

Visualizing the mind's eye: a future perspective on applications of image reconstruction from brain signals to psychiatry

Zitong Lu 

Department of Psychology, The Ohio State University, Columbus, OH 43210, USA

*Correspondence: Zitong Lu, lu.2637@osu.edu

In an era where neuroscience dances with computational advances, the power to “visualize” one’s thoughts at image-level is no longer confined to the realm of science fiction. This groundbreaking image reconstruction from brain signals (IRBS) techniques, riding the wave of deep learning and large-scale neuroimaging datasets, offer an unprecedented perspective for not only neurocognitive research but also psychiatry. In this short commentary, I provide a concise introduction to the development of IRBS technology and offer a perspective on its potential applications, current limitations, and future directions in psychiatry.

Rapid developments of image reconstruction techniques in cognitive computational neuroscience

Advances in modern neuroimaging techniques have paved the way for novel perspectives in understanding human mental processes. One of the most exciting developments in this regard is image reconstruction from brain signals (IRBS), also called “reconstructing the mind’s eye” at image-level. IRBS allows us to directly “visualize” subjective thoughts by directly translating brain activity into tangible images. Using complex algorithms and sophisticated computational models such as deep learning, researchers have successfully transformed multidimensional brain activity data measured by functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) into a two-dimensional visual image (Fig. 1). Early studies focused on using a combination of multiscale local image decoders to reconstruct observed contrast patterns (Miyawaki *et al.*, 2008). A key shortcoming of many modern machine learning models is their heavy reliance on large amounts of data, and collecting human brain data is notoriously expensive. In parallel with the development of powerful algorithms, neuroscientists have collected unprecedentedly large neuroimaging datasets to support the understanding of brain function and the development of brain–computer interfaces (Allen *et al.*, 2022; Hebart *et al.*, 2023). In the deep learning era, researchers switched to convolutional neural networks for constraint-free reconstruction of complex natural image stimuli (Shen, Horikawa, *et al.*, 2019). Also, generative adversarial networks (Khaleghi *et al.*, 2022; Mishra *et al.*, 2023; Seeliger *et al.*, 2018; Shen, Dwivedi, *et al.*, 2019; Singh *et al.*, 2023) and variational au-

toencoders (Beliy *et al.*, 2019; VanRullen & Reddy, 2019; Wakita *et al.*, 2021) have also been used for image generation. The recent surge of large language models also motivates researchers to use more semantic information in image reconstruction. Examples include visual-semantic models (Radford *et al.*, 2021), and high-performance diffusion models (Rombach *et al.*, 2022). With state-of-the-art models and rich neuroimaging data resources, a number of recent studies are able to reconstruct or generate remarkably high fidelity images from brain activity measured by fMRI (Gu *et al.*, 2023; Lin *et al.*, 2022; Ozcelik & VanRullen, 2023; Scotti *et al.*, 2023; Takagi & Nishimoto, 2023) and EEG (Lan *et al.*, 2023; Singh *et al.*, 2023; Wakita *et al.*, 2021; Zeng *et al.*, 2023). However, we need to interpret these recent reconstruction results with caution. A recent study pointed out issues in these research that the results lacked critical assessments and their overestimated performance was controversial (see more details in Shirakawa *et al.*, 2023). In addition to the reconstruction of visual stimuli (i.e. observed images) from brain activity, some studies have also successfully reconstructed subjective visual experience in the human mind (e.g. attention, memory, mental imagery, illusory perception) (Cheng *et al.*, 2023; Horikawa & Kamitani, 2022; Koide-Majima *et al.*, 2023; Lee & Kuhl, 2016; Shen, Horikawa, *et al.*, 2019; Shimizu & Srinivasan, 2022), which remains a challenge to be actively explored in the field of cognitive computational neuroscience, but also provides vast prospect in computational psychiatry.

The promise of IRBS in psychiatry

Commonly used methods such as standardized questionnaires and behavioral assessments in the diagnosis and treatment of psychiatric patients often lack comprehensive precision and can sometimes miss subtle yet crucial symptoms. Additionally, even with pharmacological treatments and psychotherapy, current approaches might not account for the individual variability in responses, leading to non-optimal treatment plans. IRBS techniques have gradually proven their efficacy in healthy human participants. We believe that IRBS techniques can be further extended to understand patients’ minds and the deficits associated with psychiatric disorders. A recent study by Cheng *et al.*, 2023 successfully reconstructed not only the world given sensory inputs but also visual illusory experiences. This study suggests that IRBS

Received: 16 August 2023; Revised: 26 September 2023; Accepted: 10 October 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of West China School of Medicine/West China Hospital (WCSM/WCH) of Sichuan University. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

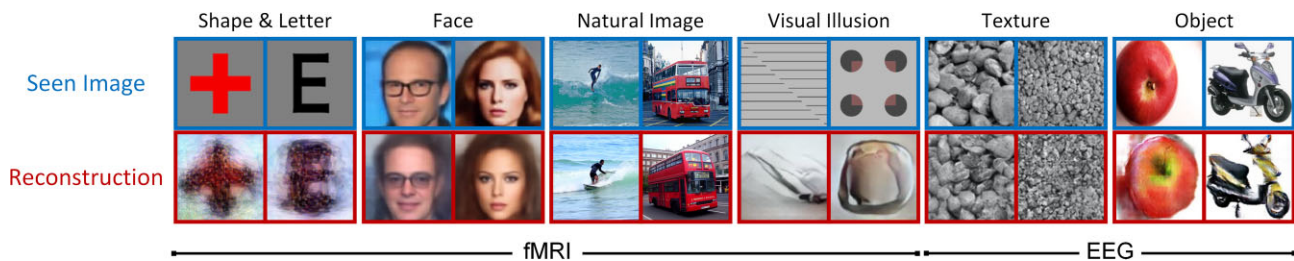


Figure 1: Examples of IRBS. From left to right: reconstructions of shapes and letters (Shen, Horikawa, et al., 2019), faces (VanRullen & Reddy, 2019), natural images (Scotti et al., 2023), and visual illusions (Cheng et al., 2023) based on fMRI, and reconstructions of textures (Wakita et al., 2021) and objects (Singh et al., 2023) based on EEG. Top row: images seen by the participant. Bottom row: reconstruction results from fMRI signals.

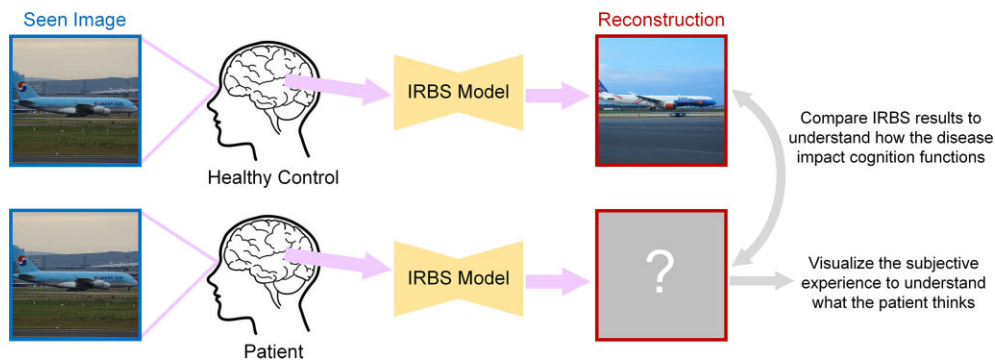


Figure 2: Potential applications of IRBS in psychiatry, including the reconstruction of images from patients' brain signals and the comparisons between the healthy control group and the patient group. The example for the seen image is from the NSD dataset (Allen et al., 2022), and the reconstruction result is from Ozelik & VanRullen (2023).

may hold immense promise for various applications in psychiatry, serving as a new window into the mind (Fig. 2). This capability could be instrumental in uncovering the elusive mechanisms underlying numerous mental disorders, which have been challenging to diagnose and treatment due to the inherent subjectivity involved.

One of the major advantages of this technology is its ability to provide a more direct and vivid visualization of subjective visual experience. Consider, for example, a condition such as schizophrenia, which is characterized by visual hallucinations and altered perceptions of reality (Liddle, 1987; López-Silva et al., 2022). Patients with these conditions often struggle to articulate their subjective experiences, and their accounts may not match the reality understood by others. Using IRBS to visualize these hallucinations provides a direct and objective description of the patient's experience. By translating these hallucinations into tangible images, clinicians can gain a more accurate and clearer understanding of what patients are thinking, which could greatly improve the diagnostic process. Similarly, for mood disorders (Struijs et al., 2021) and post-traumatic stress disorder (Ehlers et al., 2004), decoding intrusive negative thoughts or traumatic memories with IRBS can potentially allow clinicians to better understand the mental state of their patients, facilitate more effective communication, and develop individualized treatment approaches. For cognitive and memory disorders, such as Alzheimer's disease and other forms of dementia, visualizing affected memories through IRBS provides profound insights into the intricacies of cognitive decline.

Furthermore, integrating IRBS with closed-loop neurofeedback (Sitaram et al., 2017) offers promising avenues in psychiatry. Neu-

rofeedback can be tailored based on real-time IRBS, allowing patients to actively modulate their brain activity. Especially in conditions such as schizophrenia or post-traumatic stress disorder, patients can potentially "see" and "control" their hallucinations or traumatic memories, which leads to innovative therapeutic interventions.

Current limitations and future directions

While IRBS has tremendous potential, there is still a long way to go before it can be truly applied to the field of psychiatry. First, there are some inherent technical challenges of IRBS that need to be overcome. On the one hand, current IRBS techniques still have a long way to go to achieve reliable and robust reconstruction performance for medical applications. On the other hand, these techniques usually require a substantial amount of data for model training to produce clearer and more realistic images. This suggests that this technology currently requires large amounts of neuroimaging data and high computational resources, which limits its scalability. There is no doubt that collecting a large amount of neuroimaging data from patients is more challenging than from healthy individuals. Second, the effectiveness of IRBS is often influenced by individual variations in brain anatomy and function, necessitating personalized models for accurate reconstruction. Recent studies on the algorithms for inter-subject brain signal transformation (Ho et al., 2023; Lu & Golomb, 2023) open a new avenue for inter-subject brain-computer interfaces. Exploring how to augment patient neuroimaging data, whether we can predict functional data from structural data, and building cross-individual neural models are technological avenues worth

pursuing, beyond the primary challenge of accurately visualizing the mind's eye. Third, there might be different criteria currently available to evaluate the results of IRBS. However, how to select appropriate criteria to evaluate the reconstruction results under different circumstances and how to compare results between healthy controls and patients remain issues that need to be addressed. Finally, the ethical implications also warrant careful consideration. Issues of privacy, consent, and the potential misuse of personal mental images are areas that require strict regulations to protect individual rights.

Summary

The challenges also present opportunities for further research and development. IRBS represents a powerful tool for understanding and treating mental disorders in the future. As we refine the technology and navigate the ethical landscape, we envision a future in which psychiatry is enriched by the ability to visualize the "mind's eye," bridging the gap between subjective experience and objective understanding.

Author contributions

Zitong Lu (Conceptualization, Investigation, Project administration, Visualization, Writing – original draft, Writing – review & editing)

Conflict of interests

The authors declare no conflict of interest.

Acknowledgements

The author thanks Dr. Ru-Yuan Zhang from Shanghai Jiao Tong University for valuable comments on the manuscript.

References

- Allen EJ, St-Yves G, Wu Y, et al. (2022) A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nat Neurosci* **25**:116–26.
- Beliy R, Gaziv G, Hoogi A, et al. (2019) From voxels to pixels and back: self-supervision in natural-image reconstruction from fMRI. In: H Wallach, H Larochelle, A Beygelzimer, F d' Alché-Buc, E Fox, R Garnett (eds). *Advances in Neural Information Processing Systems*. Vol. **32**. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2019/file/7d2be41b1bde6ff8fe45150c37488ebb-Paper.pdf.
- Cheng F, Horikawa T, Majima K, et al. (2023) Reconstructing visual illusory experiences from human brain activity. *bioRxiv Cold Spring Harbor Laboratory*. <https://www.biorxiv.org/content/early/2023/06/15/2023.06.15.545037.1.full.pdf>.
- Ehlers A, Hackmann A, Michael T (2004) Intrusive re-experiencing in post-traumatic stress disorder: phenomenology, theory, and therapy. *Memory* **12**:403–15.
- Gu Z, Jamison K, Kuceyeski A, et al. (2023) Decoding natural image stimuli from fMRI data with a surface-based convolutional network. *Medical Imaging with Deep Learning*. <https://openreview.net/forum?id=V5vvti2Y9PA>.
- Hebart MN, Contier O, Teichmann L, et al. (2023) THINGS-data, a multimodal collection of large-scale datasets for investigating object representations in human brain and behavior. *ELife* **12**:e82580.
- Ho JK, Horikawa T, Majima K, et al. (2023) Inter-individual deep image reconstruction via hierarchical neural code conversion. *Neuroimage* **271**:120007.
- Horikawa T, Kamitani Y (2022) Attention modulates neural representation to render reconstructions according to subjective appearance. *Commun Biol* **5**:34.
- Khaleghi N, Rezaei TY, Beheshti S, et al. (2022) Visual saliency and image reconstruction from EEG signals via an effective geometric deep network-based generative adversarial network. *Electronics (Switzerland)* **11**:3637.
- Koide-Majima N, Nishimoto S, Majima K (2023) Mental image reconstruction from human brain activity keywords. *BiorXiv Cold Spring Harbor Laboratory*. <https://www.biorxiv.org/content/early/2023/03/28/2023.01.22.525062>.
- Lan Y-T, Ren K, Wang Y, et al. (2023) Seeing through the brain: image reconstruction of visual perception from human brain signals. *arXiv:2308.02510 (eess.IV)*.
- Lee H, Kuhl BA (2016) Reconstructing perceived and retrieved faces from activity patterns in lateral parietal cortex. *J Neurosci* **36**:6069–82.
- Liddle PF (1987) Schizophrenic syndromes, cognitive performance and neurological dysfunction. *Psychol Med* **17**:49–57.
- Lin S, Sprague T, Singh AK (2022) Mind Reader: reconstructing complex images from brain activities. In: S Koyejo, S Mohamed, A Agarwal, D Belgrave, K Cho, A Oh (eds). *Advances in Neural Information Processing Systems*. Vol. **35**. Curran Associates, Inc., 29624–36. https://proceedings.neurips.cc/paper_files/paper/2022/file/bee5125b773414d3d6eeb4334fbc5453-Paper-Conference.pdf.
- López-Silva P, Cavieres Á, Humpston C (2022) The phenomenology of auditory verbal hallucinations in schizophrenia and the challenge from pseudohallucinations. *Front. Psychiatry* **13**:826654.
- Lu Z, Golomb JD (2023) Generate your neural signals from mine: individual-to-individual EEG converters. In: *Proceedings of the Annual Meeting of the Cognitive Science Society* **45**. <https://escholarship.org/uc/item/5xn0885t>.
- Mishra R, Sharma K, Jha RR, et al. (2023) NeuroGAN: image reconstruction from EEG signals via an attention-based GAN. *Neural Comput Appl* **35**:9181–92.
- Miyawaki Y, Uchida H, Yamashita O, et al. (2008) Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. *Neuron* **60**:915–29.
- Ozcelik F, VanRullen R (2023) Natural scene reconstruction from fMRI signals using generative latent diffusion. *arXiv:2303.05334 (cs.CV)*.
- Radford A, Kim JW, Hallacy C, et al. (2021) Learning transferable visual models from natural language supervision. In: Marina Meila, Tong Zhang (eds). *Proceedings of Machine Learning Research. Proceedings of the 38th International Conference on Machine Learning*. Vol. **139**, pp. 8748–63. PMLR, <https://proceedings.mlr.press/v139/radford21a.html>.
- Rombach R, Blattmann A, Lorenz D, et al. (2022) High-resolution image synthesis with latent diffusion models. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10684–95.
- Scotti PS, Banerjee A, Goode J, et al. (2023) Reconstructing the mind's eye: fMRI-to-image with contrastive learning and diffusion priors. *arXiv:2305.18274 (cs.CV)*.
- Seeliger K, Güçlü U, Ambrogioni L, et al. (2018) Generative adversarial networks for reconstructing natural images from brain activity. *Neuroimage* **181**:775–85.

- Shen G, Dwivedi K, Majima K, et al. (2019) End-to-end deep image reconstruction from human brain activity. *Front Comput Neurosci* **13**:21.
- Shen G, Horikawa T, Majima K, et al. (2019) Deep image reconstruction from human brain activity. *Front Comput Neurosci* **15**:e1006633.
- Shimizu H, Srinivasan R (2022) Improving classification and reconstruction of imagined images from EEG signals. *PLoS ONE* **17**:e0274847.
- Shirakawa K, Tanaka M, Aoki S, et al. (2023) Critical Assessment of Generative AI Methods and Natural Image Datasets for Visual Image Reconstruction from Brain Activity. OSF. <https://osf.io/nmfc5/>.
- Singh P, Pandey P, Miyapuram K, et al. (2023) EEG2IMAGE: image reconstruction from EEG brain signals. arXiv:2302.10121 (cs.HC).
- Sitaram R, Ros T, Stoeckel L, et al. (2017) Closed-loop brain training: the science of neurofeedback. *Nat Rev Neurosci* **18**:86–100.
- Struijs SY, De Jong PJ, Jeronimus BF, et al. (2021) Psychological risk factors and the course of depression and anxiety disorders: a review of 15 years NESDA research. *J Affect Disord* **295**:1347–59.
- Takagi Y, Nishimoto S (2023) High-resolution image reconstruction with latent diffusion models from human brain activity. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 14453–63.
- Vanrullen R, Reddy L (2019) Reconstructing faces from fMRI patterns using deep generative neural networks. *Commun Biol* **2**:193.
- Wakita S, Orima T, Motoyoshi I (2021) Photorealistic reconstruction of visual texture from EEG signals. *Front Comput Neurosci* **15**:1–11.
- Zeng H, Xia N, Qian D, et al. (2023) DM-RE2I: a framework based on diffusion model for the reconstruction from EEG to image. *Biomed Signal Process Control* **86**:105125.