BIO Web of Conferences **237**, 03002 (2026) https://doi.org/10.1051/bioconf/202623703002
ICBB 2026

Decoding the Brain for Communication: A Review of Brain-to-Text, Brain-to-Speech, and Brain-to-Image Interfaces

Tong Gao*

Division of Biosciences, University College London, London, United Kingdom

Abstract: Brain-computer interfaces (BCIs) enable communication by directly decoding neural activity into comprehensible output forms. These interfaces can be broadly categorized into brain-to-text, brain-to-speech, and brain-to-image modalities. Brain-to-text BCIs translate neural activity into written text, facilitating communication without speech or physical interfaces. Brain-to-speech BCIs decode neural signals to synthesize spoken language, preserving personalized features such as tone and prosody, thereby enabling expressive communication. Brain-to-image BCIs reconstruct visual stimuli or imagined images from neural activity, providing a non-verbal channel for expressing complex concepts visually. This review synthesizes recent advancements, identifies ongoing challenges, and highlights the significance of these developments, emphasizing the potential to significantly improve communication and quality of life for individuals with severe motor impairments. Future work should explore minimally invasive implantation techniques, potentially combining benefits of invasive and non-invasive methods, to balance communication accuracy with user safety and accessibility.

1 Introduction

Effective communication is a fundamental aspect of human interaction, facilitating the exchange of ideas, emotions, and intentions across diverse social and professional contexts, from informal conversations to structured dialogues in meetings and conferences. Nevertheless, individuals experiencing severe motor impairments, such as those resulting from amyotrophic lateral sclerosis (ALS) or locked-in syndrome (LIS), encounter significant communication barriers. These conditions can lead to paralysis by damaging neural pathways responsible for transmitting motor signals between the brain and peripheral muscles, severely restricting speech production and the ability to manipulate conventional assistive communication devices[1, 2, 3]. To address this challenge, brain-computer interfaces (BCIs) have been developed to restore communication by directly translating neural activity into different forms of output[4, 5, 6].

BCIs are classified into invasive and non-invasive types, each with distinct advantages and limitations. Invasive BCIs require surgical implantation of intracortical microelectrode arrays or electrocorticography (ECoG) grids directly into the brain, providing high-SNR, high-resolution recordings, including single-unit activity[7, 8]. They have been applied in areas such as motor restoration, speech decoding, and brain-to-text communication [5, 6, 9]. However, they pose significant disadvantages, including surgical risks, potential immune responses, electrode degradation, and the need for long-term maintenance and recalibration [10]. In contrast, non-invasive BCIs record neural activity using external sensors such as electroencephalography (EEG), magnetoencephalography (MEG), and functional magnetic resonance imaging (fMRI), eliminating the risks associated with surgery and making them more accessible and scalable for broader applications [11, 12]. However, these systems suffer from lower signal resolution, limited frequency bandwidth, and significant interference from scalp and muscle artifacts, reducing their accuracy and real-time decoding capabilities[10].

Despite the inherent limitations of both invasive and non-invasive BCI methods, these technologies are driven by the fundamental human need to communicate [13] One application of this concept is brain-to-text communication, wherein brain activity is decoded into written text [5, 14]. Such brain-to-text BCIs could enable a user to compose emails or communicates during a text-based chat by thought alone. Beyond text-based communication, researchers are developing brain-to-speech BCIs that directly generate speech with personalized features such as tone, prosody, and voice characteristics from brain activity [2, 15, 16]. The brain-to-speech BCIs bypass the need for any typing interface by aiming to restore the full richness of spoken communication, providing an efficient and expressive channel for interaction [17]. Furthermore, brain-to-image BCIs aim to translate brain activity into visual content [4, 18]. This emerging modality would allow users to externalize mental images or concepts as pictures,

*Corresponding author: gao20030311@gmail.com

offering an alternative avenue for expression and creativity beyond verbal language [19].

All three of these brain-driven communication systems hold particular significance for individuals with severe motor impairments, who could regain crucial channels of interaction that were otherwise lost [20]. As shown in Figure 1, communication BCIs can be broadly organised into three major modalities—brain-to-text, brain-to-speech, and brain-to-image—each involving distinct methods, decoding algorithms, and applications. This review examines the current progress and future challenges associated with these three communication modalities.

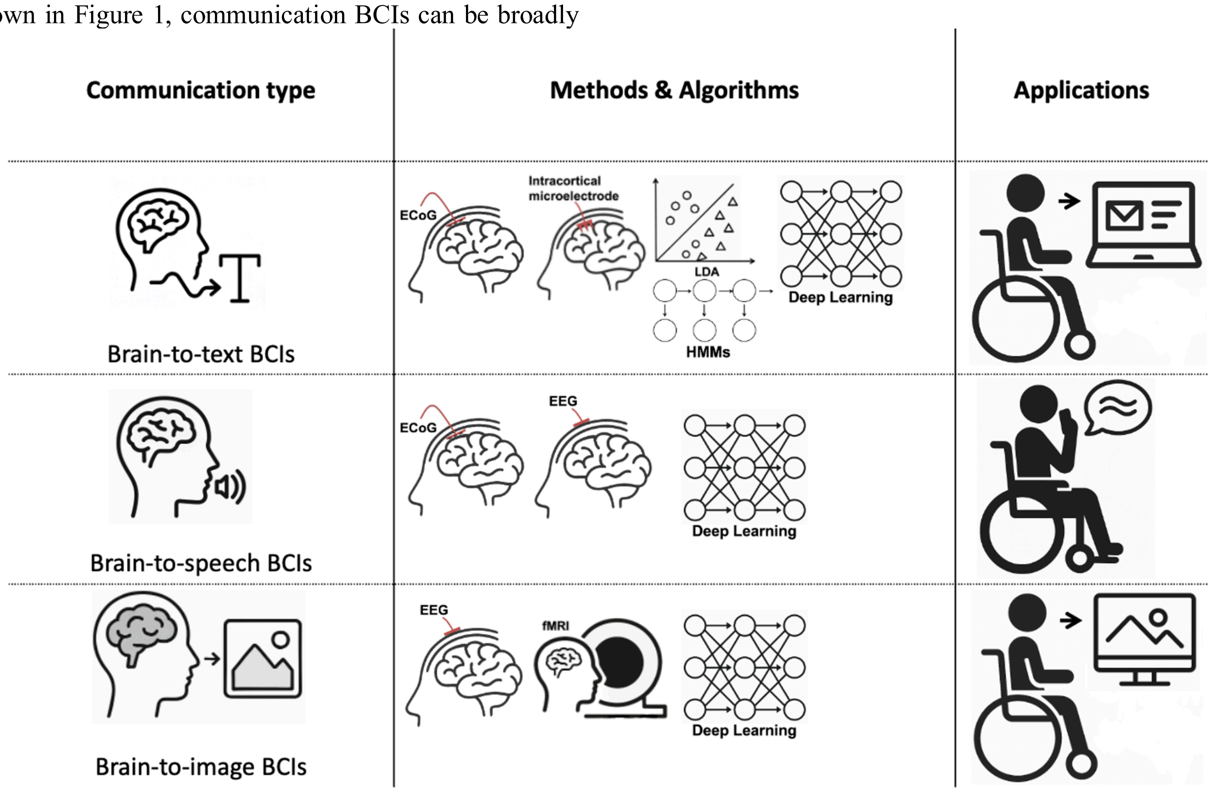

Figure 1.Overview of communication BCIs.

The figure illustrates three major types of communication BCIs–brain-to-text, brain-to-speech, and brain-to-image–together with their core methods, decoding algorithms, and applications. A portion of the illustration was conceptually generated using ChatGPT (OpenAI).

2 Brain-to-text BCIs

Over the past decade, brain-to-text BCIs have advanced significantly, aiming to restore communication for individuals with speech and motor impairments by decoding brain signals into textual representations. To achieve this, brain-to-text BCIs utilize different approaches, including cursor control-based interfaces [14], handwriting decoding [5, 21], speech-to-text conversion [22-27], phoneme-to-text translation [28], and Semantic-to-Text [29]. With these advancements, brain-to-text BCI accuracy and output speed have remarkably improved in the decade. Initial systems achieved 36% phoneme classification accuracy with an output speed of 33.6 words per minute [28], while modern intracortical BCIs have reached 97.5% accuracy [27] and an output speed of 90 characters per minute [5]. One of the key drivers behind this progress is the advancement of intracortical microelectrode array technology. Traditionally, ECoG-based BCIs use high-density electrode grids placed on the surface of the cerebral cortex to record population-level neural activity [22, ,23]. In contrast, intracortical BCIs surgically implant one to four microelectrode arrays directly into the cortical tissue, enabling access to single-neuron spiking activity and supporting significantly higher decoding accuracy and output speed [5, 14, 27]. Beyond invasive methods, non-invasive fMRI-based BCIs have been explored. For example, Tang developed an fMRI-based system that reconstructs perceived, imagined, and silent video-induced speech using a semantic-to-text approach. In addition to improvements in accuracy and speed, recent research has increasingly focused on stability and long-term usability. Traditional BCIs typically required frequent recalibration to maintain performance, posing a barrier to clinical deployment. To address this, Fan introduced a self-recalibrating intracortical BCI that maintained 93.84% accuracy over 403 days without supervised recalibration by using pseudo-labels for continual online learning. Collectively, these developments have been closely coupled with the evolution of brain-to-text BCI algorithms toward deep learning networks.

Deep learning-based algorithms significantly outperform traditional statistical models. Traditional BCI systems relied on probabilistic classifiers such as Hidden Markov Models (HMMs) and Linear Discriminant

Analysis (LDA) to map neural signals to text [14, 25]. However, HMMs are computationally and memory-intensive, which limits their efficiency—particularly on resource-constrained platforms—unless extensive optimization strategies are applied [30]. LDA, as a linear classifier, is not well-suited for handling nonlinear neural data, which limits its effectiveness in capturing the complex patterns of brain activity [31]. In contrast, deep neural networks, particularly RNNs, have demonstrated superior performance by capturing complex temporal dependencies in neural signals and enabling end-to-end learning [5, 23, 27]. For example, RNN-based decoders have significantly improved speech-to-text BCIs by leveraging sequential information to refine word predictions, while the integration of large language models (LLMs) has further enhanced the performance by capturing long-range dependencies, contextual meanings, and probabilistic relationships between words [21, 26]. These advancements have resulted in faster and more accurate text decoding, with modern BCIs achieving over 90% accuracy and real-time word prediction. However, Brain-to-text BCIs still face several limitations, including slow communication speeds, poor generalizability across users, and challenges in capturing the complexity of natural speech, as text lacks intonation and expressiveness. [12]. A promising direction is the potential integration of brain-spine interface approaches into communication BCIs. The work by Lorach introduced a real-time digital bridge that restored volitional walking by linking cortical activity to spinal cord stimulation. A similar closed-loop strategy could be applied to ALS and LIS patients, where a brainstem or facial motor BCI could restore speech production by stimulating residual speech motor pathways. Speech neuroprostheses have already demonstrated the ability to decode silent and imagined speech into intelligible sounds, with Metzger achieving direct brain-to-speech synthesis, including facial-avatar animation for expressive communication.

3 Brain-to-speech

Given the limitations of brain-to-text BCIs, researchers have increasingly turned towards brain-to-speech as alternatives to brain-to-text systems [15]. Brain-to-speech BCIs aim to restore spoken communication by decoding neural activity into audible speech, bypassing the need for text-based interfaces [12]. Recent studies have demonstrated that EEG and ECoG-based BCIs can reconstruct speech sounds and preserve individual voice characteristics, making them more intuitive for real-world communication[6, 32]. Over the past decade, brain-to-speech BCIs have advanced significantly, moving from phoneme-level reconstructions [28] to sentence synthesis [16, 33]. Accuracy has steadily improved, with speech decoding performance increasing from 47.1% word classification accuracy [33] to over 90% sentence intelligibility in recent trials[6, 16]. Similarly, output speed has progressed from an initial rate of 15.2 words per minute [33] to 78 words per minute [16], approaching natural speech rates. Several key technological advancements have driven these improvements. First, higher-density electrode arrays have enhanced neural signal capture, with modern ECoG grids covering broader speech-related cortical areas, including the motor, premotor, and inferior frontal cortices [15, 16]. Additionally, the use of more stable neural recording devices, such as intracortical implants, high-resolution ECoG, and even stereotactic EEG (sEEG), has enabled longer-term stability in speech decoding [32, 34]. Beyond hardware advancements, improvements in neural decoding algorithms have played a crucial role. Unlike brain-to-text BCIs, which have been explored for decades and initially relied on statistical models, brain-to-speech BCIs developed later, emerging primarily in the deep learning era. As a result, nearly all brain-to-speech systems from their inception have leveraged deep neural networks, bypassing the limitations of traditional probabilistic classifiers. These advancements in neural activity probing algorithms have played a crucial role in improving speech synthesis quality, with models incorporating RNNs to refine predictions and capture natural prosody and intonation [23, 35].Furthermore, self-calibrating brain-to-speech BCIs have emerged, reducing the need for frequent manual recalibration, thereby improving long-term usability [34].

4 Brain-to-image BCIs

Unlike brain-to-text and brain-to-speech BCIs, which focus on linguistic communication, brain-to-image BCIs aim to reconstruct visual information from neural activity, providing an alternative mode of expression for individuals with communication impairments. By translating brain signals into images, these systems hold potential for assisting them to convey thoughts and memories in a more intuitive and expressive manner. The primary approach to brain-to-image BCIs involves decoding neural activity, most commonly recorded via fMRI, using deep learning models to reconstruct perceived or imagined visual stimuli. [4, 18]. A common strategy in these models is to first map the fMRI signals to intermediate visual or semantic feature representations, using pretrained models such as Contrastive Language–Image Pretraining (CLIP) [36], Very Deep Variational Autoencoder (VDVAE) [37], or Vector Quantised Generative Adversarial Network (VQGAN) [38]. These de40coded features are subsequently used to condition generative models—typically latent diffusion models [18, 37] or generative adversarial networks (GANs) [38, 39]—to reconstruct images that approximate the original visual stimuli. In contrast to fMRI-based methods, visual image reconstruction from EEG signals is also feasible by decoding neural activity into semantic and latent visual features, which are subsequently used to guide image synthesis via pretrained GANs [40].

Over the past decade, significant advances have improved the quality and fidelity of reconstructed images, transitioning from low-resolution grayscale reconstructions to high-resolution, semantically accurate images with detailed textures and object structures [36, 38]. Among the key factors contributing to these improvements is the refinement of neural recording techniques, including the development of high-density fMRI scanning protocols and advancements in EEG spatial filtering [40]. Additionally, the introduction of self-supervised learning and pre-trained deep neural networks has enhanced feature extraction, allowing more robust and generalizable decoding across individuals [18]. The integration of latent diffusion models and GANs has further advanced the realism of reconstructed images by capturing hierarchical features of visual stimuli, bridging the gap between neural representation and perceptual experience [37, 38]. Recent developments have also enabled self-calibrating systems that adapt to individual neural patterns over time, reducing the need for extensive user training and improving the practical applicability of brain-to-image BCIs [36]. Despite these advancements, current decoding performance remains superior for individually tailored models compared to generalised group-level decoding, reflecting significant variability in neural representations among individuals.

5 Conclusion

This review synthesized recent progress in brain-to-text, brain-to-speech, and brain-to-image interfaces, each offering unique pathways to restore communication for individuals with severe motor impairments. While all three modalities have achieved notable improvements in decoding accuracy, output quality, and real-time performance, challenges remain in balancing signal resolution, invasiveness, and user adaptability. Text-based systems offer structured communication, speech-based systems enhance expressiveness, and image-based systems open up visual channels for ideation and emotion. Looking ahead, future research should explore the possibility of using minimally invasive implants for communication restoration, similar to those used in spinal cord interfaces for enabling movement [41] Additionally, advances in self-calibrating algorithms and personalized decoding models will be crucial for improving long-term usability and making these systems more practical for daily life.

References

1. Laureys S, Pellas F, Van Eeckhout P, et al. Progress in Brain Research. Elsevier, pp.495–611. (2005)
2. Smith, E. and Delargy, M. Locked-in syndrome. BMJ. 330(7488), pp.406–409. (2005)
3. Kübler, A. and Birbaumer, N. Brain-computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients? Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology. 119(11), pp.2658–2666. (2008)
4. Shen, G., Dwivedi, K., Majima, K., Horikawa, T. and Kamitani, Y. End-to-End Deep Image Reconstruction From Human Brain Activity. Frontiers in Computational Neuroscience. 13.(2019)
5. Willett, F.R., Avansino, D.T., Hochberg, L.R., Henderson, J.M. and Shenoy, K.V. High-performance brain-to-text communication via handwriting. Nature. 593(7858), pp.249–254.(2021)
6. Angrick M, Luo S, Rabbani Q, et al. Online speech synthesis using a chronically implanted brain–computer interface in an individual with ALS. Scientific Reports. 14(1), p.9617. (2024)
7. Daly, J.J. and Wolpaw, J.R. Brain–computer interfaces in neurological rehabilitation. The Lancet Neurology. 7(11), pp.1032–1043. (2008)
8. Wang, Yang, Yang, X., Zhang, X., Wang, Yijun and Pei, W. Implantable intracortical microelectrodes: reviewing the present with a focus on the future. Microsystems & Nanoengineering. 9(1), p.7. (2023)
9. Zhang, X., Song, Z., Li, S., Zhang, Y., Fang, T., Wang, S., Li, H., Lin, Y., Jia, J., Zhang, L., Kang, X. and Wang, H. 2021. Brain Computer Interface for the Hand Function Restoration. 9th International Winter Conference on Brain-Computer Interface (BCI), pp.1–6.(2021)
10. Xu, S., Liu, Y., Lee, H. and Li, W. Neural interfaces: Bridging the brain to the world beyond healthcare. Exploration. 4(5), p.20230146. (2024)
11. Tang, J., LeBel, A., Jain, S. and Huth, A.G. Semantic reconstruction of continuous language from non-invasive brain recordings. Nature Neuroscience. 26(5), pp.858–866. (2023)
12. Silva, A.B., Littlejohn, K.T., Liu, J.R., Moses, D.A. and Chang, E.F. The speech neuroprosthesis. Nature Reviews Neuroscience. 25(7), pp.473–492. (2024)
13. Chandler, J.A., Van der Loos, K.I., Boehnke, S., Beaudry, J.S., Buchman, D.Z. and Illes, J. Brain Computer Interfaces and Communication Disabilities: Ethical, Legal, and Social Aspects of Decoding Speech From the Brain. Frontiers in Human Neuroscience. 16.(2022)
14. Pandarinath, C., Nuyujukian, P., Blabe, C.H., Sorice, B.L., Saab, J., Willett, F.R., Hochberg, L.R., Shenoy, K.V. and Henderson, J.M. High performance communication by people with paralysis using an intracortical brain-computer interface. eLife. 6, p.e18554.(2017)
15. Herff, C., Diener, L., Angrick, M., Mugler, E., Tate, M.C., Goldrick, M.A., Krusienski, D.J., Slutzky, M.W. and Schultz, T. Intelligible Speech From Brain Activity in Motor, Premotor, and Inferior Frontal Cortices. Frontiers in Neuroscience. 13, p.1267. (2019)
16. Metzger S L, Littlejohn K T, Silva A B, et al. A high-performance neuroprosthesis for speech decoding and avatar control. Nature. 620(7976), pp.1037–1046.(2023)
17. Luo, S., Rabbani, Q. and Crone, N.E. Brain-Computer Interface: Applications to Speech Decoding and Synthesis to Augment Communication. Neurotherapeutics. 19(1), p.263. (2022)

18. Chen, Z., Qing, J., Xiang, T., Yue, W.L. and Zhou, J.H. Seeing Beyond the Brain: Conditional Diffusion Model with Sparse Masked Modeling for Vision Decoding. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp.22710–22720.(2023)
19. Vanutelli, M.E., Salvadore, M. and Lucchiari, C. BCI Applications to Creativity: Review and Future Directions, from little-c to C2. Brain Sciences. 13(4), p.665. (2023)
20. Zhang, H., Jiao, L., Yang, S., Li, H., Jiang, X., Feng, J., Zou, S., Xu, Q., Gu, J., Wang, X. and Wei, B. Brain–computer interfaces: the innovative key to unlocking neurological conditions. International Journal of Surgery. 110(9), pp.5745–5762.(2024)
21. Fan, C., Hahn, N., Kamdar, F., Avansino, D., Wilson, G.H., Hochberg, L., Shenoy, K.V., Henderson, J.M. and Willett, F.R. 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. (2023)
22. Herff, C., Heger, D., de Pesters, A., Telaar, D., Brunner, P., Schalk, G. and Schultz, T. Brain-to-text: decoding spoken phrases from phone representations in the brain. Frontiers in Neuroscience. 8. (2015)
23. Makin, J.G., Moses, D.A. and Chang, E.F. Machine translation of cortical activity to text with an encoder–decoder framework. Nature Neuroscience. 23(4), pp.575–582.(2020)
24. Sun, P., Anumanchipalli, G.K. and Chang, E.F. Brain2Char: a deep architecture for decoding text from brain recordings. Journal of Neural Engineering. 17(6), p.066015.(2020)
25. Wandelt, S.K., Bjånes, D.A., Pejsa, K., Lee, B., Liu, C. and Andersen, R.A. Online internal speech decoding from single neurons in a human participant. MedRxiv,11.02.22281775. (2022)
26. Willett F R, Kunz E M, Fan C, et al. A high-performance speech neuroprosthesis. Nature. 620(7976), pp.1031–1036.(2023)
27. Card N S, Wairagkar M, Iacobacci C, et al. An Accurate and Rapidly Calibrating Speech Neuroprosthesis. New England Journal of Medicine. 391(7), pp.609–618.(2024)
28. Mugler, E.M., Patton, J.L., Flint, R.D., Wright, Z.A., Schuele, S.U., Rosenow, J., Shih, J.J., Krusienski, D.J. and Slutzky, M.W. Direct classification of all American English phonemes using signals from functional speech motor cortex. Journal of Neural Engineering. 11(3), p.035015. (2014)
29. Tang, X., Shen, H., Zhao, S., Li, N. and Liu, J. Flexible brain–computer interfaces. Nature Electronics. 6(2), pp.109–118. (2023)
30. Ajani, T.S., Imoize, A.L. and Atayero, A.A. An Overview of Machine Learning within Embedded and Mobile Devices–Optimizations and Applications. Sensors. 21(13), p.4412. (2021)
31. Mridha, M.F., Das, S.C., Kabir, M.M., Lima, A.A., Islam, M.R. and Watanobe, Y. Brain-Computer Interface: Advancement and Challenges. Sensors. 21(17), p.5746. (2021)
32. Kohler, J., Ottenhoff, M.C., Goulis, S., Angrick, M., Colon, A.J., Wagner, L., Tousseyn, S., Kubben, P.L. and Herff, C. Synthesizing Speech from Intracranial Depth Electrodes using an Encoder-Decoder Framework. Neurons, Behavior, Data analysis, and Theory. 6(1). (2022)
33. Moses D A, Metzger S L, Liu J R, et al. Neuroprosthesis for Decoding Speech in a Paralyzed Person with Anarthria. New England Journal of Medicine. 385(3), pp.217–227. (2021)
34. Luo S, Angrick M, Coogan C, et al. Stable Decoding from a Speech BCI Enables Control for an Individual with ALS without Recalibration for 3 Months. Advanced Science. 10(35), p.2304853. (2023)
35. Anumanchipalli, G.K., Chartier, J. and Chang, E.F. Speech synthesis from neural decoding of spoken sentences. Nature. 568(7753), pp.493–498. (2019)
36. Ozcelik, F. and VanRullen, R. Natural scene reconstruction from fMRI signals using generative latent diffusion. Scientific Reports. 13(1), p.15666. (2023)
37. Lu, Y., Du, C., Zhou, Q., Wang, D. and He, H. 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. (2023)
38. Ozcelik, F., Choksi, B., Mozafari, M., Reddy, L. and VanRullen, R. Reconstruction of Perceived Images from fMRI Patterns and Semantic Brain Exploration using Instance-Conditioned GANs. 2022 International Joint Conference on Neural Networks (IJCNN), pp.1–8. (2022)
39. Shen, G., Horikawa, T., Majima, K. and Kamitani, Y. Deep image reconstruction from human brain activity. PLOS Computational Biology. 15(1), p.e1006633. (2019)
40. Shimizu, H. and Srinivasan, R. Improving classification and reconstruction of imagined images from EEG signals. PLOS ONE. 17(9), p.e0274847. (2022)
41. Lorach H, Galvez A, Spagnolo V, Martel F, Karakas S, Intering N, et al. Walking naturally after spinal cord injury using a brain–spine interface. Nature. 618(7963), pp.126–133. (2023)