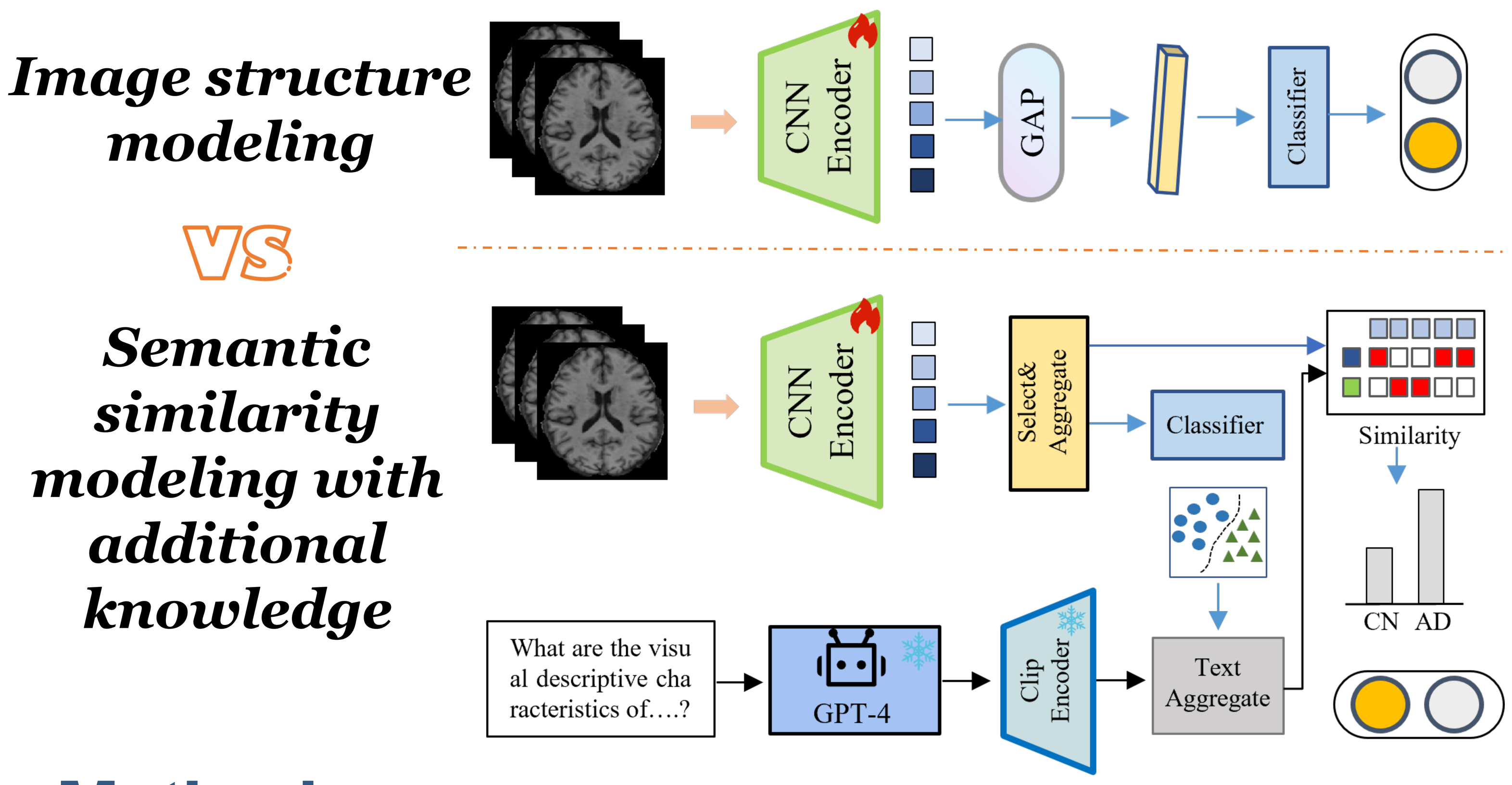
Prior-guided Prototype Aggregation Learning for Alzheimer's Disease Diagnosis

Yueqin Diao[1], Huihui Fang[2], Hanyi Yu[1], Yuning Wang[1], Yaling Tao[1], Ziyan Huang[1], Si Yong Yeo[3] , Yanwu Xu[1]

[1] School of Future Technology, South China University of Technology
[2] College of Computing and Data Science, Nanyang Technological University
[3] Lee Kong China School of Medicine, Nanyang Technological University

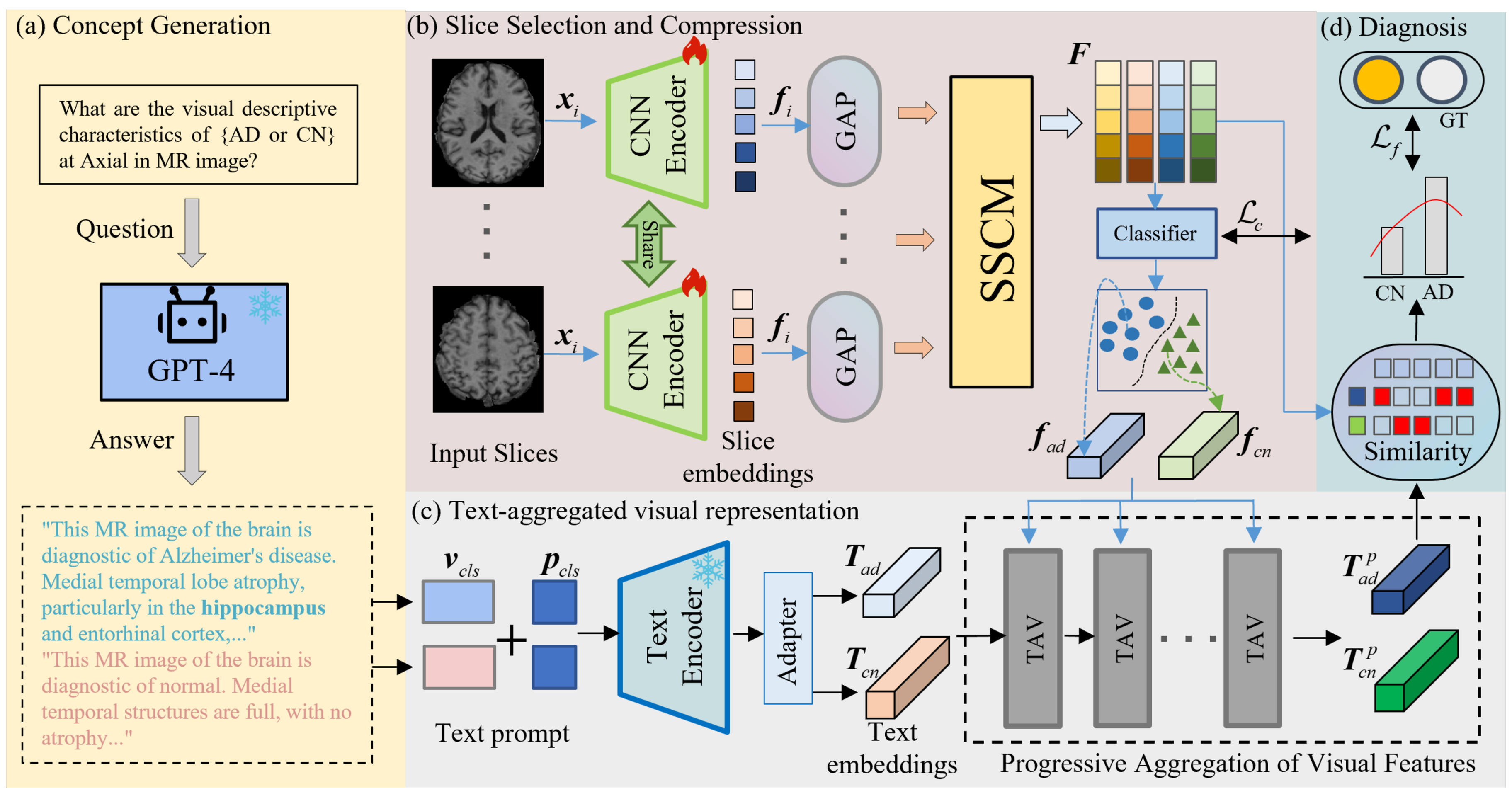
Introduction



1. Current deep learning methods primarily focus on structural magnetic resonance imaging analysis, often overlooking the **critical disease concepts** that clinicians rely on.
2. The **domain gap** between natural and medical images limits the effectiveness of CLIP-based methods.

✓ By integrating medical prior knowledge generated through LLMs with imaging features via **semantic similarity computation**, our framework enhances both diagnostic performance and interpretability in AD analysis.

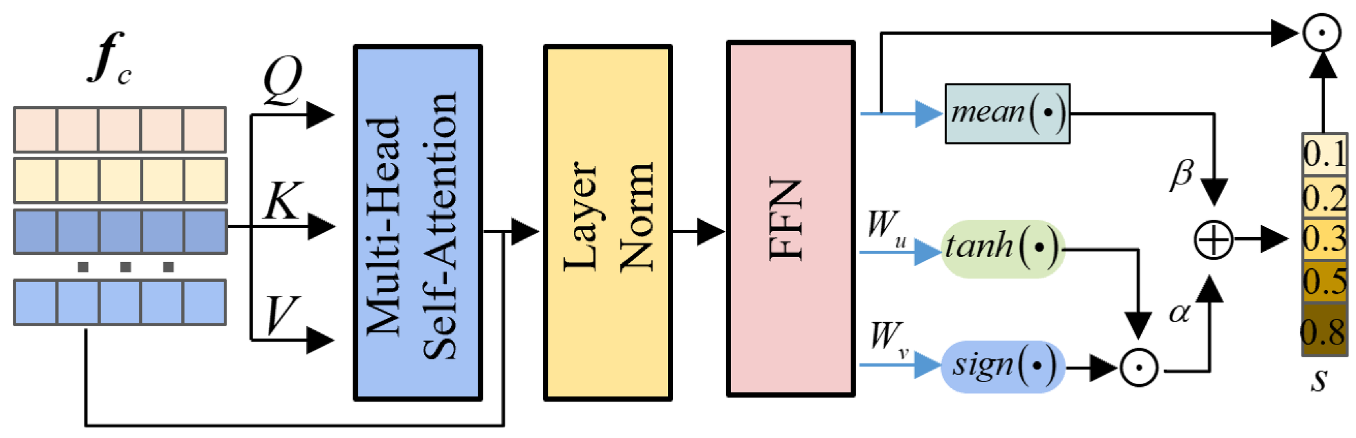
Method



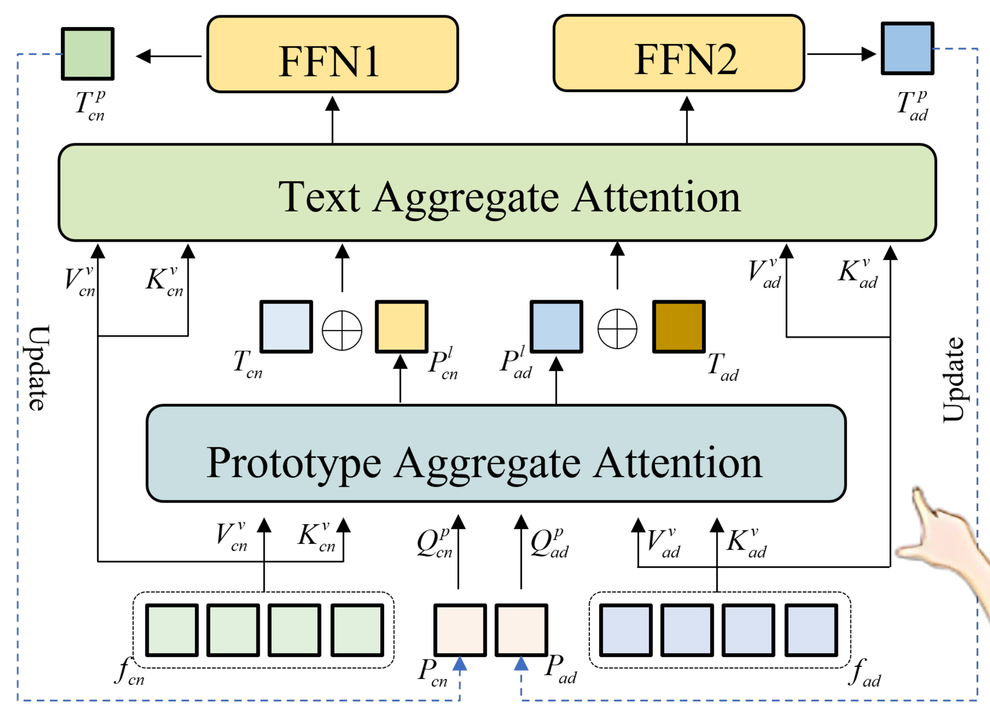
Pipeline

- We first use a LLM to extract disease-related concepts through a QA approach.
- The Slice Selection and Compression network assigns importance scores to MRI slices, selecting and aggregating key slices.
- The Text-Aggregated Visual representation network iteratively fuse visual features, reducing semantic gaps and enriching textual descriptions.

Contributions



Slice selection & compression module.



Text-aggregated visual module.

- By introducing contextual information to dynamically optimize text prompt expressions, the semantic alignment between visual and textual modalities is enhanced.

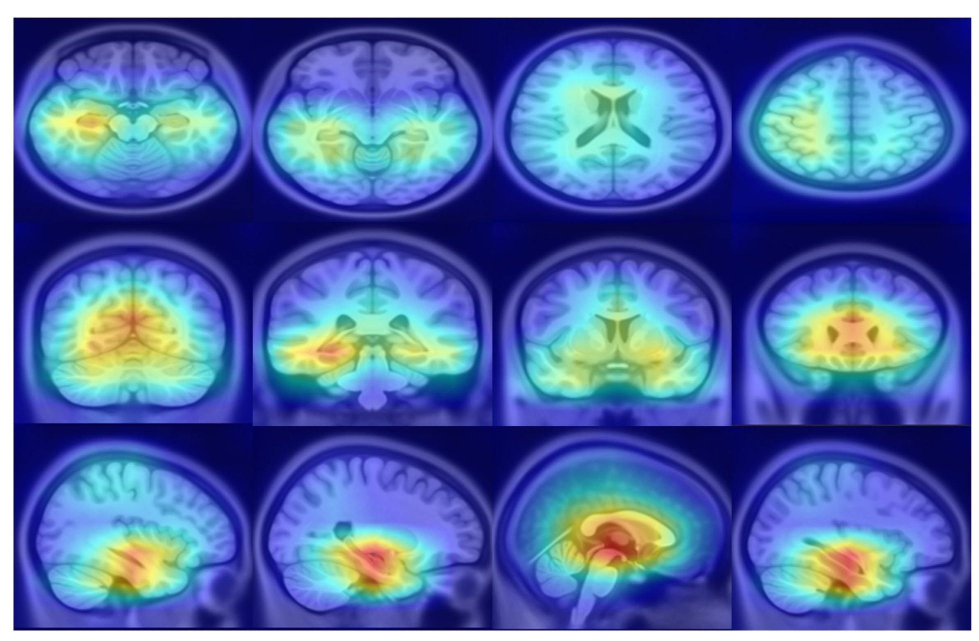
Results

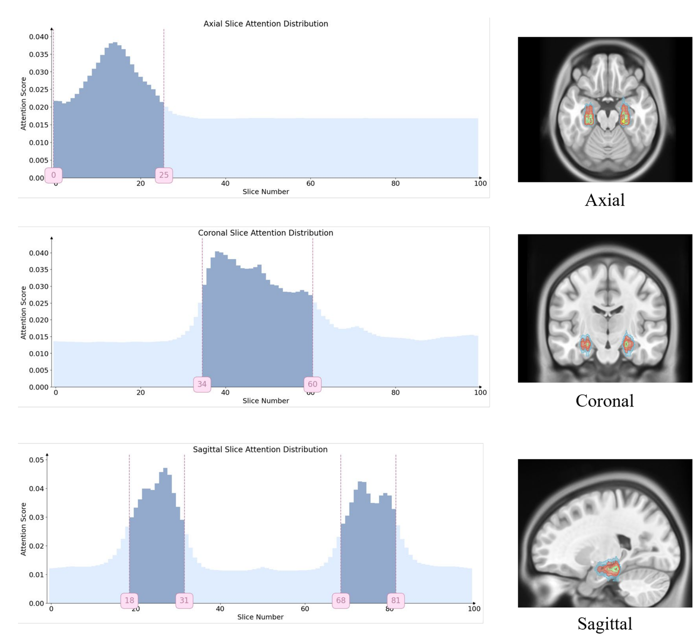
Table 1. Results of the ADNI dataset for AD diagnosis and MCI conversion prediction tasks. The gray area represents the results under zero-shot classification.

Networks	AD vs CN				sMCI vs pMCI			
	ACC	SPE	SEN	MCC	ACC	SPE	SEN	MCC
Att-Transformer	82.6	91.4	71.7	65.1	62.3	66.5	48.73	23.8
AwareNet	83.32	87.5	77.8	65.9	48.41	77.4	25.8	3.9
Majority Voting	80.4	89.7	68.9	60.5	61.4	60.1	62.9	22.9
CLIP	78.44	75.31	80.98	56.36	63.72	58.9	68.83	27.83
ViLa-MIL Low	79.19	66.53	89.94	58.16	63.09	56.44	70.13	26.79
CoOP	82.86	79.29	84.72	64.13	62.46	60.74	64.29	25.02
AXIAL	83.22	75.73	89.3	66.06	64.67	50.3	79.87	31.49
Ours	85.38	82.42	87.78	70.39	67.19	65.64	68.83	34.47

Table 2. Top 5 ROI.

Brain Area	V_r	P_r
Hippocampus (right)	1571	0.3348
Hippocampus (left)	1536	0.3272
Parahippocampal (left)	730	0.2696
Parahippocampal (right)	532	0.1965
Amygdala (left)	332	0.2012





Visualization results.

- The experimental results and visualizations demonstrate that PPAL, can effectively focus on key brain regions associated with AD. The findings highlight that vision-language models exhibit strong generalization and zero-shot capabilities under the semantic similarity paradigm .

Let's connect!

GitHub: https://github.com/diaoyq121/PPAL
Contact Information:
Yueqin Diao Email: dyq0336@163.com