iBRAIN

Intelligent BRain Analysis through Images and Networks

Brain-Inspired fMRI-to-Text Decoding via Incremental and Wrap-Up Language Modeling

Wentao Lu[1], **Dong Nie**[2], **Pengcheng Xue**[1], **Zheng Cui**[1], **Piji Li**[1], **Daoqiang Zhang**[1], **Xuyun Wen**[1,*]

[1]College of Artificial Intelligence, Nanjing University of Aeronautics and Astronautics, Nanjing, China

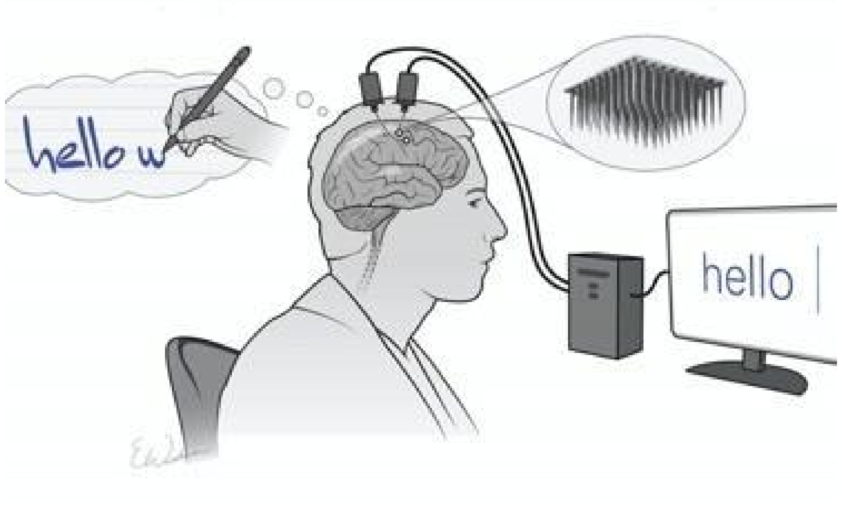
[2]ChatAlpha AI, California, USA

{luwentao,charles1231,cuizheng,pjli,dqzhang,wenxuyun}@nuaa.edu.cn

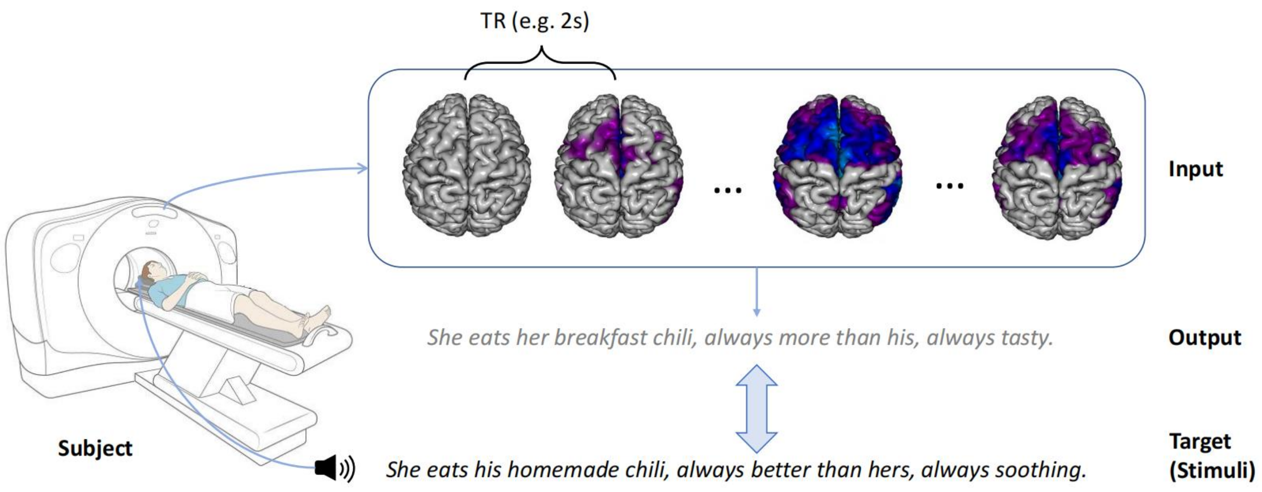
dongnie@cs.unc.edu

Background

- Brain signal-to-text decoding, which refers to **reconstructing brain signals into external linguistic stimuli** used during the signal acquisition process, is an important research direction in the field of brain-computer interface (BCI) research.
- The development of brain signal-to-text decoding can deepen our understanding of the neural system of language processing and promote the advancement of BCI.
- Currently, **fMRI** is widely used in brain signal-to-text decoding tasks due to its high spatial resolution.



Brain—to—text decoding



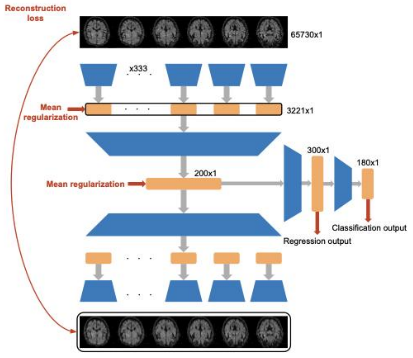
Background

Current fMRI-to-text decoding research is mainly categorized into **closed-vocabulary** and **open-vocabulary** paradigms.

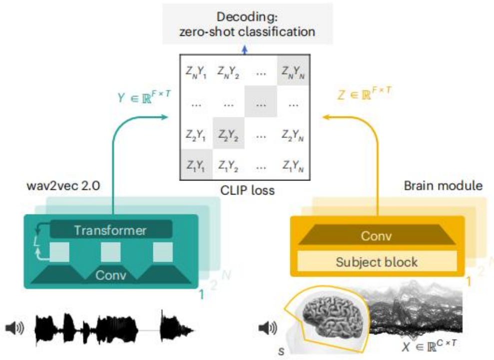
Closed-vocabulary

Early-stage fMRI-to-text decoding researches focused on closed-vocabulary sets with decoding methods including:

- **Classification models**
- **Contrastive learning**



Classification models method
(Affolter N, et al. NeurIPS, 2020..)

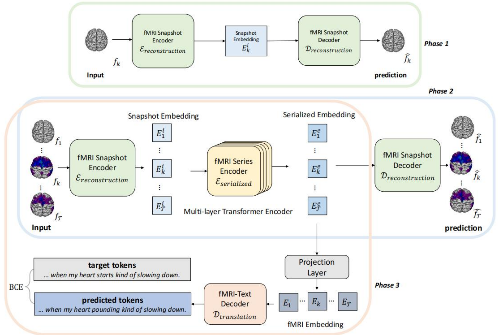


Contrastive Learning method
(Défossez A, et al. Nature Machine Intelligence, 2023.)

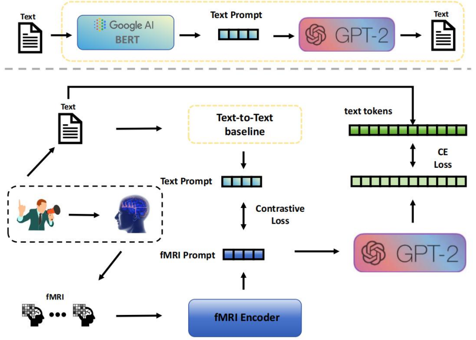
Open-vocabulary

Current fMRI-to-text decoding researches explore decoding text on open-vocabulary sets with large language models **(LLMs)**

- **UniCoRN (BART-based)**
- **BP-GPT (GPT-based)**



UniCoRN with BART
(Xi N, et al. ACL, 2023.)



BP-GPT with GPT-2
(Chen X, et al. ICASSP, 2025.)

Background

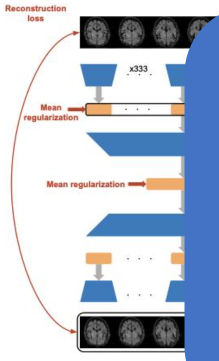
iBRAIN
Intelligent BRain Analysis through Images and Networks

Current fMRI-to-text decoding research is mainly categorized into **closed-vocabulary** and **open-vocabulary** paradigms.

Closed-vocabulary

Early-stage fMRI-to-text decoding researches focused on closed-vocabulary sets with decoding methods including:

- **Classification models**
- **Contrastive learning**



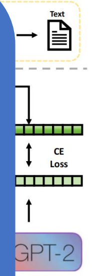
Classification m
(Affolter N, et al. NeurIPS, 2020..)

(Defossez A, et al. Nature Machine Intelligence, 2023.)

Open-vocabulary

Current fMRI-to-text decoding researches explore decoding text on open-vocabulary sets with large language models (LLMs)

- **UniCoRN (BART-based)**
- **BP-GPT (GPT-based)**



(Xi N, et al. ACL, 2023.)

(Chen X, et al. ICASSP, 2025.)

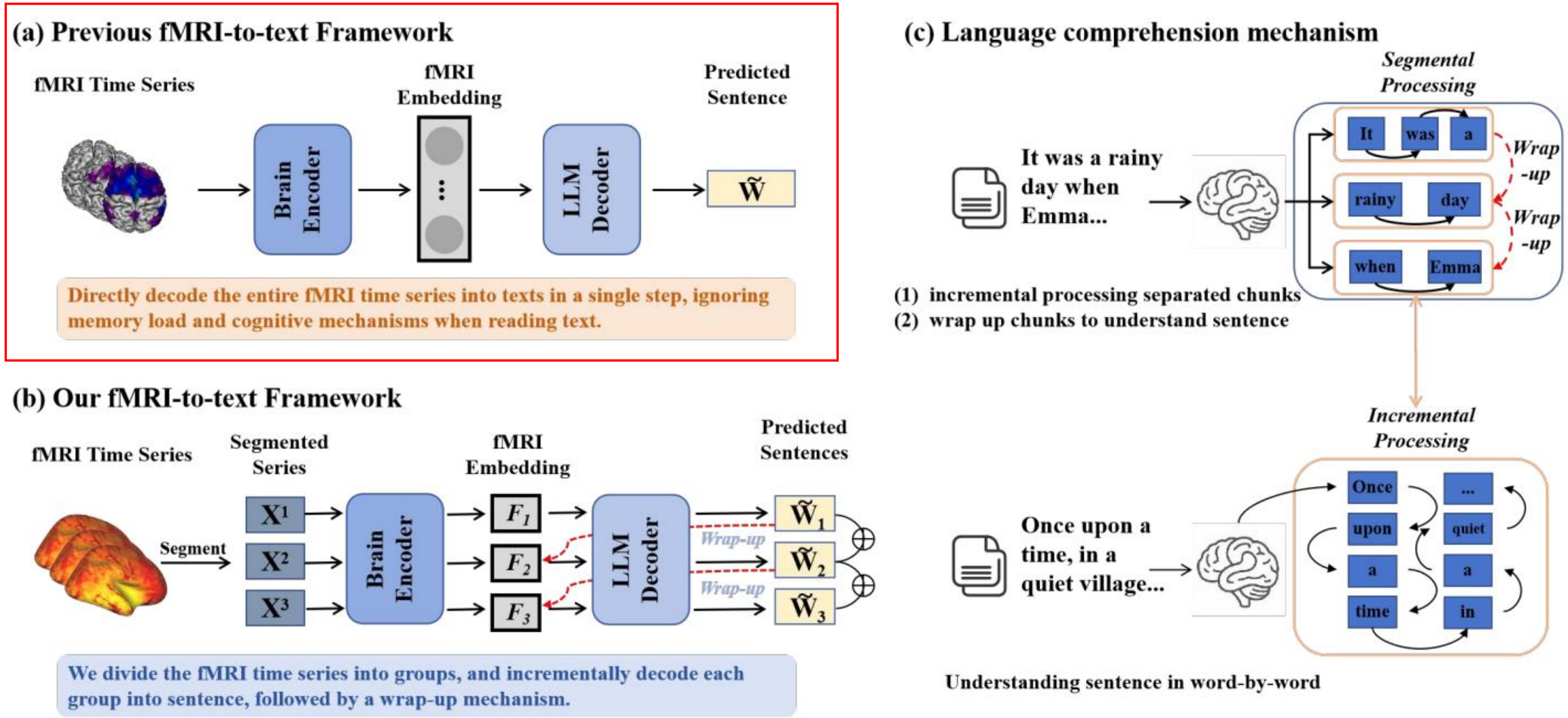
- **Early fMRI-to-text decoding methods on closed-vocabulary yield word-level output, struggle to decode sentences, and have limited application scenarios;**
- **Current fMRI-to-text decoding methods on open-vocabulary face challenges in long-form text decoding tasks.**

Contribution

Intelligent BRain Analysis through Images and Networks

- We propose a brain-inspired sequential fMRI-to-text decoding framework that mimics the human cognitive strategy of segmented and inductive language processing.

- Incremental processing enables the brain to interprete linguistic input in real time and segmental integration enables the brain to aggregate information across segments.

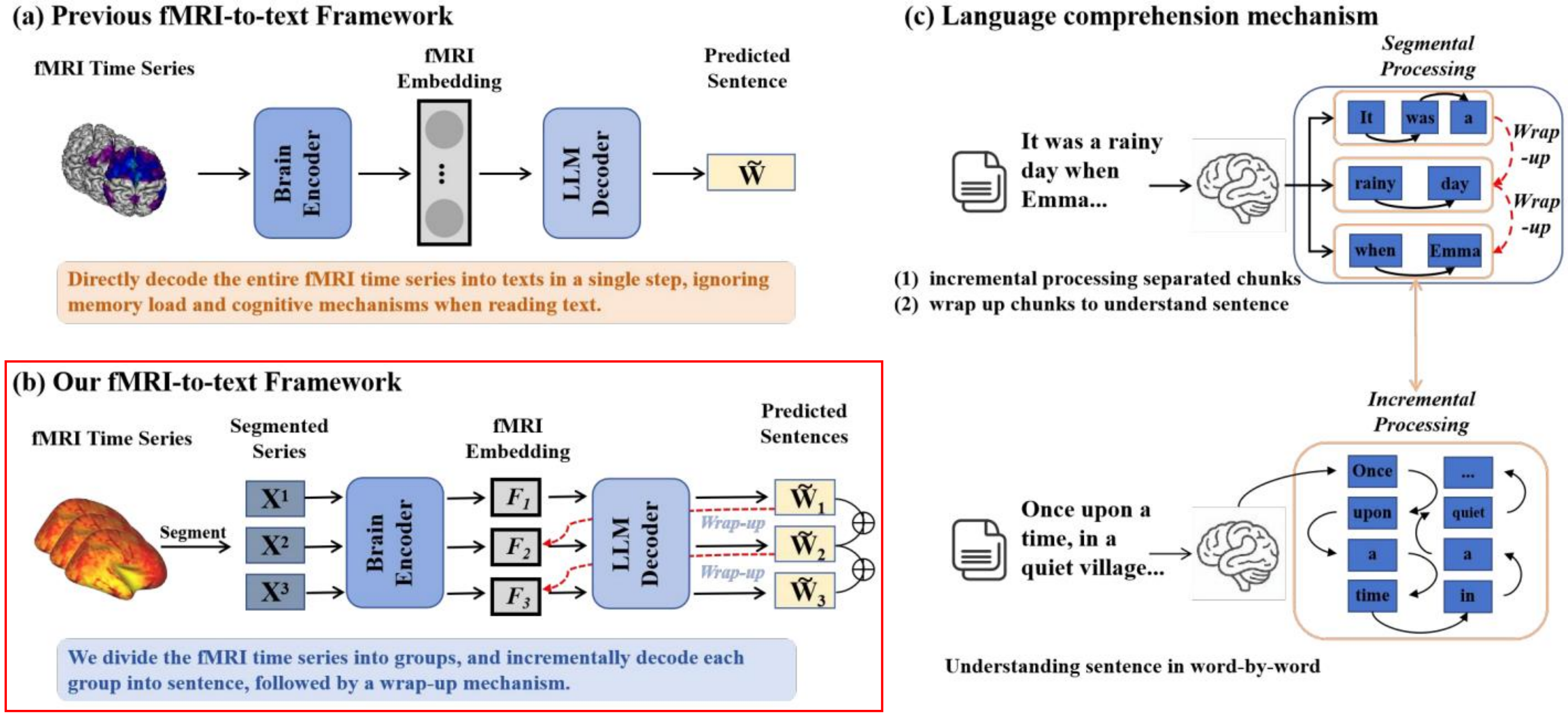


Contribution

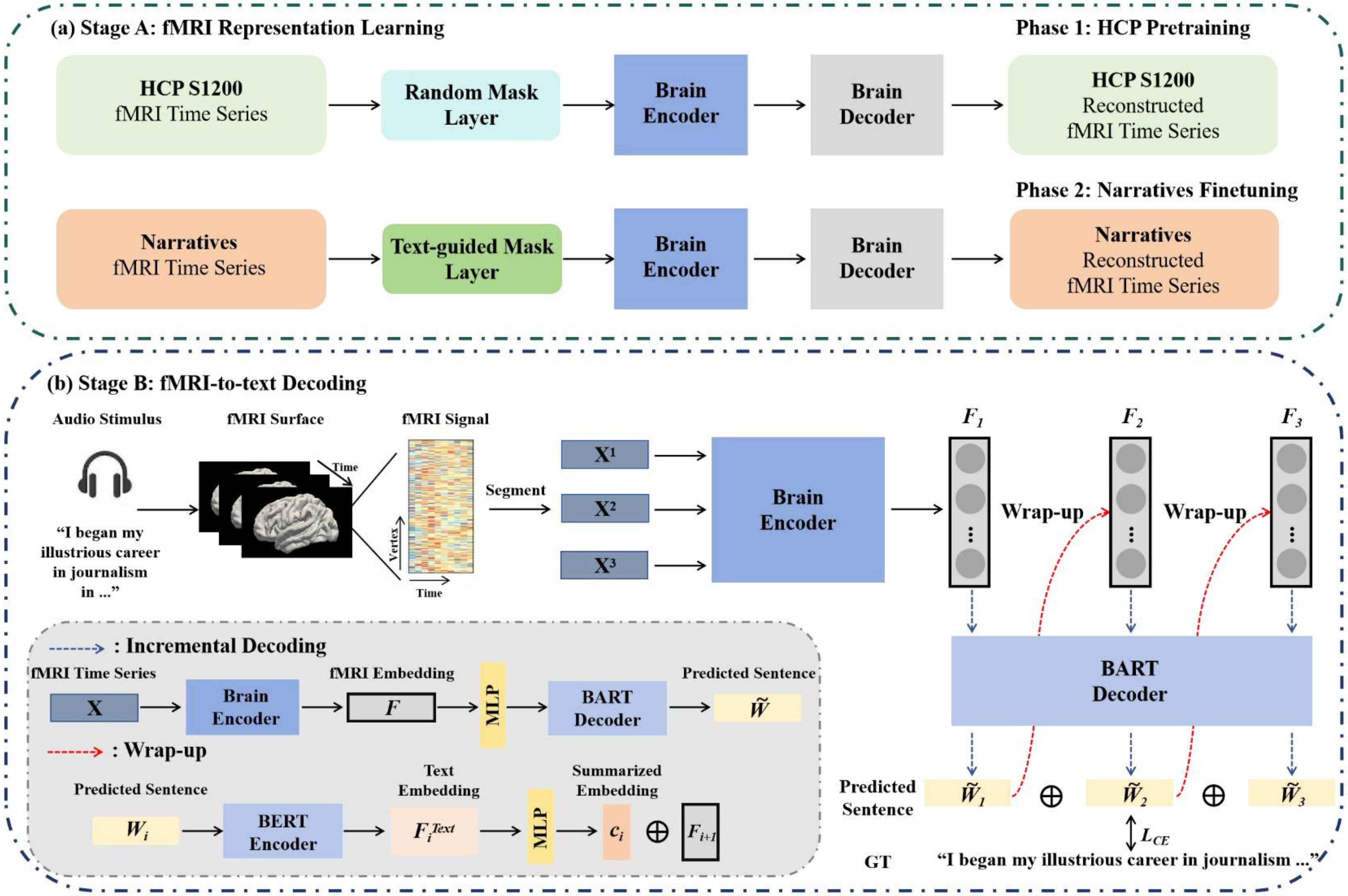
iBRAIN
Intelligent BRain Analysis through Images and Networks

- We propose a brain-inspired sequential fMRI-to-text decoding framework that mimics the human cognitive strategy of segmented and inductive language processing.

- Incremental processing enables the brain to interprete linguistic input in real time and segmental integration enables the brain to aggregate information across segments.



Method: Overview



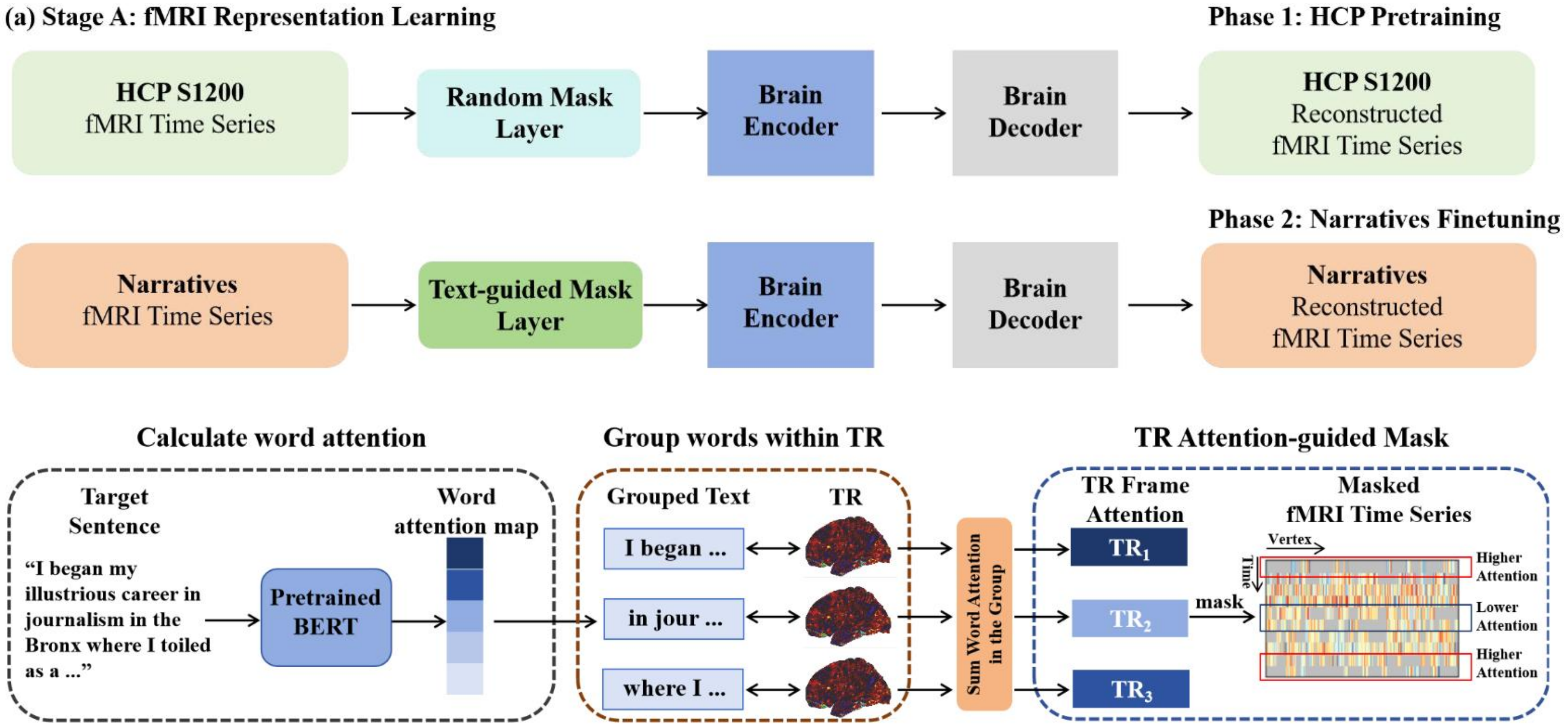
Framework of CogReader

Stage A: fMRI Representation Learning: Using Two-Stage Training Strategy to pre-train the Brain Encoder to extract better fMRI representation.

Stage B: fMRI-to-text Decoding: Decode corresponding text from fMRI representation via the decoding model BART.

Method: fMRI Representation Learning

- **Two-Stage Training Strategy:** Address the issue of fMRI-Text paired data scarcity and enhance the stability and semantic information of learned fMRI representations;
- **Text-guided Masking:** Prompt the model to focus on neural activity at key time points to learn brain representations with more key textual information.



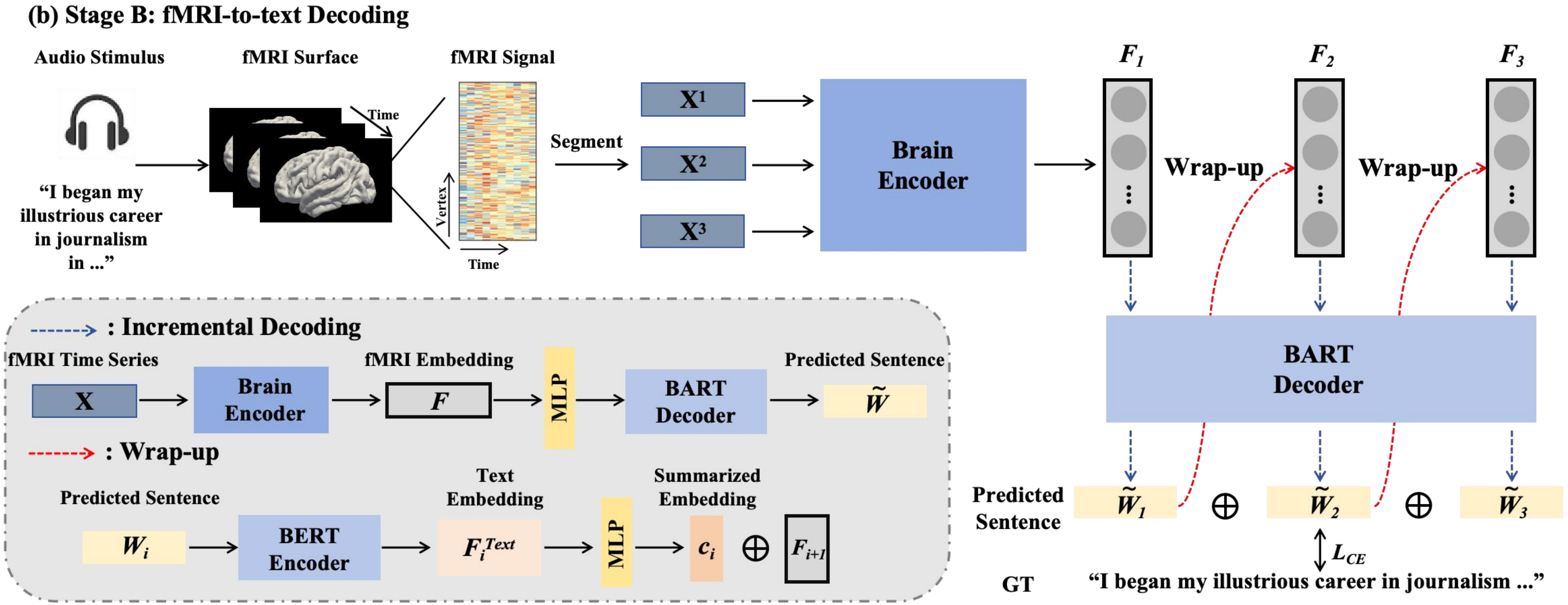
Two-stage Training Strategy

Text-guided Masking

Method: fMRI-to-text Decoding

- **Incremental Decoding within fMRI Segments:** Directly generate the associated text from the input fMRI segments.

- **Semantic Wrap-Up across fMRI Segments:** Address the potential semantic discontinuity across segments in decoded text.



Experiments

iBRAIN
Intelligent BRain Analysis through Images and Networks

Dataset:
We use two datasets to validate the effectiveness of our proposed fMRI-to-text decoding framework **CogReader**, including the **HCP S1200** dataset used in the pretraining phase, and the **Narratives** dataset used in the fine-tuning phase and decoding stage.

Experiments:

- **Comparison with State-of-the-art methods:** Compare with four SOTA methods, including UniCoRN, EEG-Text, BP-GPT and PREDFT in different input length.
- **Ablation Study:** Ablate Pretraining, Text-guided Masking, and Sequential Decoding to validate the effectiveness of each module.
- **Comparison with other fMRI Representation Learning Method:** Compare with UniCoRN to validate the effectiveness of the representation learning stage.
- **Comparison with Noise Data:** Compare with noise input to validate the effectiveness of our model.

Results: Compare with SOTA Methods

iBRAIN
Intelligent BRain Analysis through Images and Networks

Quantitative Comparison: Our method consistently outperforms all SOTA methods across all input lengths and evaluation metrics, especially in decoding long-form text sequences.

Length	Method	BLEU-N(%)				ROUGE-1(%)			BERTScore(%)		
		BLEU-1	BLEU-2	BLEU-3	BLEU-4	ROUGE-F	ROUGE-P	ROUGE-R	BERTScore-F	BERTScore-P	BERTScore-R
20TR	UniCoRN	22.9	2.5	0.3	0	20.3	19.6	21	43.9	44.2	42.8
	EEG-Text	24.6	9.3	4.4	1.9	21.9	21.1	23.4	44.6	43.9	45.4
	BP-GPT	21.6	3.8	2.5	1.7	21.6	20.9	23.4	44.1	42.1	46.3
	PREDFT	24.3	4.2	0.7	0.1	20.1	22.3	18.3	45.9	45.5	46.7
	CogReader(ours)	**25.4**	**10.5**	**4.7**	**2.6**	**23.4**	**22.6**	**24.6**	**46.3**	**45.7**	**46.9**
40TR	UniCoRN	19.1	2.3	0.5	0.1	17.8	18.2	17.6	43.8	44.8	42
	EEG-Text	20.1	7.3	3	1.3	24.4	25.1	24.7	45.4	45.8	45.5
	BP-GPT	19.9	3.6	2.3	1.5	21.1	19.3	22.9	42.6	39.4	46.1
	PREDFT	25.9	4.8	1.4	0.4	21.1	24.8	18.6	46.3	46.2	46.8
	CogReader(ours)	**31.2**	**15.3**	**10.3**	**8.2**	**29.6**	**28.7**	**30.4**	**50**	**49.3**	**51.1**
60TR	UniCoRN	18	1.7	0.2	0.4	16.5	15.9	17	43.2	43.7	42.7
	EEG-Text	22.1	8.2	3.4	1.6	28.1	29.4	28.1	47.7	47.8	47.7
	BP-GPT	19.3	3.4	1.3	0.6	19.4	19.6	19.3	41.6	38.2	45.3
	PREDFT	26.4	6.1	1.9	0.6	28.1	25.5	20.5	48.1	47.7	48.5
	CogReader(ours)	**36.2**	**20.4**	**14.7**	**12.1**	**36.2**	**35.6**	**37.2**	**53.5**	**52.6**	**54.5**

Results：Compare with SOTA Methods

iBRAIN
Intelligent BRain Analysis
through Images and Networks

Decoding Text Comparison: Our method consistently outperforms SOTA methods in both semantic information extraction and narrative consistency, such as Pie Man, feld, said and Ego sum non an bestia.

Results
Target: *was* **I** gave **the** figure a name **I** called **him** **Pie Man** capital P capital M **And I** described **him as** a cape wearing masked avenger Though **in** fact **he**'d **been** capeless **And** **said** **that as he** **fled** **the** **scene** **he** clicked **his** heels **in** rakish glee **And I** gave **him** a catchphrase **in** Latin **I** **said** **that he** ***cried*** **out** **Ego sum non an bestia** **which** means **I am** *not* an animal **Which** ***makes*** no sense
UniCoRN: **and** me gray to **the** little **Man and** no taller than my knee, with a coat too big for **his** frame lingered by **the** gate. It's **been** nigh of years **he** to **the** then **said** a case here. **The** leather sat weathered on **the** step, its lock rusted shut, but **I** knew what **was** inside: **the** girl of **the** story, **the** one who scribbled tales of dragons **in the** margins of her notebooks, had tucked her most precious drawings there before she
EEG-Text: **and** this is where it gets really interesting **I**'*m* trying to think of a title for **the** book **and I** finally come up with **an** idea **that** sounds like something **out** of one of those old pulp stories by John Wilmot who lived **in** Waukegan Illinois but **he was** actually born **in** Westchester County New York **which** is about twenty five miles outside Chicago So **I am** going to describe my novel **as The Man** Who Walks into Times Square **and** stops before **he** can get too lost
BP-GPT: **I**'*m* a teacher **and I** teach people how to make things better by using my own talents **And** so one day **I was** walking down **the** street **and** this old lady came up to me **and** she **said** You know you sound like **an** intelligent black woman So **I** told her **that**'s exactly what **I**'*m* here to do
PREDFT: **the** a **said the** girl a **man I said the Man and** use of best **and** she me **and I man** fl of Litgo be **and the** It's flirting **and** she guy of **the** eyes thelee raised **non** crate **and** she littleiving it then to of crate **non** a best owan is best me best to
Ours: **I** think **that** you realize what happened next **Pie Man** emerged from **the** late night library drop ***made*** **his** delivery **and** **fled** away ***crying*** **Ego sum non an bestia** Or that'*s* what it **said** **in** my story **in the** newspaper next day **which** ran with photos of **him** leaving **the** **scene** cape flowing behind **him** doing this **And I'm** just like praying my life doesn'*t* flash before my eyes **and** ruins

Results: Ablation Study

- The **ablation results** show a consistent improvement in performance as each module is added, validating the effectiveness of each module.

- The brain-inspired sequential decoding framework yields the most significant performance gain, demonstrating the impact of our proposed decoding approach.

Method			BLEU-N(%)				ROUGE-1(%)			BERTScore(%)		
Sequential Decoding	Pretraining	Text-guided Masking	BLEU-1	BLEU-2	BLEU-3	BLEU-4	ROUGE-F	ROUGE-P	ROUGE-R	BERTScore-F	BERTScore-P	BERTScore-R
✗	✗	✗	17.7	6.5	2.4	1.1	29.2	32.9	24.7	46.7	47.6	45
✓	✗	✗	32.5	16.5	11.1	8.9	28.2	25.8	30.6	51.1	50.3	51.8
✓	✓	✗	34.0	18.1	12.7	10.2	34.1	33.7	35.7	52.3	51.0	53.7
✗	✓	✓	21.6	7.9	3.2	1.5	26.6	29.4	25	47.4	47.7	47.2
✓	✓	✓	**36.2**	**20.4**	**14.7**	**12.1**	**36.2**	**35.6**	**37.2**	**53.5**	**52.6**	**54.5**

Results: Compare with UniCoRN

iBRAIN
Intelligent BRain Analysis through Images and Networks

Our method consistently improves decoding performance across all sequence lengths, validating the effectiveness of the proposed representation learning framework.

Length	Method	BLEU-N(%)				ROUGE-1(%)			BERTScore(%)		
		BLEU-1	BLEU-2	BLEU-3	BLEU-4	ROUGE-F	ROUGE-P	ROUGE-R	BERTScore-F	BERTScore-P	BERTScore-R
10TR	UniCoRN	18.1	2.9	0.4	0	10.5	10.2	16.6	40.2	40.1	40.4
	Ours	20.6	7	2.8	1.3	17.1	16.2	18.3	41.1	40.5	41.8
20TR	UniCoRN	22.9	2.5	0.3	0	20.3	19.6	21	43.9	44.2	42.8
	Ours	25.4	10.5	4.7	2.6	23.4	22.6	24.6	46.3	45.7	46.9
30TR	UniCoRN	20.3	2.8	0.5	0.1	18.3	18.3	18.4	41.4	41.5	41.4
	Ours	24.2	9.1	3.9	1.8	25.1	26.2	24.8	47	47.1	46.8
40TR	UniCoRN	19.1	2.3	0.5	0.1	17.8	18.2	17.6	43.8	44.8	42
	Ours	21.6	7.9	3.2	1.5	25.2	27	24.4	46.1	46.2	45.9
50TR	UniCoRN	18.9	1.9	1.8	1.1	17.3	16.8	17.4	44.8	43.9	45.7
	Ours	21	7.7	3.2	1.5	26.1	29.4	24	46.5	47.7	45.8

Results: Compare with Noise

iBRAIN
Intelligent BRain Analysis through Images and Networks

The results show that decoding performance is still higher when using real fMRI data, providing strong evidence that our proposed method is capable of extracting meaningful semantic information from fMRI time series, rather than depending solely on the memorization ability of the LLM.

Data		BLEU-N(%)				ROUGE-1(%)			BERTScore(%)		
Train	Test	BLEU-1	BLEU-2	BLEU-3	BLEU-4	ROUGE-F	ROUGE-P	ROUGE-R	BERTScore-F	BERTScore-P	BERTScore-R
Noise	Noise	27.5	9.4	4.6	1.8	25.6	26.1	25.5	48.2	48.0	48.4
Noise	fMRI	25.3	7.2	2.4	1.4	23.9	23.5	24.8	47.7	47.2	48.3
fMRI	Noise	26.8	7.7	2.7	1.2	23.9	23.1	24.9	47.9	47.4	48.5
fMRI	fMRI	**36.2**	**20.4**	**14.7**	**12.1**	**36.2**	**35.6**	**37.2**	**53.5**	**52.6**	**54.5**

Thank you！

Feel free to contact us if you have any question: ✉ wenxuyun@nuaa.edu.cn