npj | biomedical innovations

Article

https://doi.org/10.1038/s44385-026-00086-6

Symbiotic brain-machine drawing via visual brain-computer interfaces

Gao Wang[1], Yingying Huang[1,2], Lars Muckli[2] & Daniele Faccio[1]

Brain-computer interfaces (BCIs) are evolving from research prototypes into clinical, assistive, and performance enhancement technologies. Despite the rapid rise and promise of implantable technologies, there is a need for better and more capable wearable and non-invasive approaches whilst also minimising hardware requirements. We present a non-invasive BCI for iterative selection-based mind-drawing that infers a subject's internal visual intent through iterative selection of adaptive visual probes presented on a screen encoded at different flicker-frequencies and analyses the steady-state visual evoked potentials (SSVEPs). Gabor-inspired or machine-learned policies dynamically update the spatial placement of the visual probes on the screen to explore the image space and reconstruct simple imagined shapes within approximately two minutes or less using just single-channel EEG data. Additionally, by leveraging stable diffusion models, reconstructed mental images can be transformed into realistic and detailed visual representations. Whilst we expect that similar results might be achievable with e.g. eye-tracking techniques, our work shows that symbiotic human-AI interaction can increase BCI bit-rates by more than a factor 5x, providing a platform for future development of AI-augmented BCI.

Brain-computer interface (BCI) technologies enable direct communication between the brain and a computer, with applications ranging from controlling the computer as a means to regain, for example, the ability to move or navigate the world, to the actual decoding of human thought[1–5]. BCIs can be broadly divided into two paradigms. Passive BCIs infer cognitive states such as workload or fatigue, from spontaneous brain activity[6,7], while active BCIs rely on subjects intentionally modulating their neural signals to issue commands[8,9].

The ability to directly observe and decode subjective visual imagery is of fundamental importance, holding transformative potential for assistive communication, creative co-design and the study of mental health. The central limitation of existing methods is their reliance on pre-trained, data-hungry models, which act as a "dictionary" of known images[10–13]. This fundamentally restricts the creative potential of mind drawing. In contrast to direct mental imagery decoding approaches that attempt to reconstruct arbitrary visual content from neural signals alone, our method employs brain-guided collaborative visual exploration, where the system adaptively presents visual probes and the subject's selective attention guides the reconstruction process.

Following the development of a "P300 speller"–based painting system[12], a BCI painting system employing a hybrid SSVEP/P300 control approach was later proposed[13]. However, both methods rely heavily on using the SSVEP to "click" the static graphical subject interfaces (GUIs), and they lack the intelligence to predict the subject's intention[14].

Non-invasive neural decoding methods have been developed using the main brain sensing technologies, i.e. electroencephalography (EEG), functional magnetic resonance imaging (fMRI), functional near-infrared spectroscopy (fNIRS), and magnetoencephalography (MEG). In recent years, researchers have implemented neural decoding methods to monitor the brain state or to communicate with others[9,15–17] but also to reconstruct mental or visual images from brain signals. For example fMRI, with its high spatial resolution, has provided foundational insights into how visual information is encoded in the brain, has been used to reconstruct images from blood-oxygen-level-dependent (BOLD) signals[18–24].

Recent work has also highlighted the possibility to decode visual imagery via fNIRS[25]. MEG was proposed for real-time reconstruction of visual perception[26] and classification[27]. EEG is more portable than MEG, (the fewer EEG channels, the more portable) and also has high temporal resolution and has been implemented to achieve reconstruction, from visual texture to natural images, with the help of Neural Networks[26,28–37]. More recently, the 'BrainVis' approach[38] demonstrated state-of-the-art semantic fidelity reconstructions and generation quality. These mind-drawing or image-reconstruction approaches rely on high-dimensional subject data

[1]School of Physics & Astronomy, University of Glasgow, Glasgow, UK. [2]School of Psychology and Neuroscience, University of Glasgow, Glasgow, UK.
e-mail: daniele.faccio@glasgow.ac.uk

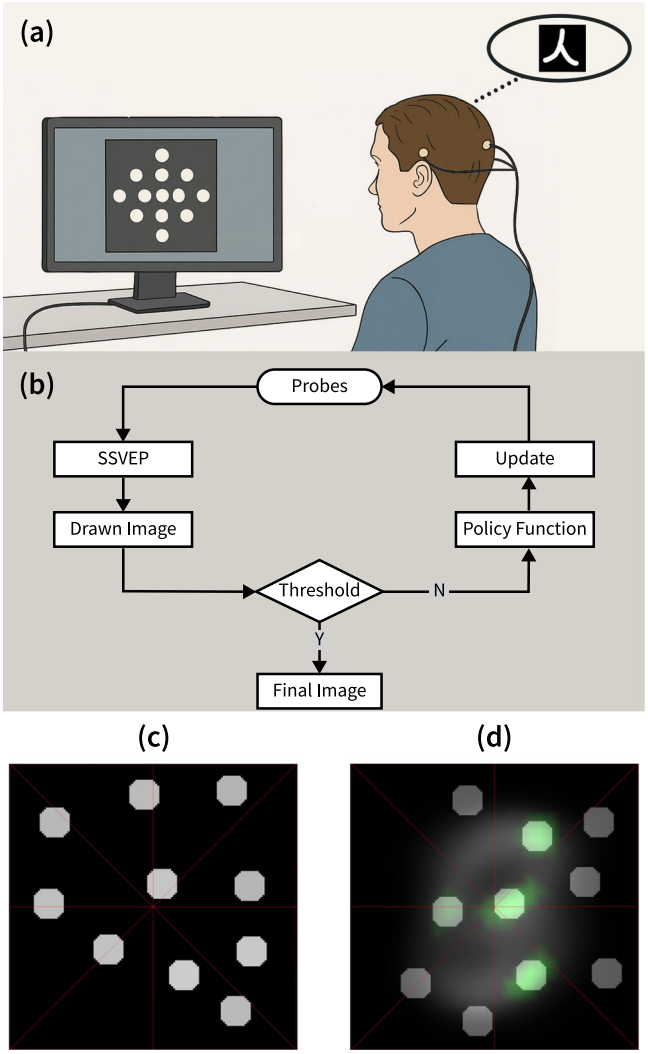

Fig. 1 | Experimental setup and workflow. a Setup of the EEG-based mind-drawing system (the picture was AI-assisted (ChatGPT, OpenAI) and subsequently redrawn and finalized by the authors). The EEG device is custom-built and consists of a headband housing three wet electrodes (using saline solution to improve contact): two placed on the temples (ground and reference) and one at the occipital Oz position. The subject selects the disc that overlaps most with their imagined image. **b** Workflow of the mind-drawing process (Y: Yes; N: No). **c** Example screen-shot stimulus under the Gabor policy function, where discs are randomly arranged and each flickers at a unique frequency. **d** Example screen-shot of the stimulus under the data-driven policy function, where discs are randomly arranged, with some representing green features and others appearing as standalone elements.

and deep learning algorithms, which are complex to implement and typically do not generalise to unseen imagery.

Despite significant progress in the field, achieving true mind-drawing in real-time with model-free and high-fidelity reconstruction of arbitrary images from imagination alone remains a challenge. A further open challenge for the field is relatively low information transfer bit-rates of BCIs, with hints that these might be fundamentally limited to ~10 bits/second[39].

Here, we propose a non-invasive neural decoding framework that overcomes key limitations of existing mind-drawing approaches. We develop reconstruction approaches based on an iterative collaborative search in image space, guided by dynamically updated sampling weights for successive visual stimuli (probes) that are presented to the subject. Crucially, our system does not rely on any pre-trained generative models. Instead, it generates the layout of subsequent visual probes in real time through a Gabor analysis of accumulated neural evidence from steady-state visual evoked potentials (SSVEPs). This allows the system to adaptively adjust its detection strategy based on the structure of mental imagery as this iteratively emerges during the drawing process, efficiently focusing resources on the most informative regions in the visual field. Our approach implements adaptive SSVEP-guided spatial localisation through iterative feedback that enables the subject and system to jointly converge on the intended image.

Preliminary human experiments demonstrate that this neuro-adaptive computational imaging framework can reconstruct simple, imagined visual shapes in approximately two minutes or less using only single-channel SSVEP data. The bit-rate of the mind-drawing process is improved by more than up to a factor 5x when compared to a simple readout of the SSVEP signal from the same device.

Results

The experimental setup is illustrated in Fig. 1a. The subject was seated in front of a computer monitor while wearing a single-channel EEG device and was instructed to select the flickering disc that overlapped most with their imagined image. Figure 1b illustrates the workflow of mind-drawing where the human is in the loop throughout the process. The disc positions are updated by the policy function based on the accumulated historical data. Importantly, the task does not require maintaining a stable mental image. Instead, subjects repeatedly identify the probe with best overlap from patterns updated every few seconds, making the cognitive task intuitive and less demanding than sustained mental imagery.

At the beginning of each iteration, 10 white visual stimulus probes (flickering discs) were displayed on a black-background screen (as shown in Fig. 1c). The subject was then instructed to focus on the region where the overlap between one presented probe and the mentally imagined object appeared to be strongest within each iteration. After 4 s of flashing stimulation (chosen based on SNR requirements: SSVEP signals require ≥2 s to stabilize, and 4 s represents the optimal accuracy-speed tradeoff for single-channel detection), the SSVEP response was recorded and the Canonical Correlation Analysis (CCA) value (a correlation coefficient ranging from 0 to 1[40]) was used to weight the selected probe/pattern that was then added to the drawn image as real-time EEG feedback to the subject displayed on the screen background in red. This process was repeated iteratively, with updated probe positions presented to the subject each time. The updated positions were determined by a policy function that determined the placement of the next set of probes (see Methods for a detailed description of the policies and how these were applied). We have tested two types of policy functions. Figure 1c shows an example stimulus under the Gabor policy function, in which discs are randomly arranged and each disc flickers at a unique frequency. Figure 1d illustrates an example stimulus under the data-driven policy function, where disks that have been selected by the subject at the previous iterations actually now represent machine-learned 'basis functions', shown in green. The final image therefore appears from the superposition of multiple basis functions as opposed to the simple discs used in the Gabor analysis approach (see details below and in Methods).

The selected probes were iteratively updated and added to the drawn-image. In our tests, we found that the optimal choice was a total of 25 iteration runs for the Gabor policy and seven iterations for the more efficient choice (but more constrained, see below) data-driven policy case. A 25-session measurement lasts 2.5 min, also including 2-s resting periods. The final drawn image $I_n(x, y)$ was then reconstructed based on the accumulated weighted pattern,

$$I_n(x,y) = \sum_{i=1}^{n} B_j P_j(x,y). \quad (1)$$

where B_j is the weight, and the pattern $P_j(x, y)$ is a Gaussian disc in the Gabor policy or is a 'basis pattern' in the data-driven policy (see Methods for a detailed description). Through iterative refinement guided by real-time EEG feedback and the selected policy function, the reconstructed image gradually converged toward the imagined object by the participant.

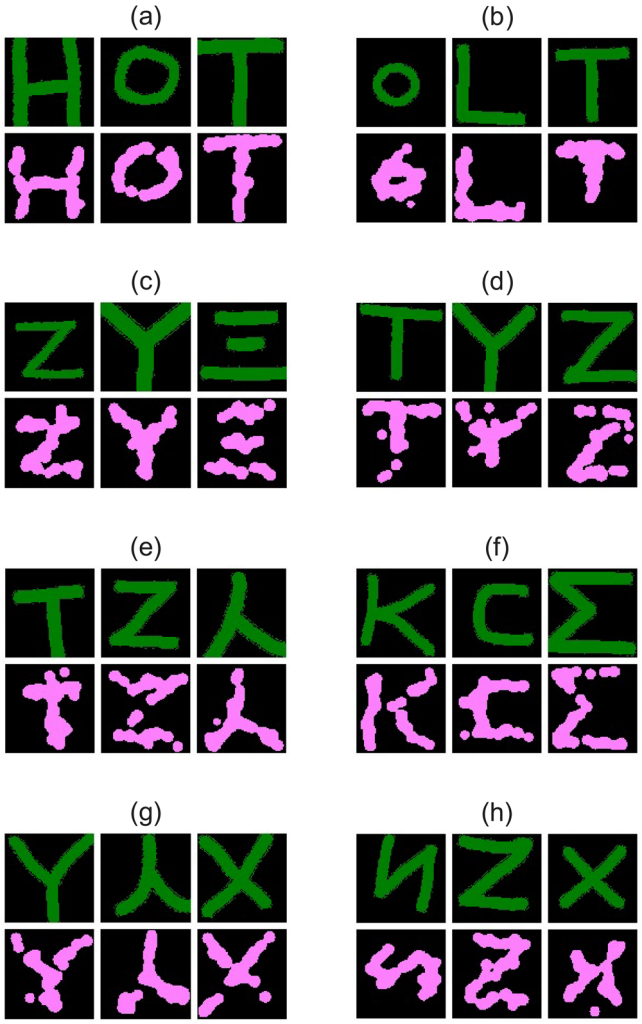

Fig. 2 | Imaging results from eight subjects, each targeting three imagined shapes. Panels are grouped three at a time, one group for each subject, i.e. panels **a** correspond to subject 1, **b** subject 2, **c** subject 3, **d** subject 4, **e** subject 5, **f** subject 6, **g** subject 7, **h** subject 8. Green-coloured images represent the handwritten target after image resizing, while magenta-coloured images indicate the reconstructed images generated by our BCI system. We calculated the cosine similarity (COSS) between each ground-truth and reconstructed image pair to quantify accuracy. The overall average COSS across all subjects was 0.76 (±0.04), indicating a good level of reconstruction fidelity.

We conducted three single-channel EEG-based mind-drawing experiments:

Experiment 1: Gabor-policy inspired mind drawing, including eight subjects with each subject drawing 3 different images, so to achieve multi-subject validation and estimate similarity and information rate analysis;

Experiment 2: Data-driven mind drawing, including one subject for different handwritten digit images, demonstrates the effect of data-driven policy as characterized by mutual information;

Experiment 3: Gabor-policy mind drawing enhanced with stable diffusion, including different drawing results under the same prompt from two different subjects, leading to different detailed images from SD's output. The varying number of participants reflects different experimental scopes: Experiment 2 serves to test test-retest reliability, whilst Experiment 1 validates robustness across multiple participants.

Gabor-inspired mind drawing

To establish a ground truth, subjects were first required to hand-draw the image they intend to imagine in the experiment and then perform the mind-drawing task using the proposed system. Our design prioritizes multi-subject validation (8 participants × 3 images) to assess robustness for real-world BCI applications, focusing on semantic alignment rather than pixel-perfect reproduction. We evaluated the performance across multiple subjects to assess the robustness of the approach. Eight healthy human subjects participated and each completed mind-drawings of three different images with 25 iterations. The results are illustrated in Fig. 2, showcasing the system's ability to reconstruct simple geometric shapes with varying degrees of accuracy across subjects. Green-coloured images represent the handwritten target after a shape-preserving transformation (i.e. rescaling, rotation, translation) so that it has same size as the mind-drawn image, shown as a magenta-coloured image. The overlap regions between the ground-truth and the reconstructed images appear in white. We calculated the cosine similarity (COSS) between each ground-truth and reconstructed image pair to quantify accuracy. The overall average COSS across all subjects was 0.76 (±0.04), indicating a good level of reconstruction fidelity.

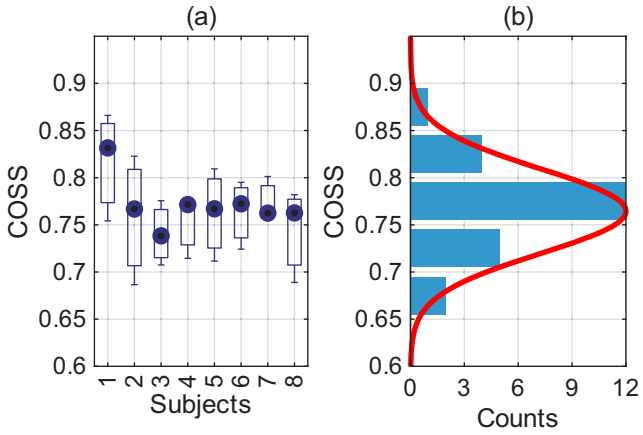

Fig. 3 | Distribution of image cosine similarity (COSS) values between target and reconstructed image pairs. a Box plot of COSS values for each subject. **b** Histogram of COSS values across all subjects.

Figure 3 shows a COSS box plot and histogram distributions that illustrates how the COSS values for each subject are relatively consistent, with most values ranging between 0.7 and 0.8. This indicates that the system is robust and generalises well across individuals. The box plot indicates that there is some variability in reconstruction accuracy between subjects, with some achieving higher COSS values than others. This variability may be attributed to individual differences in mental imagery, attention and focus during the task.

We then estimated the mutual information (MI) between the target and the reconstructed images, as MI captures the amount of shared information and provides an information-theoretic measure of reconstruction fidelity. The MI (see Methods, Eq. (3)) results for our measurements are shown in Fig. 4: the plot starts at 91 bits (corresponding to a full black image), and the MI of all subjects increases with the number of iterations and after 25 iterations reaches 222 bits, corresponding to a bit rate of 1.31 bits/second.

We estimated the more widely used Information Transfer Rate (ITR; see Methods, Eq. (4)), which provides the maximum theoretical information transfer rate (see Methods, Eq. (5)), $ITR_{max} \approx 0.83$ bit/s.

Remarkably, we find that the measured bit rate based on the mutual information between the target and reconstructed images (1.31 bits/second) is ~1.6× larger than the maximum ITR predicted for a standard BCI (0.83 bits/second)[41,42]. We attribute this to the effect of the iterative feedback and optimisation policy that is used to choose the optimal probe placement at each iteration, which in turn is based on information from the previous iteration. This indicates that our real-time brain-computer cooperation (implemented through our iterative optimisation policy) leads to a much higher bit rate and higher performance compared to what one might expect from a simple, one-way communication BCI.

Data-driven mind drawing

To further accelerate and enhance reconstruction, we leveraged prior datasets to inform probe selection, specifically by using machine-learning to create 'basis functions' that are used in place of the Gaussian discs (used in the Gabor policy case) to iteratively create the image. We therefore compromise image generalisation in favour of reconstruction speed as the use of prior datasets (that determine the specific shapes of the 'basis functions') unavoidably implies that the images that are drawn will need to resemble images in the dataset. Specifically, we conducted the second experiment with the same procedures as in Experiment 1 but using the MNIST dataset of handwritten digits with a data-driven policy function (see Methods). We emphasize that using MNIST digits was not to propose a practical digit classification system (simpler with direct selection or eye-tracking). Rather, our objective was to demonstrate significantly higher brain-AI bit-rates using a well-recognized dataset enabling direct comparison. Our iterative approach showcases single-electrode EEG interfaced with AI, dynamically exploring digit space through progressive refinement rather than requiring all options pre-displayed. The system successfully reconstructed digits based on SSVEP signals, demonstrating its ability to handle more complex visual stimuli beyond simple geometric shapes although, of course, these stimuli do need to bear resemblance to the images in the data training set. Reconstruction accuracy varied among digits, with structurally simpler digits generally yielding higher fidelity.

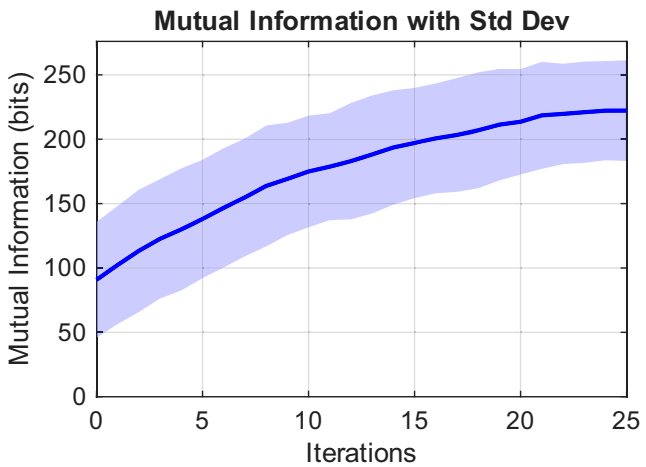

Fig. 4 | Mutual information (with standard deviation across iterations and subjects) increases with the number of iterations, although the rate of increase slows down as more iterations are added. The final MI reaches 222 bits starting from 91 bits (MI from an image with all black pixels), corresponding to an average rate of 1.31 bit/s.

As shown in Fig. 1d, selected discs are shown with the corresponding 'basis function' in green in the background. Figure 5 presents representative results for the digit "7" with increasing iterations from left to right. The top row shows, for comparison, the 'raw' image that is obtained based on Eq. (1) and showing a Gaussian disc (similarly to what was done for the Gabor policy approach) at the position of the disc from the SSVPE signal. The bottom row presents the data-driven policy image results, derived from the same process but now using the weighted 'basis functions' and near-neighbour analysis. As can be seen from the last panels in each row (f) and (l), respectively, the full data-driven pipeline provides images that are better representations of the intended image.

As more 'basis functions' are accumulated into near-neighbour analysis, the reconstructed mental image becomes progressively more detailed and structured.

The detail with which the image generation occurs is also interesting - at the early stages, the reconstruction is ambiguous, sometimes resembling a "0". With additional iterations, the image gradually clarifies, oscillating between "7" and "3," and ultimately converging to a clear "7" with high fidelity. Notably, in this experiment, the system requires only five iterations to produce a rough outline of the target digit, and six iterations to achieve a recognisable form. Each subsequent iteration further refines the reconstruction, illustrating the system's capacity for adaptive learning and progressive improvement based on accumulated SSVEP evidence. Additional representative results are provided in the SM.

We estimated the MI between the target and reconstructed images, as done in Experiment 1. As shown in Fig. 6, the MI values demonstrate the progressive improvement of reconstruction quality across iterations, reaching rates as high as 4.21 bits/s. This increase in information rate is driven by the machine-learned data-driven policy, which is introducing additional information at each iteration. However, we need to recall that this increase of bit rate from 0.83 to 4.21 bit/s (an increase of more than a factor 5×) comes at the expense of generalisability.

Generative enhancement via stable diffusion

To evaluate whether the reconstructed images could be further refined into realistic visual representations, we use the same approach as in Experiment 1, i.e., we adopt the Gabor-policy approach but then augment this with a final step that applies Stable Diffusion as a generative enhancement model. It is important to note that this Stable Diffusion step is a conditioned post-processing stage that introduces semantic priors unrelated to the EEG signals and should not be interpreted as improved neural decoding fidelity. Stable Diffusion is a generative model capable of producing high-quality images from text prompts and/or image inputs. The rationale for this was to test whether conditioning the evolving mental image would allow Stable Diffusion to complete and sharpen the reconstructions, driven by neural signals. As shown in Fig. 7, 4 different classes of images are illustrated, including a robot, a tree, a desk lamp, and an aircraft. For each class, we show

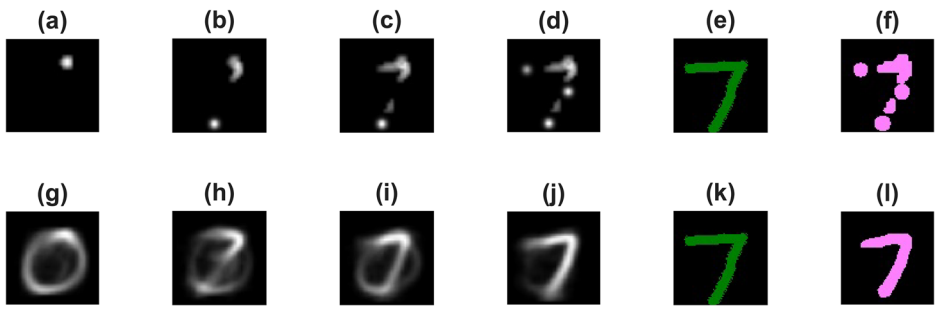

Fig. 5 | Reconstruction results for the MNIST digit "7". The first row (**a**–**d**) shows the step-by-step reconstruction of the digit from SSVEP signals by placing a Gaussian disc or 'basis functions' at the location of each disc selected by the SSVEP signal. The second row (**g**–**j**) presents the corresponding images generated by the system at each step. For simplicity, we only show the images every 2 iteration steps. Panels (**e**, **f**) and (**k**, **l**) show the binarised reconstructed and predicted images with the aligned handwritten target in different colour channels: green indicates the aligned handwritten target, magenta represents the reconstructed mental image. The COSS values for the binarised reconstructed and predicted images are 0.75 and 0.78, respectively, indicating a high degree of similarity to the target image.

two mind-drawing results under the same prompt (but different mind-drawing sessions) to recreate a new detailed image from a mind-drawn image. This approach bridges the gap between abstract neural reconstructions and photorealistic imagery, enabling applications such as creative co-design, assistive communication, and personalised content generation.

Discussion

We propose a non-invasive neural decoding approach that enables mind-drawing through an adaptive, iterative process that is based on a straightforward, single-channel EEG device.

The system's performance across multiple subjects demonstrates its robustness, with promising results in reconstructing simple visual forms. Interestingly, we found that the mutual information of the mind-drawing system is significantly larger than the maximum expected ITR. This is a result of the iterative and policy-driven optimisation process of the computer prompts, supporting the concept of true brain-machine cooperation. Our data-driven method achieves 4.21 bit/s (1.6–5× improvement over standard P300/SSVEP BCIs at 0.5–1.0 bit/s), whilst the Gabor-policy method achieves 1.31 bit/s. The data-driven approach requires only 7 iterations (~1 min) for recognizable digit reconstruction. The 4-second stimulation window ensures sufficient SNR for stable single-channel SSVEP detection. Future acceleration strategies include increasing displayed probes, optimizing probe placement, and leveraging multi-channel EEG to reduce windows to 2–3 s. Current speeds (~2 min per drawing) are acceptable for assistive communication, whilst predefined shape selection completes within seconds.

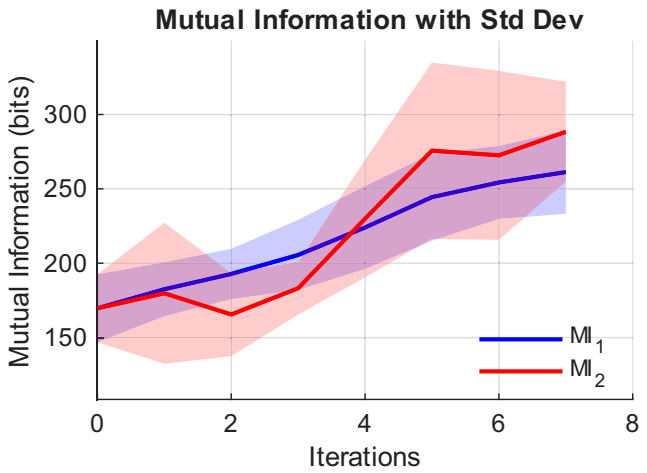

Fig. 6 | Mutual information during the reconstruction of MNIST digits, averaged over four different digit reconstructions ("7", "2", "4", and "8"). MI_1 and MI_2 denote the mutual information of the SSVEP reconstruction alone (i.e. same approach as top row in Fig. 5) and of the data-driven policy images (i.e. same approach as bottom row in Fig. 5), increasing from 170 bits to 261 bits and 288 bits, respectively. The corresponding bit rates are 3.25 bits/s and 4.21 bits/s.

We note that a similar approach could in principle also be implemented with other technologies, specifically for example with an eye tracker. However, the key point we wanted to make here is that it is possible to achieve this with EEG or neural signals alone as this might be relevant for future development towards more complex mind-driven systems. Our approach also has several advantages. First, it is a very affordable system that can be home-built for less than $100. Second, it can use SSVEP-based attention as a graded output to weight patterns (the weighting is derived directly from the intensity of the SSVEP signal), whereas an eye tracker typically provides only a binary focus position. Finally, SSVEPs can be modulated both by attention and by the focus point, while eye trackers rely solely on eye movements.

Our primary ambition with this work was to investigate what can be achieved in terms of brain-AI collaboration with the simplest possible EEG setup. While this choice ultimately limits what can be achieved, it also highlights what can still be achieved with minimal resources. The current system is optimized for simple geometric shapes, letters, and digits, which defines the scope of our contribution. Further investigation is required to either adapt the current approach to new specific tasks or, alternatively, to expand this approach into a general-purpose BCI.

The ability to infer non-verbal, open-ended spatial geometric content in real-time could provide an opportunity for new applications in neuroscience, cognitive science, and assistive communication. By prioritising subject experience, the system allows for intuitive interaction and adaptation to individual mental imagery patterns. We expect that with further developments, by focusing on overlapping regions between visual probes and imagined objects such that the system effectively captures the subject's intent, will also enable a more natural and engaging BCI experience.

Future multi-channel EEG could provide: 2–4x SNR enhancement via spatial filtering (reducing stimulus duration to 2–3 s), bit-rates of 6–8 bit/s through parallel processing, and improved robustness via information redundancy across recording sites.

Methods

Participants

Eight healthy subjects (6 males, 2 females; ages 25–35) participated in Experiment 1, one in Experiment 2, and two in Experiment 3. All were BCI-naïve except Subject 6. Participants completed 1–2 practice sessions (~5 min each) before formal experiments. All provided informed consent (University of Glasgow Ethics Committee, no. 300230193).

EEG system and pattern stimulus

We used a three-electrode EEG device, with one active electrode placed at the Oz position to capture the SSVEP signal from the primary visual cortex, a reference electrode above the left ear (M1 position), and a ground electrode above the right ear (M2 position). Participants were seated ~50–70 cm from the computer screen. The system used MATLAB for real-time data

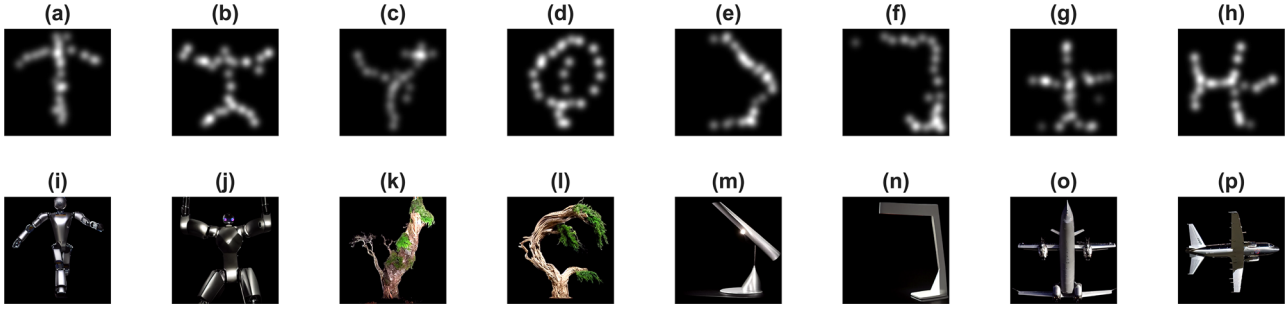

Fig. 7 | Examples of mind drawing with the Gabor policy approach that is enhanced with a final Stable Diffusion generative post-processing step. Panels **a**–**h** show examples of Gabor policy drawings. Panels **i**–**p** show the corresponding final images generated under the same SD text prompt via conditioned post-processing (these are grouped into 4 pairs, i.e. (**i**–**p**) have the same SD text prompts). Prompt and model details are provided in the SM.

acquisition and analysis, with visual stimuli presented using Psychtoolbox (PTB). EEG signals were sampled at 1 kHz via a microphone-based ADC sound card (Startech C-Media, 361 ICUSBAUDIO). The setup was implemented on a Windows 10 desktop equipped with a 480Hz 26.5-inch QHD (2560 × 1440) monitor (ROG Swift OLED PG27AQDP), working at 120 Hz or 240 Hz to ensure stable visual presentation. A square stimulus area of 1440 × 1440 pixels was centred on the screen.

At the beginning of the experiment, each subject sketched a simple imagined shape, such as a letter Y, on the canvas (details in SM) using a mouse. This image served as the target for the subsequent mental imagery task. After completion of the drawing and before the periodic flashing disc stimulus, a static image containing 10 discs was displayed on the screen shortly, and then starts flashing. The number of discs on the screen is determined by two factors: (1) the size of each disc (a 3 degree visual angle provides a robust SSVEP response) and their spacing (the centre-to-centre spacing is kept 2× the stimulus size i.e. the centre-to-centre spacing is 300 pixels). (2) the flicker frequency of each disc is chosen within the 10–19 Hz range and with a separation of 1 Hz.

After sketching the image using a drawing board with a mouse, subjects were informed to use mind drawing to draw the outline first and then go into detail, following the following strategy to enhance reconstruction accuracy and speed:

1. Focus on overlapping areas: Select the disc with the largest intersection or overlap with the imagined object. Concentrate on the overlapping region of the disc rather than its centre.
2. Prioritise outline information: Begin by marking points with larger spacing i.e. attempt to first reconstruct the large-scale features and basic structure of the object. For instance, place points on the key strokes of the object (e.g., for the letter "Z," mark points on the horizontal, diagonal, and vertical strokes).
3. Refine details: Gradually select points with smaller spacing to add finer details and progressively refine the shape of the object.

In the Gabor policy function mode, as shown in Fig. 1c, the stimulus consists of 10 discs, each flickering at a distinct frequency. These discs are placed at random locations on the screen, with each disc serving as a potential probe location for the subject's imagined image.

In the Data-driven policy function mode, as shown in Fig. 1d, the stimulus also comprises 10 discs, with some representing green features ('basis function' patterns from the MNIST dataset; details in SM) and others serving as standalone elements. The subject is instructed to select the disc whose presented pattern overlaps most closely with their imagined image. Additionally, a reconstructed image is displayed in the background throughout the experiment. Initially, this background corresponds to the dataset's average image, but it is iteratively updated based on the subject's selections and the system's ongoing reconstruction. This background serves as visual feedback and context, allowing the subject to monitor the progression of mental image reconstruction.

The CCA algorithm was applied to determine which probe the participant was focusing on. The participants focused on the disc that overlapped the most with their imagined image, eliciting an SSVEP response in the EEG signal. In the next iteration, the selected disc was weighted with its corresponding CCA value and displayed as part of the red background in the static image at the beginning of the subsequent iteration.

Gabor-inspired policy function

Gabor filters are specialised bandpass filters sensitive to textures and edges in specific directions and frequencies. A set of Gabor filters with varying orientations and scales acts as a detector array for "texture and structure." This is particularly relevant because studies have shown that the receptive fields in the primary visual cortex (V1) closely resemble Gabor functions. Thus, employing Gabor analysis to process subject-drawn images is chosen so as to emulate the early stages of human visual processing[43].

We apply a 2-step Gabor analysis (including feature kernel and probability kernel, detailed in SM) to the current estimated image, I_{pt}, which is composed of a sparse set of discrete CCA-weighted pixel points (centre of the disc), as shown in Eq. (2).

1. Extraction of Gabor features: Gaussian smoothing, $\sigma = 50$ pixels, is applied to I_{pt}, producing a blurred image I_{gs} highlighting low-frequency structure. Here, I_{gs} is equal to I_n in Eq. (1). The Gabor feature kernel, G_{feat}, is convolved with I_{gs} to extract initial structural features, I_{gg}, such as edges and stripes. We then subtract the mean of each filter channel, $\bar{I}_{gg}$, and the result is passed through a ReLU function (to retain only the positive components) followed by weighting of the original point-based estimate I_{pt}, yielding a weighted feature map I_{feat}.

2. Sampling weights: This feature map, I_{feat}, is subsequently convolved with the Gabor probability kernel, G_{prob}, producing the final probability map I_{prob} (also referred to as a sampling weight), which guides the spatial placement of probe disks in the next iteration. G_{feat}, G_{prob}, and one example of this 2-step flow is detailed in SM. Summarising the steps above in formulas, we have:

$$\begin{aligned} I_{\mathrm{gs}} &= I_{\mathrm{pt}} * \mathrm{Gaussian}, \\ I_{\mathrm{gg}} &= I_{\mathrm{gs}} * G_{\mathrm{feat}}, \\ I_{\mathrm{feat}} &= I_{\mathrm{pt}} \cdot \mathrm{ReLU}\left(I_{\mathrm{gg}} - \bar{I}_{\mathrm{gg}}\right), \\ I_{\mathrm{prob}} &= I_{\mathrm{feat}} * G_{\mathrm{prob}}. \end{aligned} \tag{2}$$

If the class of images is limited to a specific set, then as shown in the next section, a Data-driven Policy Function can be used to enhance the reconstruction process further. In this case, the Gabor feature analysis is replaced by a data-driven approach that leverages prior knowledge from a dataset, such as MNIST for handwritten digits. This allows the system to adaptively select probes based on the current estimate of the mental image.

Data-driven policy function

The data-driven policy function is designed to accelerate the reconstruction process by leveraging prior knowledge from a dataset, such as MNIST for handwritten digits.

In our experiments, we use the MNIST dataset, which contains 60,000 training and 10,000 test images of handwritten digits (each 28 × 28 pixels). Non-negative Matrix Factorisation (NNMF) decomposes these images into a set of basis patterns (atoms) and their corresponding weights. This allows images to be represented in a lower-dimensional, more efficient feature space for reconstruction. Here, we select 25 atoms from the NNMF decomposition to serve as the feature space for image representation and reconstruction, as detailed in the SM.

In our framework, atom/basis patterns are presented as probe options, similar to the previous approach of displaying 10 disc probes. As shown in Fig. 1d, some probes correspond to basis patterns (highlighted with a green feature background), while others correspond to discs themselves. The participant is asked to prioritize selection of an atom pattern. If no atom pattern overlaps with the subject's imagined image, the subject is instructed to select a normal probe (without green feature background). If there is still no overlap for any of the discs, the subject focuses on a neutral area.

The 'seed image' at the beginning of the experiment is taken simply as the average image of the whole dataset and is projected into the atom feature latent space as a vector, guiding the selection of the feature patterns.

To guide the selection of the next atom pattern, we first identify the 100 nearest MNIST images to current reconstructed image I_n in the latent space. The average of their latent vectors yields a (1×25) vector W_n, which is then used as the sampling weight for the next atom probe selection. Moreover, the prediction of the subject's mental image, obtained from W_n is displayed as a red background on the screen, visible to the subject.

Information rate calculation

For a single pixel, the mutual information can be expressed as

$$I(X;Y) = p\log_2(2p) + (1-p)\log_2\left[2(1-p)\right], \quad (3)$$

where p denotes the probability of a correct match rate between two images. If n is the number of image pixels, then the mutual information for the entire image can be estimated as MI = n $I(X;Y)$. In our experiment, p is estimated by comparing the whole image with the ground truth and calculating the correct rate for pixels.

We can also estimate the maximum theoretical information transfer rate for a BCI using the standard equation (the Wolpaw formula)[41],

$$\mathrm{ITR} = \frac{1}{T}\left(\log_2 N + P\log_2 P + (1-P)\log_2\frac{1-P}{N-1}\right) \quad (4)$$

where N is the number of targets, P is the probability of correct target identification (i.e. accuracy) and T is the average time required for a single selection. In our experiments, $N = 10$, $T = 4$ s (the time for a single iteration during which a single target is identified with probability $P = 1$), and the maximum theoretical ITR is:

$$\mathrm{ITR}_{\max} = \frac{1}{T}\log_2 N \approx 0.83\ bit/s \quad (5)$$

Data availability

Data relevant to this work is deposited at https://doi.org/10.5525/gla.researchdata.2167.

Received: 29 November 2025; Accepted: 13 April 2026;
Published online: 04 May 2026

References

1. Abdulkader, S. N., Atia, A. & Mostafa, M.-S. M. Brain computer interfacing: applications and challenges. *Egypt. Inf. J.* **16**, 213–230 (2015).
2. Bi, L., Fan, X.-A. & Liu, Y. EEG-based brain-controlled mobile robots: a survey. *IEEE Trans. Human Machine Syst.* **43**, 161–176 (2013).
3. Chen, X. et al. High-speed spelling with a noninvasive brain-computer interface. *Proc. Natl. Acad. Sci. USA* **112**, E6058-67 (2015).
4. Fouad, M.M., Amin, K.M., El-Bendary, N. & Hassanien, A.E. *Brain Computer Interface: A Review*, 3–30 (Springer International Publishing, 2015).
5. Lelievre, Y., Washizawa, Y. & Rutkowski, T. M. Single trial BCI classification accuracy improvement for the novel virtual sound movement-based spatial auditory paradigm. In *APSIPA 2013*,1–6 (IEEE, 2013).
6. Zander, T. O. & Kothe, C. Towards passive brain-computer interfaces: applying brain–computer interface technology to human–machine systems in general. *J. Neural Eng.* **8**, 025005 (2011).
7. Alimardani, M. & Hiraki, K. Passive brain-computer interfaces for enhanced human-robot interaction. *Front. Robot AI* **7**, 125 (2020).
8. Mladenovic, J. et al. Active inference as a unifying, generic and adaptive framework for a P300-based BCI. *J. Neural Eng.* **17**, 016054 (2020).
9. Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R. & Rossion, B. The steady-state visual evoked potential in vision research: a review. *J Vis.* **15**, 4 (2015).
10. Ma, Y., Liu, Y., Chen, L., Zhu, G., Chen, B. & Zheng, N. BrainCLIP: Brain Representation via CLIP for Generic Natural Visual Stimulus Decoding. *IEEE Trans Med Imaging* **44**, 3962–3972 (2025).
11. Wang, S., Liu, S., Tan, Z. & Wang, X. MindBridge: a cross-subject brain decoding framework. arXiv:2404.07850 (2024).
12. Zickler, C., Halder, S., Kleih, S. C., Herbert, C. & Kübler, A. Brain Painting. Usability testing according to the user-centered design in end users with severe motor paralysis. *Artif. Intell. Med.* **59**, 99–110 (2013).
13. Tang, Z. et al. A BCI painting system using a hybrid control approach based on SSVEP and P300. *Comput. Biol. Med.* **150**, 106118 (2022).
14. Müller, M. M. et al. Effects of spatial selective attention on the steady-state visual evoked potential in the 20-28 Hz range. *Brain Res. Cogn. Brain Res.* **6**, 249–261 (1998).
15. Nicolas-Alonso, L. F. & Gomez-Gil, J. Brain computer interfaces, a review. *Sensors* **12**, 1211–1279 (2012).
16. Wang, G. & Faccio, D. Computational ghost imaging with the human brain. *Intell. Comput.* **2**, 0014 (2023).
17. Wang, G. et al. Human-centred physical neuromorphics with visual brain-computer interfaces. *Nat. Commun.* **15**, 6393 (2024).
18. Miyawaki, Y. et al. Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. *Neuron* **60**, 915–929 (2008).
19. Du, B., Cheng, X., Duan, Y. & Ning, H. fMRI brain decoding and its applications in brain-computer interface: a survey. *Brain Sci.* **12**, 228 (2022).
20. Horikawa, T. & Kamitani, Y. Attentionally modulated subjective images reconstructed from brain activity, bioRxiv 2020.12.27.424510 (2020).
21. Le, L. et al. Brain2Pix: Fully convolutional naturalistic video frame reconstruction from brain activity. *Front Neurosci.* **16**, 940972 (2022).
22. Meng, L. & Yang, C. Dual-guided brain diffusion model: natural image reconstruction from human visual stimulus fMRI. *Bioengineering* **10**, 1117 (2023).
23. Seeliger, K., Güļü, U., Ambrogioni, L., Güļütürk, Y. & Van Gerven, M. Generative adversarial networks for reconstructing natural images from brain activity. *Neuroimage* **181**, 775–785 (2018).
24. Shen, G., Horikawa, T., Majima, K. & Kamitani, Y. Deep image reconstruction from human brain activity. *PLoS Comput. Biol.* **15**, e1006633 2019).
25. Adamic, M. et al. Progress Towards Decoding Visual Imagery via fNIRS, arXiv:2406.07662 (2024).
26. Benchetrit, Y., Banville, H. & King, J.-R. Brain decoding: toward real-time reconstruction of visual perception, arXiv:2310.19812 (2024).
27. Van De Nieuwenhuijzen, M. et al. MEG-based decoding of the spatiotemporal dynamics of visual category perception. *Neuroimage* **83**, 1063–1073 (2013).
28. Khaleghi, N. et al. Visual saliency and image reconstruction from EEG signals via an effective geometric deep network-based generative adversarial network. *Electronics* **11**, 3637 (2022).
29. Li, D. et al. Visual decoding and reconstruction via EEG embeddings with guided diffusion, NeurIPS poster 95138 (2024).
30. Pan, H., Li, Z., Fu, Y., Qin, X. & Hu, J. Reconstructing visual stimulus images from EEG signals based on deep visual representation model, arXiv:2403.06532 (2024).
31. Singh, P., Pandey, P., Miyapuram, K. & Raman, S. EEG2IMAGE: image reconstruction from EEG brain signals. In *ICASSP 2023*, 1–5 (IEEE, 2023).
32. Sokač, M., Mršić, L., Balković, M. & Brkljačić, M. Bridging Artificial Intelligence and Neurological Signals (BRAINS): A Novel Framework for Electroencephalogram-Based Image Generation 15, 405. https://www.mdpi.com/2078-2489/15/7/.
33. Wakita, S., Orima, T. & Motoyoshi, I. Photorealistic reconstruction of visual texture from EEG signals. *Front. Comput. Neurosci.* **15**, 754587 (2021).
34. Chen, Z., Qing, J., Xiang, T., Yue, W. L. & Zhou, J. H. CVPR, Seeing beyond the brain: conditional diffusion model with sparse masked modeling for vision decoding (2023).
35. Guenther, S., Kosmyna, N. & Maes, P. Image classification and reconstruction from low-density EEG. *Sci. Rep.* **14**, 16436 (2024).

36. Lee, S., Jang, S. & Jun, S. C. Exploring the ability to classify visual perception and visual imagery EEG data: toward an intuitive BCI system. *Electronics* **11**, 2706 (2022).
37. Shimizu, H. & Srinivasan, R. Improving classification and reconstruction of imagined images from EEG signals **17**, e0274847 (2022).
38. Fu, H., Wang, H., Chin, J. J. & Shen, Z. BrainVis: exploring the bridge between brain and visual signals via image reconstruction. In *ICASSP 2025*, 1–5 (IEEE, 2025).
39. Zheng, J. & Meister, M. The unbearable slowness of being:Why do we live at 10 bits/s? *Neuron* **113**, 192–204 (2025).
40. Suhendra, M. A. et al. Canonical correlation analysis and its extension for SSVEP-based BCI detection: a systematic review. *J. Penelit. Pendidik. IPA* **10**, 1027–1040 (2024).
41. Wolpaw, J., Ramoser, H., McFarland, D. & Pfurtscheller, G. EEG-based communication: improved accuracy by response verification. *IEEE Trans Rehabil Eng.* **6**, 326–333 (1998).
42. Sadeghi, S. & Maleki, A. Accurate estimation of information transfer rate based on symbol occurrence probability in brain-computer interfaces. *Biomed. Sign. Process. Control* **54**, 101607 (2019).
43. Huang, Y., Pollick, F., Liu, M. & Zhang, D. Gabor and non-Gabor neural representations are shared between visual perception and mental imagery. *J. Cogn. Neurosci.* **35**, 1045–1060 (2023).

Acknowledgements

G.W. and Y.H. acknowledge the support of the China Scholarship Council (CSC). D.F. acknowledges support from the Royal Academy of Engineering Chair in Emerging Technologies programme and the UK Engineering and Physical Sciences Research Council (grants EP/T00097X/1, EP/Y029097/1, EP/ Z533166/1).

Author contributions

G.W.: Conceptualisation, methodology and system design, software development and implementation, experimental design, data acquisition and analysis, and manuscript writing and revision; Y.H.: Experimental design, data acquisition, and manuscript editing and revision; L.M.: Manuscript revision and supervision; D.F.: Conceptualisation, manuscript revision, project administration, and funding acquisition. All authors have read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Daniele Faccio.