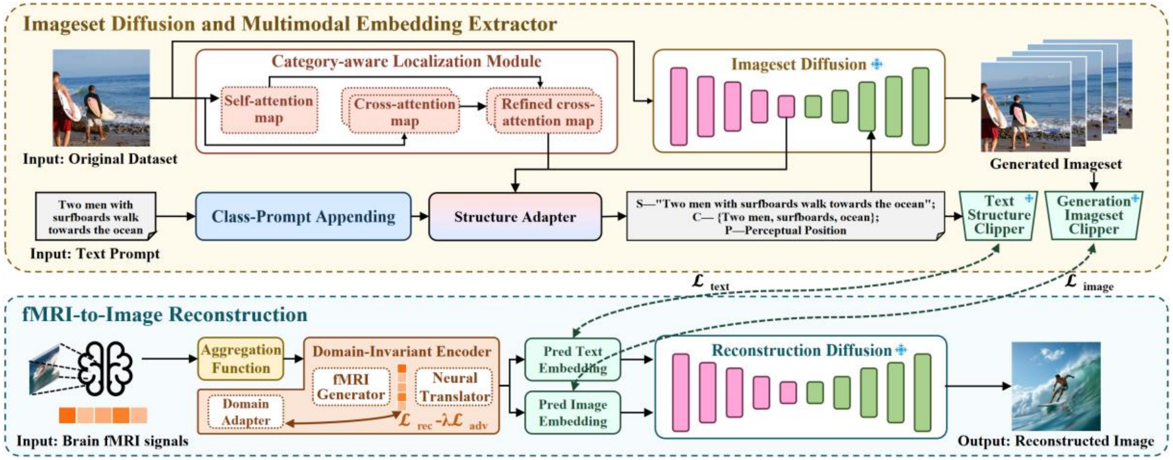
BrainCanvas: Cross-Subject Brain Decoding with Imageset Diffusion and Domain Invariance

Xin ZHOU, Tianyang DONG, Yiqiang LIAO, Wenyuan YING, Wentao DAI, Jing FAN

Frontiers of Computer Science, DOI: 10.1007/s11704-026-51973-5

Problems & Ideas

- Problems of existing fMRI-based brain decoding:
  - fMRI data is extremely scarce, and traditional augmentation distorts high-dimensional brain signals.
  - Existing models are subject-specific and fail to generalize across individuals due to inter-subject variability.
- Ideas: A unified framework (BrainCanvas) that generates semantically diverse training imagesets via diffusion models to overcome data scarcity, and learns subject-invariant representations through adversarial training to enable cross-subject brain decoding.



Schematic Diagram of BrainCanvas. Upper Section: Augments the original imageset using the Category-aware Localization Module and Class-Prompt Appending, then extracts corresponding image and text embeddings. Lower Section: Processes fMRI signals from different subjects with the Domain-Invariant Encoder to obtain shared feature representations. These embeddings are combined with a diffusion model to generate the reconstructed images.

Main Contributions

- Contributions:
  - Propose Imageset Diffusion with Class-Prompt Appending and Category-aware Localization to augment training data without distorting fMRI signals;
  - Design a Domain-Invariant Encoder with adversarial training to learn subject-shared features, enabling zero-shot cross-subject transfer;
  - Achieve SOTA on NSD in both single-subject and cross-subject settings, with cross-subject results surpassing single-subject models.

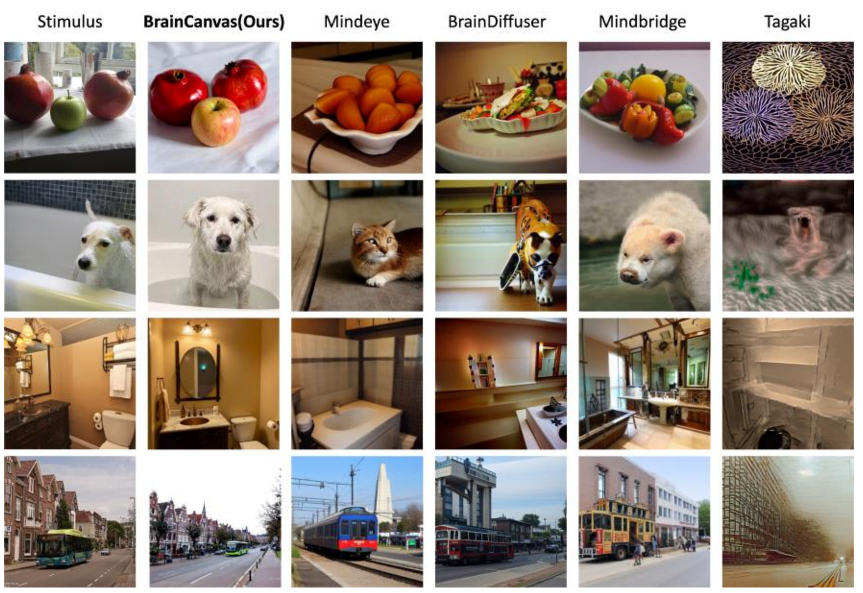


Table 1 Quantitative comparison of fMRI-to-image reconstruction models. Performance is evaluated for both single-subject and cross-subject methods, with BrainCanvas and MindBridge representing the latter. In the results, boldface indicates the best performance and underlining the second.

Type	Method	High-Level				Low-Level			
		Incep ↑	CLIP ↑	EffNet-B ↓	SwAV ↓	PixCorr ↑	SSIM ↑	Alex(2) ↑	Alex(5) ↑
Single Subject	Takagi. [24] *[CVPR'23]*	76.0%	77.0%	–	–	–	–	83.0%	83.0%
	Mind-Reader [50]*[NeurIPS'22]*	78.2%	–	–	–	–	–	–	–
	Mind-Vis [51] *[CVPR'23]*	78.8%	76.2%	0.854	0.491	0.080	0.220	72.1%	83.2%
	Gu. [42] *[MIDL'24]*	–	–	0.862	0.465	0.150	0.325	–	–
	Brain-Diffuser [6]*[SciRep'23]*	87.2%	91.5%	0.775	0.423	0.254	0.356	94.2%	96.2%
	MindEye [25]*[NeurIPS'24]*	93.8%	93.6%	**0.645**	0.377	**0.309**	0.323	92.8%	**96.9%**
	MindBridge [26]*[CVPR'24]*	92.2%	94.3%	0.713	0.413	0.148	0.259	86.9%	95.3%
	BrainCanvas (Ours)	**94.8%**	**95.6%**	0.647	**0.352**	0.291	**0.386**	**95.2%**	96.8%
Cross Subject	MindBridge [26]*[CVPR'24]*	92.4%	94.7%	0.712	0.418	0.151	0.263	87.7%	95.5%
	Wills Aligner [36]*[AAAI'25]*	94.3%	94.8%	0.649	0.373	0.271	0.328	95.8%	98.0%
	BrainGuard [37]*[AAAI'25]*	94.9%	96.4%	0.624	0.353	0.313	0.330	94.7%	97.8%
	MindEye2 [34]*[ICML'24]*	95.4%	93.0%	0.619	0.344	0.322	**0.427**	95.8%	98.1%
	UMBRAE [35]*[ECCV'24]*	91.7%	93.5%	0.700	0.393	0.283	0.341	95.5%	97.0%
	BrainCanvas (Ours)	**96.0%**	**96.5%**	**0.615**	**0.337**	**0.324**	0.401	**96.3%**	**98.4%**

Visualization results and comparison with other methods (left) and quantitative performance metrics (right) on the NSD .