

Symbiotic Brain-Machine Drawing via Visual Brain-Computer Interfaces

Gao Wang¹, Yingying Huang^{1,2}, Lars Muckli², Daniele Faccio^{1*}

¹*School of Physics & Astronomy, University of Glasgow, Glasgow, UK, G12 8QQ.*

²*School of Psychology and Neuroscience, University of Glasgow, Glasgow G12 8QB, UK*

(Dated: November 27, 2025)

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 mind-drawing that iteratively infers a subject's internal visual intent by adaptively presenting visual stimuli (probes) on a screen encoded at different flicker-frequencies and analyses the steady-state visual evoked potentials (SSVEPs). A 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 significantly increase BCI bit-rates by more than a factor 5x, providing a platform for future development of AI-augmented BCI.

Keywords: SSVEP, Human-machine collaboration, Brain-reading, Gabor filter, Brain-computer interface

I. INTRODUCTION

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.

Following the development of a "P300 speller"-based painting system [14], a BCI painting system employing a hybrid SSVEP/P300 control approach was later proposed [15]. 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 [16].

Noninvasive 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, 17–19] 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 [20–26].

Recent work has also highlighted the possibility to decode visual imagery via fNIRS [27]. MEG was proposed for real-time reconstruction of visual perception [28] and classification [29]. 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 [28, 30–39]. More recently, the 'BrainVis' approach [40] demonstrated state-of-the-art semantic fidelity reconstructions and generation quality. These mind-drawing or image-reconstruction approaches rely on high-dimensional subject data 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 [41].

* Correspondence email address: daniele.faccio@glasgow.ac.uk

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.

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.

II. RESULTS

The experimental setup is illustrated in Fig. 1(a). 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 1(b) 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. At the beginning of each iteration, 10 white visual stimulus probes (flickering discs) were displayed on a black-background screen (as shown in Fig. 1(c)). 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 seconds of flashing stimulation, the SSVEP response was recorded and the Canonical Correlation Analysis (CCA) value (a correlation coefficient ranging from 0 to 1 [42]) 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 1(c) shows an example stim-

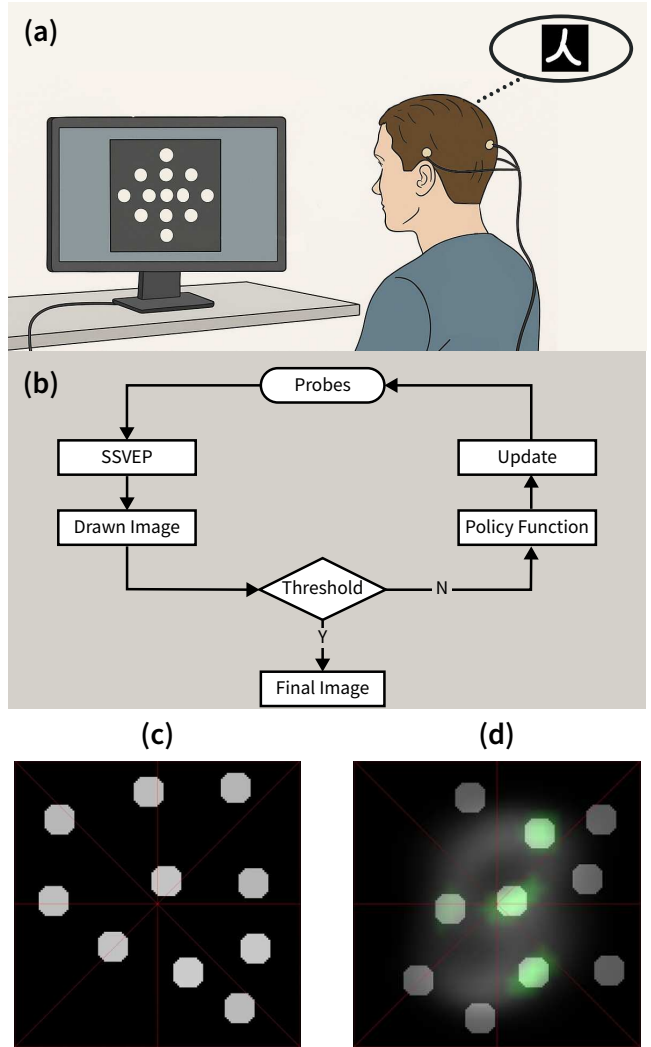


Figure 1. Experimental setup and workflow. (a) Setup of the EEG-based mind-drawing system. The EEG device is custom-built and consists of a headband housing three wet electrodes (using saline solution to improve contact): two placed on the temples (ground and reference) and one at the occipital Oz position. 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.

ulus under the Gabor policy function, in which discs are randomly arranged and each disc flickers at a unique frequency. Figure 1(d) 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 minutes, also including 2-second 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.

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.

A. 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. We evaluated the performance across multiple subjects to assess the generalizability and 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-colored images represent the handwritten target after a shape-preserving transformation (i.e. rescaling, rotation, translation) so that it has same size as

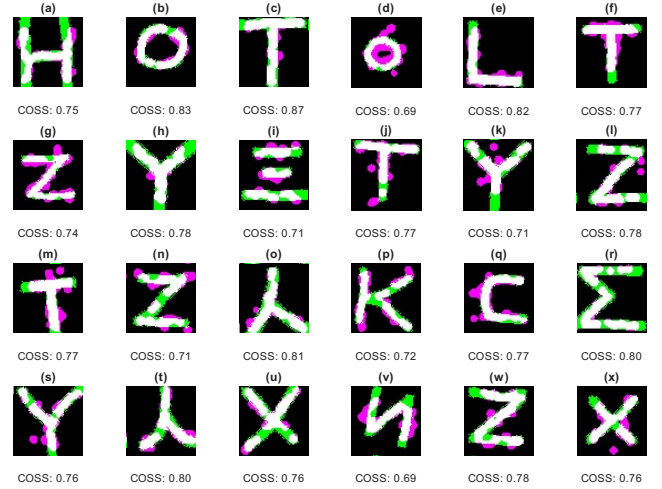


Figure 2. Imaging results from eight subjects, each targeting three imagined shapes. Panels are grouped three at a time, one group for each subject, i.e. panels (a–c) correspond to the first subject, (d–f) to the second subject, and so forth. Green-colored images represent the handwritten target after image resizing, while pink-colored images indicate the reconstructed images generated by our BCI system. Overlapped regions highlight the common areas between the two images, indicating reconstruction accuracy. We calculated the cosine similarity (COS) between each ground-truth and reconstructed image pair to quantify accuracy. The overall average COS across all subjects was $0.76 (\pm 0.04)$, indicating a good level of reconstruction fidelity.

the mind-drawn image, shown as a pink-colored image. The overlap regions between the ground-truth and the reconstructed images appear in white. We calculated the cosine similarity (COS) between each ground-truth and reconstructed image pair to quantify accuracy. The overall average COS across all subjects was $0.76 (\pm 0.04)$, indicating a good level of reconstruction fidelity.

Figure 3 shows a COS box plot and histogram distributions that illustrates how the COS 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 COS 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

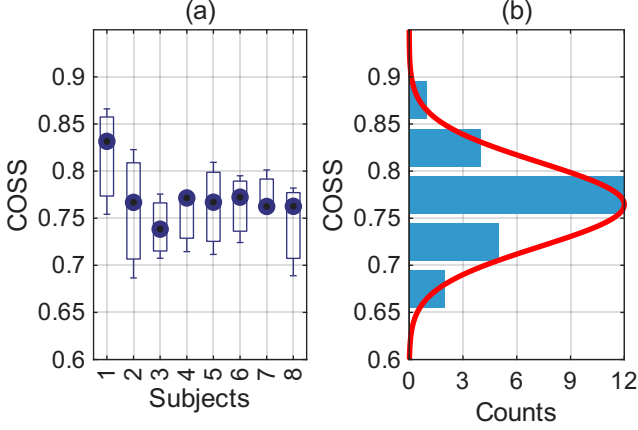


Figure 3. Distribution of image cosine similarity (COSS) values between target and reconstructed image pairs. (a) Box plot of COSS values for each subject. (b) Histogram of COSS values across all subjects.

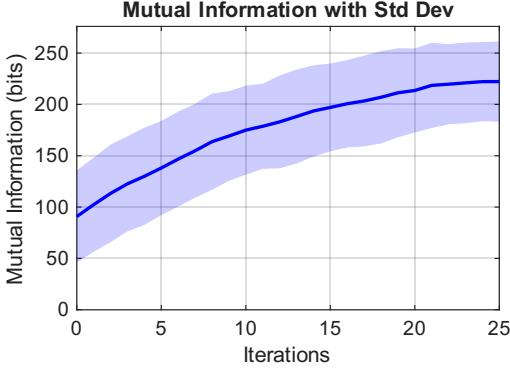


Figure 4. Mutual information (with standard deviation across iterations and subjects) increases with the number of iterations, although the rate of increase slows down as more iterations are added. The final MI reaches 222 bits starting from 91 bits (MI from an image with all black pixels), corresponding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

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 $\sim 1.6\times$ larger than the maximum ITR predicted for a standard BCI (0.83 bits/second) [43, 44]. 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 in-

formation 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.

B. 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). 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.

As shown in Fig. 1(d), 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 (h) and (p), respectively, the full data-driven pipeline provides images that are better representations of the intended image (i.e. the panels show larger 'white shaded overlap regions' in (p) compared to (h)) As more 'basis functions' are accumulated, 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 experi-

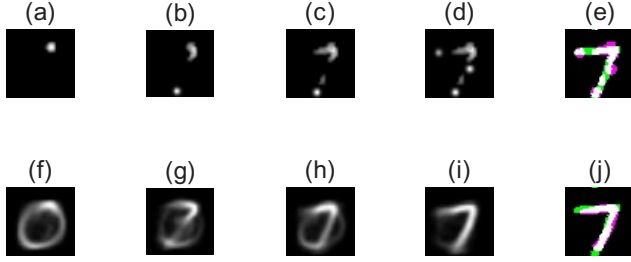


Figure 5. Reconstruction results for the MNIST digit “7.” The first row (a–d) shows the step-by-step reconstruction of the digit from SSVEP signals by placing a Gaussian disc at the location of each disc selected by the SSVEP signal. The second row (f–i) presents the corresponding images generated by the system at each step. For simplicity, we only show the images every 2 iteration steps. Panels (e) and (j) overlay the binarised reconstructed and predicted images with the aligned handwritten target in different colour channels: green indicates the aligned handwritten target, pink represents the reconstructed mental image, and overlapped regions highlight areas of agreement, demonstrating reconstruction accuracy. Grey regions indicate agreement, while colored regions highlight intensity differences. 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.

ment, 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. We need however, to recall that this increase of bit rate from 0.83 to 4.21 bit/s (an increase of more than a factor 5x) comes at the expense of generalisability.

C. Stable Diffusion enhanced mind drawing

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

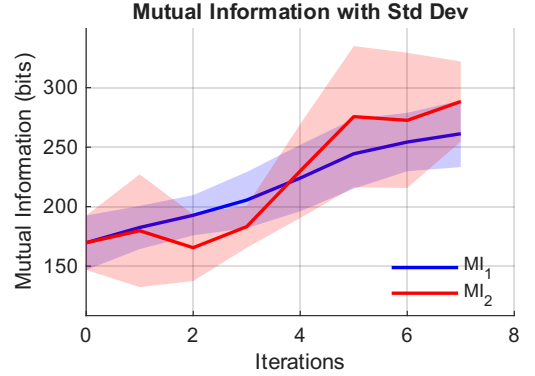


Figure 6. Mutual information during the reconstruction of MNIST digits, averaged over four different digit reconstructions (“7”, “2”, “4”, and “8”). MI₁ and MI₂ denote the mutual information of the SSVEP reconstruction alone (i.e. same approach as top row in Fig. 5) and the data driven policy images (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.

a generative enhancement model. 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 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.

III. DISCUSSION AND CONCLUSIONS

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 and generalizability, 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.

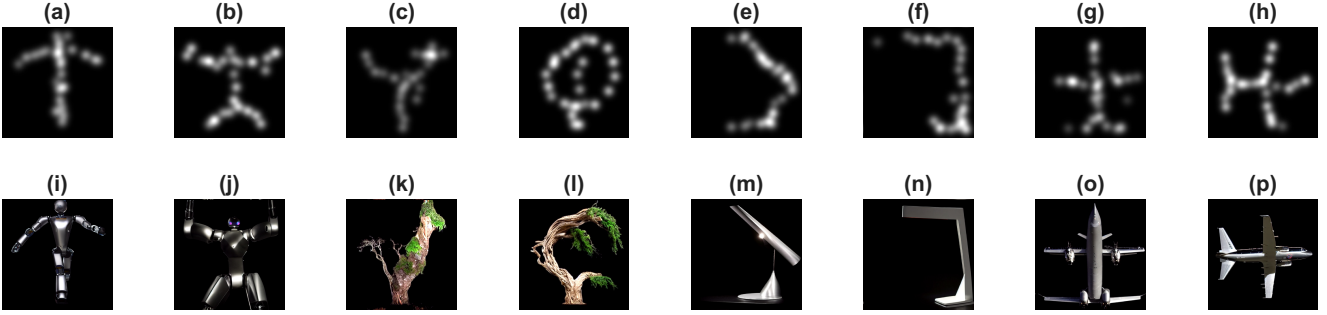


Figure 7. Examples of mind drawing with the gabor policy approach that is enhanced with a final Stable Diffusion step. Panels (a–h) show examples of Gabor policy drawings. Panels (i–p) show the corresponding final images generated under the same SD text prompt (these are grouped into 4 pairs, i.e. (i)–(j), (k)–(l), (m)–(n), (o)–(p) have the same SD text prompts). Prompt and model details are provided in the SM.

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.

The integration of Stable Diffusion further enhances the system’s capabilities, transforming abstract neural reconstructions into realistic and detailed visual representations. This bridges the gap between mental imagery and photorealistic imagery, enabling applications such as creative co-design and personalised content generation.

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.

IV. METHODS

A. 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 approximately 50–70 cm from the computer screen. The system used MATLAB for real-time data 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 x 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 start 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 2x 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, fol-

lowing 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 shown Fig 1(c), 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. 1(d), 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.

B. 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 [45].

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:

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

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.

C. 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 hand-

written 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. 1(d), 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.

D. Information Rate Calculation

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

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

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) [43],

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

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, $N = 10$, $T = 4$ seconds (the time for a single iteration during which a single target is identified with probability $P = 1$), and the maximum theoretical ITR is:

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

ACKNOWLEDGMENTS

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 (grant EP/T00097X/).

CONFLICTS OF INTEREST

All authors declare that there is no conflict of interest.

-
- [1] Abdulkader, S. N., Atia, A. & Mostafa, M.-S. M. Brain computer interfacing: Applications and challenges **16**, 213–230. URL <https://linkinghub.elsevier.com/retrieve/pii/S1110866515000237>.
 - [2] Bi, L., Fan, X.-A. & Liu, Y. EEG-Based Brain-Controlled Mobile Robots: A Survey **43**, 161–176. URL <http://ieeexplore.ieee.org/document/6461528/>.
 - [3] Chen, X. et al. High-speed spelling with a noninvasive brain-computer interface **112**. URL <https://pnas.org/doi/full/10.1073/pnas.1508080112>.
 - [4] Brain Computer Interface: A Review, 3–30 (Springer International Publishing). URL https://link.springer.com/10.1007/978-3-319-10978-7_1.
 - [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 2013 Asia-Pacific Signal and Information Processing Association Annual 1–6 (IEEE). URL <http://ieeexplore.ieee.org/document/6694317/>.
 - [6] Zander, T. O. & Kothe, C. Towards passive brain-computer interfaces: Applying brain-computer interface technology to human-machine systems in general **8**, 025005. URL <http://dx.doi.org/10.1088/1741-2560/8/2/025005>.

- [7] Alimardani, M. & Hiraki, K. Passive Brain-Computer Interfaces for Enhanced Human-Robot Interaction **7**. URL <http://dx.doi.org/10.3389/frobt.2020.00125>.
- [8] Mladenovic, J. et al. Active inference as a unifying, generic and adaptive framework for a P300-based BCI **17**, 016054. URL <http://dx.doi.org/10.1088/1741-2552/ab5d5c>.
- [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 **15**, 4. URL <http://jov.arvojournals.org/article.aspx?doi=10.1167/15.6.4>.
- [10] Benchetrit, Y., Banville, H. & King, J.-R. Brain decoding: Toward real-time reconstruction of visual perception. URL <http://arxiv.org/abs/2310.19812>.
- [11] Du, B., Cheng, X., Duan, Y. & Ning, H. fMRI Brain Decoding and Its Applications in Brain-Computer Interface: A Survey **12**, 228. URL <https://www.mdpi.com/2076-3425/12/2/228>.
- [12] Liu, Y., Ma, Y., Zhou, W., Zhu, G. & Zheng, N. Brain-CLIP: Bridging Brain and Visual-Linguistic Representation Via CLIP for Generic Natural Visual Stimulus Decoding. URL <http://arxiv.org/abs/2302.12971>.
- [13] Wang, S., Liu, S., Tan, Z. & Wang, X. MindBridge: A Cross-Subject Brain Decoding Framework. URL <http://arxiv.org/abs/2404.07850>.
- [14] 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 **59**, 99–110. URL <https://linkinghub.elsevier.com/retrieve/pii/S0933365713001206>.
- [15] Tang, Z. et al. A BCI painting system using a hybrid control approach based on SSVEP and P300 **150**, 106118. URL <https://linkinghub.elsevier.com/retrieve/pii/S0010482522008265>.
- [16] Müller, M. M. et al. Effects of spatial selective attention on the steady-state visual evoked potential in the 20–28 Hz range **6**, 249–261. URL <https://linkinghub.elsevier.com/retrieve/pii/S0926641097000360>.
- [17] Nicolas-Alonso, L. F. & Gomez-Gil, J. Brain Computer Interfaces, a Review **12**, 1211–1279. URL <https://www.mdpi.com/1424-8220/12/2/1211>.
- [18] Wang, G. & Faccio, D. Computational Ghost Imaging with the Human Brain **2**, 0014. URL <https://spj.science.org/doi/10.34133/icomputing.0014>.
- [19] Wang, G. et al. Human-centred physical neuromorphics with visual brain-computer interfaces **15**, 6393. URL <https://www.nature.com/articles/s41467-024-50775-2>.
- [20] Miyawaki, Y. et al. Visual Image Reconstruction from Human Brain Activity using a Combination of Multiscale Local Image Decoders **60**, 915–929. URL <https://linkinghub.elsevier.com/retrieve/pii/S0896627308009586>.
- [21] Du, B., Cheng, X., Duan, Y. & Ning, H. fMRI Brain Decoding and Its Applications in Brain-Computer Interface: A Survey **12**, 228. URL <https://www.mdpi.com/2076-3425/12/2/228>.
- [22] Horikawa, T. & Kamitani, Y. Attentionally modulated subjective images reconstructed from brain activity. URL <http://biorxiv.org/lookup/doi/10.1101/2020.12.27.424510>.
- [23] Le, L. et al. Brain2Pix: Fully convolutional naturalistic video reconstruction from brain activity. URL <http://biorxiv.org/lookup/doi/10.1101/2021.02.02.429430>.
- [24] Meng, L. & Yang, C. Dual-Guided Brain Diffusion Model: Natural Image Reconstruction from Human Visual Stimulus fMRI **10**, 1117. URL <http://dx.doi.org/10.3390/bioengineering10101117>.
- [25] Seeliger, K., Güçlü, U., Ambrogioni, L., Güçlütürk, Y. & Van Gerven, M. Generative adversarial networks for reconstructing natural images from brain activity **181**, 775–785. URL <https://linkinghub.elsevier.com/retrieve/pii/S105381191830658X>.
- [26] Shen, G., Horikawa, T., Majima, K. & Kamitani, Y. Deep image reconstruction from human brain activity **15**, e1006633. URL <https://dx.plos.org/10.1371/journal.pcbi.1006633>.
- [27] Adamic, M. et al. Progress Towards Decoding Visual Imagery via fNIRS. URL <https://arxiv.org/abs/2406.07662>.
- [28] Benchetrit, Y., Banville, H. & King, J.-R. Brain decoding: Toward real-time reconstruction of visual perception. URL <http://arxiv.org/abs/2310.19812>.
- [29] Van De Nieuwenhuijzen, M. et al. MEG-based decoding of the spatiotemporal dynamics of visual category perception **83**, 1063–1073. URL <https://linkinghub.elsevier.com/retrieve/pii/S1053811913008446>.
- [30] Khaleghi, N. et al. Visual Saliency and Image Reconstruction from EEG Signals via an Effective Geometric Deep Network-Based Generative Adversarial Network **11**, 3637. URL <https://www.mdpi.com/2079-9292/11/21/3637>.
- [31] Li, D. et al. Visual Decoding and Reconstruction via EEG Embeddings with Guided Diffusion. URL <https://arxiv.org/abs/2403.07721>.
- [32] Pan, H., Li, Z., Fu, Y., Qin, X. & Hu, J. Reconstructing Visual Stimulus Images from EEG Signals Based on Deep Visual Representation Model. URL <http://arxiv.org/abs/2403.06532>.
- [33] Singh, P., Pandey, P., Miyapuram, K. & Raman, S. EEG2IMAGE: Image Reconstruction from EEG Brain Signals. In ICASSP 2023 - 2023 IEEE International Conference on Acoustics and Speech Processing (ICASSP 2023). URL <https://ieeexplore.ieee.org/document/10096587/>.
- [34] 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. URL <https://www.mdpi.com/2078-2489/15/7/405>.
- [35] Wakita, S., Orima, T. & Motoyoshi, I. Photorealistic Reconstruction of Visual Texture From EEG Signals **15**. URL <https://www.frontiersin.org/articles/10.3389/fncom.2021.754587/full>.

- [36] Chen, Z., Qing, J., Xiang, T., Yue, W. L. & Zhou, J. H. Seeing Beyond the Brain: Conditional Diffusion Model with Sparse Masked Modeling for Vision Decoding. URL <https://arxiv.org/abs/2211.06956>.
- [37] Guenther, S., Kosmyna, N. & Maes, P. Image classification and reconstruction from low-density EEG **14**, 16436. URL <https://www.nature.com/articles/s41598-024-66228-1>.
- [38] Lee, S., Jang, S. & Jun, S. C. Exploring the Ability to Classify Visual Perception and Visual Imagery EEG Data: Toward an Intuitive BCI System **11**, 2706. URL <https://www.mdpi.com/2079-9292/11/17/2706>.
- [39] Shimizu, H. & Srinivasan, R. Improving classification and reconstruction of imagined images from EEG signals **17**, e0274847. URL <https://dx.plos.org/10.1371/journal.pone.0274847>.
- [40] Fu, H., Wang, H., Chin, J. J. & Shen, Z. Brain-Vis: Exploring the Bridge between Brain and Visual Signals via Image Reconstruction. In ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 1–5 (IEEE). URL <https://ieeexplore.ieee.org/document/10889805/>.
- [41] Zheng, J. & Meister, M. The unbearable slowness of being: Why do we live at 10 bits/s? **113**, 192–204. URL <https://linkinghub.elsevier.com/retrieve/pii/S0896627324008080>.
- [42] Suhendra, M. A. et al. Canonical Correlation Analysis and Its Extension for SSVEP-based BCI Detection: A Systematic Review **10**, 1027–1040. URL <https://jppipa.unram.ac.id/index.php/jppipa/article/view/9844>.
- [43] Wolpaw, J., Ramoser, H., McFarland, D. & Pfurtscheller, G. EEG-based communication: Improved accuracy by response verification **6**, 326–333. URL <http://ieeexplore.ieee.org/document/712231/>.
- [44] Sadeghi, S. & Maleki, A. Accurate estimation of information transfer rate based on symbol occurrence probability in brain-computer interfaces **54**, 101607. URL <https://linkinghub.elsevier.com/retrieve/pii/S1746809419301880>.
- [45] Huang, Y., Pollick, F., Liu, M. & Zhang, D. Gabor and Non-Gabor Neural Representations Are Shared between Visual Perception and Mental Imagery **35**, 1045–1060. URL <https://direct.mit.edu/jocn/article/35/6/1045/115627/Gabor-and-Non-Gabor-Neural-Representations-Are>.

Symbiotic Brain-Machine Drawing via Visual Brain-Computer Interfaces: Supplementary Material

Gao Wang¹, Yingying Huang^{1,2}, Lars Muckli², Daniele Faccio^{1*}

¹*School of Physics & Astronomy, University of Glasgow, Glasgow, UK, G12 8QQ.*

²*School of Psychology and Neuroscience, University of Glasgow, Glasgow G12 8QB, UK*

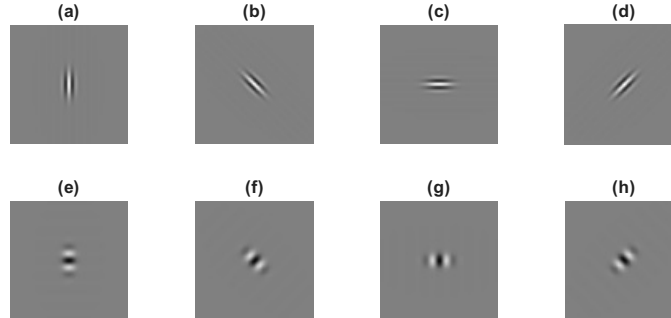
(Dated: November 27, 2025)

Supplementary Material with additional theoretical descriptions and experimental results.

Gabor-inspired Policy

We apply a set of multi-directional Gabor filters to the user’s drawing trajectory images. Given the strong resemblance between Gabor functions and the receptive fields of neurons in the primary visual cortex (V1), this analysis quantitatively captures the visual structure of the user’s creation. By identifying the Gabor filter orientation with the strongest response, the engine can dynamically infer the dominant geometric trend of the current drawing. For example, a pronounced response in the vertical direct filter indicates a strong vertical trend in the user’s creation. The inferred probability distribution is then used to adaptively position candidate targets (flashing disks) along these high-probability paths, providing real-time assistance aligned with the user’s spatial intention. Specifically, many objects, such as letters (“F”) or symbols (a cross), are composed of lines or curves.

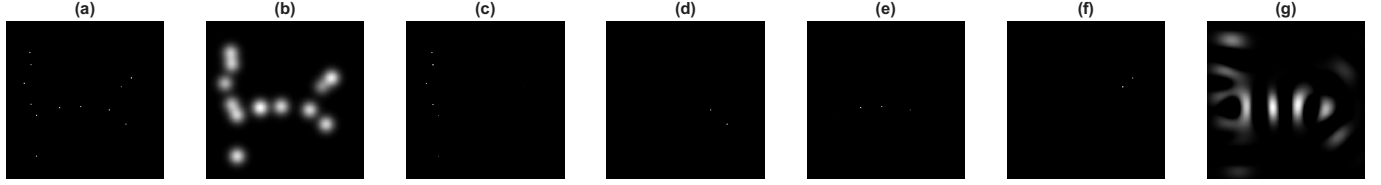
To implement this, we utilise the Gabor feature kernel (Fig. 1(a-d)) and the Gabor probability kernel (Fig. 1(e-h)). The feature kernel is effective for detecting orientation and spatial frequency, while the probability kernel is used to construct the probability map for generating subsequent probes. For example, if the vertical Gabor feature kernel (Fig. 1(a)) yields a strong response at certain locations, it suggests the presence of structural elements aligned in that direction. Consequently, the corresponding vertical Gabor probability kernel (Fig. 1(e)) assigns higher probabilities along the same direction in those regions, indicating a greater likelihood of informative content.



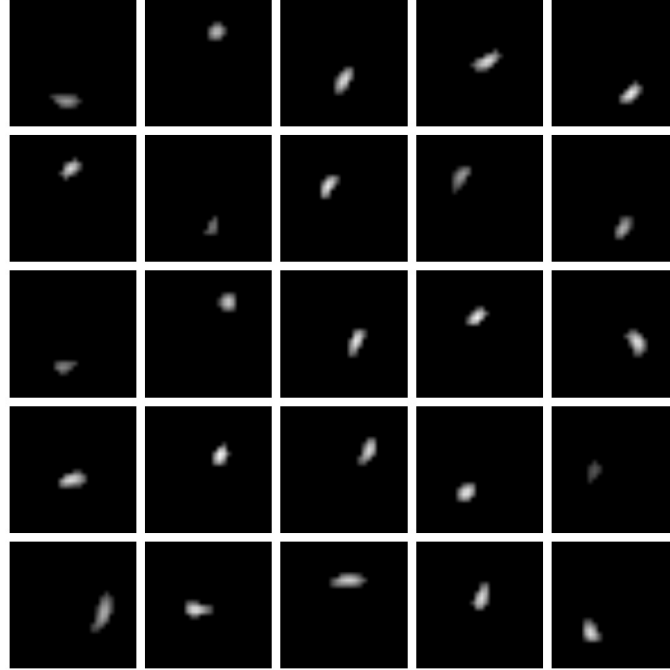
Supplementary figure 1. Gabor feature kernels with wavelength $\lambda = 2 \times$ probe size, spatial aspect ratio $= 1/2$, and orientations of 0° , 45° , 90° , and 135° . Subfigures (a), (b), (c), and (d) correspond to each orientation, respectively. Gabor probability kernels with wavelength $\lambda = 2 \times$ probe size, spatial aspect ratio $= 2$, and orientations of 90° , 135° , 180° , and 225° . Subfigures (e), (f), (g), and (h) correspond to each orientation, respectively.

Data-driven Policy

Initially, the atom/feature patterns are derived from the MNIST dataset using Non-negative Matrix Factorisation (NNMF). As shown in Fig 3, we select 25 atoms from the NNMF decomposition to serve as the feature space for image representation and reconstruction.



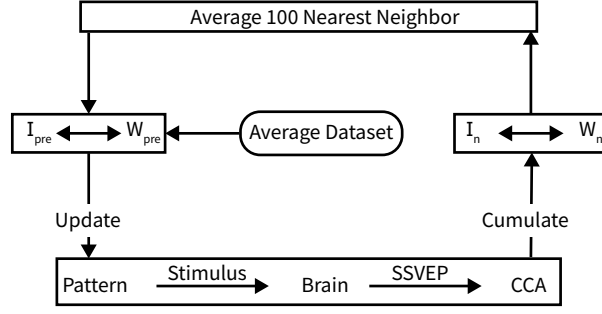
Supplementary figure 2. This is the middle of the twelfth trial in the first subject's drawing "H". (a) is discrete CCA-weighted pixel points (centre of the disc), I_{pt} , we can find 12 points on the image; (b) is Gaussian smoothing of I_{pt} , I_{gs} , to highlight low-frequency structural information; (c-f) is a weighted feature map, I_{feat} , for 4 different directions, 0, 45, 90, and 135 degrees. Different points are weighted/highlighted in various directions. g) is the probability map, I_{prob} .



Supplementary figure 3. 25 atom patterns derived from the MNIST dataset using Non-negative Matrix Factorisation (NMF). Each atom represents a unique feature extracted from the dataset.

It is important to note that the original atom patterns derived from NMF contained 40 feature patterns, and we selected 20 out of 40 feature patterns. Some of which exhibited multiple clustered features. To ensure suitability for our task, we separated these patterns, ensuring that each atom represents a single distinct feature. This preprocessing step is crucial for the subsequent reconstruction task, as it guarantees that each atom effectively captures a unique aspect of the imagined image.

The workflow for data-driven policy as shown in Fig. 4.



Supplementary figure 4. The workflow for Data-driven Policy.

Starting from the average dataset image as the I_{pre} , its latent space vector W_{pre} is used to update the pattern for stimulating the brain. The triggered SSVEP was collected and cumulated with the corresponding pattern as the reconstructed image, I_n . In the latent space, the 100 nearest neighbour image was averaged as the new I_{pre} for the next iteration.

IMAGE PROCESSING AND RESULTS

Binarisation and Registration

The reconstructed image is binarised to obtain I_{bin} by thresholding the smoothed reconstruction I_{gs} at its mean value:

$$I_{bin}(x, y) = \begin{cases} 1, & I_{gs}(x, y) > \overline{I_{gs}}, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Here, I_{gs} denotes the Gaussian-smoothed reconstruction, and $\overline{I_{gs}}$ is its mean intensity. Pixels with values above the mean are set to 1 (white), and those below are set to 0 (black), yielding a binary image for subsequent quantitative evaluation.

The subject's hand-drawn image, serving as ground truth, is spatially aligned with I_{bin} and denoted as I_{ref} . Specifically, image registration is performed using an affine transformation. The reference image is first resized to match the reconstructed image. Then, an affine (similarity) transformation is estimated to align the reference image to the binarised reconstructed image. The registered reference image is then used for quantitative comparison.

To quantitatively assess reconstruction accuracy, we employ the cosine similarity (COSS), which evaluates the degree of overlap between the binarised reconstructed image and the reference image:

$$\text{COSS}(I_{bin}, I_{ref}) = \frac{\sum_{i=1}^N I_{bin}(i) \cdot I_{ref}(i)}{\sqrt{\sum_{i=1}^N I_{bin}(i)^2} \cdot \sqrt{\sum_{i=1}^N I_{ref}(i)^2}} \quad (2)$$

where I_{bin} and I_{ref} are the binarised reconstructed and ground truth images, respectively, and N is the total number of pixels.

The binarisation of the predicted image in the Data Driven method, I_{pre} , is defined as follows. First, the threshold T is computed from the predicted image I_{pre} as

$$T = \text{mean}(I_{pre}) + \text{std}(I_{pre} \mid I_{pre} > 0), \quad (3)$$

where $\text{mean}(\cdot)$ denotes the average pixel intensity, and $\text{std}(\cdot)$ denotes the standard deviation. Specifically, $\text{std}(I_{\text{pre}} | I_{\text{pre}} > 0)$ indicates that the standard deviation is calculated only over the set of non-zero pixel values.

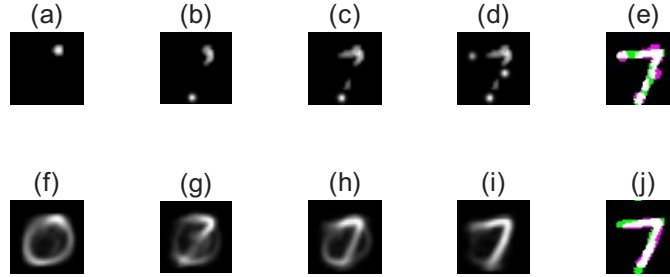
Based on this threshold, the binary image $I_{\text{pre}}^{\text{binary}}$ is obtained as

$$I_{\text{pre}}^{\text{binary}}(x, y) = \begin{cases} 1, & I_{\text{pre}}(x, y) > T, \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

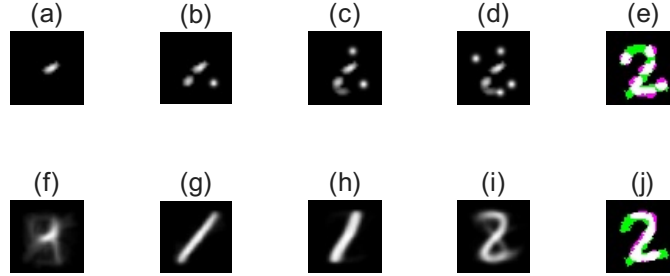
This procedure highlights salient regions in the predicted image while suppressing background noise.

MNIST Image Results

Here we show additional MNIST image results i.e. for the MNIST digits "7", "2", "4", and "8".



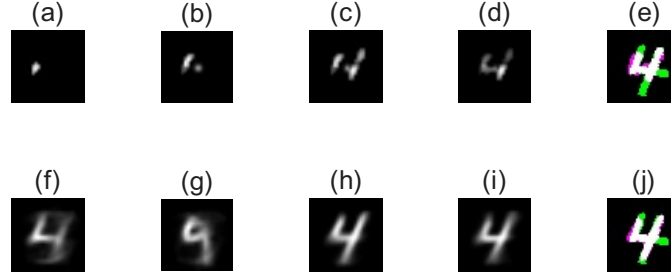
Supplementary figure 5. Reconstruction results for the MNIST digit “7.” The first row (a–g) shows the step-by-step reconstruction of the digit from SSVEP signals. The second row (i–o) presents the corresponding predicted images generated by the system at each step. Panels (h) and (p) overlay the binarised reconstructed and binarised predicted images.



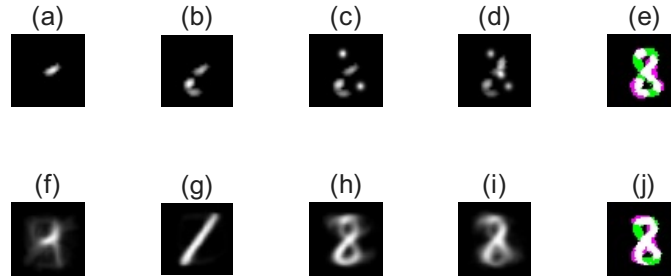
Supplementary figure 6. Reconstruction results for the MNIST digit “2.” The first row (a–g) shows the step-by-step reconstruction of the digit from SSVEP signals. The second row (i–o) presents the corresponding predicted images generated by the system at each step. Panels (h) and (p) overlay the binarised reconstructed and binarised predicted images.

Table I. COSS values for binarised reconstructed images (BRI), binarised predicted images (BPI), and predicted images (PI) across different MNIST digits.

MNIST digit	7	2	4	8
BRI	0.75	0.67	0.70	0.69
BPI	0.78	0.70	0.79	0.75
PI	0.82	0.78	0.82	0.82



Supplementary figure 7. Reconstruction results for the MNIST digit “4.” The first row (a–g) shows the step-by-step reconstruction of the digit from SSVEP signals. The second row (i–o) presents the corresponding predicted images generated by the system at each step. Panels (h) and (p) overlay the binarised reconstructed and binarised predicted images.



Supplementary figure 8. Reconstruction results for the MNIST digit “8.” The first row (a–g) shows the step-by-step reconstruction of the digit from SSVEP signals. The second row (i–o) presents the corresponding predicted images generated by the system at each step. Panels (h) and (p) overlay the binarised reconstructed and binarised predicted images.

Stable Diffusion Parameters and Results

Here we summarise the main parameters and prompts used for the stable diffusion image reconstructions. We also show additional image reconstructions and evolution of the reconstructed images during the iterative mind-drawing procedure.

Stable Diffusion checkpoint

v1-5-pruned-emaonly.safetensors [6ce0161689]

Prompt for the Robot

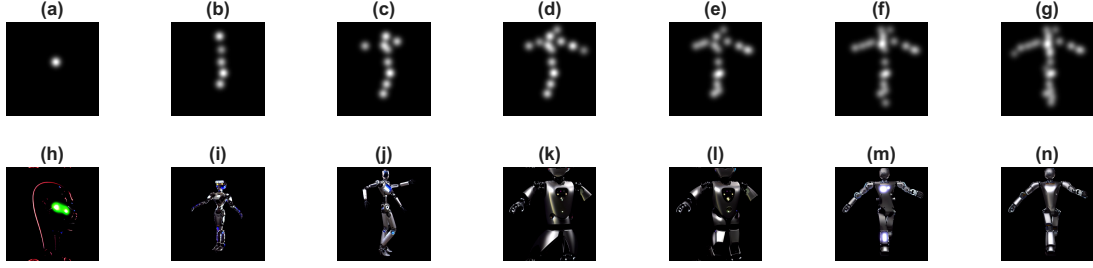
Positive prompt: A highly detailed humanoid robot, full body with head, arms, and legs, in a dynamic pose matching the input sketch. Futuristic cybernetic design with realistic metallic joints and intricate machinery. Cinematic lighting, ultra-sharp focus, 8K photorealism, anatomically correct human-like proportions, symmetrical body. Extremely detailed mechanical textures, polished metal surfaces, realistic reflections, and complex machinery.

Negative prompt: blurry, low quality, cropped, out of frame, missing limbs, extra limbs, distorted or deformed body parts, deformed hands, unnatural proportions, asymmetrical body, cartoonish, unrealistic, low resolution, overexposed, noisy background

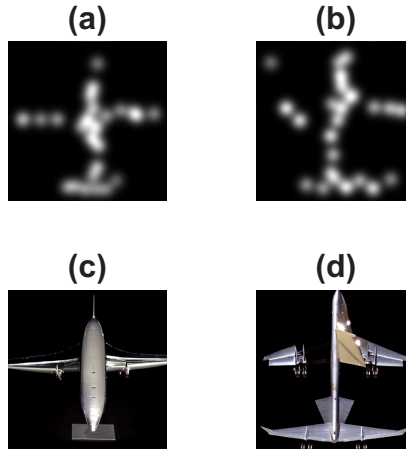
Prompt for the Tree

Positive prompt: A masterpiece, ultra-detailed, photorealistic, with a full tree growing from the ground, including visible roots, a branch trunk, and the whole tree is included, visible exposed root system anchored in natural soil with small stones and grass, thick textured trunk with deep bark cracks, knots and patches of moss

Negative prompt: cartoon, illustration, low-res, blurry, watermark, signature, text, over-saturated, pastel, painterly brush strokes, surreal elements, floating or disconnected roots, duplicated or mirrored branches, mutated/deformed shapes, extra limbs, unrealistic glossy plastic look, noisy artifacts, low detail, 3D render/artificial shading



Supplementary figure 9. Illustration of the image reconstruction process using Stable Diffusion. Panels (a–g) present the progressive refinement of mind-drawn sketches every 4 iteration steps i.e., iteration = 1,5,9..., while panels (h–n) show the corresponding photorealistic humanoid robot renderings generated by Stable Diffusion, conditioned on the sketch and a robot-related textual prompt.



Supplementary figure 10. Examples of the reconstruction process with Stable Diffusion. Panels (a–b) show the same object conceptual sketches from two subjects, whereas panels (c–d) present the corresponding high-fidelity images generated under the same prompt conditions.

Prompt for the desk lamp

Positive prompt: A full, photorealistic metal desk lamp, with the lamp at the top, a solid base at the bottom, connected by a sleek support arm. The lamp is fully visible from top to bottom, showing detailed textures of metal, reflections, and shadows. High-resolution, ultra-detailed, realistic lighting, studio-quality rendering, showing the complete structure of the lamp, including the base, support, and lampshade

Negative prompt: cartoon, low-res, blurry, text, watermark, cropped, multiple lamps, missing parts, non-metal materials, extra objects

Prompt for the aircraft

Positive prompt: A full, photorealistic metal airplane, with complete wings and tail, wings attached at the midsection of the fuselage. The entire airplane is fully visible in the frame, showing detailed textures of metal, reflections, rivets, and realistic lighting. High-resolution, ultra-detailed, realistic shadows and highlights, studio-quality rendering, complete airplane including fuselage, wings, and tail, viewed from a clear perspective. Negative prompt: cartoon, illustration, sketch, low-resolution, blurry, cut-off, incomplete, stylized, abstract, unrealistic proportions

Other Parameters

Steps: 20, Sampler: DPM++ 2M, Schedule type: Karras, CFG scale: 7, Seed: 1, Size: 512x512, Model hash: 6ce0161689, Model: v1-5-pruned-emaonly, Denoising strength: 0.75, Version: v1.10.1