Reconstructing Perceptive Images from Brain Activity by Shape-Semantic GAN

* Tao Fang
duolafang@zju.edu.cn

Yu Qi
qiyu@zju.edu.cn

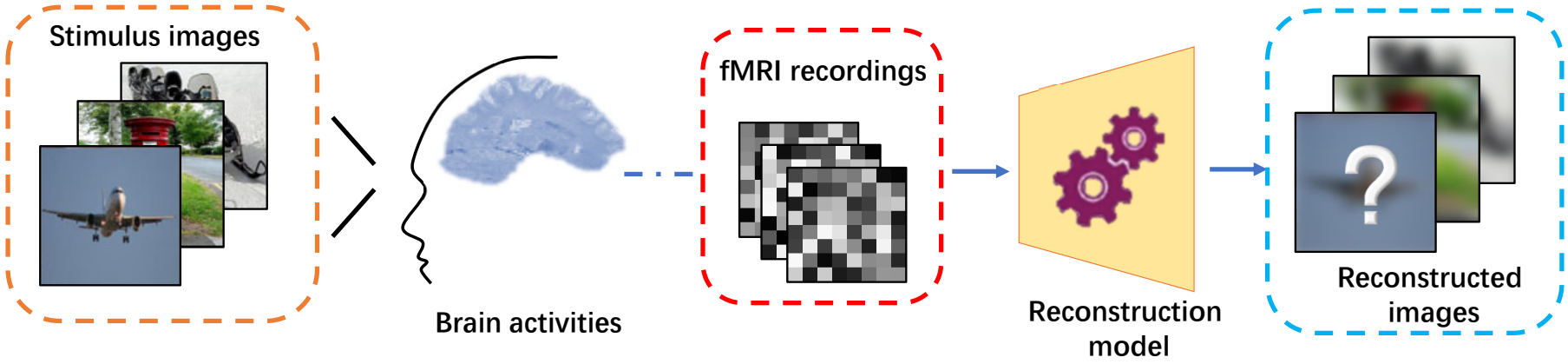
Gang Pan
gpan@zju.edu.cn

College of Computer Science and Technology,
Zhejiang University

Reconstructing Perceptive Images from Brain Activity

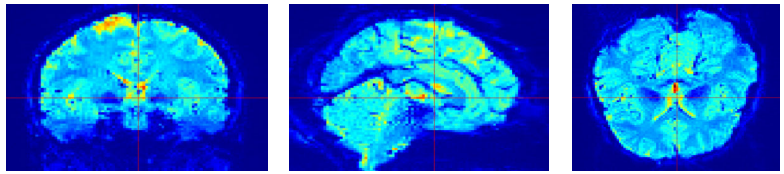
- **From : brain activities**
  - fMRI signals

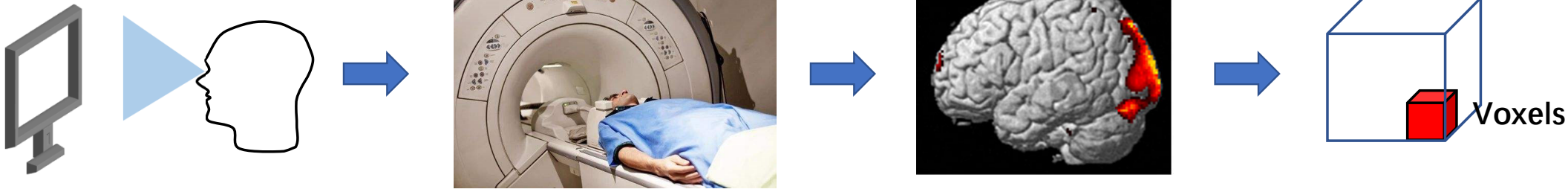
- **To : perceived images**
  - natural Images



fMRI Signals

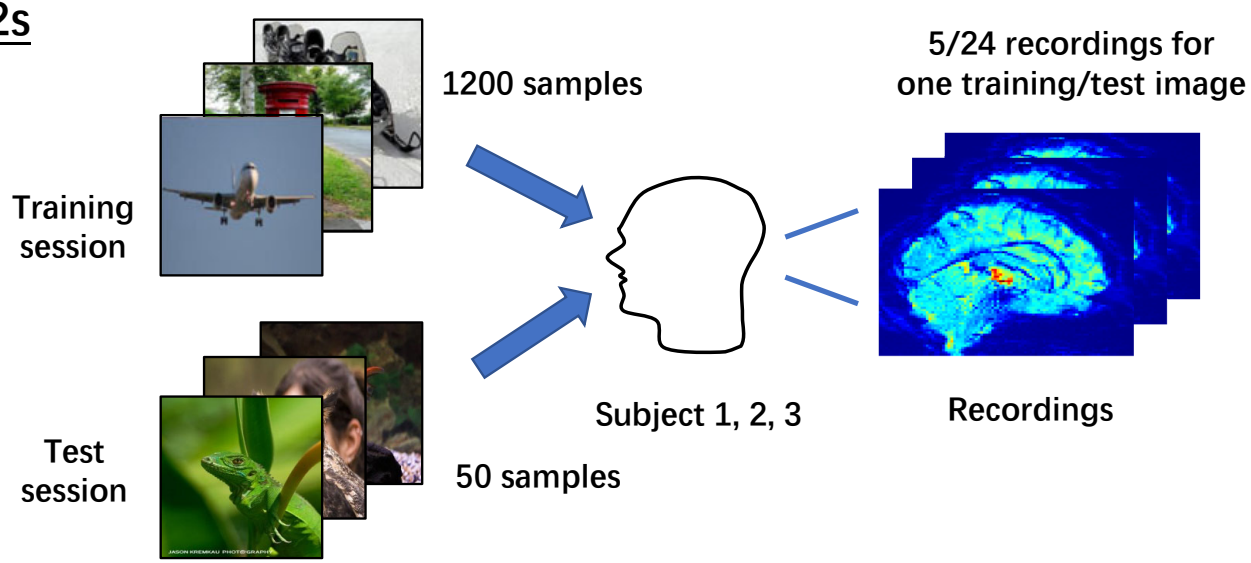
- functional Magnetic Resonance Imaging
  - Recording the variations in
    blood oxygen level dependent (BOLD)
  - fMRI scanner： 3.0 Tesla;
  - Number of slices : 76, TR : 2000 ms
  - Voxel size： 2 x 2 x 2 mm





fMRI Signals

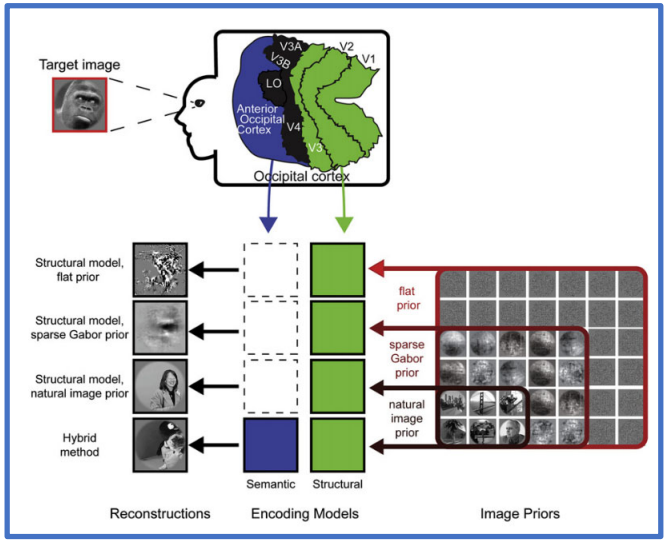
- Perception experiment
  - Stimulus images : **ImageNet** (1250 totally)
  - Center of the display
  - Block : training/test session : **8s** /**12s**
- Dataset[2]
  - Three subjects
  - Training set
    - 1200 different images
    - 5 recordings for per image
  - Test set
    - 50 different images
    - 5 recordings for per image

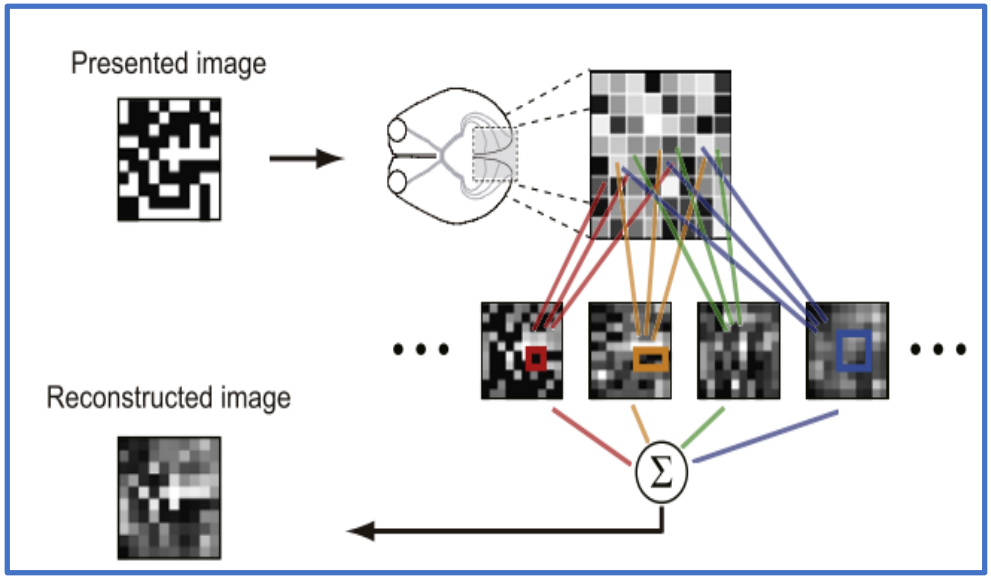


[2] Guohua Shen et al. “Deep image reconstruction from human brain activity”. In: PLoS computational biology 15.1 (2019), e1006633

Previous works

- **(a) Linear methods**[1]
  - gabor filters
  - Image patches
- (b) Basing on the pretrained CNN features[2]
- (c) End-to-end deep models[3]

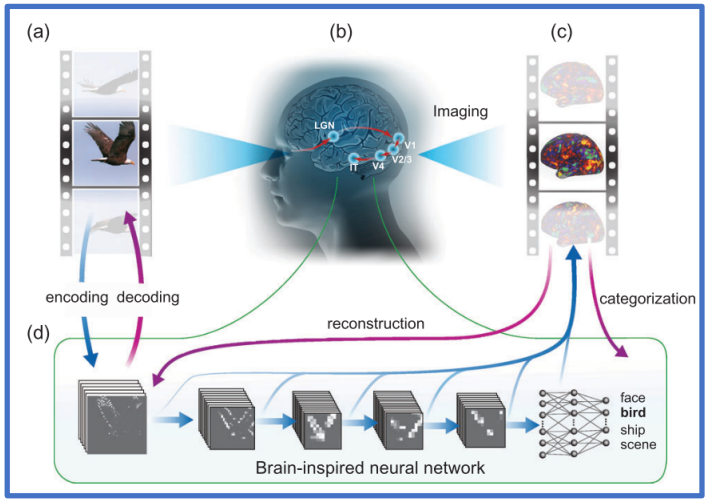


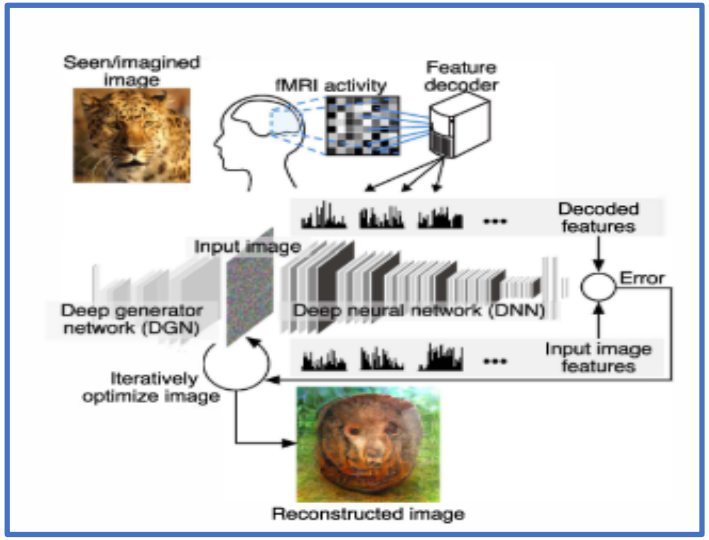


[1] Yoichi Miyawaki et al. “Visual image reconstruction from human brain activity using a combination of multiscale local image decoders”. In: Neuron 60.5 (2008), pp. 915–929.

Previous works

- (a) Linear methods[1]
- **(b) Basing on the pretrained CNN features[2]**
- (c) End-to-end deep models[3]

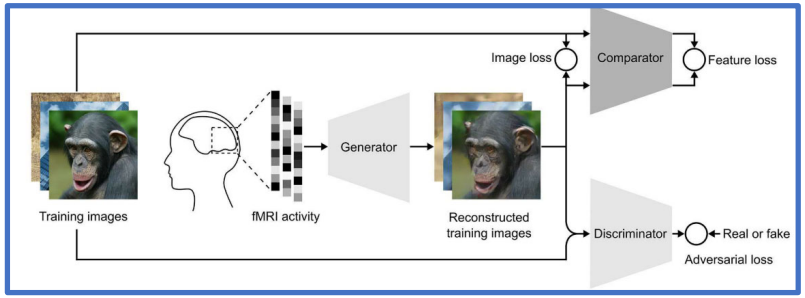




[2] Guohua Shen et al. “Deep image reconstruction from human brain activity”. In: PLoS computational biology 15.1 (2019), e1006633

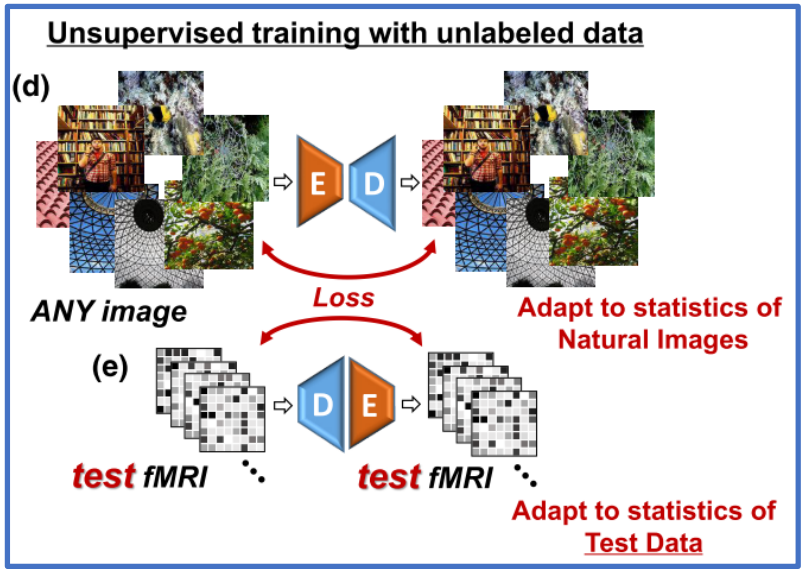
Previous works

- (a) Linear methods[1]



- (b) Basing on the pretrained CNN features[2]

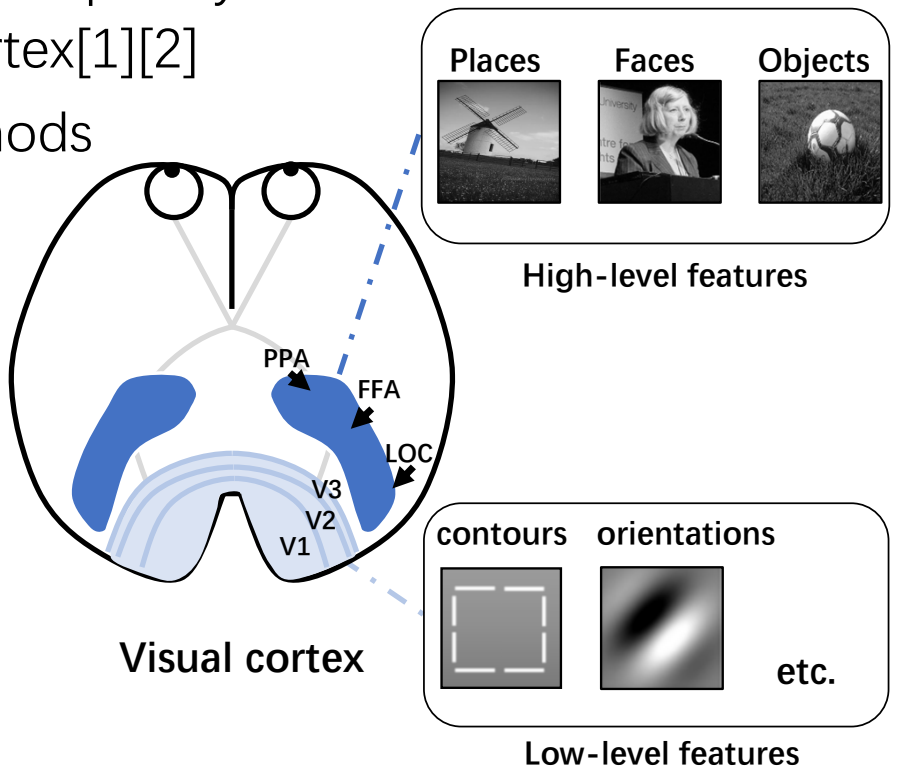
- **(c) End-to-end deep models[3]**



[3] Roman Beliy et al. “From voxels to pixels and back: Self-supervision in natural-image reconstruction from fMRI”. In: Advances in Neural Information Processing Systems. 2019, pp. 6514–6524.

Hierarchical decoding

- Previous studies:
  - A **hierarchical increase** in the complexity of representations in visual cortex[1][2]
- Designing different decoding methods for signals in different visual cortex
  - **Low-level features** from **lower visual cortex (LVC)**
  - **High-level features** from **higher visual cortex (HVC)**

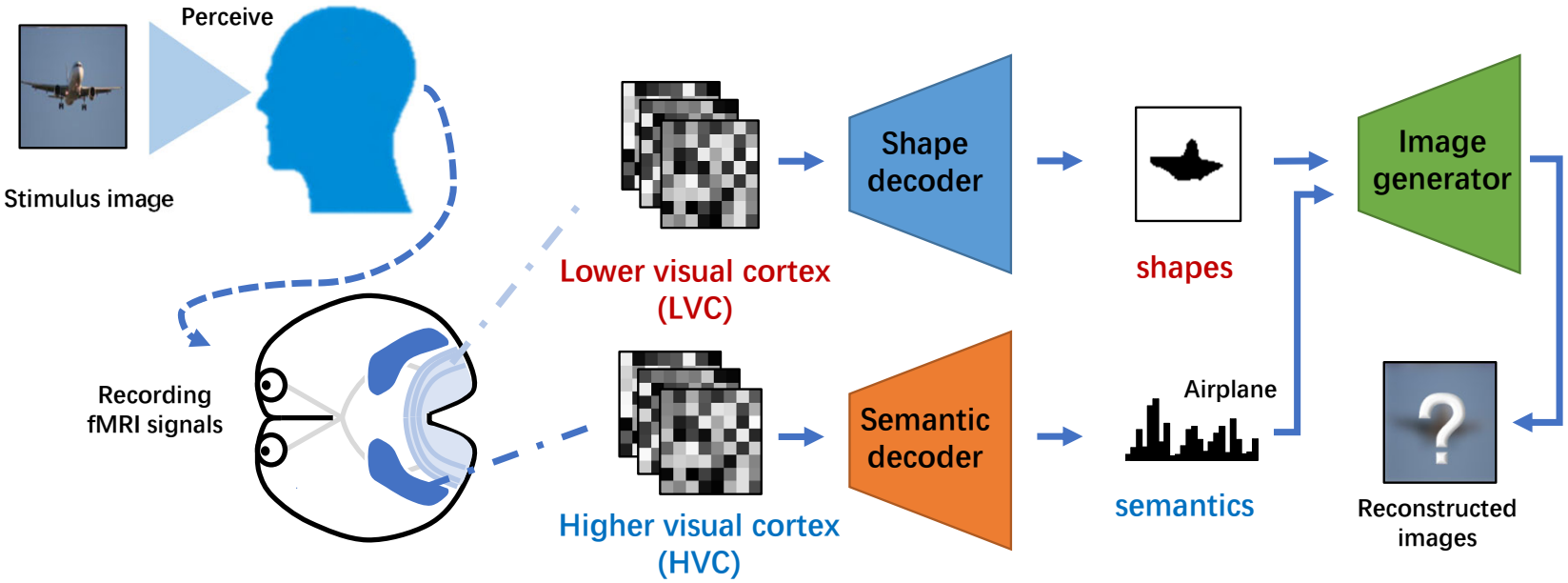


[1] Stephanie Ding et al. “Visual perception as retrospective Bayesian decoding from high- to low-level features”. In: Proceedings of the National Academy of Sciences of the United States of America 114.43 (2017), p. 201706906.

[2] Tomoyasu Horikawa and Yukiyasu Kamitani. “Generic decoding of seen and imagined objects using hierarchical visual features”. In: Nature communications 8.1 (2017), pp. 1–15.

Hierarchical decoding

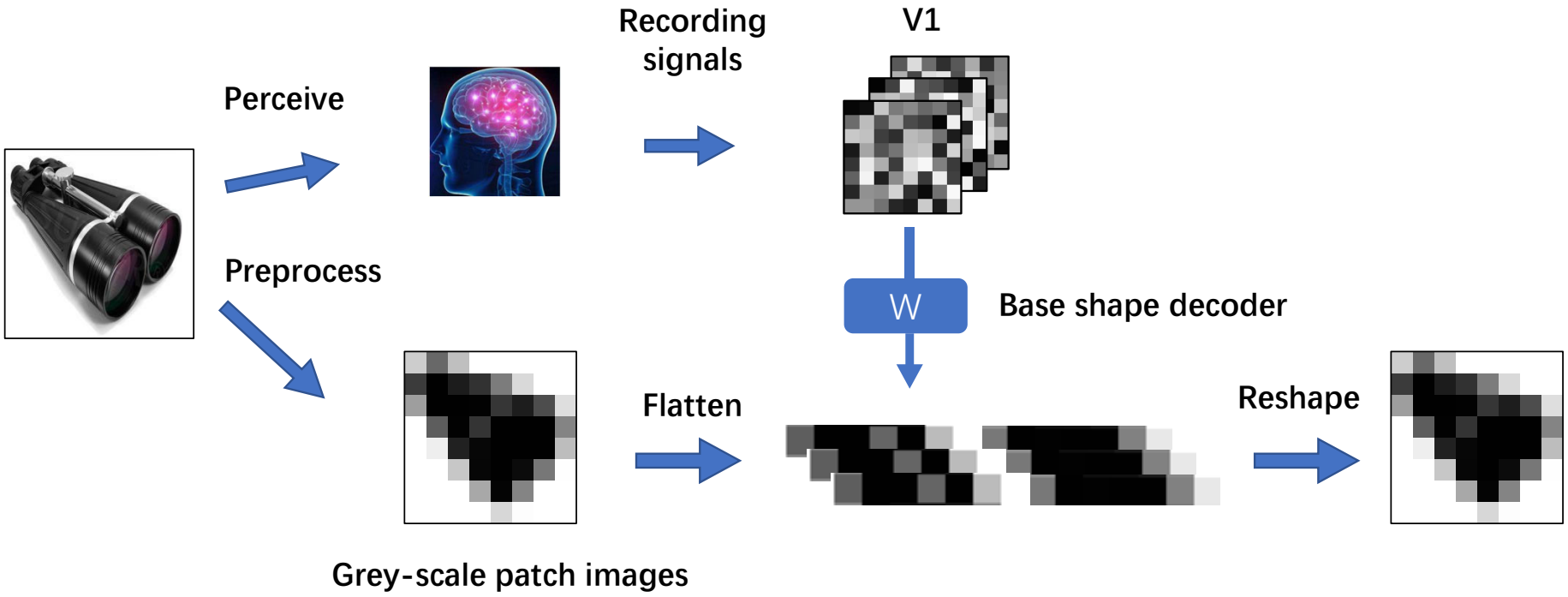
- Specialized decoding: **decoding different features from different ROIs**
  - **Lower visual cortex (LVC) - decoding low-level features (shapes)**
  - **Higher visual cortex (HVC) - decoding high-level features (shapes)**



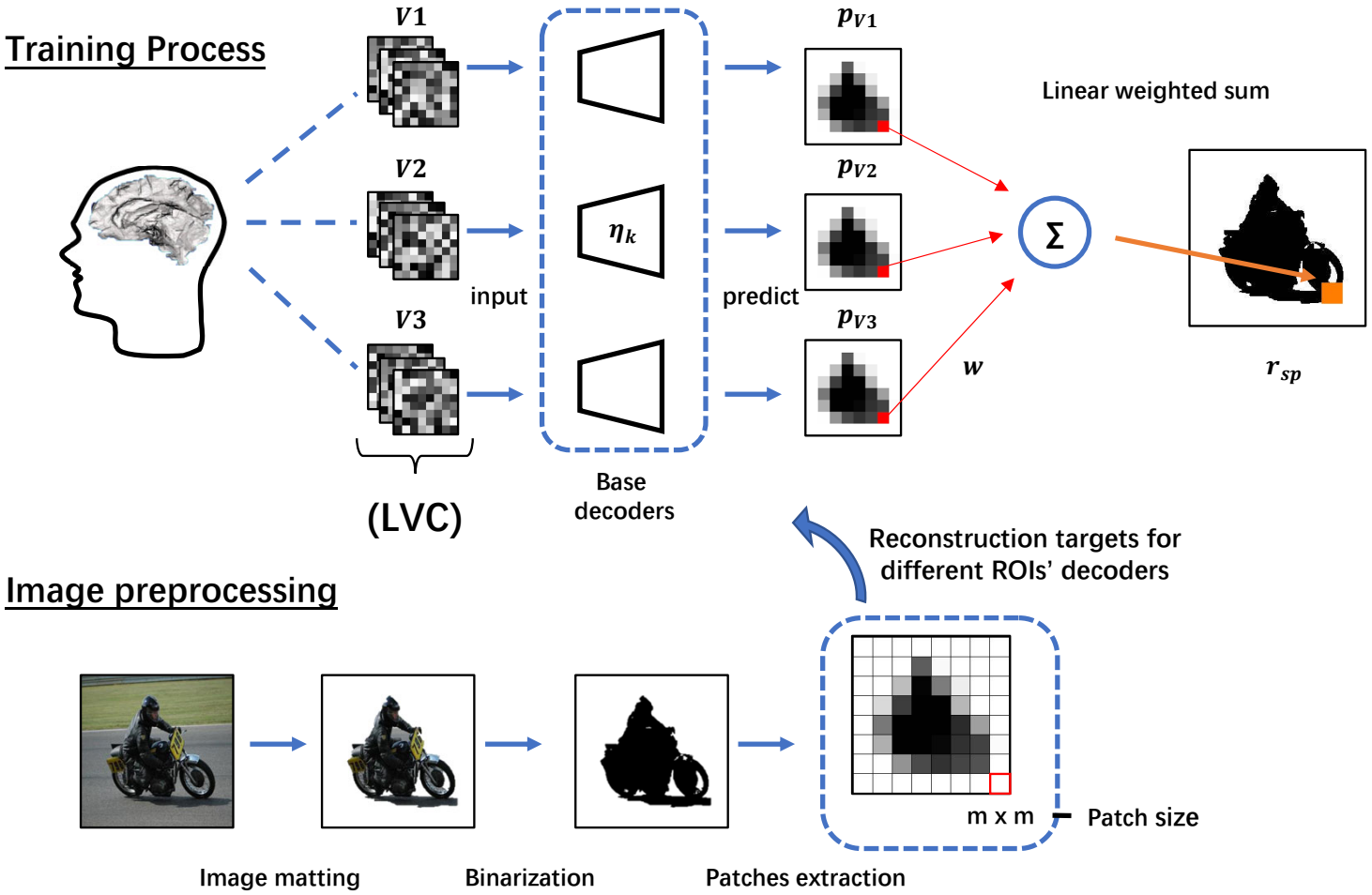
Overview : decoding shape/semantic features from lower/higher visual cortex separately.

A rough reconstruction

- Reconstructing grey-scale shapes based on linear regression



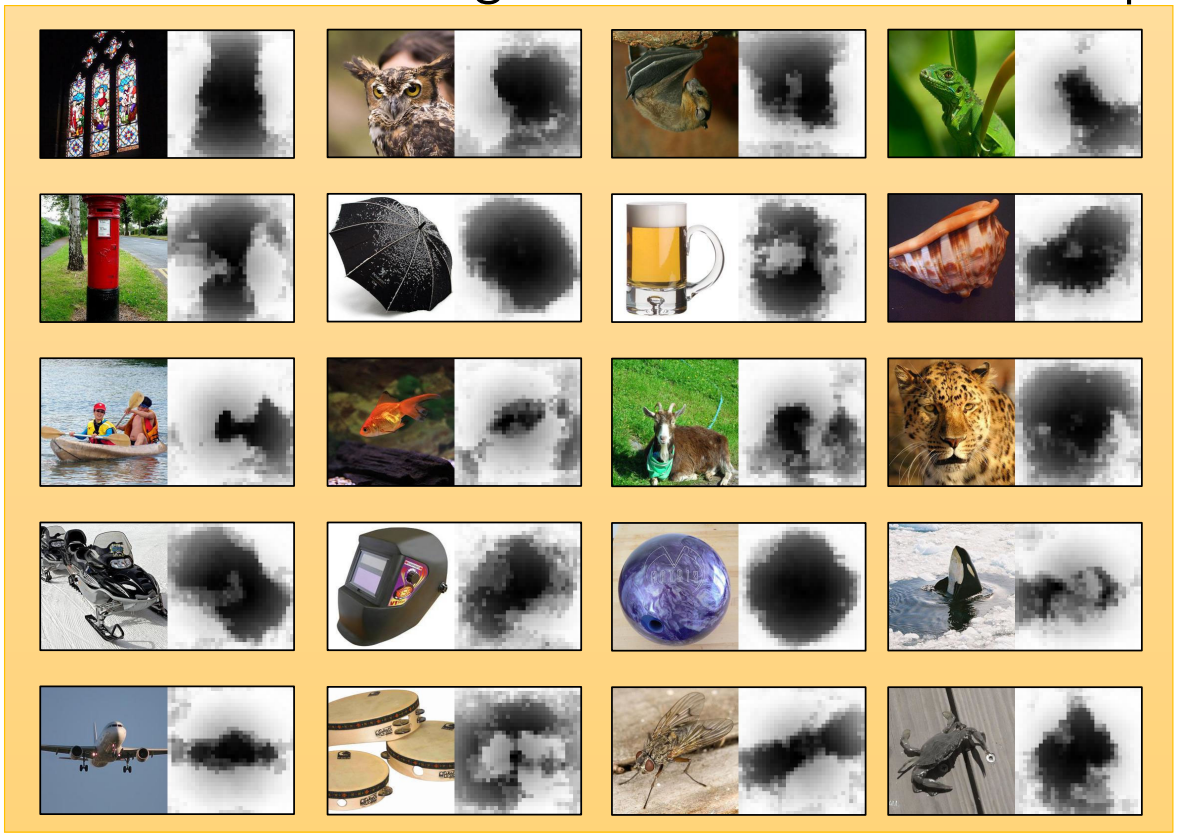
Shape decoding



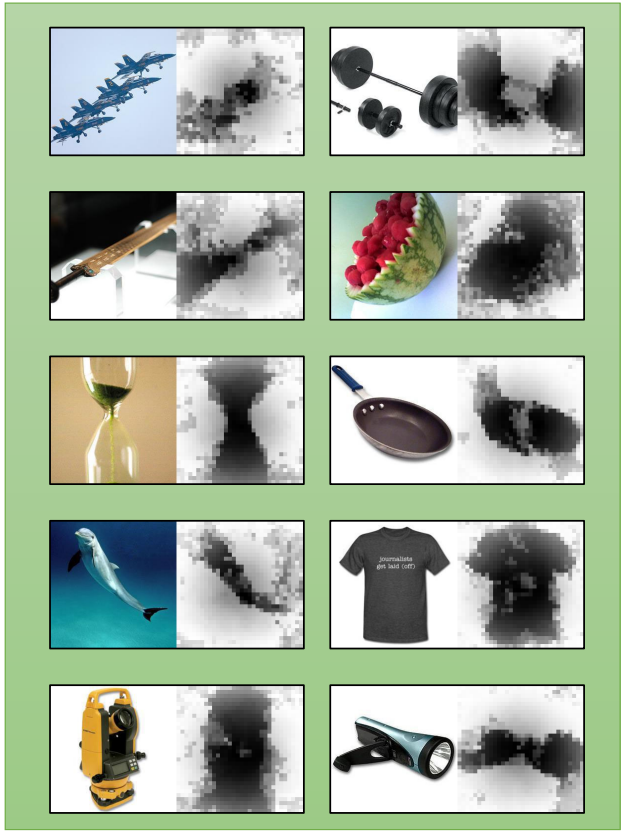
Shape decoding

- The stimulus images & the decoded shapes

Test Session (Subject 3)



Training Session (Subject 3)



Semantic Decoding

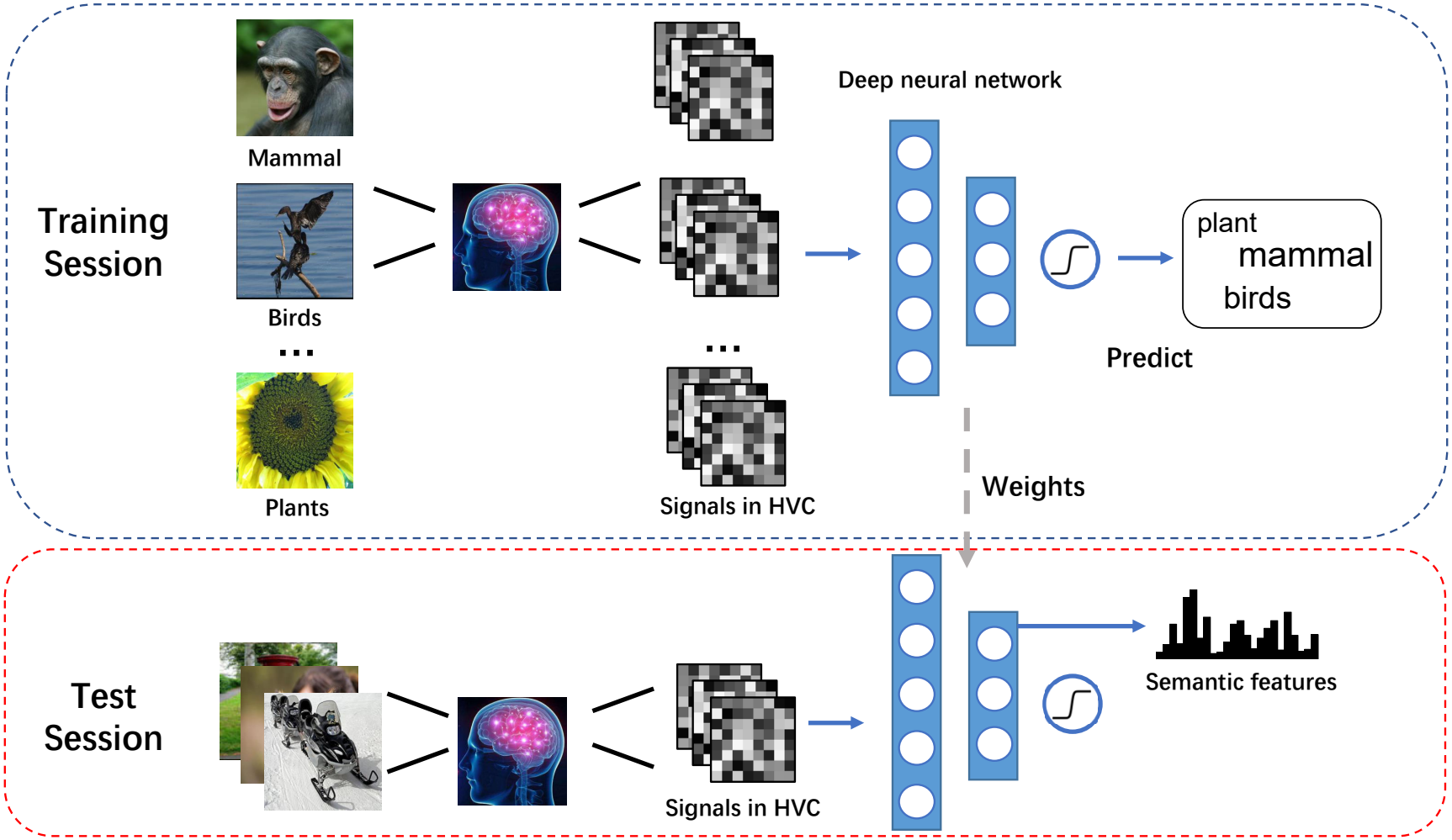
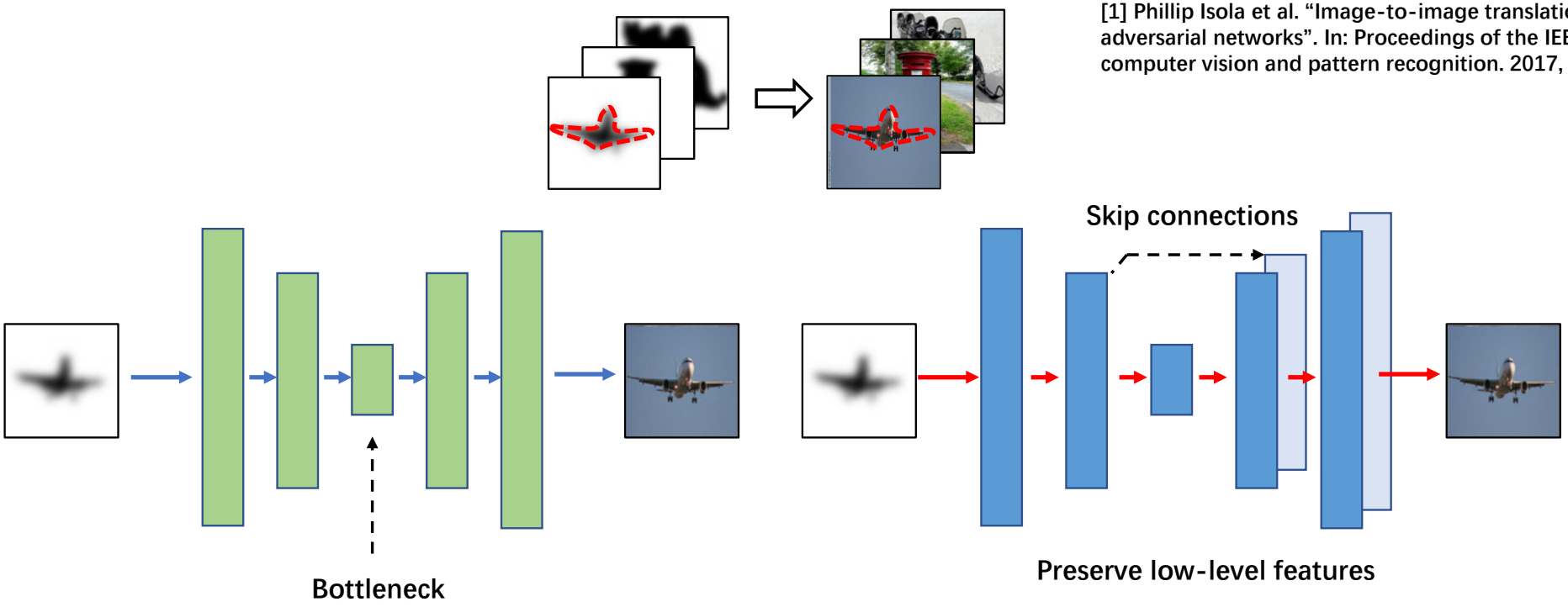


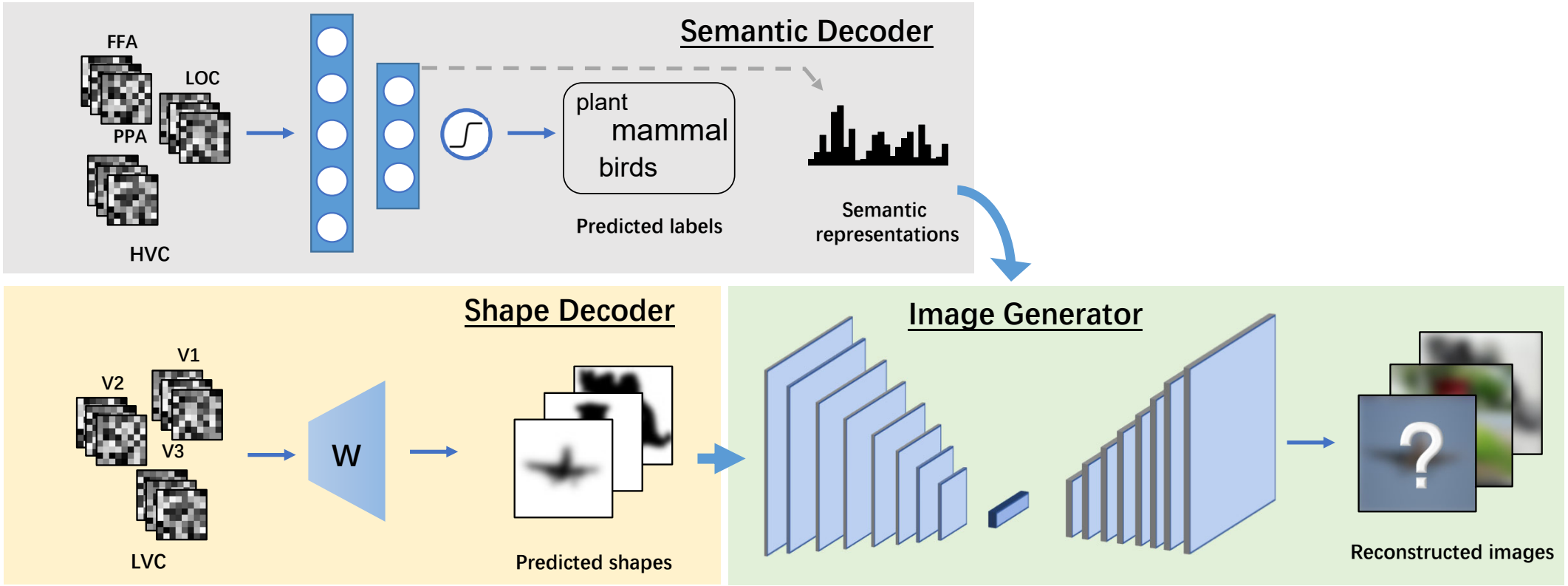
Image generator

- Reconstruction by the image translation model
  - Use the model of pix2pix[1]



[1] Phillip Isola et al. “Image-to-image translation with conditional adversarial networks”. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, pp. 1125– 1134.

Image Reconstruction



Reconstructed results

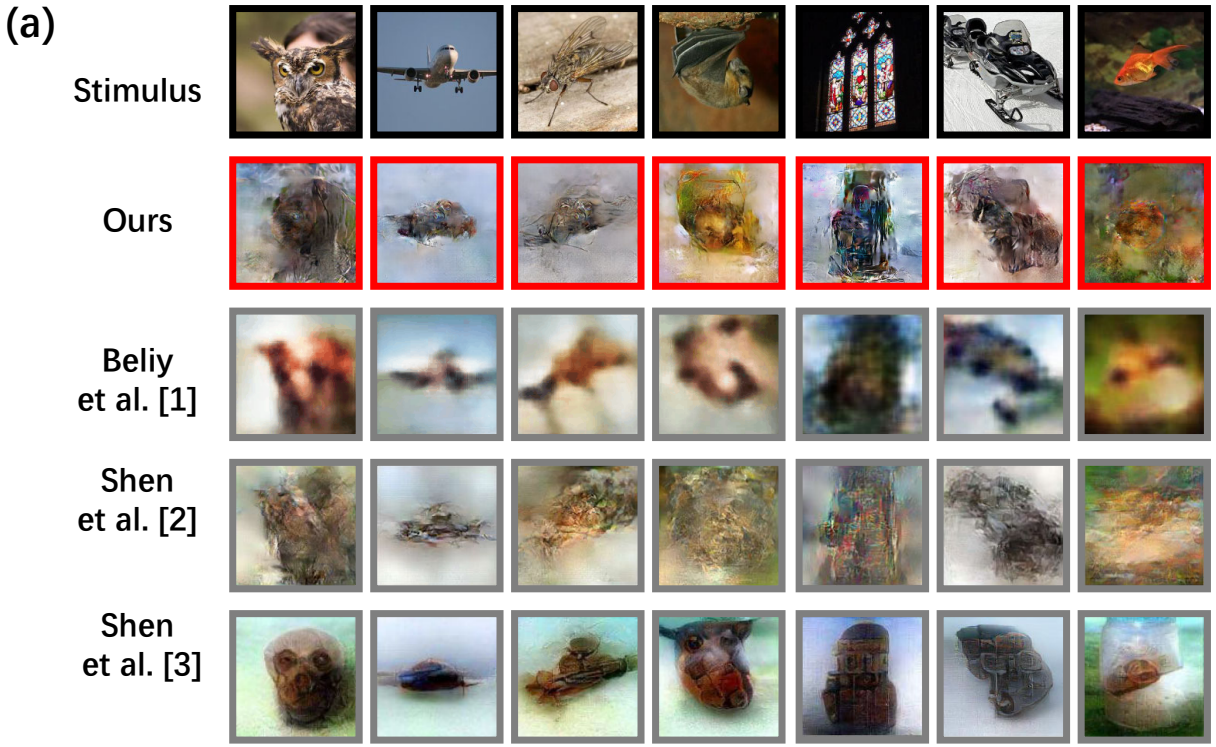
- The reconstructed images of 3 subjects

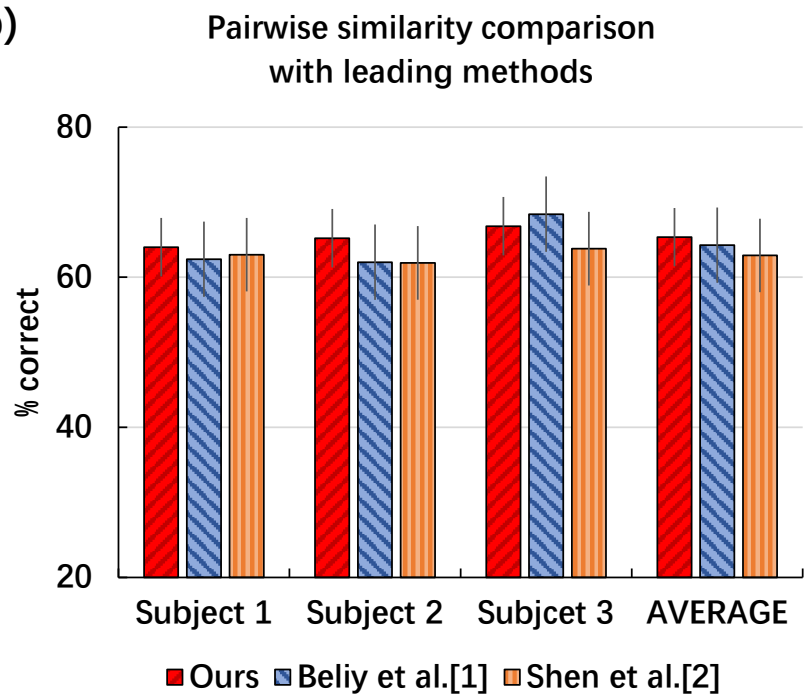
Subject 1

Subject 2

Subject 3

Reconstructed results

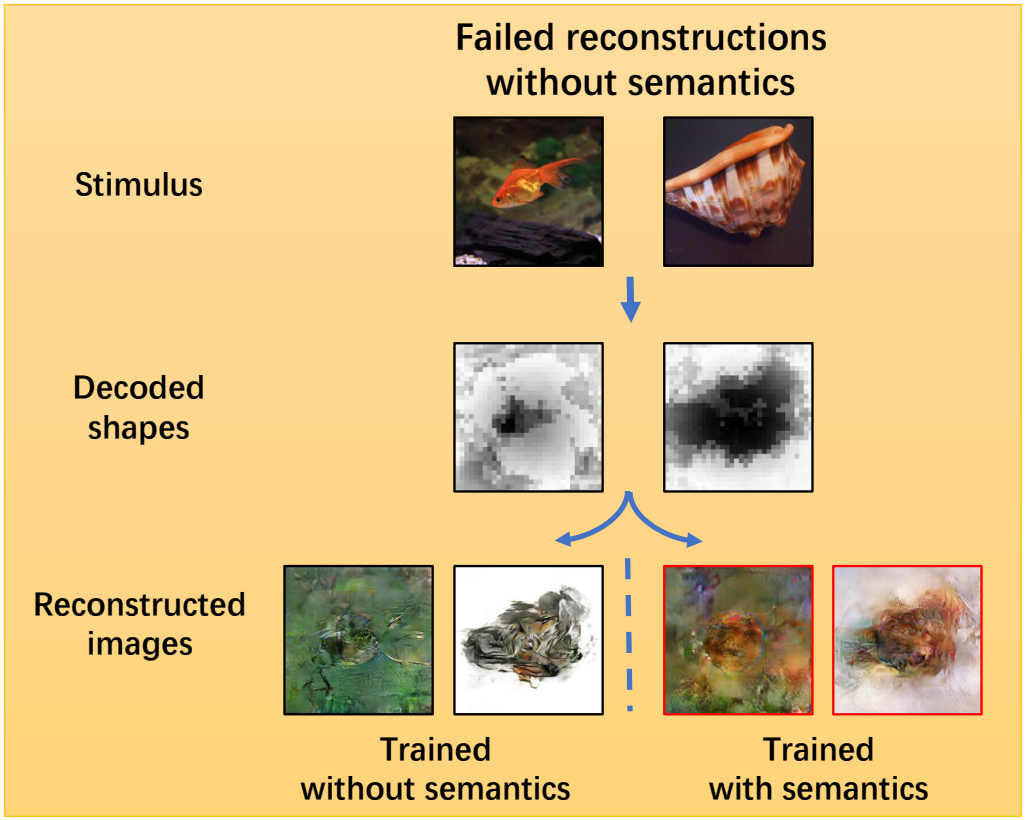


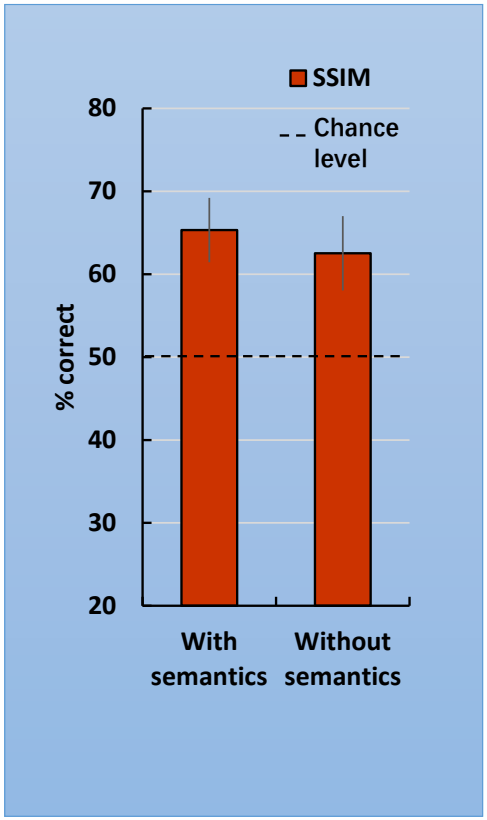


[1] Roman Beliy et al. “From voxels to pixels and back: Self-supervision in natural-image reconstruction from fMRI”. In: Advances in Neural Information Processing Systems. 2019, pp. 6514–6524.
[2] Guohua Shen et al. “End-to-end deep image reconstruction from human brain activity”. In: Frontiers in Computational Neuroscience 13 (2019).
[3] Guohua Shen et al. “Deep image reconstruction from human brain activity”. In: PLoS computational biology 15.1 (2019), e1006633.

Ablation studies

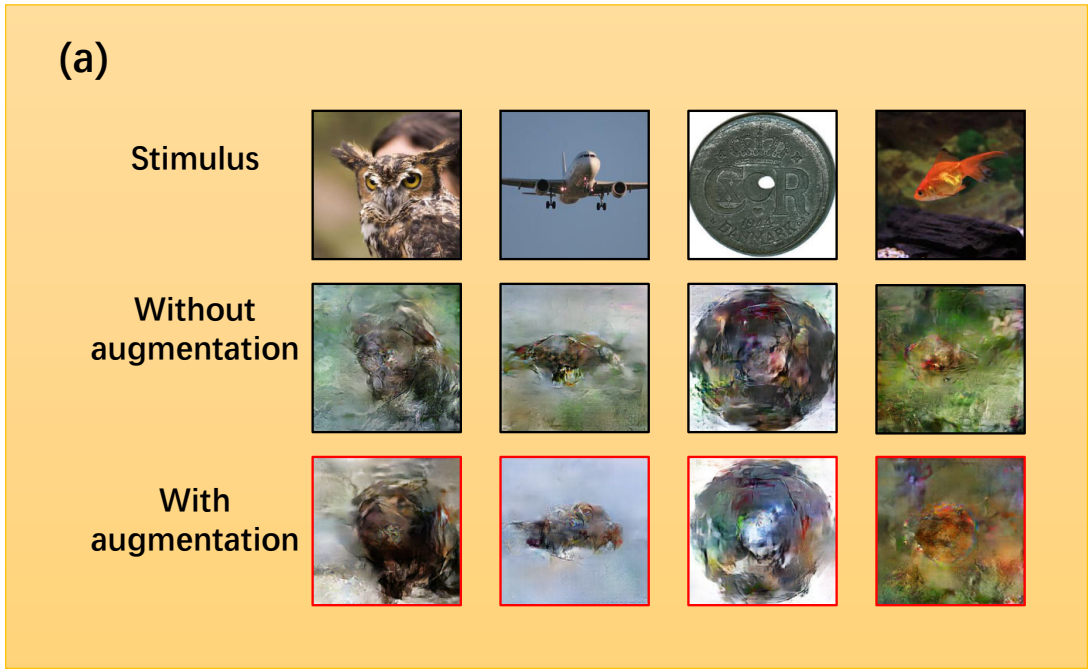
- Effectiveness of semantic information

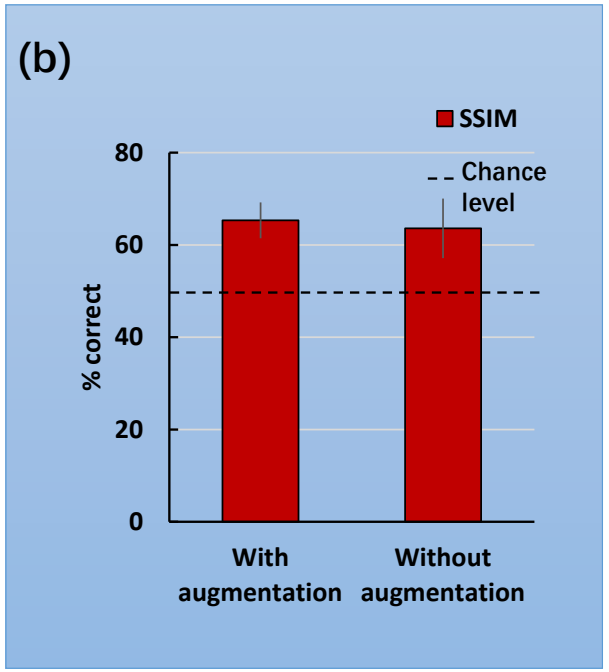




Ablation studies

- Effectiveness of Augmentation





Conclusion

- We propose a **divide and conquer** approach to break the complex visual signals into multi-level components.

- We propose a linear model based shape decoder and a DNN based semantic decoder to **decode shape and semantic information** from the lower and higher visual cortex respectively.

- We propose a GAN model to merge the decoded shape and semantic information to images, which can generate natural-looking images.

Future

- Evaluation methods
- The quality/quantity of the fMRI recordings
- Specialized decoding for different visual areas

Thanks for Your Attention