information. Information Fusion **125**, 103501 (2026)
5. Blankertz, B., et al.: The non-invasive berlin brain–computer interface: fast acquisition of effective performance in untrained subjects. NeuroImage **37**(2), 539–550 (2007)
6. Briels, C.T., et al.: Reproducibility of eeg functional connectivity in alzheimer's disease. Alzheimer's research & therapy **12**(1), 68 (2020)
7. Chen, T., Kornblith, S., Norouzi, M., Hinton, G.: A simple framework for contrastive learning of visual representations. In: International conference on machine learning. pp. 1597–1607. PmLR (2020)
8. Chu, C.J., et al.: Eeg functional connectivity is partially predicted by underlying white matter connectivity. Neuroimage **108**, 23–33 (2015)

9. Dauwels, J., Vialatte, F., Cichocki, A.: Diagnosis of alzheimer's disease from eeg signals: where are we standing? Current Alzheimer Research **7**(6), 487–505 (2010)
10. Detti, P., Vatti, G., Zabalo Manrique de Lara, G.: Eeg synchronization analysis for seizure prediction: A study on data of noninvasive recordings. Processes **8**(7), 846 (2020)
11. Faller, J., et al.: Autocalibration and recurrent adaptation: Towards a plug and play online erd-bci. IEEE Transactions on Neural Systems and Rehabilitation Engineering **20**(3), 313–319 (2012)
12. Gramfort, A., et al.: Meg and eeg data analysis with mne-python. Frontiers in Neuroinformatics **7**, 267 (2013)
13. Jiang, W., Zhao, L., liang Lu, B.: Large brain model for learning generic representations with tremendous EEG data in BCI. In: The Twelfth International Conference on Learning Representations (2024)
14. Khalighi, S., et al.: Isruc-sleep: A comprehensive public dataset for sleep researchers. Computer methods and programs in biomedicine **124**, 180–192 (2016)
15. Liu, H., et al.: An eeg motor imagery dataset for brain computer interface in acute stroke patients. Scientific Data **11**(1), 131 (2024)
16. Liu, W., et al.: Identifying similarities and differences in emotion recognition with eeg and eye movements among chinese, german, and french people. Journal of Neural Engineering **19**(2), 026012 (2022)
17. Luciw, M.D., Jarocka, E., Edin, B.B.: Multi-channel eeg recordings during 3,936 grasp and lift trials with varying weight and friction. Scientific data **1**(1), 1–11 (2014)
18. Ma, J., et al.: A large eeg dataset for studying cross-session variability in motor imagery brain-computer interface. Scientific Data **9**(1), 531 (2022)
19. Obeid, I., Picone, J.: The temple university hospital eeg data corpus. Frontiers in neuroscience **10**, 196 (2016)
20. Savran, A., et al.: Emotion detection in the loop from brain signals and facial images. In: Proc. eINTERFACE 2006. pp. 69–80 (2006)
21. Schalk, G., et al.: Bci2000: a general-purpose brain-computer interface (bci) system. IEEE Transactions on biomedical engineering **51**(6), 1034–1043 (2004)
22. Scherer, R., et al.: Individually adapted imagery improves brain-computer interface performance in end-users with disability. PloS one **10**(5), e0123727 (2015)
23. Schirrmeister, R.T., et al.: Deep learning with convolutional neural networks for eeg decoding and visualization. Human brain mapping **38**(11), 5391–5420 (2017)
24. Steyrl, D., et al.: Random forests in non-invasive sensorimotor rhythm brain-computer interfaces: a practical and convenient non-linear classifier. Biomedical Engineering/Biomedizinische Technik **61**(1), 77–86 (2016)
25. Torkamani-Azar, M., et al.: Prediction of reaction time and vigilance variability from spatio-spectral features of resting-state eeg in a long sustained attention task. IEEE journal of biomedical and health informatics **24**(9), 2550–2558 (2020)
26. Trujillo, L.: Raw EEG Data (2020). `https://doi.org/10.18738/T8/SS2NHB`, `https://doi.org/10.18738/T8/SS2NHB`
27. Trujillo, L.T., Stanfield, C.T., Vela, R.D.: The effect of electroencephalogram (eeg) reference choice on information-theoretic measures of the complexity and integration of eeg signals. Frontiers in neuroscience **11**, 425 (2017)
28. Tzallas, A.T., Tsipouras, M.G., Fotiadis, D.I.: Epileptic seizure detection in eegs using time–frequency analysis. IEEE transactions on information technology in biomedicine **13**(5), 703–710 (2009)
29. Vaswani, A., et al.: Attention is all you need. Advances in neural information processing systems **30** (2017)

30. Wang, J., et al.: CBramod: A criss-cross brain foundation model for EEG decoding. In: The Thirteenth International Conference on Learning Representations (2025)
31. Wang, X., et al.: Eeg-dino: Learning eeg foundation models via hierarchical self-distillation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 196–205. Springer (2025)
32. Yang, C., Westover, M., Sun, J.: Biot: Biosignal transformer for cross-data learning in the wild. Advances in Neural Information Processing Systems **36**, 78240–78260 (2023)
33. Yi, W., et al.: Evaluation of eeg oscillatory patterns and cognitive process during simple and compound limb motor imagery. PloS one **9**(12), e114853 (2014)
34. Zheng, W.L., et al.: Emotionmeter: A multimodal framework for recognizing human emotions. IEEE transactions on cybernetics **49**(3), 1110–1122 (2018)
35. Zheng, W.L., Lu, B.L.: Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks. IEEE Transactions on autonomous mental development **7**(3), 162–175 (2015)
36. Zhou, B., et al.: A fully automated trial selection method for optimization of motor imagery based brain-computer interface. PloS one **11**(9), e0162657 (2016)
37. Zhou, Y., et al.: CSBrain: A cross-scale spatiotemporal brain foundation model for EEG decoding. In: The Thirty-ninth Annual Conference on Neural Information Processing Systems (2025)