Assembling the Mind's Mosaic: Towards EEG Semantic Intent Decoding

Jiahe Li[1], Junru Chen[1*], Fanqi Shen[1], Jialan Yang[1], Jada Li[2], Zhizhang Yuan[1], Baowen Cheng[3], Meng Li[3,4,5], Yang Yang[1†]*

[1]Zhejiang University, [2]University of Texas at Austin, [3]Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences

[4]Guangdong-Hong Kong-Macao Greater Bay Area Center for Brain Science and Brain-Inspired Intelligence, Southern Medical University, [5]INSIDE Institute for NeuroAI

浙江大学 计算机科学与技术学院
COLLEGE OF COMPUTER SCIENCE AND TECHNOLOGY
ZHEJIANG UNIVERSITY

The Semantic Intent Decoding Framework

Limitations of Current Approaches: Traditional Speech Decoding relies heavily on motor regions, overlooking distributed semantic processing. Concept Decoding uses fixed-label classification or direct mapping into LLMs.

The SID Solution: SID addresses the trade-off between interpretability and expressiveness by defining Semantic Intent as a coherent unit of meaning formulated as a flexible set of core semantic units:

Three Core Principles of SID

1. **Compositionality:** Human communicative intent is multidimensional. The brain constructs meaning through dynamic, context-sensitive processes rather than fixed, localized representations.

2. **Continuity & Expandability:** Semantic units are decoded within an open, continuous semantic space. This parallels vector spaces in language models and allows the BCI to generalize to novel concepts.

3. **Fidelity:** Semantic fidelity keeps outputs grounded in intended concepts, while linguistic fidelity ensures the generation of plausible sentences.

BrainMosaic Architecture for SID

Compositionality → **Semantic Decomposer** → **Stage 1**

Semantic Unit Decomposition

BrainMosaic uses a *Semantic Decomposer* to transform signals into a set of semantic components. It models signals as multiple foundational semantic units.

Continuity & Expandability → **Semantic Retriever** → **Stage 2**

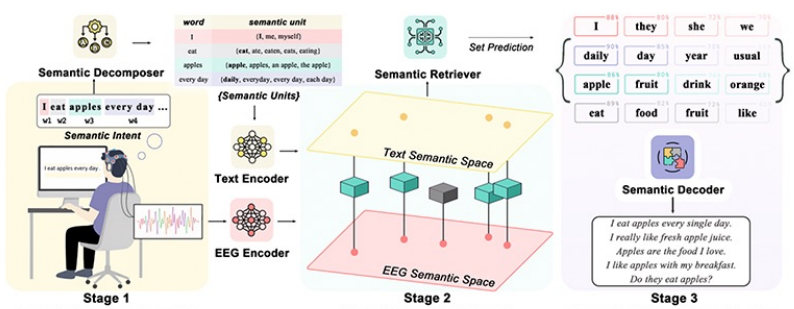
Semantic Space Alignment

Semantic Retriever bridges neural representations to an open space of conceptual representations. It aligns brain-derived units with continuous linguistic concepts.

Fidelity → **Semantic Decoder** → **Stage 3**

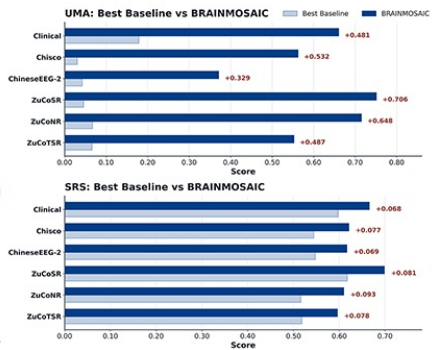
Semantic-Guided Reconstruction

Finally, the *Semantic Decoder* assembles the retrieved elements into coherent natural language. It uses LLMs to generate faithful, contextually appropriate sentences.

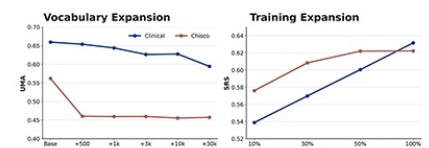


Experiments & Evaluation

- Evaluated on public multilingual EEG datasets and a private clinical Stereo-EEG dataset
- Introduced semantic-aware evaluation metrics alongside traditional NLP metrics



BrainMosaic consistently outperforms the strongest baseline across all datasets, with especially large gains in concept-level decoding accuracy.



BrainMosaic remains robust under vocabulary expansion and scales effectively with more training data, showing strong open-space generalization.