

Generative AI for brain-computer interfaces decoding: Advances, challenges and future

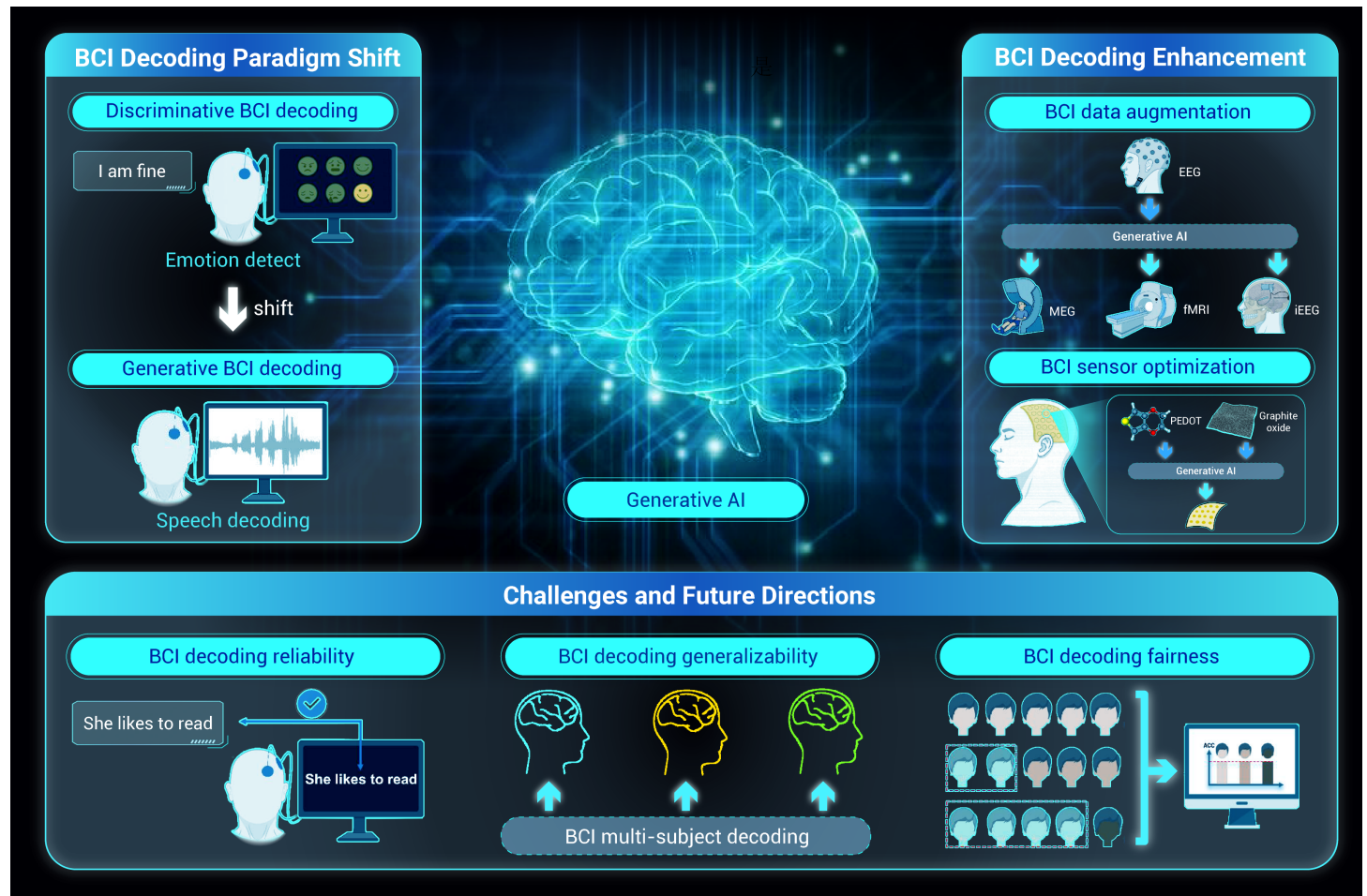
Yi Guo,^{1,2} Guanhua Ren,³ Shiqiang Ma,¹ Jun Zhou,⁴ Hasan Al Nashash,⁵ Alireza Jolfaei,⁶ Zhiguo Zhang,⁷ Guanglin Li,¹ Ye Li,^{1,9} Yong Hu,^{1,8} and Shuqiang Wang^{1,9,*}

*Correspondence: sq.wang@siat.ac.cn

Received: October 4, 2025; Accepted: January 12, 2026; Published Online: January 13, 2026; <https://doi.org/10.59717/j.xinn-med.2026.100193>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Generative BCI decoding reconstructs semantically rich content, including text and images, from brain signals.
- Generative AI boosts BCI decoding through data augmentation and sensor optimization.
- Key challenges include BCI decoding reliability, cross-subject generalizability, and fairness.

Generative AI for brain-computer interfaces decoding: Advances, challenges and future

Yi Guo,^{1,2} Guanhua Ren,³ Shiqiang Ma,¹ Jun Zhou,⁴ Hasan Al Nashash,⁵ Alireza Jolfaei,⁶ Zhiguo Zhang,⁷ Guanglin Li,¹ Ye Li,^{1,9} Yong Hu,^{1,8} and Shuqiang Wang^{1,9,*}

¹Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

²University of Chinese Academy of Sciences, Beijing 100049, China

³China National Institute of Standardization, Beijing 100191, China

⁴Zhejiang Qiangnao Keji (BrainCo) Ltd., Hangzhou 311100, China

⁵Department of Electrical Engineering, American University of Sharjah, Sharjah PO Box 26666, United Arab Emirate

⁶College of Science and Engineering, Flinders University, Adelaide SA 5042, Australia

⁷School of Intelligence Science and Engineering, College of Artificial Intelligence, Harbin Institute of Technology, Shenzhen 518055, China

⁸Department of Orthopaedics and Traumatology, The University of Hong Kong, Hong Kong SAR 999077, China

⁹ Faculty of Computer Science and Control Engineering, Shenzhen University of Advanced Technology, Shenzhen 518107, China

*Correspondence: sq.wang@siat.ac.cn

Received: October 4, 2025; Accepted: January 12, 2026; Published Online: January 13, 2026; <https://doi.org/10.59717/j.xinn-med.2026.100193>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Guo Y., Ren G., Ma S., et al. (2026). Generative AI for brain-computer interfaces decoding: Advances, challenges and future. *The Innovation Medicine* 4:100193.

Brain-computer interfaces (BCIs) establish direct communication pathways between the human brain and external devices, enabling novel modes of human-machine interaction and providing essential technological foundations for innovative clinical therapies. BCI decoding can be divided into discriminative and generative approaches. Discriminative decoding models are primarily used for predefined tasks such as emotion recognition, limited performance in reconstructing high-dimensional semantic content such as natural language or visual images. Recent advances in generative artificial intelligence (AI), including diffusion models and autoregressive Transformers, have brought generative decoding to the forefront of BCI research by enabling the reconstruction of semantically rich content from neural activity. Current studies further demonstrate that decoding visually evoked neural activity into images can achieve structural similarity index measure (SSIM) scores ranging approximately from 0.26 to 0.43. Furthermore, generative AI enhances BCI decoding through data augmentation and sensor optimization, thereby improving both performance and generalizability. Here, we review recent progress in generative AI-enabled BCI decoding, focusing on methodological advances in language and visual decoding. We also summarize its roles in data augmentation and sensor optimization, outline key challenges in decoding accuracy, cross-subject generalization and fairness, and suggest future directions such as multi-modal integration and fairness-aware frameworks.

INTRODUCTION

BCI technology, conceptualized in the 1970s, has long aimed to establish a direct communication channel between the human brain and external devices.^{1–7} It has been applied across diverse domains, including clinical rehabilitation,^{8–11} human-computer interaction,¹² and biometric identification,^{13–15} among others. Decoding refers to the process of translating neural activity into interpretable outputs, and is the central component of BCI systems. To provide a clear conceptual framework, we categorize BCI decoding into two fundamental paradigms: discriminative decoding, which maps brain signals to predefined categories, and generative decoding, which reconstructs semantically rich and open-ended content directly from neural activity. Whereas discriminative decoding is mainly limited to classification tasks such as visual category recognition^{16–19} and emotion identification,^{20–23} its closed task structures and limited semantic expressivity confine it to recognizing only predefined, discrete intentions.^{24,25} As a result, it fails to capture the continuous cognitive content naturally emerging in the brain, including language, complex visual scenes, and emotional states. In contrast, generative models aim to produce higher-level semantic content-including text,²⁶ speech,²⁷ and images,^{28,29} that better reflects the richness of brain activity. This transition represents not only a technological advance but also a conceptual shift in BCI decoding, moving systems from simple command recognition toward authentic semantic reconstruction and cognitive restoration, thereby opening the way for more natural and efficient human-computer

interaction.³⁰

Generative AI models, such as autoregressive transformers,³¹ Variational Autoencoders (VAEs),³² and diffusion models,³³ provide the foundational capabilities for generative BCI decoding and are reshaping the technological landscape of BCI. By leveraging these models, generative decoding has expanded the scope of BCI tasks, enabling the reconstruction of complex content such as natural language,^{34–36} speech,^{37,38} and images^{39–41} from brain signals, and even supporting reverse generation for bidirectional BCI.⁴² Moreover, generative AI enhances decoding performance via data augmentation,^{43–46} facilitating cross-modal generation^{47–50} and data synthesis^{51–53} to provide abundant training resources for high-performance BCI decoding models. Concurrently, it is increasingly used to assist BCI sensor optimization, improving electrode design and enabling the discovery of novel flexible materials to enhance signal acquisition quality,^{54–57} thereby strengthening the overall decoding capabilities of BCI systems.

In BCI language decoding,⁵⁸ generative AI models have been used to restore sentences that individuals hear,⁵⁹ speak,¹¹ or think,^{6,60} breaking through the limitations of traditional BCI in terms of vocabulary scale and expressive freedom. Application scenarios include assisting communication³⁴ for people with language disorders and building silent communication systems. Research cases include neural translation from microelectrode arrays (MEA) signals to text,⁶¹ as well as speech synthesis driven by electrocorticography (ECoG)^{3–62} or electroencephalogram (EEG)^{11,63} recordings. These studies not only demonstrate the strong capabilities of Generative AI in BCI language decoding but also provide new ideas for the rehabilitation of language disorder patients and the optimization of human-computer interaction in the future.⁶⁴ Through Generative AI, BCI systems can directly convert brain signals into natural language, enabling users to express their thoughts more naturally.

In BCI visual decoding, brain signals (such as fMRI,⁶⁵ EEG^{66,67} and MEG¹⁹) are modeled and combined with generative AI models to decode visual stimuli into images^{68,69} and videos,⁷⁰ which in most studies are based on externally presented visual inputs.⁷¹ In addition, generative AI has also been used to decode dreams⁷² or visual imagination,⁷³ which reflect visual content generated within the brain. Related research has shown that generative AI models can effectively improve the spatial resolution and semantic consistency of decoded images, providing important technical support for cognitive neuroscience⁷⁴ and visual prosthetics.^{75,76}

Beyond decoding algorithms, Generative AI also plays an important role in enhancing BCI decoding through two key dimensions: data augmentation and sensor optimization. On the data side, generative models enable cross-modal generation,⁴⁸ virtual neural signal synthesis,⁴⁵ and super-resolution reconstruction,^{44,77} effectively enriching training samples and improving data efficiency under limited-data regimes.^{78,79} On the hardware side, generative models are increasingly used to support the design and optimization of neural sensors. By accelerating the discovery of candidate materials and electrode architectures, they contribute to the development of sensors with

improved biocompatibility, conductivity, and signal fidelity.^{54,56}

This review first examines recent methodological advances in generative AI-enabled BCI decoding, with dedicated sections on language and visual reconstruction. It then explores the indirect yet important roles of generative AI in enhancing decoding performance through data augmentation and sensor optimization. Subsequently, it discusses three major challenges—decoding accuracy, cross-subject generalization, and fairness—together with emerging strategies to address them. The review concludes by summarizing key trends and outlining future research directions toward the practical deployment of generative AI-enabled BCIs.

THEORETICAL FOUNDATIONS OF GENERATIVE AI FOR BCI DECODING

Generative AI fundamentally differs from traditional discriminative modeling in its approach to modeling neural activity for BCI decoding. In discriminative modeling, neural signals are mapped into predefined categories, whereas generative AI learns a conditional distribution over a continuous output space and directly produces semantically meaningful content from neural activity. This modeling shift moves beyond selecting from fixed categories and instead supports the generation of rich, structured outputs driven by learned conditional distributions.^{33,80,81}

More formally, let X denote neural activity and $\mathcal{C}$ a discrete label space. A discriminative model predicts the most likely category through

$$\hat{c} = \arg \max_{c \in \mathcal{C}} P_{\theta_d}(c | X), \quad (1)$$

where θ_d denotes the model parameter set. This formulation reflects a decision process defined over a finite label set.

In contrast, generative AI models a conditional distribution over a continuous output space $\mathcal{Y}$ and produces an output by

$$\hat{Y} = \arg \max_{Y \in \mathcal{Y}} P_{\theta_g}(Y | Z), \quad Z = f_{\theta_e}(X), \quad (2)$$

where θ_g parameterizes the generative model and f_{θ_e} maps neural activity X into a latent representation Z . This formulation demonstrates how generative AI can produce semantically rich content based on neural activity.

This perspective highlights a shift from classification-style decisions toward open-ended generation. Neural activity is embedded into a latent space that provides a compact and abstract representation of neural dynamics, where embeddings with related semantic content lie in close proximity. Generative AI builds on these latent representations to synthesize language, speech, and visual content at higher semantic levels. The sections that follow build on this theoretical foundation to examine algorithmic advances, application domains and enabling technologies,^{31,33,43,45,82} as well as the challenges that remain for generative AI-enabled BCI systems.

GENERATIVE AI FOR BCI DECODING METHODS

From discriminative decoding to generative decoding for BCI

Most early BCI decoding studies focused on discriminative models that classify neural signals into predefined categories, for example in motor imagery or emotion recognition tasks.^{83–85} While such methods achieve good accuracy in restricted settings, their reliance on closed output spaces prevents them from addressing richer forms of cognition.⁸⁶

Recent work has therefore explored generative decoding, which directly reconstructs open-ended content such as text, speech, or images from brain activity.^{33,82,87} Enabled by advances in generative artificial intelligence, these models can generate high-dimensional outputs that align more closely with natural cognitive processes.^{31,81} Demonstrations in natural language,^{82,87} continuous speech,⁸⁸ and visual imagery⁸⁹ illustrate the potential of this direction, motivating the following sections on language and visual decoding.

Overall, the shift from discriminative to generative decoding not only enhances the expressive range of BCI systems, but also moves the field toward authentic semantic reconstruction and cognitive restoration (Figure 1).^{31,33} Table 1 summarizes representative generative decoding tasks together with their reported performance metrics.

BCI language decoding with generative AI

Language decoding in BCI refers to the process of translating neural activ-

ity into interpretable linguistic outputs, typically in the form of natural language text or synthesized speech (Figure 2). This capability is fundamental for restoring communication in individuals with severe paralysis or speech impairments. Traditional approaches have been constrained by predefined vocabularies or fixed linguistic categories, which limit the expressive capacity and generalizability of BCI systems. Recent advances in generative AI, particularly the development of large language models (LLMs), have fundamentally transformed the field by enabling the direct mapping of complex brain signals to open-ended linguistic outputs, including continuous text generation and speech synthesis.

BCI text decoding. Text decoding has emerged as a fundamental research direction in BCI language decoding, aiming to reconstruct text directly from neural activity. Current studies span both invasive and non-invasive signal acquisition modalities.

Invasive BCIs typically leverage ECoG or MEA to record cortical neural activity. Early efforts focused predominantly on discriminative classification tasks, such as intent-driven spelling devices^{93,94} and word-pair classification.⁹⁵ With the introduction of generative AI models, invasive approaches have gradually evolved from simple classification to open-vocabulary text reconstruction. Notably, several studies employing high-density ECoG arrays have achieved continuous speech-to-text reconstruction, significantly reducing word error rates and demonstrating real-time decoding of natural language.^{96–100} The adoption of representation learning and transformer-based generative models has further enabled large-vocabulary and sentence-level neural text decoding, substantially advancing the prospects of neural speech prostheses.^{101,102} Moreover, the transfer of motor-related decoding paradigms to language applications has yielded important breakthroughs, including real-time decoding of imagined handwriting and inner speech signals, highlighting the stability of neural representations and the adaptive capabilities of current models.^{6,61,64}

In parallel, non-invasive BCI text decoding has also witnessed rapid progress, primarily utilizing fMRI, MEG, and EEG. Compared with invasive recordings, these modalities are safer and more scalable but inherently suffer from lower signal-to-noise ratios (SNRs), which pose significant challenges for fine-grained language decoding. Early research in this domain mainly addressed the discriminative classification of basic semantic categories, analyzing neural responses to different linguistic stimuli and revealing the distributed nature of semantic representations in the brain.^{59,103} Subsequently, decoding efforts expanded from coarse semantic classes to individual words and phrases, marking initial steps toward reconstructing continuous natural language semantics.^{104,105}

Building on these developments, the emergence of generative AI further accelerated non-invasive text decoding. Leveraging autoregressive LLMs such as GPT, non-invasive approaches have achieved unprecedented progress. Notably, these advances include the first demonstrations of open-vocabulary and continuous text generation from brain signals.^{106–108} To address challenges posed by the low SNR and substantial inter-subject variability of EEG data, recent work has explored multimodal fusion, cross-subject representation alignment, and self-supervised contrastive learning, yielding notable gains in open-vocabulary EEG text decoding.^{26,109,110} In parallel, strategies involving cross-modal semantic encoding and joint training with language models have effectively bridged representation gaps, enhancing both the semantic consistency and generalizability of decoded text.^{111–113} For fMRI and MEG, end-to-end natural text generation frameworks have enabled direct mapping from brain signals to text embedding spaces, resulting in marked improvements in decoding accuracy and semantic relevance.^{27,86,114} In particular, the combination of MEG's high temporal and spatial resolution with transformer models has achieved real-time, sentence-level continuous decoding, highlighting the considerable practical potential of non-invasive BCIs.^{115,116}

More recently, pioneering studies have explored multilingual decoding^{117,118} as well as reconstructing semantically relevant textual descriptions of images from neural responses to visual stimuli,^{119–121} further expanding the application scenarios and task boundaries of BCI text decoding. At the methodological level, end-to-end LLM-based frameworks demonstrate superior performance in multimodal representation fusion and generalization tasks, while the integration of audiovisual semantic fusion⁸⁴ and multi-task joint decoding

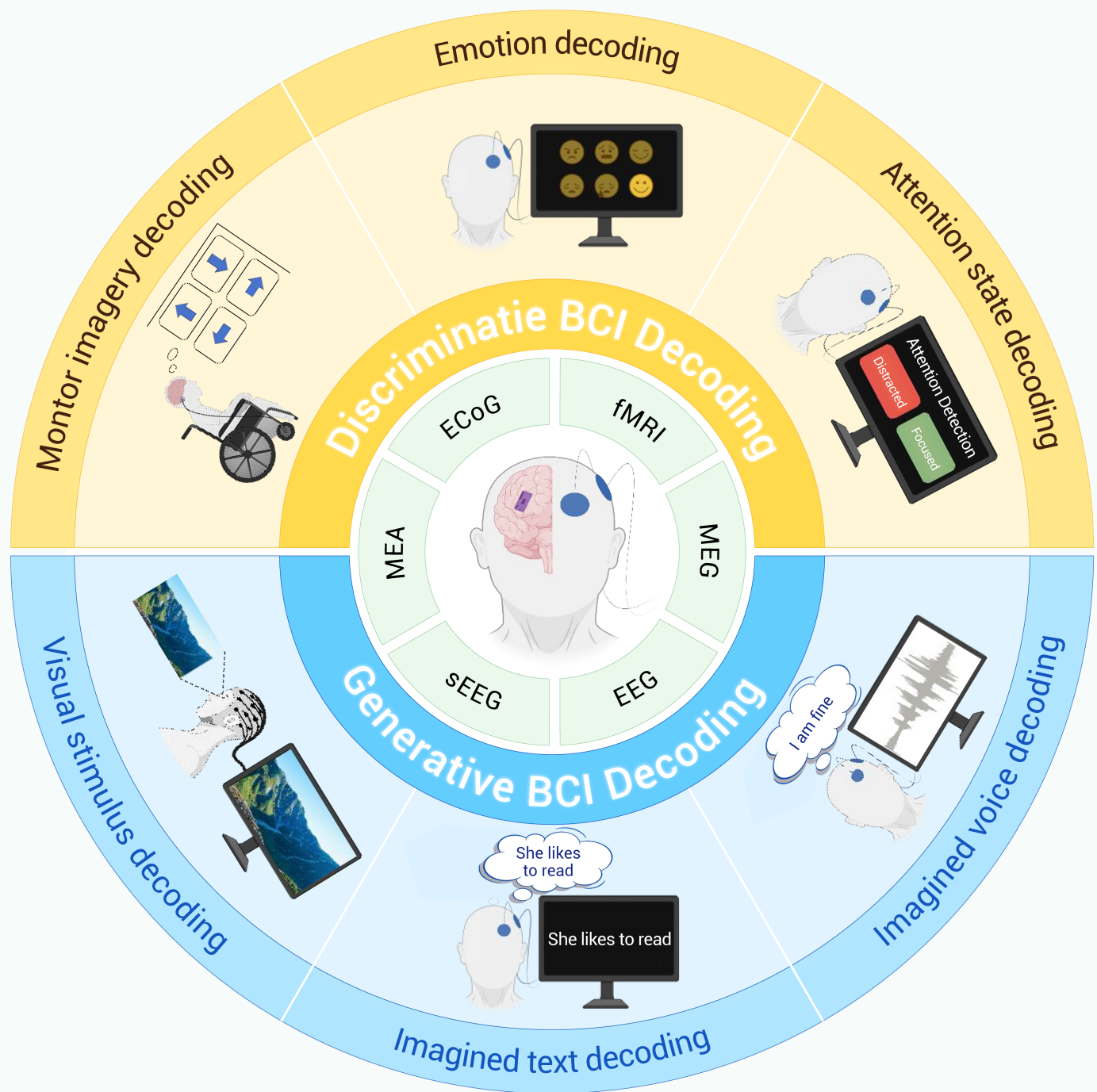


Figure 1. Schematic illustration of discriminative and generative brain decoding Discriminative decoding classifies brain activity into predefined categories. Generative decoding reconstructs open-ended outputs, such as text or images, directly from brain signals.

strategies^{122,123} has broadened applicability to a wider range of real-world scenarios. Advanced neural prosthetic systems now enable multimodal synchronous control centered on text decoding, while also integrating speech and virtual avatar facial animations,^{37,124} underscoring the rapidly expanding frontiers of BCI text decoding.

BCI audio decoding. In recent years, audio decoding via BCIs has become a key technology for restoring communication in individuals with neurological impairments. Unlike text decoding, which can be converted into generic speech via text-to-speech synthesis, audio decoding directly reconstructs natural speech signals with speaker-specific characteristics such as timbre

and prosody, thereby enabling more expressive and personalized communication. Research in this domain spans both invasive and non-invasive recording modalities, leveraging advanced generative AI techniques to reconstruct clear and intelligible speech from neural activity.

Invasive approaches have achieved substantial progress in audio decoding. Early studies used ECoG or MEA to record cortical activity and employed linear regression or classical algorithms to decode and synthesize vowels and basic speech units.^{58,88,125} With the introduction of deep learning, including convolutional and recurrent neural networks, speech quality has markedly improved.^{126–128} The integration of transformer-based encoders with neural

Table 1. Representative studies of generative BCI decoding with performance metrics.

Task Type	Neural Modality	Year	Generative Decoding Method	Performance Metrics	Reference
Text Decoding	ECoG	2023	HuBERT-guided decoding	WAcc = 0.74	Metzger et al. ³⁷
Text Decoding	fMRI	2023	LLM-based semantic decoding	WAcc = 0.08 BERTScore = 0.81	Tang et al. ³¹
Audio Decoding	MEA	2025	Autoregressive Transformer	WAcc = 0.72 MCD = 2.90	Wairagkar et al. ¹¹
Image Decoding	fMRI	2024	CLIP-aligned diffusion	SSIM = 0.43 CLIP = 0.93	Scotti et al. ⁹⁰
Image Decoding	fMRI	2024	Subject-invariant diffusion	SSIM = 0.26 CLIP = 0.94	Wang et al. ⁹¹
Video Decoding	fMRI	2024	Temporal video diffusion	SSIM = 0.39 CLIP = 0.73	Gong et al. ⁹²

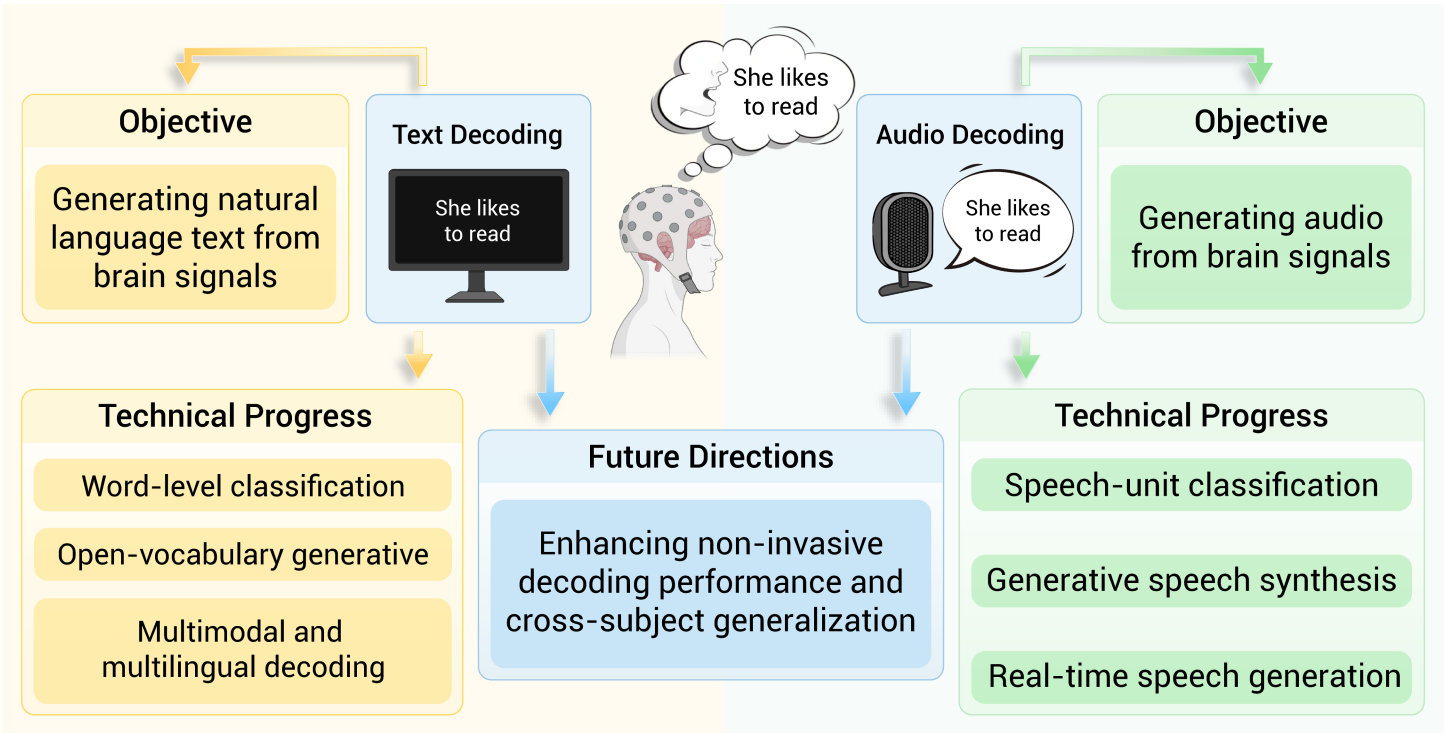


Figure 2. Schematic of BCI language decoding Generative models enable text and audio decoding, while improving non-invasive performance and cross-subject generalization remains challenging.

vocoders has enabled the accurate reconstruction of natural spoken sentences directly from ECoG signals,^{129–132} and transformer models have demonstrated potential for transfer from overt to imagined speech decoding tasks.⁶⁰

Building on these advances, generative AI frameworks have emerged as the foundation of state-of-the-art end-to-end audio decoding, excelling in real-time synthesis, naturalness, and the preservation of speaker-specific vocal features.^{9,38} Neural prosthetic systems leveraging SIEG, ECoG or MEA now achieve real-time speech synthesis for communication restoration,^{133–135} with extensions to prosodic modulation¹³⁴ and even the restoration of singing in a pioneering study.¹¹ Importantly, these systems have demonstrated clinical feasibility and long-term stability in implant trials.^{136,137} Beyond cortical signals, recent findings indicate that subcortical structures such as the thalamus also carry rich language-related information,^{138,139} underscoring the potential of combining cortical and subcortical signals to further advance generative audio decoding.¹²⁷

In non-invasive BCIs, the low SNR of externally recorded signals poses inherent challenges for audio decoding. Early work focused on reconstructing speech envelopes and basic acoustic features, and subsequent deep nonlinear models improved quality and intelligibility beyond linear

baselines.^{140,141} To model long, continuous speech, multiscale fusion networks and state space models have been developed to process extended EEG sequences and reconstruct mel-spectrograms.¹⁴² More recently, generative AI has been introduced into non-invasive EEG-based audio decoding, leveraging pre-trained text-to-speech systems to achieve superior alignment between EEG and audio representations and to directly generate natural, high-quality speech.^{143,144} In parallel, generative AI frameworks that integrate temporal convolutional feature extraction with language modeling have enabled the online reconstruction of continuous, natural sentences from EEG signals.¹⁴⁵ Beyond EEG-based audio decoding, fMRI has also been employed for non-invasive brain-to-speech reconstruction, where deep auditory features extracted from brain activity are combined with generative AI models to reconstruct the sounds actually perceived by the subject.¹⁴⁶ Collectively, these advances highlight the emerging central role of generative AI in bridging non-invasive neural signals and complex speech generation, laying the groundwork for practical, portable, and cost-effective BCIs.^{143,147}

Overall, the rapid development of audio decoding in BCIs is driven by deep learning, and especially by the introduction of generative AI. Both invasive and non-invasive research lines have demonstrated significant clinical promise, substantially broadening communication options for patients with neurologi-

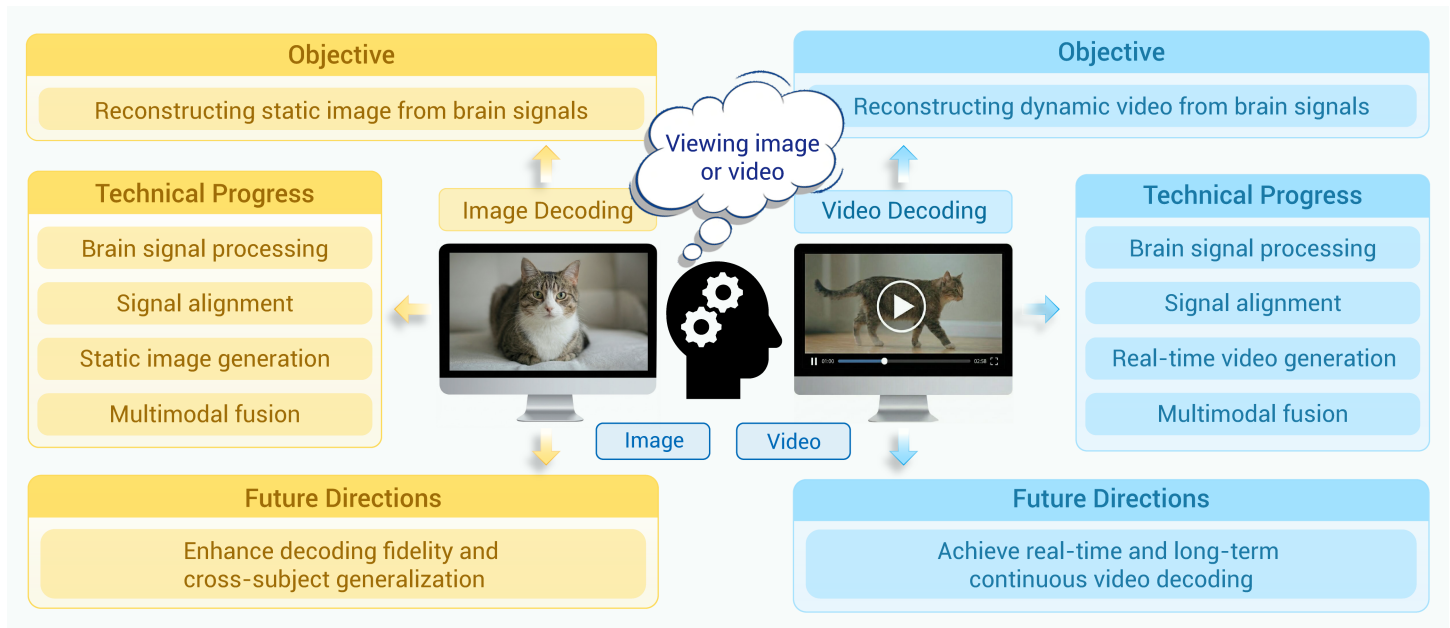


Figure 3. Schematic of visual decoding from brain signals Progress in generative modeling drives improvements in fidelity, semantic consistency, and cross-modal integration.

cal injuries, and advancing BCI technology toward more natural, real-time, and emotionally expressive human-machine interaction.

BCI visual decoding with generative AI

Recent advances in visual decoding leverage generative AI techniques to substantially enhance reconstruction fidelity and semantic consistency. This progress spans multiple domains, including static image decoding and dynamic video decoding, collectively facilitating richer, more accurate visual content reconstruction from brain signals (Figure 3). Most existing studies in this area decode neural responses to externally presented visual stimuli, as these paradigms provide stable and well-controlled experimental conditions.⁷¹ Research in this area has been driven primarily by non-invasive modalities such as fMRI and EEG, which offer broader accessibility and the availability of large-scale datasets. While invasive recordings provide higher spatial/temporal fidelity, their use is limited by surgical risks and clinical indications; consequently, non-invasive modalities currently dominate human studies, with invasive work more common in specific clinical or animal research contexts.¹⁴⁸ Overall, the landscape of BCI visual decoding has thus far been shaped predominantly by non-invasive approaches.

As illustrated in Figure 4, the transformations along the visual pathway map closely onto the encoder, embedding, and decoder stages of brain signal visual decoding. From a representational perspective, the latent features exploited by generative AI exhibit a strong correspondence with the functional hierarchy of the visual cortex. Structure-sensitive components align more closely with responses in early visual areas such as V1, whereas semantic-rich latent dimensions better predict activity in downstream ventral temporal cortex.^{149,150} This hierarchical correspondence offers a compact explanation for how generative models recover both fine spatial structure and high-level semantics from non-invasive neural signals.

BCI image decoding. Image decoding is a core direction within BCI decoding research, aiming to reconstruct static visual content from brain signals and capture an individual's perceptual experience at a specific moment.^{151–153} This includes both simple and complex image reconstruction tasks, and serves as a valuable tool for probing human visual perception and the neural mechanisms underlying the processing of static visual stimuli.^{82,154,155} In addition to its neuroscientific value, image decoding also holds creative potential—enabling, for example, the translation of imagined visual scenes into tangible images, thereby augmenting artistic workflows. The development of BCI-based image decoding can be traced back to early studies that focused on reconstructing simple visual stimuli from neural activity, such as black-and-white patterns,¹⁵⁶ handwritten digits,¹⁵⁷ and human faces.^{158,159} These pioneering efforts demonstrated that general shapes and spatial information could

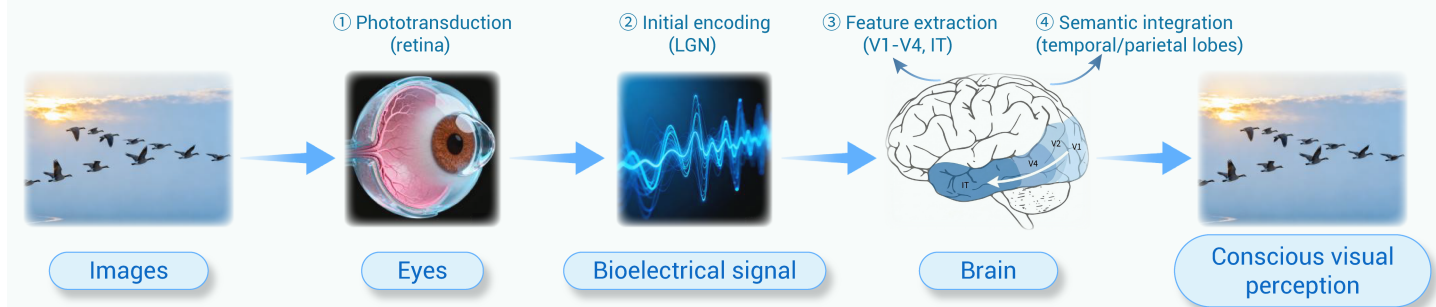
be inferred from brain signals, laying the groundwork for subsequent advances toward decoding richer and more semantically complex visual content.^{160,161}

With the rapid development of deep learning technologies, the field of BCI image decoding has witnessed remarkable progress.^{162,163} Deep learning-based methods possess a more powerful capacity for feature capturing and integration.^{16,164} They can learn to map nerve activities in the visual cortex to the visual stimuli observed during the perceptual process through joint training of generation and adversarial learning. By training a deep generative adversarial network and combining it with a perceptual loss term, the model is able to reconstruct images that are highly similar to the original stimulus images. Due to the difficulty in obtaining paired brain signal and visual stimulus image training samples, many methods apply semi-supervised,^{165,166} self-supervised¹⁶⁷ and unsupervised¹⁶⁸ techniques to the pre-training process of generative models to enhance the models' visual representation capabilities.

In recent years, generative AI models, particularly GANs and diffusion models, have emerged as key technologies for decoding brain activity and reconstructing perceived images. GANs have demonstrated their ability to generate corresponding images from brain signals (e.g., thoughts of numbers, characters, or objects), providing strong evidence that brain signals encode visual cognitive content.^{169,170} For EEG signals, researchers have developed more advanced GAN frameworks. For instance, combining stacked LSTM to learn compact and noise-free representations of EEG data has been used to generate specific visual stimuli.^{171,172} Specialized architectures, such as the Dual-Condition and Lateralization-Supported GAN (DCLS-GAN), have further enhanced the ability to visualize image-evoked thoughts from EEG signals.¹⁷³ In the fMRI domain, GANs have been successfully applied to address inter-subject heterogeneity by reconstructing visual images from fMRI data acquired while participants perceive images, significantly improving accuracy.^{66,152} By integrating techniques such as VAEs, GANs have achieved robust decoding from fMRI data and even accomplished tasks like gender classification.¹⁵⁹ GANs that incorporate semantic feature similarity as a conditioning signal, such as the Similarity-Conditioned GAN (SC-GAN), have significantly improved the similarity between reconstructed images and the original stimuli.¹⁷⁴ Collectively, these works highlight the substantial potential of GANs in enhancing the precision and interpretability of brain decoding.

Diffusion models, with their powerful generative capabilities and flexibility, have shown remarkable performance in brain decoding, especially in fMRI image reconstruction. Some methods have designed multi-stage processes, such as first aligning semantics and then generating details,^{175,176} or combining multiple generative models, such as masked autoencoders in combina-

A Hierarchical visual information processing in the brain



B Hierarchical visual information decoding from brain signals

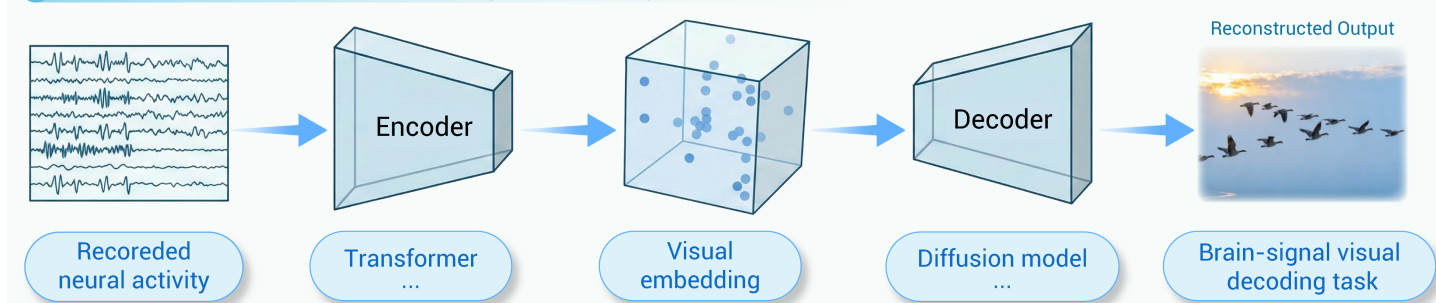


Figure 4. Visual information flow and its correspondence to brain signal visual decoding (A) Visual information progresses from phototransduction in the retina to contrast coding in the lateral geniculate nucleus, then to feature extraction in visual areas V1 to V4 and inferotemporal cortex, and ultimately to semantic perception in higher cortical regions. (B) corresponding visual decoding pipeline, in which neural activity is transformed through an encoder, an intermediate visual embedding, and a generative decoder to reconstruct perceptual content, reflecting a hierarchical organization analogous to biological visual processing.

tion with latent diffusion models,⁸⁷ or VDVAE coupled with BLIP and a universal diffusion model,^{29,177} to hierarchically recover visual and semantic details. Other approaches have employed data-driven methods to synthesize images that activate specific brain regions;¹⁷⁸ improved classification and retrieval performance by linearly mapping to semantic features;¹⁷⁹ and achieved state-of-the-art performance in image retrieval and reconstruction tasks by integrating contrastive learning and diffusion priors.^{33,90} In addition, to alleviate the problem of insufficient training data, some methods have adopted self-supervised learning (especially masked modeling) to learn robust and information-rich fMRI representations.^{87,180} Diffusion models have also been applied to EEG signal decoding. By leveraging the strong transferability of pre-trained text-to-image diffusion models and combining techniques such as temporal masked signal modeling to address signal noise and individual variability, high-quality image generation from EEG has been achieved.¹⁸¹ Other diffusion-based frameworks have further improved reconstruction quality by optimizing internal mechanisms and introducing edge estimation strategies.^{81,182}

Furthermore, the further development of multimodal data fusion technology is also an important direction for the future.^{183,184} By combining more modalities of data, the precision and robustness of image decoding can be further improved.¹⁸⁵ At the same time, refining cross-subject decoding technology is also a focus of future research. Developing cross-subject decoding models with strong generalizability will help promote the widespread application of image decoding technology.^{91,186} For instance, combining fMRI signals with image data can reconstruct visual content.^{175,187–189} The combination of these modalities provides a richer set of information, thereby enhancing the accuracy and reliability of decoding. By combining brain signals with other modalities (e.g., text descriptions, clinical information, etc.), richer information can be provided for visual decoding, which better constrains the generated visual content and thereby helps the generative models to reconstruct high-quality images from brain activity. Feature disentanglement and semantic representation are crucial for interpreting neural responses and have played an important role in shaping the high-level neural representations underlying visual perception.^{42,168,190}

BCI video decoding. Brain-evoked signals encode not only static images but also temporally continuous dynamic scenes.^{89,179} Effectively extracting

and decoding these temporal representations promises new insights into the dynamic mechanisms of the human visual system and will advance the next generation of brain-computer video interfaces (BCVIs).¹⁹¹ One of the most challenging tasks in this domain is video information decoding, whose goal is to classify perceived dynamic stimuli¹⁹² or to reconstruct individual frames/short clips from fMRI-measured brain activity.¹⁹³ In recent years, researchers have made significant progress in this field, proposing a variety of innovative methods to improve the quality and efficiency of video reconstruction. For example, NeuralFlix employs spatial and temporal enhancements to learn fMRI representations and utilizes a diffusion model enhanced by prior noise to generate videos.¹⁹⁴ Additionally, Brain Netflix accurately generates two seconds and three seconds video clips from brain activities of different participants and datasets through multi-dataset and multi-subject training.¹⁹⁵ These studies not only demonstrate the potential for reconstructing dynamic visual experiences from fMRI signals but also provide an important technical foundation for the development of future BCVIs.

Recently, three-dimensional convolutional networks (3D-CNNs) and two-stream spatiotemporal networks^{196,197} have been shown to generate hierarchical spatiotemporal features that bridge dynamic visual stimuli and brain signals. To mitigate the scarcity of paired training data, researchers first pre-train a 3D-CNN on large-scale video datasets such as Kinetics-400 and Moments-in-Time, and then use the resulting spatiotemporal features as the decoding target for reconstructing natural video clips.

In recent years, the rapid advancement of generative AI has spurred an explosive growth in methods that decode brain signals to generate videos. Diffusion models, owing to their powerful generative capabilities, have been widely adopted for brain-based video decoding tasks. One line of work¹⁹⁸ employs a cascaded diffusion backbone conditioned on text prompts to synthesize high-frame-rate, high-fidelity videos from fMRI, adding motion-consistency regularizers that markedly reduce temporal jitter. NeuroClips⁹² inserts spatiotemporal Transformer blocks into the diffusion trunk, using local-global attention for smooth frame transitions and enabling minute-long sequence reconstruction. Others¹⁷⁶ couple diffusion priors with adversarial discriminators and cross-frame contrastive losses to refine dynamic details and achieve the first faithful recovery of natural motion scenes. To harness richer multimodal cues, certain methods⁷⁰ introduce text-image-video tri-

modal alignment, allowing users to steer reconstruction via language instructions. To mitigate label scarcity, self-supervised diffusion pipelines have been proposed that reconstruct long videos from fMRI without paired annotations. EEG2Video¹⁹⁹ pioneers joint training of noisy EEG signals with text-conditioned diffusion, achieving 512×512 dynamic-scene decoding through cross-modal alignment and laying the groundwork for portable brain-video interfaces.

Compared to images, videos contain richer information and more complex scene transitions, which pose significant challenges for decoding extensive visual stimuli from brain signals. Some studies have attempted to reconstruct high-quality video content from brain activity by integrating brain signals with other semantic information.⁸³ Multimodal data joint decoding offers significant advantages for brain decoding research. First, by integrating data from different modalities, it provides a more comprehensive view of brain activity and enables more accurate capture of complex brain activity patterns.⁸³ Second, multimodal decoding enhances the robustness and accuracy of decoding, with the complementary nature of different modality data compensating for the shortcomings of single-modality data.²⁰⁰ Moreover, multimodal decoding can reveal the intrinsic connections between different modality data, helping researchers better understand how the brain processes and integrates information from different senses.²⁰¹ This integrative approach offers new perspectives and tools for in-depth research into the neural mechanisms of the brain. These studies have further optimized the performance of multimodal decoding and provided technical support for real-time applications of video decoding.

Recent advances in generative decoding also point toward potential clinical relevance for visual rehabilitation. Although current retinal prostheses can partially restore vision in disorders such as age-related macular degeneration, their perceptual quality remains constrained by the resolution and signal capacity of the implant.²⁰² Generative AI may help overcome these limitations by transforming degraded or low-bandwidth visual input into simplified and more interpretable representations, as demonstrated by PVGAN for prosthetic vision enhancement.²⁰³ At the same time, high-channel cortical stimulation has shown that the visual cortex can integrate structured electrical patterns into coherent shape percepts, supporting the feasibility of combining prosthetic stimulation with generative AI reconstruction in future rehabilitation approaches.²⁰⁴ Building on the ability of generative models to produce content across modalities, reconstructed visual information can be converted into auditory signals that assist individuals with severe visual impairment in interpreting their surroundings. There is also potential for the same information to be conveyed as tactile patterns generated through peripheral or cortical electrotactile stimulation, providing an additional non-visual pathway when auditory or residual visual cues are insufficient. Together, these possibilities suggest that generative decoding could support visual prosthetics, perceptual training, residual-function assessment and sensory-substitution approaches in patients with damage to the eyes or early visual cortex.

GENERATIVE AI ENABLES BCI DECODING THROUGH DATA AUGMENTATION AND SENSOR OPTIMIZATION

Beyond its direct use in BCI decoding algorithms, generative AI is also enhancing BCI decoding by improving upstream components, including data augmentation and support for sensor optimization (Figure 5). The following sections detail these two directions.

Generative AI-enabled data augmentation for BCI decoding

As summarized in Table 2, each neural recording modality used in BCI systems presents specific strengths and limitations that directly influence decoding performance. Non-invasive techniques such as EEG and fNIRS are portable and cost-effective, but have limited spatial resolution and are sensitive to artefacts. MEG offers high temporal precision yet requires a magnetically shielded environment and is non-portable, while fMRI provides excellent spatial detail but has low temporal resolution and limited applicability in real-time settings. Invasive methods including ECoG, SEEG and MEA achieve high signal quality and access to deep brain structures, but involve substantial surgical risks, high procedural costs and restricted data availability. These trade-offs constrain the scalability, robustness and generalizability of BCI decoding.

Generative AI is increasingly leveraged to improve the quality of BCI data and thereby enhance decoding performance. Neural signal super-resolution reconstructs high-resolution signals from low-resolution recordings in both the temporal and spatial domains, with recent studies showing that generative AI can recover finer spatial organization and more stable temporal structure from low-density or low-quality neural recordings, leading to improved performance in downstream decoding tasks.^{44,53,205} Cross-modal generation transforms data from low-cost acquisition modalities into high-cost modalities while integrating information across multiple sources, typically by mapping signals into a shared latent space and reconstructing target modalities with richer information content, allowing accessible sensors to approximate the representational quality of high-cost systems and thereby strengthening decoding performance.^{47,49} Data synthesis generates realistic neural activity patterns under data-scarce scenarios, where generative models learn the underlying distribution of neural signals and produce additional samples that enhance model robustness and generalization in low-resource settings.^{45,52,206–207} Together, these approaches address complementary limitations in data acquisition, ultimately leading to more reliable BCI decoding.

Generative AI-enabled sensor optimization for BCI decoding

Apart from data augmentation, generative models are also being applied to BCI sensor optimization, where they serve as assistive tools for the discovery and refinement of electrode materials and architectures. In this setting, generative AI can learn patterns linking microstructural features to recording performance and generate synthetic variations that explore alternative material compositions as well as different electrode layouts. By accelerating the in silico evaluation of candidate materials and spatial configurations, generative AI supports more efficient identification of designs that improve conductivity, biocompatibility, and signal fidelity in neural interfaces.^{54,56} Generative AI-enabled sensor optimization ultimately aims to enhance the SNR of neural recordings and thereby improve the overall performance of BCI decoding.

CHALLENGES AND FUTURE DIRECTIONS OF GENERATIVE AI FOR BCI DECODING

Despite the transformative potential of generative AI in expanding the capabilities of BCI decoding, several critical challenges remain unresolved in practice.²⁰⁸ Chief among these are the limited decoding accuracy, poor cross-subject generalizability of BCI models, and fairness in BCI decoding. This section examines each of these challenges in turn and outlines potential directions for future research. Figure 6 conceptually summarizes these core challenges and the directions driven by generative AI.

BCI decoding reliability constraints

BCI decoding reliability remains limited, with performance largely determined by the neural signal acquisition modality.⁹⁹ Invasive BCIs can deliver high-fidelity recordings and enable more precise decoding, but their clinical use is restricted by ethical concerns and the risks of surgical implantation.²⁰⁹ Non-invasive BCIs are safer and more scalable, yet their decoding reliability suffers from low SNRs and reduced spatial resolution in external recordings. For example, in BCI text decoding, invasive ECoG systems can achieve character-level accuracies corresponding to a word accuracy (WAcc) of about 0.79,¹⁰⁰ whereas non-invasive fMRI decoding yields much lower performance, with WAcc around 0.08.³¹ Improving reliability under these constraints remains a central challenge, particularly for non-invasive systems where poor signal quality continues to be a major factor limiting stable decoding performance.

To improve the reliability of BCI decoding, generative AI leverages data augmentation techniques to enhance neural data quality.^{47–50} For instance, it could be explored for the online replacement of bad or missing EEG channels, a problem that frequently occurs in high-density EEG recordings and whose mitigation could help stabilize signal quality.²¹⁰ Enhancing data quality is particularly critical for applications where decoding robustness directly affects system performance. This capability would be especially valuable for scenarios such as biometric identification, where high reliability is essential and generative augmentation offers a promising way to strengthen the robustness of EEG-based identifiers.^{13–15}

Beyond issues related to BCI data quality, model hallucination and seman-

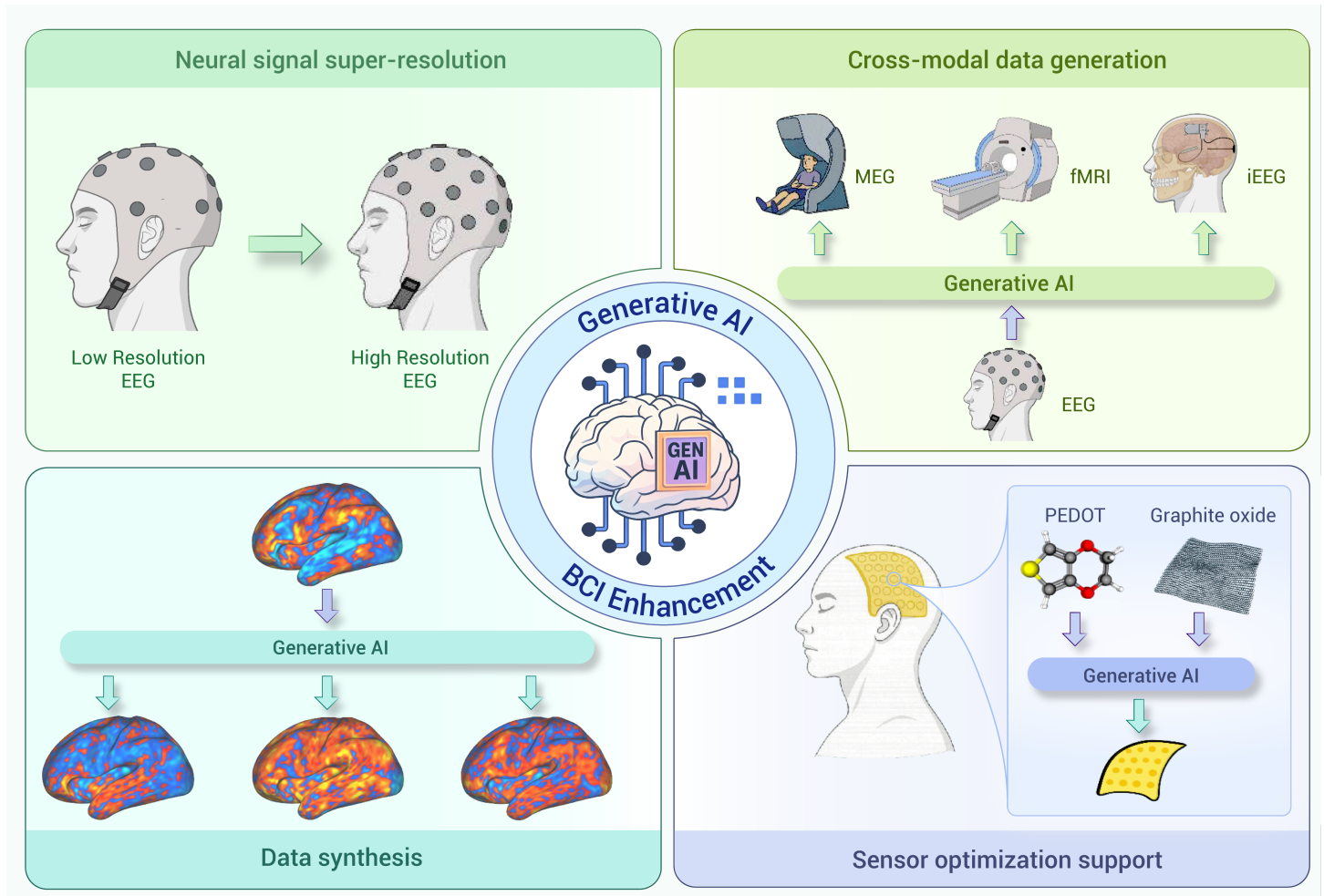


Figure 5. Schematic of generative AI-enabled enhancements in BCI decoding Generative artificial intelligence drives multidimensional improvements in neural decoding, including neural signal super-resolution, cross-modal data generation, data synthesis, and sensor optimization, collectively addressing key challenges in the reliability and scalability of BCI decoding.

Table 2. Summary of key strengths and limitations of representative BCI data modalities, and the potential of generative AI to mitigate these limitations.

Modalities	Invasive	SNR	Portable	Cost	Temporal Resolution	Spatial Resolution
EEG	Non-invasive	Low	Portable	Low	High	Low
MEG	Non-invasive	Moderate	Non-portable	High	High	Moderate
fNIRS	Non-invasive	Moderate	Portable	Moderate	Low	Low
fMRI	Non-invasive	Moderate	Non-portable	High	Low	High
ECoG	Invasive	High	Non-portable	High	High	High
SEEG	Invasive	High	Non-portable	High	High	High
MEA	Invasive	High	Non-portable	High	High	High

tic gaps also affect the reliability of generative BCI decoding. Hallucination occurs when the model generates outputs that appear coherent but are not genuinely supported by the underlying neural activity, reflecting an overreliance on learned priors. One potential direction for reducing this risk is to incorporate generative constraints that restrict model behavior to more principled representation manifolds.²¹¹ The semantic gap refers to the difficulty of ensuring that the latent features learned during decoding correspond to the true cognitive states they are intended to capture, which similarly undermines decoding reliability. Reducing this gap requires representation strategies that preserve behaviorally meaningful information rather than relying solely on statistical compression.⁷⁴

Achieving greater decoding reliability will require generative architectures

that can model the complex and dynamic patterns of brain activity, enabling more stable performance in challenging tasks. Yet the same advances that improve reliability may also increase the risk of re-identification, as models capable of enhancing individual-specific signal features could inadvertently turn neural recordings into biometric data. Under General Data Protection Regulation (GDPR), such data are strictly regulated as special category information, making it essential that progress towards higher reliability goes hand in hand with safeguards for privacy and security, ultimately paving the way for reliable and safe BCIs.

BCI cross-subject generalization challenges

Limited generalizability across individuals remains another major bottleneck for the widespread deployment of BCI decoding. Inter-subject differ-

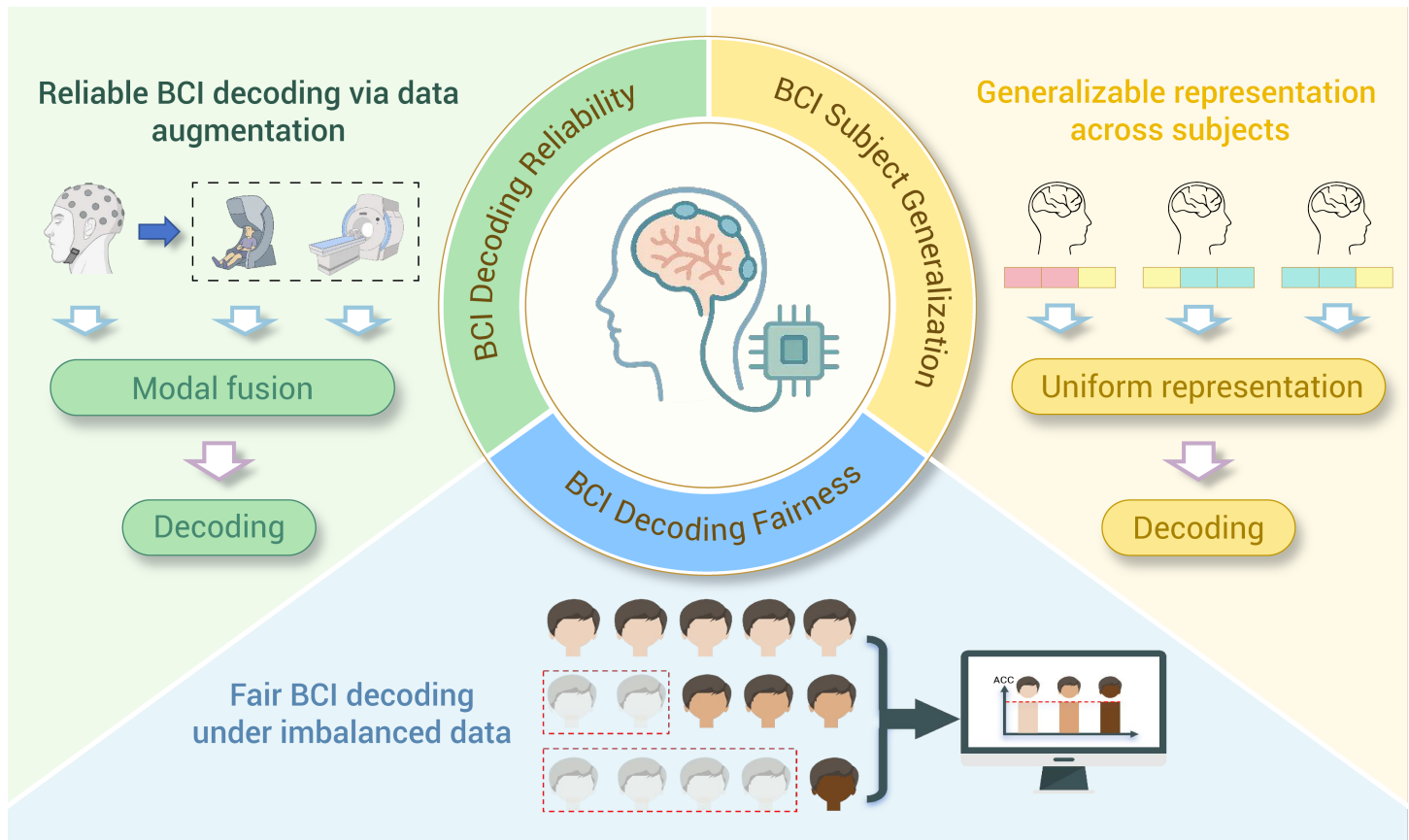


Figure 6. Schematic of key limitations in BCI decoding, encompassing decoding reliability, cross-subject generalization, and BCI decoding fairness.

ences in neural characteristics often cause decoding models to perform poorly on new users or under different conditions.^{31,212} In clinical populations, these differences are further amplified by heterogeneous and progressive pathology, where neural responses can evolve over time under the influence of disease progression, treatment, and rehabilitation. Ensuring that decoding model remain stable under such longitudinal changes is essential for their use in real-world healthcare settings. This sensitivity to user-specific characteristics means that models trained on one group or session may not transfer well to others, undermining the scalability of BCI technology. Generative AI offers a promising strategy to mitigate this challenge by learning the latent structure of neural data and extracting representations that are invariant across different subjects.

In practical terms, generative AI can be trained to learn abstract features of neural signals that emphasize patterns shared among individuals while de-emphasizing idiosyncratic, subject-specific details. By capturing the underlying structure of brain activity, a generative model can help filter out irrelevant variations—such as signal drift, electrode placement differences, or individual physiological noise—and retain the information that is most relevant for the decoding task. This approach has been increasingly recognized as a promising direction for improving cross-subject generalization in BCI systems.^{90,104,213–214} In essence, generative BCI decoding methods aim to build a kind of universal latent representation of neural data that remains consistent across users, thereby enabling models to transfer more effectively to new individuals without extensive retraining.

Despite the progress made with generative approaches, current models still fall short of achieving robust cross-subject generalization at scale. Many techniques remain limited in their ability to adapt to highly heterogeneous user populations and often require lengthy, subject-specific calibration sessions to maintain performance. These difficulties are further exacerbated by the data-hungry nature of generative AI, which require sufficiently diverse neural datasets to learn representations that generalize across individuals. Bridging this gap is an important future direction for the field. One priority is to develop generative frameworks that incorporate few-shot learning and

domain adaptation, allowing models to quickly adjust to new users with minimal data.²¹⁵ Another complementary approach is personalized pretraining strategies, where generative models are pre-trained on large, diverse neuroimaging datasets and then fine-tuned to individual variations.²¹⁶ Advancements in these areas are essential for building scalable and reliable BCI systems that can operate effectively across diverse individuals and real-world settings without sacrificing decoding accuracy.

Fairness challenges in BCI decoding

Fairness is emerging as a critical challenge for generative BCI decoding. Studies in medical AI have shown that imbalances in demographic representation, inconsistencies in data collection protocols, and the underrepresentation of certain groups can lead to uneven model performance across populations.^{217–218} Similar issues are increasingly observed in BCI decoding. Neural datasets are often collected from relatively small or homogeneous participant groups under specific experimental paradigms, meaning that systemic differences in neural signals related to gender, age, ethnicity, or other demographic factors may not be adequately represented. Consequently, BCI decoding models can display substantial performance disparities, with notably lower accuracy in underrepresented cohorts. These disparities raise serious concerns about the fairness and inclusiveness of current BCI systems, and data imbalance has become a key barrier to their scalable and equitable deployment.²¹⁹ Ensuring consistent decoding performance across all user groups remains a major challenge for the field.

Generative AI offers a promising avenue for mitigating fairness-related issues in BCI decoding. By learning the latent distribution of neural data, generative models can produce diverse, high-quality synthetic neural signals that supplement real datasets, helping to fill gaps in data coverage for underrepresented populations without the burden of large-scale new data collection. Such augmentation can create more demographically balanced training datasets and improve fairness by narrowing performance gaps across user groups.²¹⁹

However, using generative AI to enhance fairness also introduces chal-

lenges. Concerns remain regarding the safety, controllability, and fidelity of neural data produced by generative AI, as synthetic signals must accurately reflect real neural patterns without introducing new biases or artifacts. Future efforts should focus on robust validation of generated data and on integrating generative modeling with privacy-preserving, collaborative strategies. For instance, multi-center collaborations and federated learning could aggregate knowledge from diverse populations without centralizing sensitive data. These approaches represent a critical step toward building a fair, reliable, and socially responsible BCI ecosystem that benefits all user groups.²²⁰

CONCLUSIONS

The rapid development of Generative AI is reshaping the core capabilities of BCI systems.²²¹ By integrating deeply with BCI, generative models have driven advances in decoding paradigms, data augmentation, and sensor optimization, thereby enhancing the generalization and usability of BCI systems. To translate these advances into real-world impact, the future development of generative AI-enabled BCI requires multidisciplinary collaboration.²²² Computer scientists, neuroscientists, and materials scientists must work together to optimize algorithms, interpret brain signals, and innovate hardware, while policymakers, ethicists, and sociologists should ensure sustainability and social acceptance. Only through such joint efforts can generative AI-enabled BCI technology achieve widespread application in medical rehabilitation, human-computer collaboration, and neuroscience, ultimately contributing to human welfare.

REFERENCES

- Xu S., Liu Y., Lee H., et al. (2024). Neural interfaces: Bridging the brain to the world beyond healthcare. *Exploration* **4**:20230146. DOI:10.1002/exp.20230146
- Chaudhary U., Birbaumer N. and Ramos-Murguialday A. (2016). Brain-computer interfaces for communication and rehabilitation. *Nat. Rev. Neurol.* **12**:513–525. DOI:10.1038/nrneurol.2016.113
- Schwemmer M.A., Skomrock N.D., Sederberg P.B., et al. (2018). Meeting brain-computer interface user performance expectations using a deep neural network decoding framework. *Nat. Med.* **24**:1669–1676. DOI:10.1038/s41591-018-0171-y
- Degenhart A.D., Bishop W.E., Oby E.R., et al. (2020). Stabilization of a brain-computer interface via the alignment of low-dimensional spaces of neural activity. *Nat. Biomed. Eng.* **4**:672–685. DOI:10.1038/s41551-020-0542-9
- Ding Y., Udumpanyawit C., Zhang Y., et al. (2025). EEG-based brain-computer interface enables real-time robotic hand control at individual finger level. *Nat. Commun.* **16**:1–20. DOI:10.1038/s41467-025-61064-x
- Kunz E.M., Krasa B.A., Kamdar F., et al. (2025). Inner speech in motor cortex and implications for speech neuroprostheses. *Cell* **188**:4658–4673. DOI:10.1016/j.cell.2025.06.015
- Wei Q., Cui M., Liu Z., et al. (2025). Integrating statistical design and inference: A roadmap for robust and trustworthy medical AI. *Innov. Med.* **3**:100145. DOI:10.59717/j.xinn-med.2025.100145
- Schalk G., McFarland D.J., Hinterberger T., et al. (2004). BCI2000: A general-purpose brain-computer interface (BCI) system. *IEEE Trans. Biomed. Eng.* **51**:1034–1043. DOI:10.1109/tbme.2004.827072
- Littlejohn K.T., Cho C.J., Liu J.R., et al. (2025). A streaming brain-to-voice neuroprosthesis to restore naturalistic communication. *Nat. Neurosci.* **28**:902–912. DOI:10.1038/s41593-025-01905-6
- Santhanam G., Ryu S.I., Yu B.M., et al. (2006). A high-performance brain-computer interface. *Nature* **442**:195–198. DOI:10.1038/nature04968
- Wairagkar M., Card N.S., Singer-Clark T., et al. (2025). An instantaneous voice-synthesis neuroprosthesis. *Nature* **644**:145–152. DOI:10.1038/s41586-025-09127-3
- Willsey M.S., Shah N.P., Avansino D.T., et al. (2025). A high-performance brain-computer interface for finger decoding and quadcopter game control in an individual with paralysis. *Nat. Med.* **31**:96–104. DOI:10.1038/s41591-024-03341-8
- Del Pozo-Banos M., Alonso J.B., Ticay-Rivas J.R., et al. (2014). Electroencephalogram subject identification: A review. *Expert Syst. Appl.* **41**:6537–6554. DOI:10.1016/j.eswa.2014.05.013
- Hu Y., Sun L., Mao X., et al. (2024). EEG data augmentation method for identity recognition based on spatial-temporal generating adversarial network. *Electronics* **13**:4310. DOI:10.3390/electronics13214310
- Jolfaei A., Wu X.W. and Muthukumarasamy V. (2013). On the feasibility and performance of pass-thought authentication systems. *2013 Fourth International Conference on Emerging Security Technologies* pp:33–38. DOI:10.1109/est.2013.12
- Horikawa T. and Kamitani Y. (2017). Generic decoding of seen and imagined objects using hierarchical visual features. *Nat. Commun.* **8**:15037. DOI:10.1038/ncomms15037
- Fares A., Zhong S.h. and Jiang J. (2019). EEG-based image classification via a regionlevel stacked bi-directional deep learning framework. *BMC Med. Inform. Decis. Making* **19**:268. DOI:10.1186/s12911-019-0967-9
- Mukherjee P., Das A., Bhunia A.K. et al. (2019). Cogni-net: Cognitive feature learning through deep visual perception. *2019 IEEE International Conference on Image Processing* pp:4539–4543. DOI:10.1109/icip.2019.8803717
- Song Y., Wang Y., He H., et al. (2025). Recognizing natural images from EEG with language-guided contrastive learning. *IEEE Trans. Neural Netw. Learn. Syst.* **36**:15896–15910. DOI:10.1109/tnnls.2025.3562743
- Patel P., B S. and Annavarapu R.N. (2025). Application of supervised machine learning models in human emotion classification using Tsallis entropy as a feature. *J. Big Data* **12**:126. DOI:10.1186/s40537-025-01177-8
- Fu K., Du C., Wang S., et al. (2022). Multi-view multi-label fine-grained emotion decoding from human brain activity. *IEEE Trans. Neural Netw. Learn. Syst.* **35**:9026–9040. DOI:10.1109/tnnls.2022.3217767
- Huang Z., Du C., Li C., et al. (2025). Identifying the hierarchical emotional areas in the human brain through information fusion. *Inf. Fusion* **113**:102613. DOI:10.1016/j.inffus.2024.102613
- Jin M., Du C., He H., et al. (2024). PGCN: Pyramidal graph convolutional network for EEG emotion recognition. *IEEE Trans. Multimedia* **26**:9070–9082. DOI:10.1109/tmm.2024.3385676
- Silversmith D.B., Abiri R., Hardy N.F., et al. (2021). Plug-and-play control of a brain-computer interface through neural map stabilization. *Nat. Biotechnol.* **39**:326–335. DOI:10.1038/s41587-020-0662-5
- Huang D., Qian K., Fei D.Y., et al. (2012). Electroencephalography (EEG)-based brain-computer interface (BCI): A 2-D virtual wheelchair control based on event-related desynchronization/synchronization and state control. *IEEE Trans. Neural Syst. Rehabil. Eng.* **20**:379–388. DOI:10.1109/tnsre.2012.2190299
- Huang S., Wang Y. and Luo H. (2025). CCSUMSP: A cross-subject Chinese speech decoding framework with unified topology and multi-modal semantic pre-training. *Inf. Fusion* **119**:103022. DOI:10.1016/j.inffus.2025.103022
- Zhao X., Sun J., Wang S. et al. (2024). Mapguide: A simple yet effective method to reconstruct continuous language from brain activities. *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* pp:3822–3832. DOI:10.18653/v1/2024.naacl-long.211
- de la Torre-Ortiz C., Spapé M.M., Ravaja N., et al. (2024). Cross-subject EEG feedback for 'implicit image generation. *IEEE Trans. Cybern.* **54**:6105–6117. DOI:10.1109/tcyb.2024.3406159
- Meng L. and Yang C. (2023). Dual-guided brain diffusion model: Natural image reconstruction from human visual stimulus fMRI. *Bioengineering* **10**:1117. DOI:10.3390/bioengineering10101117
- Deng C., Li X. and Dai J. (2023). Challenges for translating implantable brain-computer interface to medical device. *Innov. Med.* **1**:100040. DOI:10.59717/j.xinn-med.2023.100040
- Tang J., LeBel A., Jain S., et al. (2023). Semantic reconstruction of continuous language from non-invasive brain recordings. *Nat. Neurosci.* **26**:858–866. DOI:10.1038/s41593-023-01304-9
- Ozdenizci O., Wang Y., Koike-Akino T. et al. (2019). Transfer learning in brain-computer interfaces with adversarial variational autoencoders. *9th International IEEE EMBS Conference on Neural Engineering* pp:207–210. DOI:10.1109/ner.2019.8716897
- Scotti P., Banerjee A., Goode J., et al. (2023). Reconstructing the mind's eye: fMRI-toimage with contrastive learning and diffusion priors. *Advances in Neural Information Processing Systems* **36**:24705–24728. DOI:10.48550/arXiv.2305.18274
- Jeong J.H., Cho J.H., Lee B.H., et al. (2022). Real-time deep neurolinguistic learning enhances noninvasive neural language decoding for brain-machine interaction. *IEEE Trans. Cybern.* **53**:7469–7482. DOI:10.1109/tcyb.2022.3211694
- Antonello R., Vaidya A. and Huth A. (2023). Scaling laws for language encoding models in fMRI. *Advances in Neural Information Processing Systems* **36**:21895–21907. DOI:10.48550/arXiv.2305.11863
- Ye Z., Ai Q., Liu Y., et al. (2025). Generative language reconstruction from brain recordings. *Commun. Biol.* **8**:346. DOI:10.1038/s42003-025-07731-7
- Metzger S.L., Littlejohn K.T., Silva A.B., et al. (2023). A high-performance neuroprosthesis for speech decoding and avatar control. *Nature* **620**:1037–1046. DOI:10.1038/s41586-023-06443-4
- Liu Y., Zhao Z., Xu M., et al. (2023). Decoding and synthesizing tonal language speech from brain activity. *Sci. Adv.* **9**:eadh0478. DOI:10.1126/sciadv.adh0478
- Tian Z., Quan R., Ma F., et al. (2025). BRAINGUARD: Privacy-preserving multisubject artifact reconstructions from brain activities. *Proceedings of the AAAI Conference on Artificial Intelligence* **39**:14414–14422. DOI:10.1609/aaai.v39i13.33579
- Song Y., Liu B., Li X. et al. (2024). Decoding natural images from EEG for object recognition. *12th International Conference on Learning Representations* pp:15475–15492. DOI:10.48550/arXiv.2308.13234
- Xia W., De Charette R., Oztireli C. et al. (2024). Dream: Visual decoding from reversing human visual system. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* pp:8226–8235. DOI:10.1109/wacv57701.2024.00804
- Kumar S., Sumers T.R., Yamakoshi T., et al. (2024). Shared functional specialization in transformer-based language models and the human brain. *Nat. Commun.*

- 15:5523. DOI:10.1038/s41467-024-49173-5
43. Ko W., Jeon E., Yoon J.S., et al. (2022). Semi-supervised generative and discriminative adversarial learning for motor imagery-based brain-computer interface. *Sci. Rep.* **12**:4587. DOI:10.1038/s41598-022-08490-9
44. Wang S., Zhou T., Shen Y., et al. (2025). Generative AI enables EEG super-resolution via spatio-temporal adaptive diffusion learning. *IEEE Trans. Consum. Electron.* **71**:1034–1045. DOI:10.1109/tce.2025.3528438
45. Vetter J., Macke J.H. and Gao R. (2024). Generating realistic neurophysiological time series with denoising diffusion probabilistic models. *Patterns* **5**:101047. DOI:10.1016/j.patter.2024.101047
46. Kwon J. and Im C.H. (2022). Novel signal-to-signal translation method based on StarGAN to generate artificial EEG for SSVEP-based brain-computer interfaces. *Expert Syst. Appl.* **203**:117574. DOI:10.1016/j.eswa.2022.117574
47. Yao W., Lyu Z., Mahmud M., et al. (2025). CATD: Unified representation learning for EEG-to-fMRI cross-modal generation. *IEEE Trans. Med. Imaging* **44**:2757–2767. DOI:10.1109/tmi.2025.3550206
48. Li Y., Wang Y., Lei B., et al. (2025). SCDM: Unified representation learning for EEG-to-fNIRS cross-modal generation in MI-BCIs. *IEEE Trans. Med. Imaging* **44**:2384–2394. DOI:10.1109/tmi.2025.3532480
49. Hu M., Chen J., Jiang S., et al. (2022). E2SGAN: EEG-to-SEEG translation with generative adversarial networks. *Front. Neurosci.* **16**:971829. DOI:10.3389/fnins.2022.971829
50. Antoniadis A., Spyrou L., Martin-Lopez D., et al. (2018). Deep neural architectures for mapping scalp to intracranial EEG. *Int. J. Neural Syst.* **28**:1850009. DOI:10.1142/s0129065718500090
51. Zhang R., Zeng Y., Tong L., et al. (2022). ERP-WGAN: A data augmentation method for EEG single-trial detection. *J. Neurosci. Meth.* **376**:109621. DOI:10.1016/j.jneumeth.2022.109621
52. Fahimi F., Dosen S., Ang K.K., et al. (2020). Generative adversarial networks-based data augmentation for brain-computer interface. *IEEE Trans. Neural Netw. Learn. Syst.* **32**:4039–4051. DOI:10.1109/tnnls.2020.3016666
53. Tang Y., Chen D., Liu H., et al. (2022). Deep EEG superresolution via correlating brain structural and functional connectivities. *IEEE Trans. Cybern.* **53**:4410–4422. DOI:10.1109/tcyb.2022.3178370
54. Gayon-Lombardo A., Mosser L., Brandon N.P., et al. (2020). Pores for thought: generative adversarial networks for stochastic reconstruction of 3D multi-phase electrode microstructures with periodic boundaries. *npj Comput. Mater.* **6**:82. DOI:10.1038/s41524-020-0340-7
55. Sanchez-Lengeling B. and Aspuru-Guzik A. (2018). Inverse molecular design using machine learning: Generative models for matter engineering. *Science* **361**:360–365. DOI:10.1126/science.aat2663
56. Manica M., Born J., Cadow J., et al. (2023). Accelerating material design with the generative toolkit for scientific discovery. *npj Comput. Mater.* **9**:69. DOI:10.1038/s41524-023-01028-1
57. Dan Y., Zhao Y., Li X., et al. (2020). Generative adversarial networks (GAN) based efficient sampling of chemical composition space for inverse design of inorganic materials. *npj Comput. Mater.* **6**:84. DOI:10.1038/s41524-020-00352-0
58. Guenther F.H., Brumberg J.S., Wright E.J., et al. (2009). A wireless brain-machine interface for real-time speech synthesis. *PLoS one* **4**:e8218. DOI:10.1371/journal.pone.0008218
59. Tang J. and Huth A.G. (2025). Semantic language decoding across participants and stimulus modalities. *Curr. Biol.* **35**:1023–1032. DOI:10.1016/j.cub.2025.01.024
60. Komeiji S., Mitsuhashi T., Iimura Y., et al. (2024). Feasibility of decoding covert speech in ECoG with a transformer trained on overt speech. *Sci. Rep.* **14**:11491. DOI:10.1101/2024.02.05.578911
61. Fan C., Hahn N., Kamdar F., et al. (2023). Plug-and-play stability for intracortical brain-computer interfaces: a one-year demonstration of seamless brain-to-text communication. *Advances in Neural Information Processing Systems* **36**:42258–42270. DOI:10.48550/arXiv.2311.03611
62. Metzger S.L., Liu J.R., Moses D.A., et al. (2022). Generalizable spelling using a speech neuroprosthesis in an individual with severe limb and vocal paralysis. *Nat. Commun.* **13**:6510. DOI:10.1038/s41467-022-33611-3
63. Proix T., Delgado Saa J., Christen A., et al. (2022). Imagined speech can be decoded from low-and cross-frequency intracranial EEG features. *Nat. Commun.* **13**:48. DOI:10.1038/s41467-021-27725-3
64. Qi Y., Zhu X., Xiong X., et al. (2025). Human motor cortex encodes complex handwriting through a sequence of stable neural states. *Nat. Hum. Behav.* **9**:1260–1271. DOI:10.1038/s41562-025-02157-x
65. Miliotou E., Kyriakis P., Hinman J.D., et al. (2023). Generative decoding of visual stimuli. *International Conference on Machine Learning* **202**:24775–24784. <https://proceedings.mlr.press/v202/miliotou23a/miliotou23a.pdf>.
66. Liu A., Jing H., Liu Y., et al. (2024). Hidden states in LLMs improve EEG representation learning and visual decoding. *European Conference on Artificial Intelligence* pp:2130–2137. DOI:10.3233/faia240732
67. Huang S., Sun L., Yousefnezhad M., et al. (2023). Functional alignment-auxiliary generative adversarial network-based visual stimuli reconstruction via multi-subject fMRI. *IEEE Trans. Neural Syst. Rehabil. Eng.* **31**:2715–2725. DOI:10.1109/tnsr.2023.3283405
68. Shimizu S., Ota A. and Nakane A. (2025). Reviving intentional facial expressions: An interface for ALS patients using brain decoding and image-generative AI. *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* pp:1–10. DOI:10.1145/3706599.3719775
69. Mahajan G., Jeevan R., Divija L. et al. (2024). Deciphering EEG waves for the generation of images. *2024 12th International Winter Conference on Brain-Computer Interface* pp:1–6. DOI:10.1109/bci60775.2024.10480501
70. Yeung J., Luo A.F., Sarch G. et al. (2025). Reanimating images using neural representations of dynamic stimuli. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* pp:5331–5343. DOI:10.1109/cvpr52734.2025.00502
71. Allen E.J., St-Yves G., Wu Y., et al. (2022). A massive 7t fmri dataset to bridge cognitive neuroscience and artificial intelligence. *Nat. Neurosci.* **25**:116–126. DOI:10.1038/s41593-021-00962-x
72. Horikawa T., Tamaki M., Miyawaki Y., et al. (2013). Neural decoding of visual imagery during sleep. *Science* **340**:639–642. DOI:10.1126/science.1234330
73. Breedlove J.L., St-Yves G., Olman C.A., et al. (2020). Generative feedback explains distinct brain activity codes for seen and mental images. *Curr. Biol.* **30**:2211–2224. DOI:10.1016/j.cub.2020.04.014
74. Ritchie J.B., Kaplan D.M. and Klein C. (2019). Decoding the brain: Neural representation and the limits of multivariate pattern analysis in cognitive neuroscience. *British J. Philos. Sci.* **70**:581–607. DOI:10.1093/bjps/axx023
75. Wang S., Jiang C., Yu Y., et al. (2025). Tellurium nanowire retinal nanoprosthesis improves vision in models of blindness. *Science* **388**:eadu2987. DOI:10.1126/science.adu2987
76. Kravitz D.J., Saleem K.S., Baker C.I., et al. (2011). A new neural framework for visuospatial processing. *Nat. Rev. Neurosci.* **12**:217–230. DOI:10.1038/nrn3008
77. Kwon M., Han S., Kim K., et al. (2019). Super-resolution for improving EEG spatial resolution using deep convolutional neural network—feasibility study. *Sensors* **19**:5317. DOI:10.3390/s19235317
78. Bao G., Zhang Q., Gong Z., et al. (2025). Wills Aligner: Multi-subject collaborative brain visual decoding. *Proceedings of the AAAI Conference on Artificial Intelligence* **39**:14194. DOI:10.1609/aaai.v39i13.33554
79. Li D., Du C., Wang S., et al. (2021). Multi-subject data augmentation for target subject semantic decoding with deep multi-view adversarial learning. *Inf. Sci.* **547**:1025–1044. DOI:10.1016/j.ins.2020.09.012
80. Ozcelik F. and VanRullen R. (2023). Natural scene reconstruction from fMRI signals using generative latent diffusion. *Sci. Rep.* **13**:15666. DOI:10.1038/s41598-023-42891-8
81. Takagi Y. and Nishimoto S. (2023). High-resolution image reconstruction with latent diffusion models from human brain activity. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* pp:14453–14463. DOI:10.1109/cvpr52729.2023.01389
82. Du C., Fu K., Li J., et al. (2023). Decoding visual neural representations by multimodal learning of brain-visual-linguistic features. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**:10760–10777. DOI:10.1109/tpami.2023.3263181
83. Shen G., Zhao D., He X., et al. (2024). Neuro-vision to language: Enhancing brain recording-based visual reconstruction and language interaction. *Advances in Neural Information Processing Systems* **37**:98083–98110. DOI:10.52202/079017-3113
84. Zhang A., Wang B., Wu X. et al. (2025). A novel multimodal method for decoding speech perception from brain activities. *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing* pp:1–5. DOI:10.1109/icassp49660.2025.10889637
85. Luo S., Angrick M., Coogan C., et al. (2023). Stable decoding from a speech BCI enables control for an individual with ALS without recalibration for 3 months. *Adv. Sci.* **10**:2304853. DOI:10.1002/adv.202304853
86. Xi N., Zhao S., Wang H. et al. (2023). Unicorn: Unified cognitive signal reconstruction bridging cognitive signals and human language. *Proceedings of the Annual Meeting of the Association for Computational Linguistics* pp:13277–13291. DOI:10.18653/v1/2023.acl-long.741
87. Guo J., Yi C., Li F. et al. (2024). MindLDM: Reconstruct visual stimuli from fMRI using latent diffusion model. *2024 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications* pp:1–6. DOI:10.1109/civemsa58715.2024.10586647
88. Wilson G.H., Stavisky S.D., Willett F.R., et al. (2020). Decoding spoken english from intracortical electrode arrays in dorsal precentral gyrus. *J. Neural Eng.* **17**:066007. DOI:10.1088/1741-2552/abbef
89. Ferrante M., Boccato T., Ozcelik F. et al. (2023). Multimodal decoding of human brain activity into images and text. *Advances in Neural Information Processing Systems* vol. 243 pp:87–101. <https://proceedings.mlr.press/v243/ferrante24a>.
90. Scotti P.S., Tripathy M., Villanueva C.K.T. et al. (2024). MindEye2: Shared-subject models enable fMRI-to-image with 1 hour of data. *Proceedings of the 41st International Conference on Machine Learning* pp:44038–44059. DOI:10.48550/arXiv.2403.11207
91. Wang S., Liu S., Tan Z. et al. (2024). Mindbridge: A cross-subject brain decoding framework. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* pp:11333–11342. DOI:10.1109/cvpr52733.2024.01077
92. Gong Z., Bao G., Zhang Q., et al. (2024). NeuroClips: Towards high-fidelity and

- smooth fMRI-to-video reconstruction. *Advances in Neural Information Processing Systems* **37**:51655–51683. DOI:10.52202/079017-1636
93. Vansteensel M.J., Pels E.G., Bleichner M.G., et al. (2016). Fully implanted brain-computer interface in a locked-in patient with ALS. *N. Engl. J. Med.* **375**:2060–2066. DOI:10.1056/nejmoa1608085
 94. Willett F.R., Avansino D.T., Hochberg L.R., et al. (2021). High-performance brain-to-text communication via handwriting. *Nature* **593**:249–254. DOI:10.1038/s41586-021-03506-2
 95. Martin S., Brunner P., Iturrate I., et al. (2016). Word pair classification during imagined speech using direct brain recordings. *Sci. Rep.* **6**:25803. DOI:10.1038/srep25803
 96. Herff C., Heger D., De Pestiers A., et al. (2015). Brain-to-text: decoding spoken phrases from phone representations in the brain. *Front. Neurosci.* **8**:141498. DOI:10.3389/fnins.2015.00217
 97. Moses D.A., Leonard M.K., Makin J.G., et al. (2019). Real-time decoding of question-and-answer speech dialogue using human cortical activity. *Nat. Commun.* **10**:3096. DOI:10.1038/s41467-019-10994-4
 98. Moses D.A., Metzger S.L., Liu J.R., et al. (2021). Neuroprosthesis for decoding speech in a paralyzed person with anarthria. *N. Engl. J. Med.* **385**:217–227. DOI:10.1056/nejmoa2027540
 99. Duraivel S., Rahimpour S., Chiang C.H., et al. (2023). High-resolution neural recordings improve the accuracy of speech decoding. *Nat. Commun.* **14**:6938. DOI:10.1038/s41467-023-42555-1
 100. Zhang D., Wang Z., Qian Y., et al. (2024). A brain-to-text framework for decoding natural tonal sentences. *Cell Rep.* **43**:114924. DOI:10.1016/j.celrep.2024
 101. Makin J.G., Moses D.A. and Chang E.F. (2020). Machine translation of cortical activity to text with an encoder-decoder framework. *Nat. Neurosci.* **23**:575–582. DOI:10.1038/s41593-020-0608-8
 102. Sun P., Anumanchipalli G.K. and Chang E.F. (2020). Brain2Char: a deep architecture for decoding text from brain recordings. *J. Neural Eng.* **17**:066015. DOI:10.1088/1741-2552/abc742
 103. Simanova I., Hagoort P., Oostenveld R., et al. (2014). Modality-independent decoding of semantic information from the human brain. *Cereb. Cortex* **24**:426–434. DOI:10.1093/cercor/bhs324
 104. Defossez A., Caucheteux C., Rapin J., et al. (2023). Decoding speech perception ' from non-invasive brain recordings. *Nat. Mach. Intell.* **5**:1097–1107. DOI:10.1038/s42256-023-00714-5
 105. Chen Q., Wang Y., Wang F., et al. (2025). Decoding text from electroencephalography signals: a novel hierarchical gated recurrent unit with masked residual attention mechanism. *Eng. Appl. Artif. Intel.* **139**:109615. DOI:10.1016/j.engappai.2024.109615
 106. Wang Z. and Ji H. (2022). Open vocabulary electroencephalography-to-text decoding and zero-shot sentiment classification. *Proceedings of the AAAI Conference on Artificial Intelligence* **36**:5350–5358. DOI:10.1609/aaai.v36i5.20472
 107. Duan Y., Zhou J., Wang Z., et al. (2023). Dewave: Discrete encoding of EEG waves for EEG to text translation. *Advances in Neural Information Processing Systems* **36**:9907–9918. DOI:10.48550/arXiv.2309.14030
 108. Chen X., Du C., Liu C., et al. (2025). BP-GPT: Auditory neural decoding using fmriprompted llm. *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing* pp:1–5. DOI:10.1109/icassp49660.2025.10890142
 109. Liu H., Hajialilgol D., Antony B. et al. (2024). EEG2Text: Open vocabulary EEG-to-text translation with multi-view transformer. *2024 IEEE International Conference on Big Data* pp:1824–1833. DOI:10.1109/bigdata62323.2024.10825980
 110. Zhou J., Duan Y., Chang Y.C., et al. (2024). BELT: bootstrapped EEG-to-language training by natural language supervision. *IEEE Trans. Neural Syst. Rehabil. Eng.* **32**:3278–3288. DOI:10.1109/tnrsre.2024.3450795
 111. Amrani H., Micucci D. and Napoletano P. (2024). Deep representation learning for open vocabulary electroencephalography-to-text decoding. *IEEE J. Biomed. Health Inform.* pp:1–12. DOI:10.1109/JBHI.2024.3416066
 112. Tao Y., Liang Y., Wang L. et al. (2025). See: Semantically aligned EEG-to-text translation. *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing* pp:1–5. DOI:10.1109/icassp49660.2025.10889658
 113. Wang J., Song Z., Ma Z. et al. (2024). Enhancing EEG-to-text decoding through transferable representations from pre-trained contrastive EEG-text masked autoencoder. *Proceedings of the Annual Meeting of the Association for Computational Linguistics* pp:7278–7292. DOI:10.18653/v1/2024.acl-long.393
 114. Luo K. (2024). Real-time open-vocabulary sentence decoding from MEG signals using transformers. *2024 International Conference on Image Processing, Computer Vision and Machine Learning* pp:460–463. DOI:10.1109/icicml63543.2024.10957766
 115. Boyko M., Druzhinina P., Kormakov G. et al. (2024). MEGFormer: enhancing speech decoding from brain activity through extended semantic representations. *International Conference on Medical Image Computing and Computer-Assisted Intervention* pp:281–290. DOI:10.1007/978-3-031-72069-7_27
 116. Wang B., Xu X., Zhang L. et al. (2024). Semantic reconstruction of continuous language from meg signals. *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing* pp:2190–2194. DOI:10.1109/icassp48485.2024.10448281
 117. Guo Y., Dong Y., Ng M.K.P. et al. (2025). A pre-trained framework for multilingual brain decoding using non-invasive recordings. *arXiv Preprint arXiv:2506.03214*. DOI:10.48550/arXiv.2506.03214
 118. Silva A.B., Liu J.R., Metzger S.L., et al. (2024). A bilingual speech neuroprosthesis driven by cortical articulatory representations shared between languages. *Nat. Biomed. Eng.* **8**:977–991. DOI:10.1038/s41551-024-01207-5
 119. Huang W., Yang P., Tang Y., et al. (2024). From sight to insight: A multi-task approach with the visual language decoding model. *Inf. Fusion* **112**:102573. DOI:10.1016/j.inffus.2024.102573
 120. Han J., Gong K., Zhang Y. et al. (2024). Onellm: One framework to align all modalities with language. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* pp:26584–26595. DOI:10.1109/cvpr52733.2024.02510
 121. Chen J., Qi Y., Wang Y., et al. (2025). Mindgpt: Interpreting what you see with non-invasive brain recordings. *IEEE Trans. Image Process.* **34**:3281–3293. DOI:10.1109/tip.2025.3572784
 122. Tang J., Yang Y., Zhao Q., et al. (2024). Visual guided dual-spatial interaction network for fine-grained brain semantic decoding. *IEEE Trans. Instrum. Meas.* **73**:2534214. DOI:10.1109/tim.2024.3480232
 123. Ikegawa Y., Fukuma R., Sugano H., et al. (2024). Text and image generation from intracranial electroencephalography using an embedding space for text and images. *J. Neural Eng.* **21**:036019. DOI:10.1088/1741-2552/ad417a
 124. Willett F.R., Kunz E.M., Fan C., et al. (2023). A high-performance speech neuroprosthesis. *Nature* **620**:1031–1036. DOI:10.1101/2023.01.21.524489
 125. Ramsey N.F., Salari E., Aarnoutse E.J., et al. (2018). Decoding spoken phonemes from sensorimotor cortex with high-density ECoG grids. *Neuroimage* **180**:301–311. DOI:10.1016/j.neuroimage.2017.10.011
 126. Angrick M., Herff C., Mugler E., et al. (2019). Speech synthesis from ECoG using densely connected 3d convolutional neural networks. *J. Neural Eng.* **16**:036019. DOI:10.1088/1741-2552/ab0c59
 127. Herff C., Diener L., Angrick M., et al. (2019). Generating natural, intelligible speech from brain activity in motor, premotor, and inferior frontal cortices. *Front. Neurosci.* **13**:1267. DOI:10.3389/fnins.2019.01267
 128. Berezutskaya J., Freudenburg Z.V., Vansteensel M.J., et al. (2023). Direct speech reconstruction from sensorimotor brain activity with optimized deep learning models. *J. Neural Eng.* **20**:056010. DOI:10.1088/1741-2552/ace8be
 129. Anumanchipalli G.K., Chartier J. and Chang E.F. (2019). Speech synthesis from neural decoding of spoken sentences. *Nature* **568**:493–498. DOI:10.1038/s41586-019-1119-1
 130. Shigemori K., Komeiji S., Mitsuhashi T. et al. (2023). Synthesizing speech from ECoG with a combination of transformer-based encoder and neural vocoder. *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing* pp:1–5. DOI:10.1109/icassp49357.2023.10097004
 131. Ticha M.B.B., Ran X., Roussel P. et al. (2024). A vision transformer architecture for overt speech decoding from ECoG data. *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* pp:1–4. DOI:10.1109/embs3108.2024.10781877
 132. Chen X., Wang R., Khalilian-Gourtani A., et al. (2024). A neural speech decoding framework leveraging deep learning and speech synthesis. *Nat. Mach. Intell.* **6**:467–480. DOI:10.1038/s42256-024-00824-8
 133. Meng K., Goodarzi F., Kim E., et al. (2023). Continuous synthesis of artificial speech sounds from human cortical surface recordings during silent speech production. *J. Neural Eng.* **20**:046019. DOI:10.1088/1741-2552/ace7f6
 134. Wu X., Wellington S., Fu Z., et al. (2024). Speech decoding from stereoelectroencephalography (sEEG) signals using advanced deep learning methods. *J. Neural Eng.* **21**:036055. DOI:10.1088/1741-2552/ad593a
 135. Angrick M., Ottenhoff M.C., Diener L., et al. (2021). Real-time synthesis of imagined speech processes from minimally invasive recordings of neural activity. *Commun. Biol.* **4**:1055. DOI:10.1038/s42003-021-02578-0
 136. Angrick M., Luo S., Rabbani Q., et al. (2024). Online speech synthesis using a chronically implanted brain-computer interface in an individual with ALS. *Sci. Rep.* **14**:9617. DOI:10.1038/s41598-024-60277-2
 137. Card N.S., Wairagkar M., Iacobacci C., et al. (2024). An accurate and rapidly calibrating speech neuroprosthesis. *N. Engl. J. Med.* **391**:609–618. DOI:10.1056/nejmoa2314132
 138. Wu H., Cai C., Ming W., et al. (2024). Speech decoding using cortical and subcortical electrophysiological signals. *Front. Neurosci.* **18**:1345308. DOI:10.3389/fnins.2024.1345308
 139. Feng C., Cao L., Wu D., et al. (2025). Acoustic inspired brain-to-sentence decoder for logossyllabic language. *Cyborg Bionic Syst.* **6**:0257. DOI:10.34133/cbsystems.0257
 140. Accou B., Vanthornhout J., hamme H.V., et al. (2023). Decoding of the speech envelope from EEG using the VLAAl deep neural network. *Sci. Rep.* **13**:812. DOI:10.1038/s41598-022-27332-2
 141. Van Dyck B., Yang L. and Van Hulle M.M. (2023). Decoding auditory EEG responses using an adapted wavenet. *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing* pp:1–2. DOI:10.1109/icassp49357.2023.10095420
 142. Fan C., Zhang S., Zhang J. et al. (2025). SSM2Mel: State space model to reconstruct Mel spectrogram from the EEG. *ICASSP 2025-2025 IEEE International Conference*

- on Acoustics, Speech and Signal Processing pp:1–5. DOI:10.1109/icassp49660.2025.10888785
143. Qi D., Kong L., Yang L. et al. (2023). Audiodiffusion: Generating high-quality audios from EEG signals: Reconstructing audio from EEG signals. *2023 4th International Symposium on Computer Engineering and Intelligent Communications* pp:344–348. DOI:10.1109/isceic59030.2023.10271237
144. Xiong W., Ma L. and Li H. (2025). Synthesizing intelligible utterances from EEG of imagined speech. *Front. Neurosci.* **19**:1565848. DOI:10.3389/fnins.2025.1565848
145. Lee J.W., Lee S.H., Lee Y.E. et al. (2023). Sentence reconstruction leveraging contextual meaning from speech-related brain signals. *2023 IEEE International Conference on Systems, Man, and Cybernetics* pp:3721–3726. DOI:10.1109/smc53992.2023.10394348
146. Park J.Y., Tsukamoto M., Tanaka M., et al. (2025). Natural sounds can be reconstructed from human neuroimaging data using deep neural network representation. *PLoS Biol.* **23**e3003293. DOI:10.1371/journal.pbio.3003293
147. Craik A., Dial H. and Contreras-Vidal J.L. (2025). Continuous and discrete decoding of overt speech with scalp electroencephalography (EEG). *J. Neural Eng.* **22**:026017. DOI:10.1088/1741-2552/ad8d0a
148. Dado T., Papale P., Lozano A., et al. (2024). Brain2GAN: Feature-disentangled neural encoding and decoding of visual perception in the primate brain. *PLoS Comput. Biol.* **20**e1012058. DOI:10.1371/journal.pcbi.1012058
149. Doerig A., Kietzmann T.C., Allen E., et al. (2025). High-level visual representations in the human brain are aligned with large language models. *Nat. Mach. Intell.* **7**:1220–1234. DOI:10.1038/s42256-025-01072-0
150. Wang H., Zhang Q., Wang L. et al. (2025). Neurons: Emulating the human visual cortex improves fidelity and interpretability in fMRI-to-video reconstruction. *Proceedings of the IEEE/CVF International Conference on Computer Vision* pp:18367–18376. DOI:10.48550/arXiv.2503.11167
151. Xia R., Yin C. and Li P. (2024). Decoding the echoes of vision from fmri: Memory disentangling for past semantic information. *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* pp:2040–2052. DOI:10.18653/v1/2024.emnlp-main.122
152. Meng L. and Yang C. (2024). Semantics-guided hierarchical feature encoding generative adversarial network for visual image reconstruction from brain activity. *IEEE Trans. Neural Syst. Rehabil. Eng.* **32**:1267–1283. DOI:10.1109/tnsre.2024.3377698
153. Dado T., Güçlütürk Y., Ambrogioni L., et al. (2022). Hyperrealistic neural decoding for reconstructing faces from fMRI activations via the GAN latent space. *Sci. Rep.* **12**:141. DOI:10.1038/s41598-021-03938-w
154. Quan R., Wang W., Tian Z. et al. (2025). Psychometry: An omnifit model for image reconstruction from human brain activity. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* pp:233–243. DOI:10.1109/cvpr52733.2024.00030
155. Wu H., Li Q., Zhang C. et al. (2025). Bridging the vision-brain gap with an uncertainty aware blur prior. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* pp:2246–2257. DOI:10.1109/cvpr52734.2025.00215
156. 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**:915–929. DOI:10.1016/j.neuron.2008.11.004
157. Du C., Du C., Huang L., et al. (2018). Reconstructing perceived images from human brain activities with bayesian deep multiview learning. *IEEE Trans. Neural Netw. Learn. Syst.* **30**:2310–2323. DOI:10.1109/tnnls.2018.2882456
158. Güçlütürk Y., Güçlü U., Seeliger K., et al. (2017). Reconstructing perceived faces from brain activations with deep adversarial neural decoding. *Advances in Neural Information Processing Systems* **30**:4249–4260.
159. VanRullen R. and Reddy L. (2019). Reconstructing faces from fMRI patterns using deep generative neural networks. *Commun. Biol.* **2**:193. DOI:10.1038/s42003-019-0438-y
160. Du C., Du C., Huang L., et al. (2020). Conditional generative neural decoding with structured cnn feature prediction. *Proceedings of the AAAI Conference on Artificial Intelligence* **34**:2629–2636. DOI:10.1609/aaai.v34i03.5647
161. Naselaris T., Prenger R.J., Kay K.N., et al. (2009). Bayesian reconstruction of natural images from human brain activity. *Neuron* **63**:902–915. DOI:10.1016/j.neuron.2009.09.006
162. Belyi R., Gaziv G., Hoogi A., et al. (2019). From voxels to pixels and back: Self-supervision in natural-image reconstruction from fMRI. *Advances in Neural Information Processing Systems* **32**:6517–6527. DOI:10.48550/arXiv.1907.02431
163. Shen G., Horikawa T., Majima K., et al. (2019). Deep image reconstruction from human brain activity. *PLOS Comput. Biol.* **15**e1006633. DOI:10.1371/journal.pcbi.1006633
164. Luo J., Cui W., Liu J., et al. (2022). Visual image decoding of brain activities using a dual attention hierarchical latent generative network with multiscale feature fusion. *IEEE Trans. Cogn. Develop. Syst.* **15**:761–773. DOI:10.1109/tcds.2022.3181469
165. Akamatsu Y., Harakawa R., Ogawa T., et al. (2020). Brain decoding of viewed image categories via semi-supervised multi-view bayesian generative model. *IEEE Trans. Signal Process.* **68**:5769–5781. DOI:10.1109/tsp.2020.3028701
166. Li D., Du C. and He H. (2020). Semi-supervised cross-modal image generation with generative adversarial networks. *Pattern Recognit.* **100**:107085. DOI:10.1016/j.patcog.2019.107085
167. Gaziv G., Belyi R., Granot N., et al. (2022). Self-supervised natural image reconstruction and large-scale semantic classification from brain activity. *NeuroImage* **254**:119121. DOI:10.1016/j.neuroimage.2022.119121
168. Zhou Q., Du C., Li D., et al. (2025). Interpretable visual neural decoding with unsupervised semantic disentanglement. *Mach. Intell. Res.* **22**:553–570. DOI:10.1007/s11633-023-1484-y
169. Tirupattur P., Rawat Y.S., Spampinato C. et al. (2018). Thoughtviz: Visualizing human thoughts using generative adversarial network. *Proceedings of the 26th ACM International Conference on Multimedia* pp:950–958. DOI:10.1145/3240508.3240641
170. Palazzo S., Spampinato C., Kavasidis I. et al. (2017). Generative adversarial networks conditioned by brain signals. *Proceedings of the IEEE international conference on computer vision* pp:3410–3418. DOI:10.1109/iccv.2017.369
171. Kavasidis I., Palazzo S., Spampinato C. et al. (2017). Brain2Image: Converting brain signals into images. *Proceedings of the 25th ACM International Conference on Multimedia* pp:1809–1817. DOI:10.1145/3123266.3127907
172. Ahmadi H., Gassemi F. and Moradi M.H. (2024). Visual image reconstruction based on EEG signals using a generative adversarial and deep fuzzy neural network. *Biomed. Signal Proces.* **87**:105497. DOI:10.1016/j.bspc.2023.105497
173. Fares A., Zhong S.H. and Jiang J. (2020). Brain-media: A dual conditioned and lateralization supported GAN (DCLS-GAN) towards visualization of image-evoked brain activities. *Proceedings of the 28th ACM International Conference on Multimedia* pp:1764–1772. DOI:10.1145/3394171.3413858
174. Huang W., Yan H., Wang C., et al. (2020). Perception-to-image: Reconstructing natural images from the brain activity of visual perception. *Ann. Biomed. Eng.* **48**:2323–2332. DOI:10.1007/s10439-020-02502-3
175. Lu Y., Du C., Zhou Q. et al. (2023). Minddiffuser: Controlled image reconstruction from human brain activity with semantic and structural diffusion. *Proceedings of the 31st ACM International Conference on Multimedia* pp:5899–5908. DOI:10.1145/3581783.3613832
176. Lu Y., Du C., Wang C. et al. (2025). Animate your thoughts: Reconstruction of dynamic natural vision from human brain activity. *13th International Conference on Learning Representations* pp:23255–23293. DOI:10.48550/arXiv.2405.03280
177. Li J., Li D., Xiong C. et al. (2022). Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation. *International Conference on Machine Learning* pp:12888–12900. DOI:10.48550/arXiv.2201.12086
178. Luo A., Henderson M., Wehbe L., et al. (2023). Brain diffusion for visual exploration: Cortical discovery using large scale generative models. *Advances in Neural Information Processing Systems* **36**:75740–75781.
179. Ferrante M., Boccato T., Passamonti L., et al. (2024). Retrieving and reconstructing conceptually similar images from fmri with latent diffusion models and a neuro-inspired brain decoding model. *J. Neural Eng.* **21**:046001. DOI:10.1088/1741-2552/ad593c
180. Chen Z., Qing J., Xiang T. et al. (2023). Seeing beyond the brain: Conditional diffusion model with sparse masked modeling for vision decoding. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* pp:22710–22720. DOI:10.1109/cvpr52729.2023.02175
181. Bai Y., Wang X., Cao Y.P. et al. (2024). Dreamdiffusion: High-quality EEG-to-image generation with temporal masked signal modeling and CLIP alignment. *Computer Vision–ECCV 2024: 18th European Conference* pp:472–488. DOI:10.1007/978-3-031-72751-1_27
182. Liu Q., Zhu H., Chen N., et al. (2024). Mind-bridge: reconstructing visual images based on diffusion model from human brain activity. *Signal Image Video Process.* **18**:953–963. DOI:10.1007/s11760-024-03207-z
183. Zhou Q., Du C., Wang S. et al. (2024). CLIP-MUSED: CLIP-guided multi-subject visual neural information semantic decoding. *12th International Conference on Learning Representations* pp:43466–43482. DOI:10.48550/arXiv.2402.08994
184. Ma Y., Liu Y., Chen L., et al. (2025). BrainCLIP: Brain representation via CLIP for generic natural visual stimulus decoding. *IEEE Trans. Med. Imaging* **44**:3962–3972. DOI:10.1109/TMI.2025.3537287
185. Bhalerao S.V. and Pachori R.B. (2024). Automated classification of cognitive visual objects using multivariate swarm sparse decomposition from multichannel EEG-MEG signals. *IEEE Trans. Hum.-Mach. Syst.* **54**:455–464. DOI:10.1109/thms.2024.3395153
186. Gong Z., Zhang Q., Bao G., et al. (2025). Mindtuner: Cross-subject visual decoding with visual fingerprint and semantic correction. *Proceedings of the AAAI Conference on Artificial Intelligence* **39**:14247–14255. DOI:10.1609/aaai.v39i13.33560
187. Luo A.F., Henderson M.M., Tarr M.J. et al. (2024). BrainSCUBA: Fine-grained natural language captions of visual cortex selectivity. *12th International Conference on Learning Representations* pp:1744–1780. DOI:10.48550/arXiv.2310.0442
188. Zhao Y., Dong G., Zhu L., et al. (2025). Memory recall: Retrieval-augmented mind reconstruction for brain decoding. *Inf. Fusion* **123**:103280. DOI:10.1016/j.inffus.2025.103280
189. Huo J., Wang Y., Wang Y. et al. (2024). Neuropictor: Refining fMRI-to-image reconstruction via multi-individual pretraining and multi-level modulation. *European Conference on Computer Vision* pp:56–73. DOI:10.1007/978-3-031-72983-6_4
190. Liu M. and Kobayashi I. (2024). Do feature representations from different language

- models affect accuracy of brain encoding models' predictions? *2024 IEEE International Conference on Systems, Man, and Cybernetics* pp:2766–2771. DOI:10.1109/smc54092.2024.10831584
191. Singh A.K., Wang Y.K., King J.T., et al. (2020). Extended interaction with a BCI video game changes resting-state brain activity. *IEEE Trans. Cogn. Develop. Syst.* **12**:809–823. DOI:10.1109/tcds.2020.2985102
 192. Nishimoto S., Vu A.T., Naselaris T., et al. (2011). Reconstructing visual experiences from brain activity evoked by natural movies. *Curr. Biol.* **21**:1641–1646. DOI:10.1016/j.cub.2011.08.031
 193. Hanke M., Baumgartner F.J., Ibe P., et al. (2014). A high-resolution 7-Tesla fMRI dataset from complex natural stimulation with an audio movie. *Sci. Data* **1**:14003. DOI:10.1038/sdata.2014.3
 194. Sun J., Li M. and Moens M.F. (2025). Neuralflix: A simple while effective framework for semantic decoding of videos from non-invasive brain recordings. *Proceedings of the AAAI Conference on Artificial Intelligence* **39**:7096–7104. DOI:10.1609/aaai.v39i7.32762
 195. Fosco C., Lahner B., Pan B. et al. (2024). Brain netflix: Scaling data to reconstruct videos from brain signals. *European Conference on Computer Vision* pp:457–474. DOI:10.1007/978-3-031-73347-5_26
 196. Simonyan K. and Zisserman A. (2014). Two-stream convolutional networks for action recognition in videos. *Advances in Neural Information Processing Systems* **27**:568–576. DOI:10.48550/arXiv.1406.2199
 197. Carreira J. and Zisserman A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* pp:6299–6308. DOI:10.1109/cvpr.2017.502
 198. Chen Z., Qing J. and Zhou J.H. (2023). Cinematic mindscapes: High-quality video reconstruction from brain activity. *Advances in Neural Information Processing Systems* **36**:24841–24858. DOI:10.48550/arXiv.2305.11675
 199. Liu X.H., Liu Y.K., Wang Y., et al. (2024). EEG2video: Towards decoding dynamic visual perception from EEG signals. *Advances in Neural Information Processing Systems* **37**:72245–72273. DOI:10.52202/079017-2306
 200. Lin J., Chen H., Fan Y. et al. (2025). Multi-layer visual feature fusion in multimodal llms: 1167 Methods, analysis, and best practices. *Proceedings of the IEEE/CVF Conference on 1168 Computer Vision and Pattern Recognition* pp:4156–4166. DOI:10.1109/cvpr52734.2025.00393
 201. Zhang B., Fang Y., Ren T. et al. (2022). Multimodal analysis for deep video understanding with video language transformer. *Proceedings of the 30th ACM International Conference on Multimedia* pp:7165–7169. DOI:10.1145/3503161.3551600
 202. Holz F.G., Le Mer Y., Muqit M.M. et al. (2025). Subretinal photovoltaic implant to restore vision in geographic atrophy due to AMD. *N. Engl. J. Med.* **394**:232–242. DOI:10.1056/nejmoa2501396
 203. Elnabawy R.H., Abdennadher S., Hellwich O., et al. (2022). PVGAN: a generative adversarial network for object simplification in prosthetic vision. *J. Neural Eng.* **19**:056007. DOI:10.1088/1741-2552/ac8acf
 204. Chen X., Wang F., Fernandez E., et al. (2020). Shape perception via a high-channelcount neuroprosthesis in monkey visual cortex. *Science* **370**:1191–1196. DOI:10.1126/science.abd7435
 205. Luo T.j., Fan Y., Chen L., et al. (2020). EEG signal reconstruction using a generative adversarial network with wasserstein distance and temporal-spatial-frequency loss. *Front. Neuroinf.* **14**:15. DOI:10.3389/fninf.2020.00015
 206. Aznan N.K.N., Atapour-Abarghouei A., Bonner S. et al. (2019). Simulating brain signals: Creating synthetic EEG data via neural-based generative models for improved SSVEP classification. *2019 International Joint Conference on Neural Networks* pp:1–8. DOI:10.1109/ijcnn.2019.8852227
 207. Xie J., Chen S., Zhang Y., et al. (2021). Combining generative adversarial networks and multi-output CNN for motor imagery classification. *J. Neural Eng.* **18**:046026. DOI:10.1088/1741-2552/abec65
 208. Mao M., Komes D., Zhao S., et al. (2025). Non-invasive neuromodulation assisted by exogenous stimuli-responsive nanoplateforms for Alzheimer's disease and Parkinson's disease therapy. *Innov. Med.* **3**:100121. DOI:10.59717/j.xinn-med.2025.100121
 209. Yuste R., Goering S., Arcas B.A.Y., et al. (2017). Four ethical priorities for neurotechnologies and AI. *Nature* **551**:159–163. DOI:10.1038/551159a
 210. Bahador N., Jokelainen J., Mustola S., et al. (2021). Reconstruction of missing channel in electroencephalogram using spatiotemporal correlation-based averaging. *J. Neural Eng.* **18**:056045. DOI:10.1088/1741-2552/ac23e2
 211. An D., Guo Y., Lei N. et al. (2020). AE-OT: A new generative model based on extended semi-discrete optimal transport. *8th International Conference on Learning Representations* pp: 8996–9014. <https://openreview.net/pdf?id=HkldyTNYwH>.
 212. Anderson A.J., McDermott K., Rooks B., et al. (2020). Decoding individual identity from brain activity elicited in imagining common experiences. *Nat. Commun.* **11**:5916. DOI:10.1038/s41467-020-19630-y
 213. Li C., Qian X., Wang Y. et al. (2024). Enhancing cross-subject fMRI-to-video decoding with global-local functional alignment. *European Conference on Computer Vision* pp:353–369. DOI:10.1007/978-3-031-73010-8_21
 214. Wen S., Yin A., Furlanello T., et al. (2023). Rapid adaptation of brain-computer interfaces to new neuronal ensembles or participants via generative modelling. *Nat. Biomed. Eng.* **7**:546–558. DOI:10.1038/s41551-021-00811-z
 215. Karpowicz B.M., Ali Y.H., Wimalasena L.N., et al. (2025). Stabilizing brain-computer interfaces through alignment of latent dynamics. *Nat. Commun.* **16**:4662. DOI:10.1038/s41467-025-59652-y
 216. Ahuja C. and Sethia D. (2025). SS-EMERGE - self-supervised enhancement for multidimension emotion recognition using GNNs for EEG. *Sci. Rep.* **15**:14254. DOI:10.1038/s41598-025-98623-7
 217. Chen R.J., Wang J.J., Williamson D.F., et al. (2023). Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nat. Biomed. Eng.* **7**:719–742. DOI:10.1038/s41551-023-01056-8
 218. Yang Y., Zhang H., Gichoya J.W., et al. (2024). The limits of fair medical imaging AI in real-world generalization. *Nat. Med.* **30**:2838–2848. DOI:10.1038/s41591-024-03113-4
 219. Yao W., Chen X. and Wang S. (2025). Empowering functional neuroimaging: A pretrained generative framework for unified representation of neural signals. *arXiv Preprint arXiv:2506.02433*. DOI:10.48550/arXiv.2506.02433
 220. Liu R., Chen Y., Li A., et al. (2024). Aggregating intrinsic information to enhance BCI performance through federated learning. *Neural Netw.* **172**:106100. DOI:10.1016/j.neunet.2024.106100
 221. Dai J., Xu H., Chen T., et al. (2025). Artificial intelligence for medicine 2025: Navigating the endless frontier. *Innov. Med.* **3**:100120. DOI:10.59717/j.xinn-med.2025.100120
 222. Huang T., Xu H., Wang H., et al. (2023). Artificial intelligence for medicine: Progress, challenges, and perspectives. *Innov. Med.* **1**:10030. DOI:10.59717/j.xinn-med.2023.100030

FUNDING AND ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China under Grant 2023YFC2506902, National Natural Science Foundations of China under Grant 62172403 and 12326614. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

Conceptualization, Yi Guo and Shuqiang Wang; Literature survey, Yi Guo, Guanhua Ren, Shiqiang Ma, Jun Zhou, Hasan Al Nashash, Alireza Jolfaei, Zhiguo Zhang, Guanglin Li, Ye Li, and Yong Hu; Writing-original draft, Yi Guo and Shiqiang Ma; Writing-review & editing, Yi Guo, Guanhua Ren, Shiqiang Ma, Shuqiang Wang, Jun Zhou, Hasan Al Nashash, Alireza Jolfaei, Zhiguo Zhang, Guanglin Li, Ye Li, and Yong Hu; Funding acquisition, Shuqiang Wang; Supervision, Shuqiang Wang. All authors contributed to and approved the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DATA AND MATERIALS AVAILABILITY

Not applicable.