S3: Spiking Neurons as an Isolating Segmenter for Brain Signal Decoding

Qian Zheng*†, Ming Chen*, Sha Zhao, Shi Gu, Peng Lin, De Ma, Huajin Tang, Gang Pan†

The State Key Lab of Brain-Machine Intelligence, Zhejiang University
College of Computer Science and Technology, Zhejiang University
{qianzheng, mchen88, szhao, gus, penglin, made, htang, gpan}@zju.edu.cn

Abstract

Recent brain decoding studies have primarily emphasized the development of brain decoders, while largely neglecting the segmentation step. Existing methods typically adopt fixed-length segmentation, which might overlook subject- or task-level variability and disrupt temporal patterns within brain signals. To address this gap, we propose **S3**, which leverages spiking neurons as an isolating segmenter for brain signal decoding. **S3** segments brain signals adaptively, considering subject- and task-level variability while preserving intrinsic temporal patterns of brain signals. It exploits the unique reset mechanism of spiking neurons to isolate previous irrelevant temporal patterns during the generation of each segmentation point. To optimize **S3** for enhancing task performance in the absence of segmentation labels, we develop an optimization method where segmentation pseudo-labels are created with a stochastic-greedy algorithm to optimize them, while circumventing gradient blockade between **S3** and task performance. Experiments on 10 downstream tasks across 13 public datasets demonstrate that **S3** consistently outperforms existing methods, validating its effectiveness, generalizability and interpretability.

Code — https://github.com/MChen808/S3

1 Introduction

Brain computer interface (BCI) is a widely-adopted technique to understand how human brains respond to real-world stimuli, which supports extensive medical research. Recently, growing attention has been directed to decoding brain signals, especially electroencephalogram (EEG), to support various downstream BCI tasks (Zheng et al. 2024; Duan et al. 2023). A typical brain decoding workflow involves four components: input data, segmentation, brain decoder and optimization (Défossez et al. 2023; Benchetrit, Banville, and King 2024). Current mainstream efforts focus on improving brain decoders, aimed at addressing subject-level variability and task-specific characteristics. (Mentzelopoulos et al. 2024; Zheng et al. 2024). An emerging direction is the EEG foundation model, which learns generalized representations across a wide range of subjects and tasks (Jiang, Zhao, and Lu 2024; Wang et al. 2025). These methods consistently exhibit superior performance across various downstream BCI tasks.

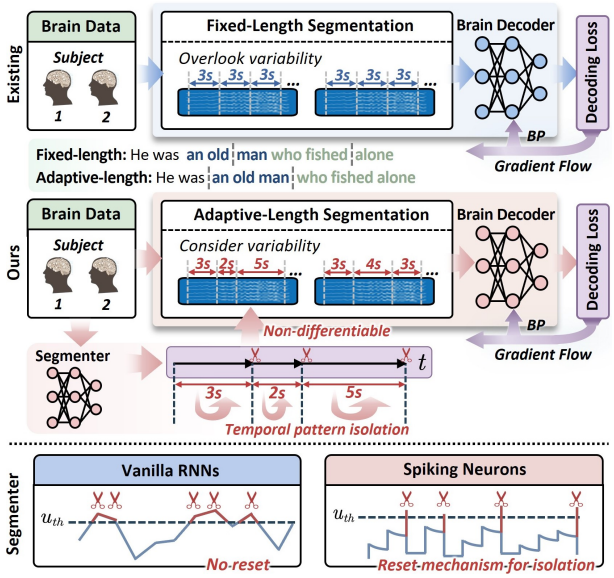

Figure 1: **Top:** Comparison of existing brain decoding workflow with ours. Fixed-length segmentation ignores variability and disrupts semantic structures, while our model adaptively segments brain signals to consider and preserve them. **Bottom:** Comparison of vanilla RNNs and spiking neurons as segmenters. Unlike RNNs, spiking neurons reset themselves at each segmentation point, isolating previous irrelevant temporal patterns.

However, these newly published works, representing the current state-of-the-art (SOTA), commonly adopt fixed-length segmentation, which poses two limitations: i) Oversight of subject- or task-level variability. Despite full considerations into subject- or task-level variability in brain decoders, they overlook it during segmentation. This oversight can lead to large performance deviations across subjects or tasks, as exemplified by An et al. (2024) and Défossez et al. (2023). ii) Disruption of temporal patterns. The existence of temporal patterns within brain signals has been proved, such

*These authors contributed equally.
†Corresponding author.

as semantic comprehension and syntactic structures (Hewitt and Manning 2019; Broderick et al. 2018). These structures often exhibit temporal continuity and length variability, potentially disrupted by rigid fixed-length segmentation.

To address these limitations, it is crucial to design an adaptive segmentation method. Some works have explored this direction but achieved limited success. For example, Song et al. (2024) and Benchetrit, Banville, and King (2024) perform a grid search to identify segments that maximize task performance. However, such fully empirical approaches offer limited interpretability and do not provide methodological guidance. Earlier studies have explored interpretable adaptive segmentation especially for EEG by leveraging autocorrelation functions to detect boundaries between distinct temporal patterns (e.g., amplitude or frequency) (Michael and Houchin 1979; Praetorius, Bodenstein, and Creutzfeldt 1977). These data-driven approaches enhance interpretability by relating the identified boundaries to intrinsic signal dynamics. However, these methods cannot be optimized for improving performance of downstream BCI tasks.

To bridge this gap, we aim to develop an adaptive segmentation method offering both superior task performance and strong interpretability. However, two challenges arise: **i) A segmenter supporting temporal pattern isolation.** Early segmentation studies have proved that non-stationary brain signals exhibit isolated stationary temporal patterns (Agarwal and Gotman 1999; Barlow et al. 1981). Based on this insight, the determination of each segmentation point should rely exclusively on the current temporal pattern, isolating the influence of previous irrelevant patterns. Achieving this requires, first, a temporal model capable of processing brain signals sequentially to generate segmentation points in order. Second, once a segmentation point is generated, all information before it must be excluded, thereby enabling temporal pattern isolation. While standard temporal models such as vanilla recurrent neural networks (RNNs) naturally support sequential processing, they fail to exclude past information. **ii) Optimization with gradient blockade.** In most brain decoding studies, labels for temporal pattern isolation are unavailable. Consequently, the segmenter should be optimized based on the decoding loss or task performance. However, in the brain decoding workflow, the segmenter performs signal segmentation in a non-differentiable manner, which blocks the gradient-based optimization through the decoding loss.

To address challenge i), we propose **S3**, which leverages **S**piking neurons as an isolating **S**egmenter for brain **S**ignal decoding. Spiking neurons sequentially accumulate input features into their membrane potentials and emit spikes when the threshold is exceeded. These spikes represent segmentation points generated in order. Importantly, after firing, these neurons reset their membrane potentials, isolating previous information during each segmentation decision. Therefore, these neurons act as an isolating segmenter, which can further comprise a spiking neural network (SNN) to be optimized for improved task performance. To address challenge ii), we propose a stochastic-greedy pseudo-label optimization to optimize the segmenter **S3**. We initially construct segmentation pseudo-labels that indicate expected segmentation points. Subsequently, we introduce a stochastic-greedy algorithm which leverages the task performance to optimize these pseudo-labels. These optimized pseudo-labels are subsequently used to supervise **S3**, thus circumventing gradient blockade between **S3** and task performance.

To sum up, our contributions are as follows:

- We propose **S3**, which leverages **spiking neurons as an isolating segmenter** for brain signal decoding. This model considers subject- and task-level variability, and preserves temporal patterns within brain signals, two aspects overlooked or disrupted by current SOTA works. Additionally, it enables temporal pattern isolation during segmentation, which standard temporal models fail in.
- We develop an **optimization** method, where segmentation pseudo-labels are created with a stochastic-greedy algorithm to optimize them for improving task performance. These optimized pseudo-labels then supervise the segmenter **S3**, thus circumventing gradient blockade between **S3** and task performance.
- We evaluate our model with SOTA brain decoders on **10 downstream BCI tasks** using **13 public datasets**. Our model presents consistent performance improvements and exhibits strong interpretability.

2 Related Work

Brain Decoder. Recent SOTA brain decoders focus on improving generalization across diverse subject- and task-specific configurations. For instance, EEG foundation models are trained on large-scale EEG datasets (Jiang, Zhao, and Lu 2024; Wang et al. 2025). Other models introduce innovative architectures, such as subject-level blocks to address individual variability (Mentzelopoulos et al. 2024). However, these studies overlook subject- or task-level variability during segmentation. In contrast, our model segments brain signals in a subject- and task-aware manner.

Fixed-length segmentation. Current brain decoding studies mostly adopt fixed-length segmentation. All subjects sharing the same segmentation method definitely ignores variability and might cause high deviations in subject-level results (An et al. 2024). Different tasks somewhat ajust segmentation lengths: longer for stable patterns and shorter for rapidly changing patterns (Khalighi et al. 2016; Brennan and Hale 2019). Yet, these length settings lack evidence of true optimality. To replace this rigid segmentation method, we propose segmenting adaptively to address subject- and task-level variability.

Adaptive-length segmentation. Traditional adaptive segmentation methods for EEG are fully data-driven, thus unable to optimize task performance (Haddad and Najafizadeh 2019; Azami, Anisheh, and Hassanpour 2014; Barlow et al. 1981). RNNs and SNNs are both reliable techniques to capture temporal patterns, which can achieve learnable adaptive segmentation on sequential data (Li et al. 2022; Adama, Lotfi, and Ranson 2021; Wang et al. 2024; Cao et al. 2024). However, the reset mechanism of SNNs uniquely enables

temporal pattern isolation, compatible with the local stationarity assumption of brain signals (Siuly et al. 2019). Motivated by this, we leverage spiking neurons to build a SNN for adaptive segmentation.

Spiking neurons for time-varying data. Spiking neurons encode temporal information of time-varying data in the discrete spikes, enabling event-driven, low-latency and temporally heterogeneous computation. This intrinsic temporal dynamics makes SNNs particularly effective for processing time-varying data (Hu et al. 2023; Hu, Tang, and Pan 2023; Gao et al. 2025; Yan et al. 2025; Gao, Jiang, and Zhang 2025) while deploying on asynchronous neuromorphic hardware (Ma et al. 2024; Xiao et al. 2025; Ming, Zonghua, and Gang 2018). These properties align naturally with the non-stationary and event-sparse characteristics of brain signals, motivating us to leverage spiking neurons for efficient and temporally precise neural data analysis.

3 Methodology

Brain Decoding Workflow

Consider brain signals denoted by $\mathbf{X} \in \mathbb{R}^{C\times T}$ with corresponding task labels $\mathbf{Y} \in \mathbb{R}^{F\times T}$, where C is the number of channels, T is the number of time points, and F is feature dimension. We define the segmentation method and brain decoder as $\mathcal{S}$ and $\mathcal{D}$, respectively. The final output contains a decoding loss $\mathcal{L}$ for optimization, and downstream task performance p as the evaluation metric:

$$\mathcal{L}(\mathcal{S},\mathcal{D}) = f_{\mathcal{L}}\big(\mathcal{D}(\mathbf{X}_{\mathcal{S}}), \mathbf{Y}_{\mathcal{S}}\big)\,, \tag{1}$$

$$p(\mathcal{S},\mathcal{D}) = f_{p}\big(\mathcal{D}(\mathbf{X}_{\mathcal{S}}), \mathbf{Y}_{\mathcal{S}}\big)\,, \tag{2}$$

where $\mathbf{X}_{\mathcal{S}}$ and $\mathbf{Y}_{\mathcal{S}}$ represent brain signals and task labels segmented by the segmentation method $\mathcal{S}$. Functions $f_{\mathcal{L}}$ and f_p are head functions based on task types. Existing works typically employ fixed-length segmentation:

$$\mathcal{S}_f = \big\{j\Delta t\big\}_{j=0}^{\lfloor T/\Delta t\rfloor}\,, \tag{3}$$

to segment the signals in a non-differentiable manner:

$$\begin{aligned}\mathbf{X}_{\mathcal{S}_f} &= \big\{\mathbf{X}_t | \mathcal{S}_f[i] < t \leqslant \mathcal{S}_f[i+1]\big\}_{i=1}^{|\mathcal{S}_f|-1}\,,\\ \mathbf{Y}_{\mathcal{S}_f} &= \big\{\mathbf{Y}_t | \mathcal{S}_f[i] < t \leqslant \mathcal{S}_f[i+1]\big\}_{i=1}^{|\mathcal{S}_f|-1}\,,\end{aligned} \tag{4}$$

where Δt denotes the manually defined segmentation length and the symbol $|\cdot|$ represents set cardinality.

In our paper, we propose $Segmenter$ that generates an adaptive segmentation method $\mathcal{S}_a$, to replace $\mathcal{S}_f$. We also develop an optimization method to optimize this $Segmenter$ for improving task performance $p(\mathcal{S},\mathcal{D})$, despite the gradient blockade between them.

S3: Spiking Neurons as an Isolating Segmenter

Here, we introduce implementation details of the $Segmenter$, which takes brain signals $\mathbf{X}$ as input and output the adaptive segmentation method $\mathcal{S}_a$, denoted as:

$$\mathcal{S}_a = Segmenter(\mathbf{X}) = STS \circ SNN \circ Event(\mathbf{X}), \tag{5}$$

which contains three components, introduced as below.

Event-based representation. To leverage spiking neurons for segmentation like Wang et al. (2024) and Cao et al. (2024), we transform the brain signals $\mathbf{X}$ into an event stream as:

$$\mathcal{E} = \{e_i\}_{i\in\mathbb{N}^+} = \{(c_i, p_i, t_i)\}_{i\in\mathbb{N}^+}\,. \tag{6}$$

An event e_i is triggered at channel $c_i \in [1, C]$ and time point $t_i \in [1, T]$ if the signal magnitude change exceeds a pre-defined threshold with polarity $p_i \in \{1, -1\}$ reflecting the changing direction. Following the mainstream research, we downsample event data into voxel-grid representations (Bardow, Davison, and Leutenegger 2016). Given the downsampling time interval t_0, we obtain the representation $\mathbf{G}$ defined as:

$$\mathbf{G}(c,p,t) = \sum_{e_i\in\mathcal{E}} \delta(c_i - c, p_i - p)\cdot \mathbb{1}_{(t\cdot t_0 - t, t\cdot t_0]}(t_i)\,. \tag{7}$$

Here, $c \in [1, C]$, $p \in \{-1, 1\}$ and $t \in [1, T_0]$, where $T_0 = \lfloor T/t_0\rfloor$. The function $\delta(x, y)$ represents the Dirac delta function, and $\mathbb{1}_A(x)$ is the indicator function. Note that the time point $t \in [1, T_0]$ in $\mathbf{G}$ does not correspond to the time point $t \in [1, T]$ in $\mathbf{X}$. The mapping between them, defined as $\theta(\cdot)$, is: $t_{\mathbf{X}} = \theta(t_{\mathbf{G}}) = (t_{\mathbf{G}} - 1) * t_0 + 1$.

Eventually, we obtain event-based representations $\mathbf{G}$ spanning the whole time axis. We denote this whole conversion process as:

$$\mathbf{G} \in \mathbb{R}^{C\times 2\times T_0} = Event(\mathbf{X})\,. \tag{8}$$

Construction of an SNN with spiking neurons. After preparing the brain signals as event-based representations, we construct an SNN with spiking neurons to process these representations and output spikes.

The Leaky-Integrate-and-Fire (LIF) model (Burkitt 2006; Hunsberger and Eliasmith 2015) is a widely used neuron model in SNNs. The discrete LIF equations are given by:

$$\begin{aligned} u_t &= v_{t-1} + \frac{1}{\tau}(I_t - v_{t-1})\,,\\ s_t &= H(u_t - u_{th})\,,\\ v_t &= u_t(1 - s_t) + u_{reset}\,. \end{aligned} \tag{9}$$

The spiking neuron temporally receives and accumulates the input current I_t into its membrane potential u_t while decaying at a constant rate τ. When reaching the threshold u_{th}, a spike is emitted and the potential is cleared and reset to u_{reset}. Therefore, the neuron emits spikes sequentially, and excludes accumulated features before each spike once one is generated. This can enforce temporal pattern isolation for the generation of each spike.

With the representations $\mathbf{G} = [\mathbf{G}_t]_{t=1}^{T_0}$ forwarded into our SNN in a multi-step manner, we firstly encode subject-level information through a per-subject 1×1 convolutional block, inspired by Défossez et al. (2023). Note that this block will remain fixed and untrained for subject-agnostic training pipelines that emit subject-level information. Then we perform stacked convolutional blocks to refine features. The refined features are finally linearly projected, and processed by a LIF node to produce outputs over T_0 time points, denoted as:

$$\{s_t\}_{t=1}^{T_0} = SNN(\mathbf{G}) \tag{10}$$

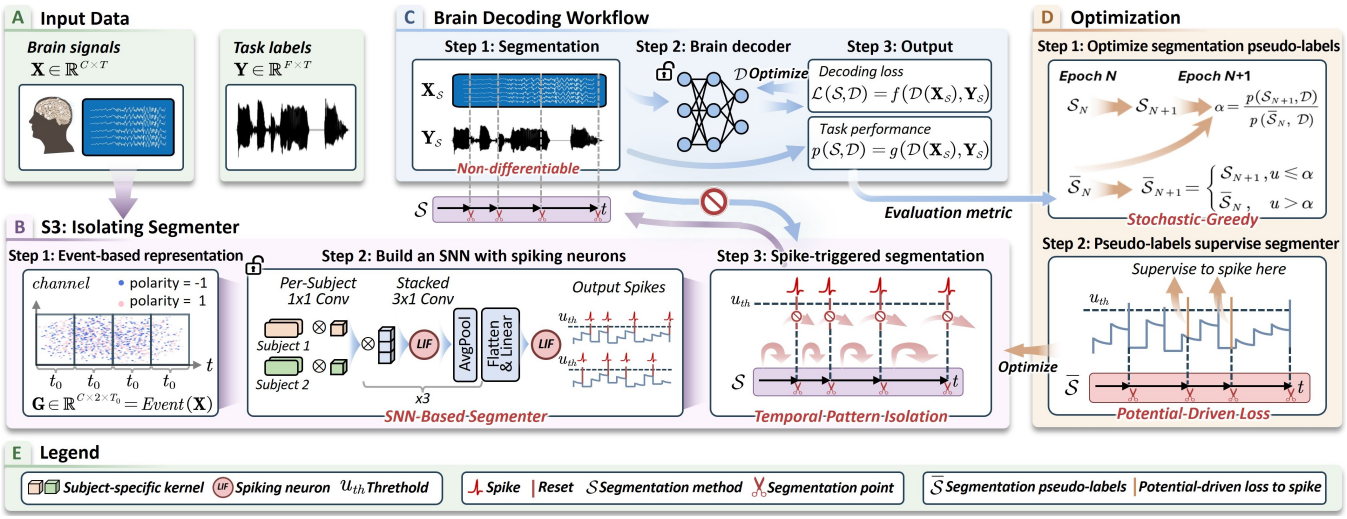

Figure 2: Overview of our method. **A)** Input data includes brain signals and task labels (e.g., audio). **B) S3**: Brain signals are converted into event-based representations and processed by an SNN to output spikes. Each spike triggers a segmentation point, formulating an adaptive segmentation method $\mathcal{S}$. The reset mechanism of spiking neurons excludes previous features, ensuring temporal pattern isolation for the generation of each segmentation point. **C)** Brain Decoding Workflow: Brain signals and task labels are segmented by $\mathcal{S}$, then used to compute a decoding loss and task performance. The decoding loss optimizes the brain decoder through backpropagation. **D)** Optimization: Pseudo-labels, representing expected segmentation points, are optimized for improving task performance via a stochastic-greedy algorithm. These pseudo-labels supervise **S3** to spike at these points.

Spike-triggered Segmentation (STS). Anytime the SNN emits a spike at time point t (i.e., $s_t = 1$), we trigger a segmentation point. Therefore, given the SNN outputs $\{s_t\}_{t=1}^{T_0}$, we collect all segmentation points based on the spiking time points. These segmentation points formulate an adaptive segmentation method, denoted as:

$$\begin{aligned}\mathcal{S}_a &= STS\big(\{s_t\}_{t=1}^{T_0}\big) \\ &= \{0\} \cup \{\theta(t) \mid s_t = 1, t \in [1, T_0]\} \cup \{T\}\,.\end{aligned} \tag{11}$$

Through this, we leverage the spikes to perform spike-triggered segmentation on brain signals. Unlike fixed-length segmentation, our approach allows the segmentation lengths within subjects to vary adaptively based on the underlying neural structures.

Stochastic-Greedy Pseudo-Label Optimization

Here, we describe how we optimize the whole framework. Our brain decoding workflow can be viewed as three steps:

$$\begin{aligned}&\text{Step 1: } \mathcal{S}_a = Segmenter(\mathbf{X}) \\ &\text{Step 2: } \mathbf{X}_{\mathcal{S}_a} = \mathcal{S}_a \circ \mathbf{X}\,, \mathbf{Y}_{\mathcal{S}_a} = \mathcal{S}_a \circ \mathbf{Y} \\ &\text{Step 3: } \mathcal{L}(\mathcal{S}_a, \mathcal{D}) = f_{\mathcal{L}}\big(\mathcal{D}(\mathbf{X}_{\mathcal{S}_a}), \mathbf{Y}_{\mathcal{S}_a}\big) \\ &\qquad\quad p(\mathcal{S}_a, \mathcal{D}) = f_p\big(\mathcal{D}(\mathbf{X}_{\mathcal{S}_a}), \mathbf{Y}_{\mathcal{S}_a}\big)\end{aligned}$$

If all steps were differentiable, the loss $\mathcal{L}$ could be backpropagated to optimize both $Segmenter$ and decoder $\mathcal{D}$. However, Step 2 is **non-differentiable**, creating a gradient blockade. As a result, only $\mathcal{D}$ can be optimized via standard backpropagation, while $Segmenter$ remains unaffected. To circumvent this gradient blockade, we introduce the stochastic-greedy pseudo-label optimization.

Pseudo-labels to supervise Segmenter. We create pseudo-labels at epoch $N \in \mathbb{N}^+$ as $\bar{\mathcal{S}}_N$ to compensate for the lack of segmentation labels. We denote the segmentation method produced by $Segmenter$ at epoch N as $\mathcal{S}_N$. These pseudo-labels, indicating expected segmentation points, supervise $Segmenter$ to spike at these time points via a potential-driven loss from Cao et al. (2024). Through this supervision, the segmentation method $\mathcal{S}_N$ produced by $Segmenter$ aligns with these pseudo-labels $\bar{\mathcal{S}}_N$. Therefore, our target becomes optimizing the pseudo-labels $\bar{\mathcal{S}}_N$ to maximize the task performance, ensuring their effectiveness. To realize it, we propose a stochastic-greedy algorithm, introduced below.

A stochastic-greedy algorithm to optimize pseudo-labels. Aimed at progressively optimizing pseudo-labels for improving task performance, we update them via a greedy algorithm, which prioritizes the more effective candidate during each update. Through this, we can push the pseudo-labels to an empirical optimum that maximizes task performance. However, a standard greedy search would probably lead to premature convergence, yielding suboptimal performance. Therefore, inspired by the Metropolis-Hastings algorithm (Metropolis et al. 1953; Hastings 1970), our approach implements a stochastic-greedy algorithm for pseudo-label updates, that is:

$$\bar{\mathcal{S}}_{N+1} = \begin{cases} \mathcal{S}_{N+1}\,, u \leqslant \alpha \text{ (accept)}\,, \\ \bar{\mathcal{S}}_N\,, \quad u > \alpha \text{ (reject)}\,, \end{cases} \tag{12}$$

where

$$\alpha = \frac{p(\mathcal{S}_{N+1}, \mathcal{D})}{p(\bar{\mathcal{S}}_N, \mathcal{D})} \quad \text{and} \quad u \sim \mathcal{U}(0, 1). \tag{13}$$

Methods	Mumtaz2016 (Mental Disorder)			MentalArithmetic (Stress)		
	Bal Acc	AUC-PR	AUROC	Bal Acc	AUC-PR	AUROC
BIOT	93.58 ± 0.52	97.36 ± 0.34	97.58 ± 0.42	68.75 ± 1.86	60.04 ± 1.95	75.36 ± 1.44
LaBraM	94.09 ± 0.79	97.98 ± 0.93	97.82 ± 0.57	69.09 ± 1.25	59.99 ± 1.55	77.21 ± 0.93
LaBraM + **S3**	95.10 ± 0.71	98.52 ± 0.77	98.63 ± 0.59	70.88 ± 1.10	63.03 ± 1.27	79.33 ± 0.83
CBraMod	95.60 ± 0.56	99.23 ± 0.32	99.21 ± 0.25	72.56 ± 1.32	62.67 ± 0.99	79.05 ± 0.73
CBraMod + **S3**	**96.89** ± 0.67	**99.24** ± 0.43	**99.23** ± 0.22	**74.89** ± 1.40	**66.45** ± 1.12	**83.03** ± 0.99
	TUAB (Abnormal)			SHU-MI (Motor Imagery)		
	Bal Acc	AUC-PR	AUROC	Bal Acc	AUC-PR	AUROC
BIOT	79.59 ± 0.57	87.92 ± 0.23	88.15 ± 0.43	61.79 ± 1.83	67.70 ± 1.19	66.09 ± 1.27
LaBraM	82.49 ± 0.25	92.21 ± 0.15	91.56 ± 0.17	61.66 ± 1.92	67.61 ± 0.83	66.04 ± 0.91
LaBraM + **S3**	82.51 ± 0.28	92.21 ± 0.14	91.77 ± 0.17	63.02 ± 1.33	69.38 ± 0.91	69.72 ± 0.80
CBraMod	82.89 ± 0.22	**92.58** ± 0.08	92.27 ± 0.11	63.70 ± 1.51	71.39 ± 0.88	69.88 ± 0.68
CBraMod + **S3**	**83.03** ± 0.40	92.28 ± 0.20	**92.34** ± 0.21	**64.35** ± 0.97	**72.36** ± 0.99	**70.17** ± 0.84

Table 1: Results for binary classification tasks (unit: %).

u is a uniform random number from 0 to 1. α is the acceptance ratio, positively correlated with segmentation effectiveness. We quantify the segmentation effectiveness through task performance $p(\mathcal{S}, \mathcal{D})$ to prioritize candidates that maximize it.

In this way, we will always accept the higher-performance candidate as new pseudo-labels (i.e., $\alpha \geqslant 1 > u$), thus achieving "greedy". Meanwhile, candidates that degrade task performance can still be accepted but only "stochastically". This stochastic-greedy algorithm exploits high-performance segmentation methods as pseudo-labels to supervise $Segmenter$, while allowing search for alternative segmentation candidates via controlled stochastic sampling to prevent premature convergence. Finally, $Segmenter$ is optimized to maximize task performance, successfully circumventing gradient blockade between them.

Iterative optimization of Segmenter and brain decoder. To optimize the full framework, we propose a two-stage iterative training algorithm that jointly optimizes $Segmenter$ and brain decoder $\mathcal{D}$. In each iteration, we first fix $\mathcal{D}$ and update $Segmenter$ to produce a segmentation method $\mathcal{S}$ that maximizes task performance $p(\mathcal{S}, \mathcal{D})$. Then, we freeze $Segmenter$ and optimize $\mathcal{D}$ using the decoding loss $\mathcal{L}(\mathcal{S}, \mathcal{D})$ via backpropagation.

4 Experiment

Experimental Setup

Downstream BCI tasks and datasets. To comprehensively evaluate the performance of our method, we select 10 downstream BCI tasks with 13 datasets. Task and dataset summaries are presented in Table 2. Note that numbers of samples are counted before optimization, and might change during optimization of segmentation. Details about datasets, head functions and evaluation metrics are displayed in Appendix C and Appendix D.

SOTA brain decoders. To evaluate the compatibility and effectiveness of **S3** with SOTA brain decoders, we select the recently proposed EEG foundation model, **CBraMod** (Wang et al. 2025), and denote our implementation upon it as **CBraMod+S3**. For Task I, we benchmark **S3** against three SOTA brain decoders which are high-performing on speech-related decoding: **BrainMagick** (Défossez et al. 2023), **Du-IN** (Zheng et al. 2024), and **DeWave** (Duan et al. 2023). We also include **Seegnificant** (Mentzelopoulos et al. 2024) which emphasizes subject-level variability. For Tasks II–X, we choose two well-known EEG foundation models, **LaBraM** (Jiang, Zhao, and Lu 2024) and **BIOT** (Yang, Westover, and Sun 2023), following the experimental protocol of Wang et al. (2025). For some tasks, we additionally implement **LaBraM+S3** to verify the generalizability of our model across different foundation models. During semantic analysis, we incorporate **U-Time** (Perslev et al. 2019), a RNN-based segmentation model for reference.

BCI Tasks	Datasets	Samples	Head
I. Speech Perception	Brennan2019	10,080	zero-shot
	Broderick2019	13,300	zero-shot
II. Sleep Staging	ISRUC	89,240	5-class
III. Emotion Recognition	FACED	10,332	9-class
	SEED-V	117,744	5-class
IV. Imagined Speech	BCIC2020-3	6,000	5-class
V. Event Type	TUEV	112,491	6-class
VI. Motor Imagery	PhysioNet-MI	9,837	4-class
	SHU-MI	11,988	binary
VII. Abnormal Detection	TUAB	409,455	binary
VIII. Mental Disorder	Mumtaz2016	7,143	binary
IX. Mental Stress	MentalArithmetic	1,707	binary
X. Vigilance Estimation	SEED-VIG	20,355	regression

Table 2: Overview of downstream BCI tasks and datasets.

Overall Performance against SOTA Brain Decoders

Superiority. We report performance for all datasets comparing against SOTA brain decoders in Table 1, Table 3 and in Appendix Table 6. Across all datasets, our model consistently outperforms these brain decoders across nearly all evaluation metrics. Notably, even when implemented with the current leading brain decoder, CBraMod, our model still yields further performance gains. In some datasets such as SEED-VIG and Broderick2019 where subject-level variabil-

Methods	Brennan2019 (Speech)		Broderick2019 (Speech)	
	Top-10@50	10@All	Top-10@50	10@All
BrainMagick	75.18 ± 2.69	26.33 ± 2.57	76.17 ± 1.53	19.66 ± 0.69
Du-IN	80.84 ± 2.55	26.93 ± 2.54	78.33 ± 1.01	20.88 ± 0.51
DeWave	76.32 ± 3.01	20.83 ± 3.12	74.80 ± 1.88	16.79 ± 0.61
Seegnificant	74.81 ± 3.32	22.79 ± 2.90	73.89 ± 1.76	19.05 ± 0.69
CBraMod	81.23 ± 2.96	22.01 ± 2.07	75.36 ± 1.34	20.89 ± 0.68
CBraMod + **S3**	**85.73** ± 2.53	**29.31** ± 1.46	**82.02** ± 0.73	**22.55** ± 0.51

Methods	SEED-VIG (vigilance estimation)		
	Correlation	R2 Score	RMSE ↓
BIOT	59.96 ± 1.82	25.43 ± 0.73	27.42 ± 0.29
LaBraM	59.31 ± 0.98	24.32 ± 0.85	27.62 ± 0.48
LaBraM + **S3**	61.38 ± 0.90	27.67 ± 0.71	27.01 ± 0.55
CBraMod	64.59 ± 0.98	33.65 ± 0.68	25.87 ± 0.39
CBraMod + **S3**	**68.01** ± 0.67	**37.53** ± 0.61	**25.10** ± 0.30

Table 3: Results for zero-shot prediction tasks (left) and regression tasks (right) (unit: %).

Datasets	Adaptive (s)	STD/Mean (%)	Fixed (s)
Brennan2019	3.1 ± 1.2	38.7*	3
Broderick2019	3.3 ± 1.6	48.5*	3
ISRUC	29.3 ± 1.0	3.4	30
FACED	31.0 ± 13.1	42.3*	30
SEED-V	1.0 ± 0.4	40.0*	1
BCIC2020-3	3.1 ± 0.3	9.7	3
TUEV	4.9 ± 2.1	42.9*	5
PhysioNet-MI	4.1 ± 1.5	36.6*	4
SHU-MI	3.9 ± 0.5	12.8	4
TUAB	9.1 ± 4.5	49.5*	10
Mumtaz2016	3.0 ± 0.2	6.7	3
MentalArithmetic	2.9 ± 0.1	3.4	3
SEED-VIG	8.2 ± 4.8	58.5*	8

Table 4: Segmentation lengths of adaptive- and fixed-length methods. The symbol * marks STD exceeding 30% of Mean, indicating aggresive searching during adaptive-length segmentation.

ity is prominent, our approach achieves significant improvements, demonstrating its ability to effectively capture and adapt to inherent subject-level variability within brain signals. For other datasets, our model maintains strong performance of CBraMod while offering different levels of enhancements.

Generalizability. In addition to CBraMod, we also implement our model on another well-recognized SOTA EEG foundation model, LaBraM, to evaluate its generalizability. We report results of LaBraM+**S3** in Table 1 and Table 3. Across all datasets, our model consistently improves performance, demonstrating its effectiveness upon LaBraM. Therefore, our method exhibits a certain degree of generalizability across SOTA EEG foundation models.

Effectiveness of Adaptive Segmentation across Tasks and Subjects

Cross-task. To analyze task-specific variability, we report the average lengths and standard deviations (STDs) of data segments in Table 4, comparing them with the fixed-length settings in Défossez et al. (2023) and Wang et al. (2025). Adaptive mean lengths are generally consistent with fixed-length segmentation. This indicates that our model inherits and exploits segmentation preferences of the brain decoder. In contrast, STDs vary significantly by task. This reflects that our model adopts different searching strategies for effective segmentation points: 8/13 tasks with STDs exceeding 30% of the mean indicate aggressive searching, while others reflects conservative searching. For example, sleep staging EEG (i.e., ISRUC) with stable patterns yields low STDs, while speech EEG (i.e., Broderick2019) with diverse patterns yields high STDs. This behavior highlights our model's ability to adapt to inter-task variability, to more effectively capture intra-task neural characteristics.

Cross-subject. To investigate subject-specific variability, we randomly select six subjects from each of three datasets. These datasets all exhibit notable subject-level variability, which has been studied by other works (Défossez et al. 2023; Pérez-Velasco et al. 2022; Hwang et al. 2021). We visualize the distributions of segmentation points and report subject-level task performance of the selected subjects in Figure 3. Our analysis reveals two key findings. First, the selected subjects within each task exhibit distinct patterns, indicating that our model effectively performs subject-aware segmentation. Second, by presenting subject-level performances, we demonstrate the benefits of adaptive segmentation. Specifically, our model achieves higher minimum, maximum, and median performance values, indicating consistent improvements across most subjects. Additionally, the significantly reduced interquartile range (IQR) suggests that we push subjects toward a more robust optimum, minimizing subject-level performance variability. This result highlights our model's capability of adaptively adjusting segmentation methods to manage inter-subject variability, aimed at robustly capturing intra-subject neural features.

High-Level Semantic Comprehension of Adaptive Segmentation

To illustrate the interpretability of our adaptive segmentation, we examine the predictive power of decoded representation segments for various levels of semantic features in the speech perception task. Following Défossez et al. (2023), we train a Ridge linear regression model to predict semantic features using the decoded representation segments generated by the brain decoder. We consider three levels of semantic features from low to high: word frequency (Speer 2022), word embedding (Honnibal and Montani 2015), sentence embedding (Schwenk and Douze 2017). Using five-fold cross-validation, we compute the Pearson correlation between the predicted and ground-truth semantic features, with results shown in Figure 4. Across both datasets, our model consistently achieves significantly higher correlations for high-level semantic features (i.e., sentence embedding) while showing little change for low-level features (i.e., word

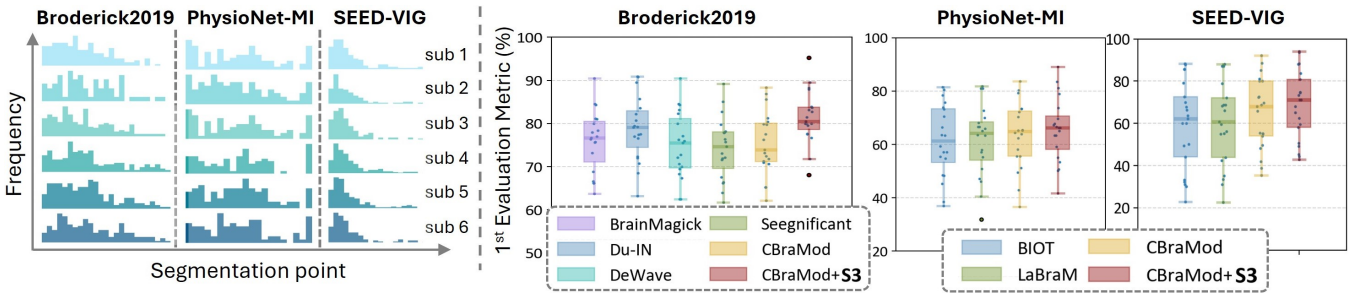

Figure 3: Adaptive segmentation across subjects. **Left:** Distribution of segmentation points for six randomly selected subjects from each of three datasets, illustrating the model's ability to implement subject-aware segmentation. **Right:** Subject-level task performance showing that adaptive segmentation achieves higher minimum, maximum, and median values, along with a notably reduced interquartile range (IQR), indicating improved robustness and performance across subjects.

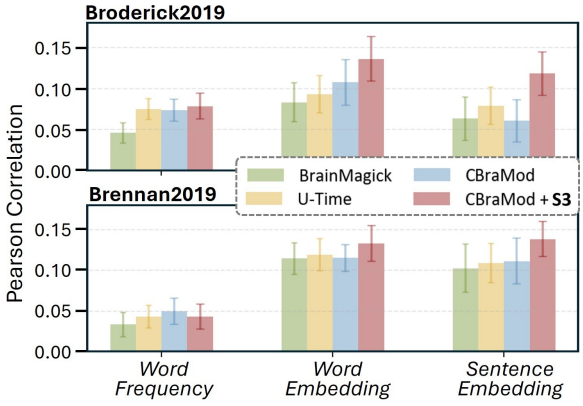

Figure 4: Predictive power of decoded representation segments for different levels of semantic features in the speech perception task. Error bars are the standard error of the mean across subjects.

Methods	Top-10@50	p value	Top-10@All	p value
SimpleCNN + **S3**	75.21 ± 1.79		12.13 ± 0.75	
w/o per-subject conv	71.73 ± 1.59	0.01	10.03 ± 0.72	<0.01
w/o stochastic-greedy	69.03 ± 1.37	0.02	8.35 ± 0.57	<0.01
w/o reset mechanism	62.12 ± 1.39	<0.01	4.73 ± 0.42	<0.01
SimpleCNN	57.66 ± 2.05	<0.01	1.79 ± 0.39	<0.01
CBraMod + **S3**	82.02 ± 0.73		22.55 ± 0.51	
w/o per-subject conv	81.03 ± 0.77	0.07	22.17 ± 0.47	0.26
w/o stochastic-greedy	79.93 ± 0.87	0.07	21.87 ± 0.53	0.37
w/o reset mechanism	77.13 ± 0.89	<0.01*	21.30 ± 0.57	0.14
CBraMod	75.36 ± 1.34	0.04*	20.89 ± 0.68	0.33

Table 5: Results for ablation analysis on Broderick2019 dataset (unit: %), with two-sample t-test showing statistical significance. The symbol * marks statistical significance (i.e., $p < 0.05$) during CBraMod ablation analysis.

frequency).

Given that sentence-level embeddings are known to capture complex semantic and syntactic structures (Caucheteux and King 2022; Hewitt and Manning 2019), the results show that our model achieves high-level semantic comprehension during segmentation without relying on any language-related information, showing strong interpretability.

Ablation Analysis

To assess the contribution of each component in our model, we perform an ablation study. We adopt two backbones: SimpleCNN and CBraMod. SimpleCNN is only composed of stacked convolutional blocks, which can exclude interference from advanced architectures of CBraMod. We analyze them on the Broderick2019 dataset, which exhibits rich subject-level variability and temporal patterns during speech listening. We systematically ablate the model in the following order: remove the per-subject convolutional block; replace the stochastic-greedy algorithm with a standard greedy algorithm; remove the reset mechanism of spiking neurons; remove the adaptive segmentation. After implementing each modified version, we perform two-sample t-tests to assess the statistical significance of performance degradation. Results are summarized in Table 5.

For SimpleCNN, all the removals lead to significant performance drops ($p < 0.05$). For CBraMod, despite consistent performance drops, only a subset of the removals (marked with *) indicate statistical significance. This is primarily due to the inherent performance contributions of CBraMod's architecture, as well as the relatively small difference in results between CBraMod + **S3** and CBraMod alone. Nevertheless, the results supports the conclusion that each component of our model contributes to improved task performance.

5 Conclusion

We propose **S3**, which leverages spiking neurons as an isolating segmenter for brain signal decoding. Here, **S3** is proposed to adaptively segment brain signals, considering subject- and task-level variability while preserving intrinsic brain temporal patterns. Extensive experimental results across various tasks verify superior effectiveness, generalizability and interpretability.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (No. 2024YDLN0005), in part by the National Natural Science Foundation of China (62376247, 62436008), in part by Fundamental Research Funds for the Central Universities (226-2025-00057), in part by the grants from Key R&D Program of Zhejiang (2022C01048), and in part by Chongqing Natural Science Foundation Innovation and Development Joint Fund (No. CSTB2025NSCQ-LZX0160).

References

Adama, D. A.; Lotfi, A.; and Ranson, R. 2021. Adaptive Segmentation and Sequence Learning of human activities from skeleton data. *Expert Systems with Applications*.

Agarwal, R.; and Gotman, J. 1999. Adaptive segmentation of electroencephalographic data using a nonlinear energy operator. In *1999 IEEE International Symposium on Circuits and Systems*.

An, S.; Kang, M.; Kim, S.; Chikontwe, P.; Shen, L.; and Park, S. H. 2024. Subject-Adaptive Transfer Learning Using Resting State EEG Signals for Cross-Subject EEG Motor Imagery Classification. In *Medical Image Computing and Computer Assisted Intervention*.

Azami, H.; Anisheh, S. M.; and Hassanpour, H. 2014. An Adaptive Automatic EEG Signal Segmentation Method Based on Generalized Likelihood Ratio. In *Artificial Intelligence and Signal Processing*.

Bardow, P.; Davison, A. J.; and Leutenegger, S. 2016. Simultaneous Optical Flow and Intensity Estimation from an Event Camera. In *Proc. of Computer Vision and Pattern Recognition*.

Barlow, J. S.; Creutzfeldt, O. D.; Michael, D.; Houchin, J.; and Epelbaum, H. 1981. Automatic adaptive segmentation of clinical EEGs. *Electroencephalography and Clinical Neurophysiology*.

Benchetrit, Y.; Banville, H.; and King, J.-R. 2024. Brain decoding: toward real-time reconstruction of visual perception. In *Proc. of International Conference on Learning Representations*.

Brennan, J. R.; and Hale, J. T. 2019. Hierarchical structure guides rapid linguistic predictions during naturalistic listening. *PLOS ONE*.

Broderick, M. P.; Anderson, A. J.; Liberto, G. M. D.; Crosse, M. J.; and Lalor, E. C. 2018. Electrophysiological Correlates of Semantic Dissimilarity Reflect the Comprehension of Natural, Narrative Speech. *Current Biology*.

Burkitt, A. N. 2006. A Review of the Integrate-and-fire Neuron Model: I. Homogeneous Synaptic Input. *Biological Cybernetics*.

Cao, J.; Sun, M.; Wang, Z.; Cheng, H.; Zhang, Q.; Zhou, S.; and Xu, R. 2024. Spiking Neural Network as Adaptive Event Stream Slicer. In *Advances in Neural Information Processing Systems*.

Caucheteux, C.; and King, J.-R. 2022. Brains and algorithms partially converge in natural language processing. *Communications Biology*.

Chen, J.; Wang, X.; Huang, C.; Hu, X.; Shen, X.; and Zhang, D. 2023. A Large Finer-grained Affective Computing EEG Dataset. *Scientific Data*.

Duan, Y.; Zhou, J.; Wang, Z.; Wang, Y.-K.; and Lin, C.-t. 2023. DeWave: Discrete Encoding of EEG Waves for EEG to Text Translation. In *Advances in Neural Information Processing Systems*.

Défossez, A.; Caucheteux, C.; Rapin, J.; Kabeli, O.; and King, J.-R. 2023. Decoding speech perception from non-invasive brain recordings. *Nature Machine Intelligence*.

Gao, R.; Jiang, C.; and Zhang, Y. 2025. Recurrent spiking neural networks as models of the entorhinal–hippocampal system for path integration: Grid cells and beyond. *Neurocomputing*.

Gao, S.; Zhu, R.; Qin, Y.; Tang, W.; and Zhou, H. 2025. Sg-snn: a self-organizing spiking neural network based on temporal information. *Cognitive Neurodynamics*.

Haddad, A. E.; and Najafizadeh, L. 2019. Source-Informed Segmentation: A Data-Driven Approach for the Temporal Segmentation of EEG. *IEEE Transactions on Biomedical Engineering*.

Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*.

Hewitt, J.; and Manning, C. D. 2019. A Structural Probe for Finding Syntax in Word Representations. In Burstein, J.; Doran, C.; and Solorio, T., eds., *Proc. of North American Chapter of the Association for Computational Linguistics: Human Language Technologies*.

Honnibal, M.; and Montani, I. 2015. spaCy: Industrial-strength Natural Language Processing in Python. Software available from https://spacy.io.

Hu, Y.; Tang, H.; and Pan, G. 2023. Spiking Deep Residual Networks. *IEEE Transactions on Neural Networks and Learning Systems*.

Hu, Y.; Zheng, Q.; Jiang, X.; and Pan, G. 2023. Fast-SNN: Fast Spiking Neural Network by Converting Quantized ANN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Hunsberger, E.; and Eliasmith, C. 2015. Spiking Deep Networks with LIF Neurons. *ArXiv preprint*. ArXiv:1510.08829.

Hwang, S.; Park, S.; Kim, D.; Lee, J.; and Byun, H. 2021. Mitigating Inter-Subject Brain Signal Variability FOR EEG-Based Driver Fatigue State Classification. *IEEE International Conference on Acoustics, Speech and Signal Processing*.

Jeong, J.-H.; Cho, J.-H.; Lee, Y.-E.; Lee, S.-H.; Shin, G.-H.; Kweon, Y.-S.; Millán, J. d. R.; Müller, K.-R.; and Lee, S.-W. 2022. 2020 International brain–computer interface competition: A review. *Frontiers in Human Neuroscience*.

Jiang, W.; Zhao, L.; and Lu, B.-l. 2024. Large Brain Model for Learning Generic Representations with Tremendous EEG Data in BCI. In *Proc. of International Conference on Learning Representations*.

Khalighi, S.; Sousa, T.; Santos, J. M.; and Nunes, U. 2016. ISRUC-Sleep: A comprehensive public dataset for

sleep researchers. *Computer Methods and Programs in Biomedicine*.

Li, J.; Li, J.; Zhu, L.; Xiang, X.; Huang, T.; and Tian, Y. 2022. Asynchronous Spatio-Temporal Memory Network for Continuous Event-Based Object Detection. *IEEE Transactions on Image Processing*.

Liu, W.; Qiu, J.-L.; Zheng, W.-L.; and Lu, B.-L. 2022. Comparing Recognition Performance and Robustness of Multimodal Deep Learning Models for Multimodal Emotion Recognition. *IEEE Transactions on Cognitive and Developmental Systems*.

Ma, D.; Jin, X.; Sun, S.; Li, Y.; Wu, X.; Hu, Y.; Yang, F.; Tang, H.; Zhu, X.; Lin, P.; and Pan, G. 2024. Darwin3: a large-scale neuromorphic chip with a novel ISA and on-chip learning. *National Science Review*.

Ma, J.; Yang, B.; Qiu, W.; Li, Y.; Gao, S.; and Xia, X. 2022. A large EEG dataset for studying cross-session variability in motor imagery brain-computer interface. *Scientific Data*.

Mentzelopoulos, G.; Chatzipantazis, E.; Ramayya, A.; Hedlund, M.; Buch, V.; Daniilidis, K.; Kording, K.; and Vitale, F. 2024. Neural decoding from stereotactic EEG: accounting for electrode variability across subjects. In *Advances in Neural Information Processing Systems*.

Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; and Teller, E. 1953. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*.

Michael, D.; and Houchin, J. 1979. Automatic EEG analysis: A segmentation procedure based on the autocorrelation function. *Electroencephalography and Clinical Neurophysiology*.

Min, J.; Wang, P.; and Hu, J. 2017. Driver fatigue detection through multiple entropy fusion analysis in an EEG-based system. *PLOS ONE*.

Ming, Z.; Zonghua, G. U.; and Gang, P. a. N. 2018. A Survey of Neuromorphic Computing Based on Spiking Neural Networks. *Chinese Journal of Electronics*.

Mumtaz, W. 2016. MDD Patients and Healthy Controls EEG Data (New). *figshare*. 10.6084/m9.figshare.4244171.v2.

Obeid, I.; and Picone, J. 2016. The Temple University Hospital EEG Data Corpus. *Frontiers in Neuroscience*.

Perslev, M.; Jensen, M.; Darkner, S.; Jennum, P. J. r.; and Igel, C. 2019. U-Time: A Fully Convolutional Network for Time Series Segmentation Applied to Sleep Staging. In *Advances in Neural Information Processing Systems*, 4415–4426.

Praetorius, H. M.; Bodenstein, G.; and Creutzfeldt, O. D. 1977. Adaptive segmentation of EEG records: a new approach to automatic EEG analysis. *Electroencephalography and Clinical Neurophysiology*, 42.

Pérez-Velasco, S.; Santamaría-Vázquez, E.; Martínez-Cagigal, V.; Marcos-Martínez, D.; and Hornero, R. 2022. EEGSym: Overcoming Inter-Subject Variability in Motor Imagery Based BCIs With Deep Learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*.

Radford, A.; Kim, J. W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; Krueger, G.; and Sutskever, I. 2021. Learning Transferable Visual Models From Natural Language Supervision. In *Proc. of International Conference on Machine Learning*.

Schalk, G.; McFarland, D.; Hinterberger, T.; Birbaumer, N.; and Wolpaw, J. 2004. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering*.

Schwenk, H.; and Douze, M. 2017. Learning Joint Multilingual Sentence Representations with Neural Machine Translation. In *Proc. of Workshop on Representation Learning for NLP*.

Siuly, S.; Bajaj, V.; Sengur, A.; and Zhang, Y. 2019. An Advanced Analysis System for Identifying Alcoholic Brain State Through EEG Signals. *International Journal of Automation and Computing*.

Song, Y.; Liu, B.; Li, X.; Shi, N.; Wang, Y.; and Gao, X. 2024. Decoding Natural Images from EEG for Object Recognition. In *Proc. of International Conference on Learning Representations*.

Speer, R. 2022. rspeer/wordfreq: v3.0. *Zenodo*. 10.5281/zenodo.7199437.

Wang, J.; Zhao, S.; Luo, Z.; Zhou, Y.; Jiang, H.; Li, S.; Li, T.; and Pan, G. 2025. CBraMod: A Criss-Cross Brain Foundation Model for EEG Decoding. In *Proc. of International Conference on Learning Representations*.

Wang, Z.; Wang, Z.; Li, H.; Qin, L.; Jiang, R.; Ma, D.; and Tang, H. 2024. EAS-SNN: End-to-End Adaptive Sampling and Representation for Event-Based Detection with Recurrent Spiking Neural Networks. In *Proc. of European Conference on Computer Vision*.

Xiao, Y.; Liu, Y.; Zhang, B.; Chen, P.; Zhu, H.; He, E.; Zhao, J.; Huo, W.; Jin, X.; Zhang, X.; Jiang, H.; Ma, D.; Zheng, Q.; Tang, H.; Lin, P.; Kong, W.; and Pan, G. 2025. Bio-plausible reconfigurable spiking neuron for neuromorphic computing. *Science Advances*.

Yan, J.; Wang, C.; Ma, D.; Tang, H.; Zheng, Q.; and Pan, G. 2025. Training High Performance Spiking Neural Network by Temporal Model Calibration. *Proc. of International Conference on Machine Learning*.

Yang, C.; Westover, M.; and Sun, J. 2023. BIOT: Biosignal Transformer for Cross-data Learning in the Wild. In *Advances in Neural Information Processing Systems*.

Zheng, H.; Wang, H.-T.; Jiang, W.-B.; Chen, Z.-T.; He, L.; Lin, P.-Y.; Wei, P.-H.; Zhao, G.-G.; and Liu, Y.-Z. 2024. Du-IN: Discrete units-guided mask modeling for decoding speech from Intracranial Neural signals. In *Advances in Neural Information Processing Systems*.

Zyma, I.; Tukaev, S.; Seleznov, I.; Kiyono, K.; Popov, A.; Chernykh, M.; and Shpenkov, O. 2019. Electroencephalograms during Mental Arithmetic Task Performance. *Data*.