Revealing neurocognitive and behavioral patterns through unsupervised manifold learning of dynamic brain data

Zhou, Zixia; Liu, Junyan; Wu, Wei Emma; Fang, Ruogu; Liu, Sheng; Wei, Qingyue; Yan, Rui; Guo, Yi; Tao, Q.; More Authors

DOI
10.1038/s43588-025-00911-9

Publication date
2025

Document Version
Final published version

Published in
Nature Computational Science

Citation (APA)
Zhou, Z., Liu, J., Wu, W. E., Fang, R., Liu, S., Wei, Q., Yan, R., Guo, Y., Tao, Q., & More Authors (2025). Revealing neurocognitive and behavioral patterns through unsupervised manifold learning of dynamic brain data. *Nature Computational Science*, *5*(12), 1238-1252. https://doi.org/10.1038/s43588-025-00911-9

nature computational science

Article

https://doi.org/10.1038/s43588-025-00911-9

Revealing neurocognitive and behavioral patterns through unsupervised manifold learning of dynamic brain data

Received: 9 September 2024

Accepted: 24 October 2025

Published online: 4 December 2025

Check for updates

Zixia Zhou[1], Junyan Liu[1], Wei Emma Wu[1], Ruogu Fang[2], Sheng Liu[1], Qingyue Wei[1,3], Rui Yan[1,3], Yi Guo[4], Qian Tao[5], Yuanyuan Wang[4], Md Tauhidul Islam[1] & Lei Xing[1,3,6]

Dynamic brain data are becoming increasingly accessible, providing a gateway to understanding the inner workings of the brain in living participants. However, the size and complexity of the data pose a challenge in extracting meaningful information across various data sources. Here we introduce a generalizable unsupervised deep manifold learning for exploration of neurocognitive and behavioral patterns. Unlike existing methods that extract patterns directly from the input data, the proposed brain-dynamic convolutional-network-based embedding (BCNE) captures brain-state trajectories by analyzing temporospatial correlations within the data and applying manifold learning. The results demonstrate that BCNE effectively delineates scene transitions, underscores the involvement of different brain regions in memory and narrative processing, distinguishes dynamic learning processes and identifies differences between active and passive behaviors. BCNE provides an effective tool for exploring general neuroscience inquiries or individual-specific patterns.

The brain is a highly complex and dynamic organ, continuously adapting and responding to both external stimuli and internal processes. Recent advances in neuroimaging and electrophysiological techniques, from noninvasive whole-brain methods such as functional magnetic resonance imaging (fMRI)[1] and electroencephalography[2] to invasive intracranial recordings with implanted electrodes in nonhuman primates, have enabled the study of neural activity across diverse spatial and temporal scales. These diverse data sources have provided effective means for illuminating the neural bases of cognitive and behavioral processes, such as problem solving, learning, decision-making and consciousness[3–5].

Dynamic brain data often involve extended, nonperiodic cognitive and behavioral processes, further increasing the complexity of the data. When considering the entire time series, the dynamic response of the brain across different recording sites is typically depicted as a matrix, where rows and columns correspond to spatial and temporal variables, respectively. The inherent complexity of the data with intertwined temporospatial relationships, compounded by the sparse and noisy nature of the data, poses a formidable challenge in cognitive and behavioral pattern discovery and analysis of brain behavior[6–9]. Computationally, dimensionality reduction and visualization techniques are often used to project high-dimensional (HD) vectors representing the brain's spatial responses at given time points into lower dimensions sequentially, obtaining temporal brain state trajectories (Fig. 1a). Common dimensionality reduction methods, such as Uniform Manifold Approximation and Projection (UMAP)[10], *t*-distributed stochastic neighbor embedding

[1]Department of Radiation Oncology, Stanford University, Stanford, CA, USA. [2]J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL, USA. [3]Institute of Computational and Mathematical Engineering, Stanford University, Stanford, CA, USA. [4]Department of Electronic Engineering, School of Information Science and Technology, Fudan University, Shanghai, China. [5]Department of Imaging Physics, Delft University of Technology, Delf, the Netherlands. [6]Department of Electrical Engineering, Stanford University, Stanford, CA, USA. e-mail: tauhid@stanford.edu; lei@stanford.edu

https://doi.org/10.1038/s43588-025-00911-9

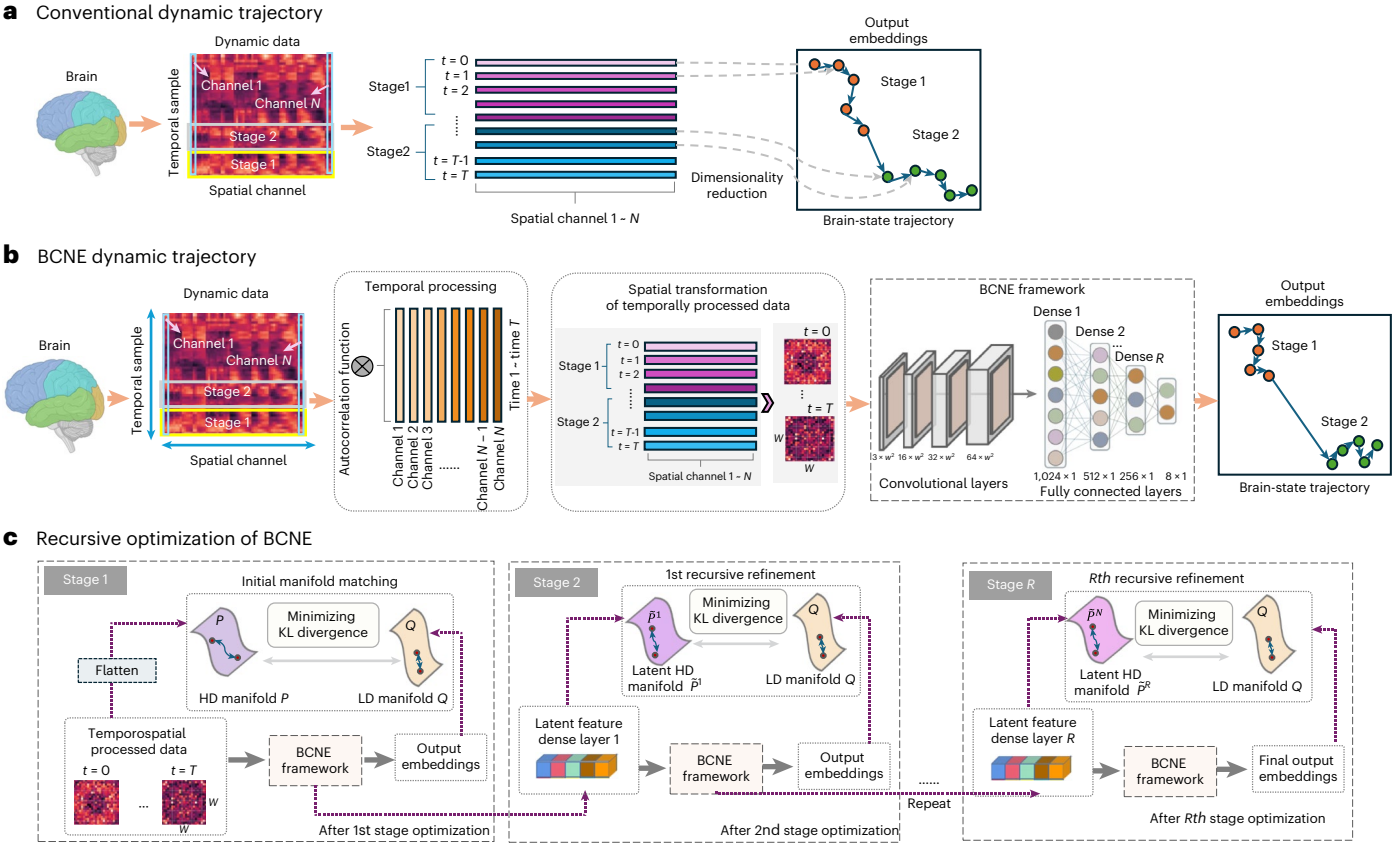


Fig. 1 | Schematic of the brain-dynamic trajectory visualization pipeline. **a**, Workflow of a conventional dimensionality reduction approach for understanding the evolvement of brain responses during behavioral transitions. **b**, Workflow of the proposed BCNE: original dynamic brain data (dimension $N \times T$) are first processed through temporal and spatial modules to produce structured images (dimension $N \times w \times w$). These structured images are then fed into the BCNE framework, which progressively optimizes the brain-state trajectory visualization by minimizing the KL divergence at each stage. **c**, Detailed illustration of the progressive BCNE network optimization process. At each recursion step r ($r = 1, \ldots, R$), latent features obtained from dense layer r become the input data representation for the subsequent stage. At each stage, the BCNE network is optimized by minimizing the KL divergence between the reference distribution (from the initial embedding P to the progressively refined embedding $\tilde{P}^R$) and the current embedding distribution Q. This process is repeated, with each stage incorporating a larger temporal or spatial receptive field, until the final stage R is reached.

(t-SNE)[11] and potential of heat-diffusion for affinity-based transition embedding (PHATE)[12], treat brain activity as a series of instantaneous snapshots, which often yield fragmented and spurious embeddings. Recent efforts have improved dynamic brain data visualization by explicitly incorporating temporal continuity, notably temporal PHATE (T-PHATE)[13], which utilizes temporal context for trajectory denoising, and CEBRA[14], which uses contrastive learning for meaningful structure identification from neural recordings. Our proposed brain-dynamic convolutional-network-based embedding (BCNE) approach advances dynamic brain data analysis by combining the strength of T-PHATE in temporal signal processing and the deterministic advantages of deep neural network-based techniques such as CEBRA. Specifically, BCNE explicitly denoises temporal signals within each spatial channel using an autocorrelation-based affinity matrix, as inspired by T-PHATE. To incorporate the interactions among recording channels, an image representation is introduced for each time point, which maps the responses of the spatial channels at that time point to image pixels in a manner that the intrinsic interchannel relationships are uniquely encoded through the image contextual pattern (see Fig. 1b and the Methods for details). These structured images are then processed by a convolutional neural network (CNN)-based framework, providing unsupervised dimensionality reduction of temporospatial data by progressively minimizing the Kullback–Leibler (KL) divergence between the pairwise similarity distributions in their HD and low-dimensional (LD) representation spaces (Fig. 1c). BCNE thus generates structured two-dimensional (2D) embeddings that effectively integrate temporal dynamics with spatial relationships, substantially improving the detection and interpretation of subtle cognitive and behavioral patterns.

Specifically, we introduce a principled and broadly applicable deep-learning-based visualization framework, termed BCNE, designed for a wide range of dynamic brain data (Fig. 1b,c). BCNE leverages both temporal continuity and the correlational structure among different recording channels, such as fMRI voxels and electrophysiological electrode sites, to uncover meaningful trajectories that reflect underlying cognitive or behavioral states. Although it is well established that interactions across distinct brain states and recording locations are crucial for characterizing cognitive function, these interactions have seldom been fully harnessed for deeper pattern discovery. In BCNE, we begin by constructing a temporospatial representation of the multivariate time series data that captures both temporal dependencies and spatial correlations. From this representation, we apply an unsupervised convolutional framework to learn deterministic mappings that yield structured, interpretable trajectories. To further enhance our ability to uncover subtle patterns, we implement a recursive, manifold-based optimization strategy. This iterative refinement process progressively incorporates deeper-level constraints from the latent representation[15–18], enhancing the model's fidelity and enabling the extraction of more nuanced trajectories that reveal subtle underlying brain activity. It is important to emphasize that BCNE's deterministic nature produces consistent trajectories for identical stimuli under similar conditions,

supporting reproducibility, identifying individualized patterns of brain activity and offering a robust framework for analyzing dynamic brain functions.

The proposed BCNE technique is showcased using several well-established datasets sourced from diverse origins, including the Sherlock fMRI blood oxygen level-dependent (BOLD) dataset, the rat hippocampus dataset and the macaque dataset. Across all datasets, BCNE consistently outperforms traditional HD data interpretation tools in accurately delineating cognitive or behavioral stages within dynamic processes, offering versatility and noteworthy insights into brain dynamics that are unattainable with existing techniques.

Results

BCNE is a deterministic deep learning framework for visualizing dynamic brain data by jointly modeling temporal dependencies and spatiotemporal interactions across channels. Its methodological pipeline unfolds as follows, with details given in the Methods. (1) Temporal processing (Fig. 1b): To emphasize salient temporal dependencies while suppressing high-frequency noise, we calculate the autocorrelation function for each recording channel, determine the lag threshold at which correlations sharply decline and generate a lag-weighted average signal representation. (2) Generation of structured image representation of spatial–channel responses (Fig. 1b): to incorporate the interactions among recording channels, we introduce a schematically meaningful image representation for each time point, which maps the channel responses at that time point to image pixels such that inter-channel relationships are uniquely encoded in the resulting contextual image pattern[19]. (3) Feature extraction from the generated images via BCNE (Fig. 1b,c): the interactive features encoded in the image patterns of the data are extracted through a CNN, termed BCNE, to generate an LD embedding of the imaging data by minimizing the KL divergence between the joint-probability matrices of the HD input data and their LD embeddings (dashed box on the left of Fig. 1c). (4) Recursive refinement of the outcome (Fig. 1c): to further improve the resulting trajectory of the brain-dynamic data, the initial HD probability matrices used for network optimization are replaced by feature vectors extracted from the first dense layer of BCNE (Fig. 1c). Subsequently, BCNE fine-tunes the KL divergence between the joint probability matrices, yielding a refined LD embedding.

Task-related associations across regions from fMRI BOLD data

We analyzed the Sherlock dataset using various embedding approaches to demonstrate BCNE's potential for elucidating fMRI BOLD signal dynamics during viewing of the BBC Sherlock movie (Fig. 2a). We focused on four regions of interest (ROIs): the early visual cortex (EV), high visual cortex (HV), early auditory cortex (EA) and posterior medial cortex (PMC) (Fig. 2b). These ROIs were selected on the basis of prior studies[13,20,21] showing that they generated robust and reliable signals under naturalistic audiovisual stimulation. Each region demonstrated distinct activity levels during the task. Figure 2c,d illustrates visualizations produced by BCNE and several standard dimensionality reduction techniques, annotated with 39 movie-scene-related categories. Traditional methods (principal component analysis (PCA), t-SNE, UMAP and PHATE) offered limited clarity in capturing temporal and contextual aspects of brain activity. By incorporating temporal structure, T-PHATE showed noticeable differences among the four regions: EV and PMC exhibited comparatively less variability in perception and comprehension processes, whereas HV and EA revealed stronger trajectories. However, T-PHATE's performance was sometimes unstable (Extended Data Figs. 1 and 2), struggling to capture detailed temporal dynamics consistently. For CEBRA, we specifically used CEBRA-Time (Methods), finding that, and found that, although it could produce richer results in three dimensions, its 2D embeddings conveyed less detail. By contrast, BCNE excelled at embedding dynamic signals with stable and interpretable patterns. As shown in Fig. 2c,d, BCNE sharpened the visualization, highlighting intricate morphological trajectories within each ROI more effectively than other methods. Notably, BCNE visualizations revealed that EV and PMC embeddings were more spread out, while EA and HV showed tighter, continuous trajectories. We next evaluated downstream classification performance by assigning the dimension-reduced embeddings to 39 scene categories using 10-fold cross-validation and a k-nearest neighbors (KNN) classifier (Fig. 2e–g). BCNE achieved the highest accuracy across all four ROIs, improving on T-PHATE by 81.74% (averaged across participants and ROIs).

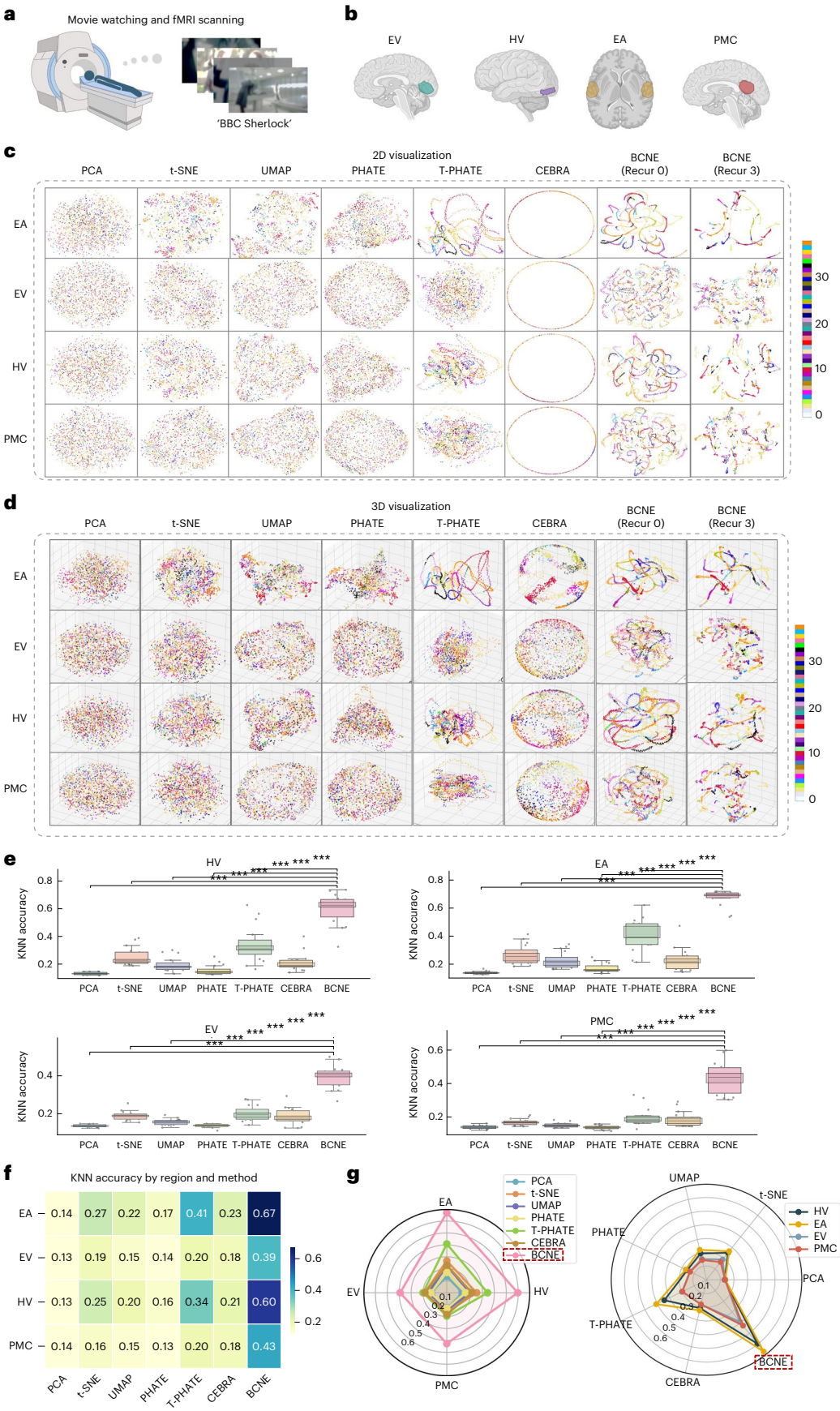
Further, deeper latent representations from BCNE (Fig. 3a,b) made cognitive and behavioral transitions even more distinct, underscoring its ability to capture changes in perception and comprehension throughout the movie. We also evaluated whether BCNE's performance is sensitive to model structure and hyperparameters by comparing different architectural and parameter variants. The results showed that BCNE was robust to these changes and achieved the best performance with manageable computational costs (Extended Data Figs. 3–6 and Supplementary Table 1). Finally, we performed behaviorally rated event-boundary analyses to assess how well each method's embeddings captured meaningful cognitive transitions[13,22,23]. This analysis measured alignment between the embeddings and event boundaries identified by an independent behavioral rater (Methods). As shown in Fig. 3f–h, across 16 participants and four ROIs, BCNE outperformed PCA, t-SNE, UMAP, PHATE, T-PHATE and CEBRA in capturing these transitions. Although CEBRA slightly exceeded BCNE in PMC, its variance was notably larger, and BCNE remained superior in HV, EA and EV. Overall, these results confirmed that BCNE preserved both local (intra-event) coherence and interpretable across-event transitions, offering a comprehensive perspective on cognitive trajectories.

Learning stage transitions from hippocampal dynamics

We applied BCNE to a rat hippocampus dataset to examine neural changes during learning. In this experiment, rats learned to traverse a linear track for a water reward (Fig. 4a), allowing us to observe how hippocampal activity evolved as the animals memorized an optimal path.

Fig. 2 | Visualization and analysis of the Sherlock fMRI dataset. All results are averaged over three random seeds, with the mean value used as the final evaluation criterion. **a**, Schematic illustration of the fMRI BOLD signal acquisition process during viewing of the BBC Sherlock movie, involving 16 participants. Movie-related images within the figure have been intentionally obscured to comply with copyright restrictions. **b**, ROIs selected for embedding and visualization. **c**, Comparative 2D visualizations generated by PCA, t-SNE, UMAP, PHATE, T-PHATE, CEBRA, BCNE (Recur0) and BCNE (Recur3) for four ROIs: EV, HV, EA and PMC, with data colored by movie scene. **d**, Comparative 3D visualizations generated by the same methods and colored by movie scene. **e**–**g**, Boxplots, heatmap and radar charts showing the average KNN classifier accuracy for the movie scene classification (39 categories) task based on 2D embeddings generated by PCA, t-SNE, UMAP, PHATE, T-PHATE, CEBRA and BCNE (Recur3) across the four ROIs. For all boxplots, the center line indicates the median, box limits denote the interquartile range (IQR; 25th–75th percentiles), whiskers indicate the full data range and individual points represent each participant ($n = 16$ per group). The shaded rectangle within each box highlights the dense region (40th–60th percentiles), using a darkened box color to emphasize central tendency. Overlaid dot plots display individual participant values. The unit of analysis is a single human participant; all n values refer to distinct participants. No technical replicates were used. Statistical comparisons between groups were performed using two-sided independent t-tests unless otherwise indicated. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. For all pairwise comparisons, the t-statistic, degrees of freedom and exact two-sided P values are provided in the Source data. Only significant comparisons are annotated in the plots. Box colors correspond to the methods indicated on the x axis. Unless otherwise specified, the same statistical testing and boxplot conventions are applied in Figs. 3 and 4.

https://doi.org/10.1038/s43588-025-00911-9



https://doi.org/10.1038/s43588-025-00911-9

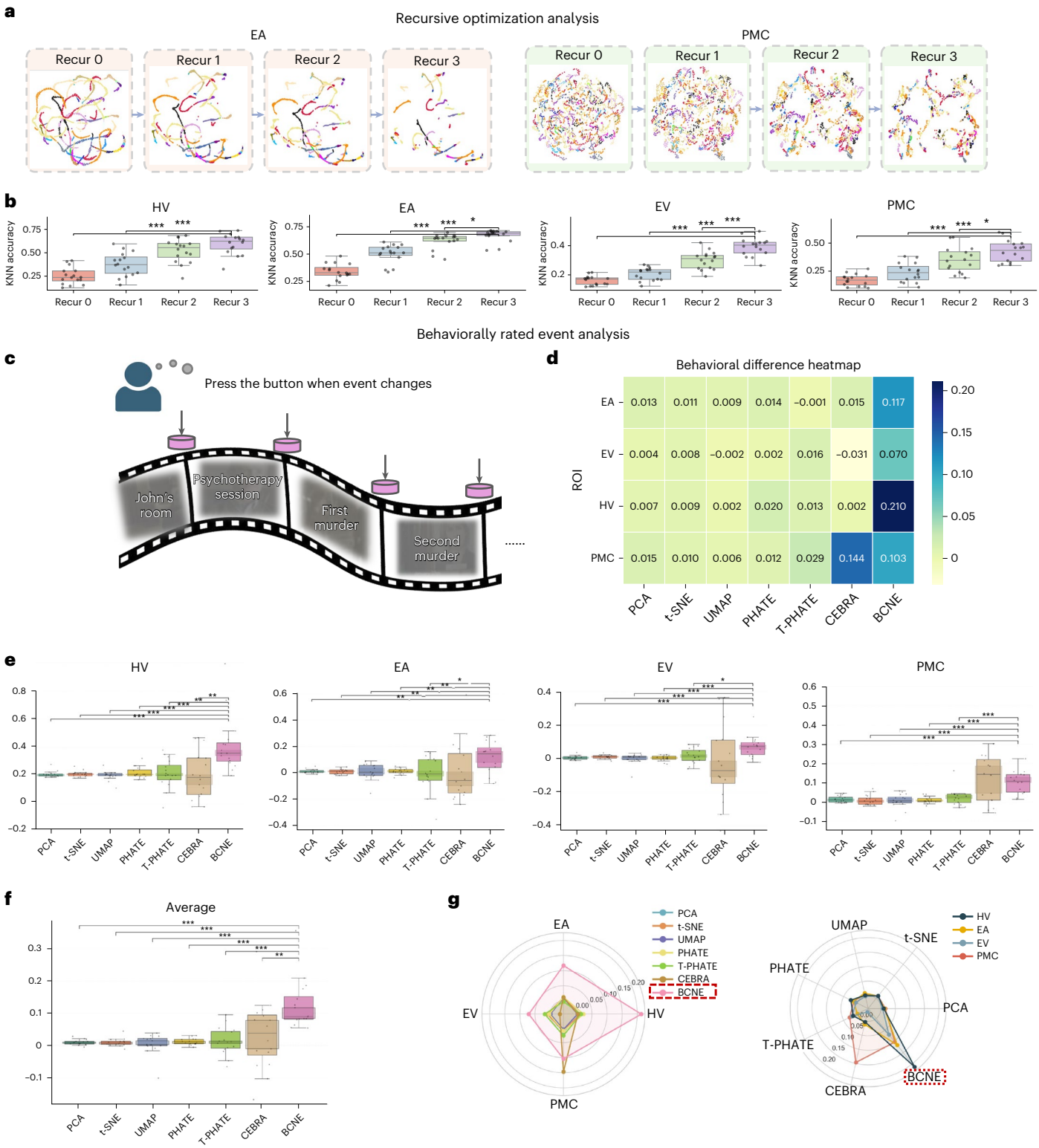


Fig. 3 | Evaluation of the BCNE framework and behaviorally rated event analysis of the Sherlock fMRI dataset. a, Examples of 2D visualizations generated by BCNE with varying recursion depths (Recur 0–3). Colors correspond to the same scheme defined in Fig. 2c. **b**, Boxplots showing KNN classifier accuracy calculated from 2D embeddings generated by BCNE at different recursion depths ($n = 16$ per group). **c**, Description of the process by which behavioral raters generated event segmentation labels. **d**, Heatmaps showing behaviorally rated event-segmentation scores across 2D embeddings generated by PCA, t-SNE, UMAP, PHATE, T-PHATE, CEBRA and BCNE for the four ROIs. **e**, Boxplots of event-segmentation scores for each ROI, including statistical significance testing across methods ($n = 16$ per group). **f**, Boxplots of the averaged event-segmentation scores across ROIs ($n = 16$ per group). **g**, Radar charts summarizing event-segmentation performance across methods for the four ROIs.

https://doi.org/10.1038/s43588-025-00911-9

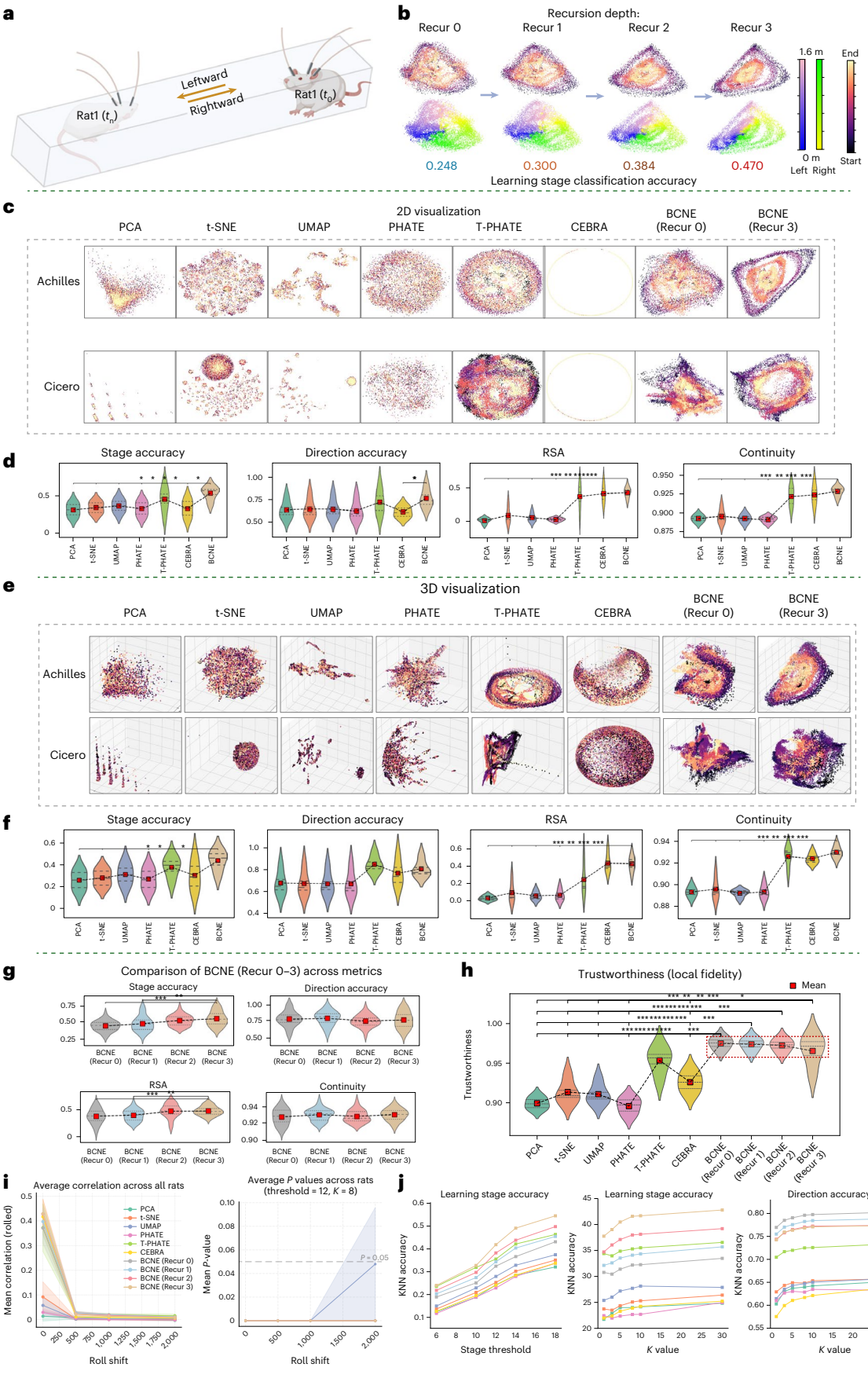


Fig. 4 | Results of the rat hippocampus dataset. All results are averaged over three random seeds, with the mean value used as the final evaluation criterion. **a**, Schematic illustration of the rat hippocampus dataset acquisition process. Electrophysiological data were collected while four rats traversed a 1.6-m linear track in either a leftward or rightward direction. **b**, Example of 2D visualizations generated by BCNE with varying recursion depths, with classification accuracy for learning stages indicated below each visualization. The upper visualization is colored by learning stage, while the lower visualization uses two distinct color bars to differentiate movement positions for leftward and rightward motions. **c**, Two-dimensional visualizations for two randomly selected rats (named Achilles and Cicero) generated by PCA, t-SNE, UMAP, PHATE, T-PHATE, CEBRA and BCNE (Recur 0 and Recur 3). **d**, Violin plots of stage accuracy, direction accuracy, RSA and continuity for the 2D embeddings. **e**, 3D visualizations of the embeddings corresponding to the conditions described in **c**. **f**, Violin plots summarizing the distribution of quantitative metrics derived from the 3D embeddings **d**. **g**, Violin plots showing stage accuracy, direction accuracy, RSA and continuity, calculated using 2D BCNE with varying recursion depths (Recur 0–3). **h**, Trustworthiness comparison calculated by 2D embeddings across all methods. **i**, Average correlation analysis using roll-shift testing calculated by 2D embeddings across all rats. Correlation values and *P* values are plotted. **j**, Analysis of learning stage accuracy calculated by 2D embeddings across different stage thresholds and the influence of *k* values on KNN classifier performance. Violin plots show data from $n = 4$ biological replicates per group (four individual rats). No adjustments for multiple comparisons were made. For **i** and **j**, line colors correspond to the methods indicated in the boxplots and are defined consistently across all panels. Panel **a** created with BioRender.com.

First, we assessed how increasing the recursive depth in BCNE affected the clarity of learning stage separation. With deeper recursion (Recur 0 → Recur 3), BCNE more clearly distinguished temporal phases in the neural data (Fig. 4b). We quantified these differences by dividing the navigation data into discrete learning stages and classifying the embeddings with a KNN classifier. Accuracy rose from 0.248 (Recur 0) to 0.470 (Recur 3), showing that BCNE could uncover meaningful structure related to behavioral learning. We then compared BCNE with other methods—PCA, t-SNE, UMAP, PHATE, T-PHATE and CEBRA—using both 2D and three-dimensional (3D) embeddings (Fig. 4c,e). For instance, with PCA, rat Achilles showed minimal separation of time points, while t-SNE and UMAP often produced scattered points with little temporal continuity. PHATE's outputs were typically circular with limited dynamic information. T-PHATE occasionally captured some temporal structure, but results varied across rats. CEBRA's 2D embeddings were largely indistinct, although 3D ones showed some improvement. By contrast, BCNE reliably captured dynamic variations, forming clear trajectories and demonstrating robustness across animals. Extended evaluations on additional rat datasets and systematic ablation studies further validate the robustness of BCNE and its ability to uncover patterns that are otherwise obscured. (Extended Data Figs. 7 and 8).

Quantitatively, BCNE achieved the highest classification accuracy for learning-stage identification among all tested rats (Fig. 3d,f). Additional analyses, including direction accuracy, representational similarity analysis (RSA) with spatial labels and embedding continuity, all favored BCNE in two dimensions. For 3D embeddings, T-PHATE had slightly higher direction accuracy, and CEBRA performed marginally better in RSA, but BCNE led in all other metrics. A systematic evaluation of BCNE's recursive optimization showed that deeper recursion generally improved learning-stage differentiation, while other metrics remained stable (Fig. 4g). For local structure preservation, BCNE embeddings consistently yielded higher trustworthiness scores than other methods (Fig. 4h–j). As recursion depth increased, BCNE captured more global structure but lost some fine detail, leading to a minor decrease in trustworthiness.

To further validate the method, a roll-shift analysis showed that all embeddings—except for UMAP—maintained low *P* values across time shifts, indicating statistical reliability. UMAP's performance degraded as time shifts increased. Finally, we tested different criteria for dividing learning stages (by varying reward frequency and *k* values in KNN classification). Under all tested conditions, BCNE remained the top performer, accurately delineating learning stages and tracking the evolution of the rats' navigational strategies. Regardless of the KNN neighborhood size, BCNE achieved the highest accuracy, consistently identifying and following dynamic hippocampal processes.

Discrimination between active and passive behavior

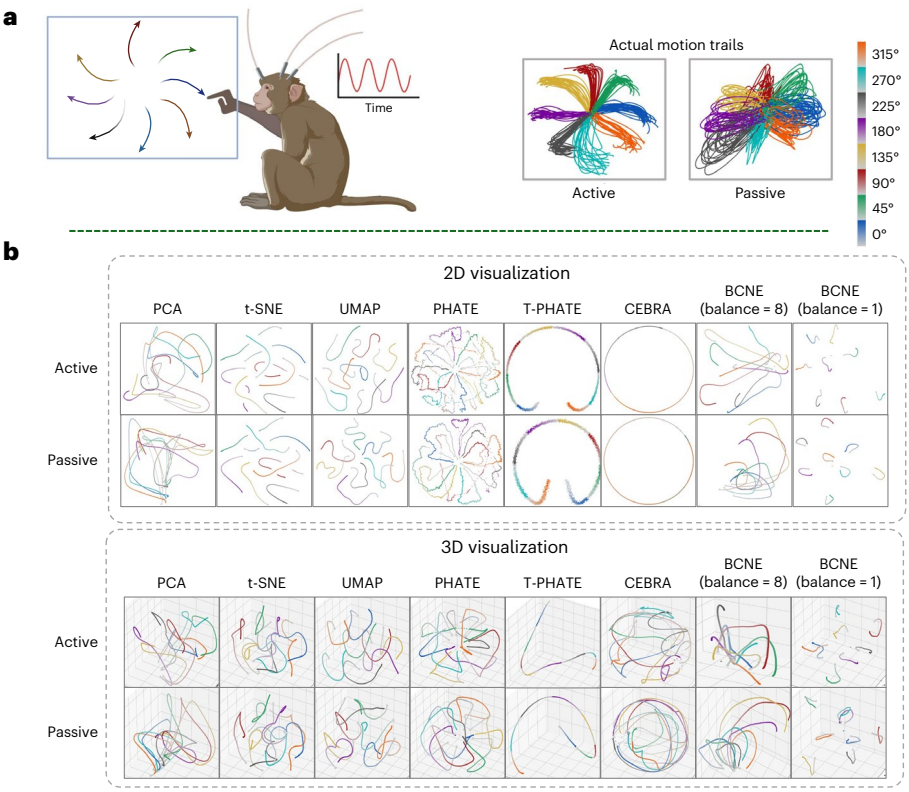
We next evaluated BCNE on a macaque sensorimotor dataset in which a macaque performed center-out arm movements along eight distinct directions, under both active and passive conditions (Fig. 5a). To visualize these data, we first averaged trials sharing the same target angle, then generated 2D and 3D embeddings separately for the active and passive conditions (Fig. 5b). Notably, BCNE was the only method that clearly differentiated between active and passive states in the resulting LD space, whereas PCA, t-SNE, UMAP, PHATE and T-PHATE all produced similar or overlapping embeddings for the two modes. In previous studies[24,25], the experimental results showed that neurons in area 2 represented limb states differently depending on the movement type and that the encoding of passive movements differed systematically from active reaches. This aligned with our observations: the BCNE-derived embeddings produced distinct trajectory shapes under active and passive conditions, reflecting differences in how neural signals vary in response to movement type. When comparing the trajectories generated by BCNE from brain data, we observed that active movements produced more linear paths, reflecting a direct reach toward a specific angle, while passive movements resulted in more curved or circular patterns, aligning closely with the actual motion trails.

To evaluate the alignment between neural embeddings and actual limb positions, we performed RSA and roll-shift tests across all eight reach directions (Fig. 5c,d). The results were visualized using a heatmap that compared active and passive RSA scores across different embedding methods. As shown, RSA values were generally higher in the active mode, indicating a more explicit encoding of position-related information during voluntary movements. Among the methods, BCNE achieved the highest RSA scores in the active mode, with averages of 0.973 (2D) and 0.930 (3D). T-PHATE ranked second, with correlation values of 0.918 (2D) and 0.924 (3D). By contrast, RSA scores in the passive mode were largely unaffected by increasing roll-shifts (Fig. 5d), implying that fewer task-related signals were present during passive movements and, consequently, that limb position was less distinctly represented in the neural embeddings.

Lastly, to quantify how well each embedding preserved angle-specific information, we next applied a KNN classifier to predict one of eight reaching directions under both active and passive

Fig. 5 | Results of the macaque dataset. All results are averaged over three random seeds, with the mean value used as the final evaluation criterion. **a**, Illustration of the macaque dataset acquisition process. Data were collected as a macaque performed active or passive movements in eight directions. The actual position of the macaque's arm was recorded and displayed. **b**, The 2D and 3D visualizations of active and passive modes, generated by PCA, t-SNE, UMAP, PHATE, T-PHATE, CEBRA and BCNE. The color bar transitions gradually from light gray to vibrant colors, representing the progression of time during directional reaching. **c**, Circular plots and heatmaps showing the RSA between actual positions for each angle and the embeddings generated by BCNE and all compared methods. **d**, Line charts for roll-shift testing calculated by 2D embeddings. **e**, Bar plots showing the reaching angle classification accuracy based on KNN classifiers, using 2D embeddings generated by different methods, for both active and passive modes. Panel **a** created with BioRender.com.

https://doi.org/10.1038/s43588-025-00911-9



conditions (Fig. 5e). For this eight-class task, chance performance was only 12.5%. Using both 2D and 3D embeddings, t-SNE and BCNE achieved near-perfect accuracy, while UMAP and T-PHATE also performed well, exceeding 90% accuracy. By contrast, the remaining methods yielded substantially lower scores. These results demonstrated that BCNE not only visually segregated distinct movement types but also retained precise, angle-specific neural dynamics that supported highly reliable decoding.

Discussion

BCNE provides a robust framework for uncovering dynamic neural patterns, with one of its major strengths being the ability to disentangle continuous temporal evolution from abrupt behavioral transitions. Unlike methods that cluster time points primarily on the basis of proximity, BCNE constructs embeddings that emphasize correlation-driven similarity rather than mere adjacency. This feature ensures that transitions between distinct cognitive or behavioral states are faithfully represented in the LD manifold. The flexibility introduced by adjusting recursion depth allows BCNE to reveal the dynamics at multiple levels of granularity, ranging from broad continuous trajectories to discrete, event-like clusters. This multiresolution capacity is especially valuable for exploratory analyses of complex brain dynamics, where the underlying states may manifest across different temporal scales.

Compared with traditional parametric manifold-learning methods such as t-SNE or UMAP, BCNE's temporospatially integrated CNN backbone and recursive refinement provide substantial improvements in embedding quality and interpretability. Unlike multilayer perceptron (MLP)-based approaches, which often fail to capture intricate spatiotemporal features, BCNE leverages convolutional structures to encode fine-grained spatial patterns while preserving global trajectory continuity. The scalability of BCNE also makes it more suitable for large-scale datasets than alternatives such as T-PHATE, which often suffer from exponential increases in computational cost as the dataset grows. This computational efficiency is critical for dynamic brain data, which frequently span long recording periods with thousands of time points.

The recursive optimization strategy adopted by BCNE addresses a long-standing limitation of direct dimensionality reduction methods, which tend to preserve only local neighborhoods at the expense of global structure. By progressively compressing information through intermediate latent representations[15–18], BCNE aligns local relationships while maintaining the global topology of the underlying manifold. This hierarchical compression, combined with similarity constraints, helps to mitigate the risk of local minima and ensures stable convergence during training. Such a principled design allows BCNE to reveal both fine-scale and large-scale patterns, offering a more holistic view of neural dynamics that is difficult to achieve with single-stage dimensionality reduction.

From a conceptual perspective, BCNE emphasizes unsupervised discovery rather than supervised classification. Brain activity is highly heterogeneous and often exhibits emergent patterns that cannot be adequately captured by predefined labels. Supervised approaches risk biasing the analysis toward known concepts, potentially obscuring previously unexplored dynamics. By contrast, BCNE's unsupervised strategy enables open-ended exploration, allowing the intrinsic structure of the data to emerge without prior constraints. This approach is particularly advantageous for datasets where behavioral annotations are limited, ambiguous or subjective.

Despite these strengths, BCNE has several limitations. First, the architecture and hyperparameters of BCNE were tuned using only three representative datasets. The generalizability of these design choices to data with markedly different complexity, noise characteristics or sampling density remains uncertain. In particular, datasets with extreme temporal sparsity or irregular sampling may require modifications to the temporal affinity matrix or alternative weighting strategies to ensure robust embeddings. Second, the current design of the temporal affinity matrix prioritizes local correlations and introduces a drop-off threshold to maintain computational efficiency. While this design improves scalability, it may fail to capture long-range dependencies when behavioral states recur after extended intervals. Future work could incorporate adaptive correlation models, such as attention-based mechanisms or graph neural networks, to better account for these long-range interactions. Third, the convergence behavior of BCNE is influenced by heuristic parameters inherited from classical manifold-learning methods, including fixed perplexity, binary search iterations and tolerance values. These parameters may not be universally optimal, and dataset-specific calibration could be required to achieve optimal results. In addition, BCNE's deep network training can encounter instability under extreme conditions such as HD low-sample regimes or datasets with severe class imbalance, potentially leading to gradient vanishing or explosion[26]. Strategies such as adaptive learning rates, regularization or meta-learning could improve stability and generalizability. Another limitation concerns interpretability in higher-dimensional embeddings. Although this study focuses on 2D and 3D representations for visualization purposes, BCNE is inherently capable of producing embeddings in arbitrary dimensions. Systematic exploration of these higher-dimensional embeddings, particularly for predictive or diagnostic tasks, is a promising but unexamined direction. Furthermore, while BCNE effectively preserves the geometric structure of the input data, its embeddings are not directly optimized for downstream predictive accuracy. Integrating self-supervised objectives or hybrid learning schemes could bridge this gap.

Several promising avenues could enhance BCNE's future utility. First, extending BCNE to handle multimodal neural data—such as simultaneous fMRI, electroencephalography and calcium imaging recordings—could provide more comprehensive brain-state mappings. Incorporating cross-modal alignment techniques would substantially improve its utility in interpreting complex neural phenomena. Second, validating BCNE across diverse experimental conditions and varied participant populations could further establish its general applicability. Testing its effectiveness on datasets representing pathological conditions, such as epilepsy, neurodegenerative diseases or psychiatric disorders, may uncover potential clinical diagnostic and prognostic applications. Trajectory-based analyses could provide valuable biomarkers for monitoring disease progression or therapeutic responses. Third, refining correlation analysis within BCNE to better incorporate long-range temporal dependencies constitutes a crucial next step. Recent advancements in transformer architectures or graph convolutional networks could effectively model these dependencies while maintaining computational feasibility. Fourth, exploring BCNE's capacity for real-time dynamic functional connectivity analysis could facilitate applications in closed-loop neurofeedback or brain–computer interface systems. Real-time applications necessitate computational efficiency and interpretability—attributes inherent to BCNE's batch-wise training strategy and unsupervised manifold construction. Lastly, systematically comparing BCNE against state-of-the-art self-supervised and contrastive learning techniques could clarify the relative advantages of manifold-learning versus contrastive objectives in capturing neural dynamics. Incorporating uncertainty estimation into BCNE could further enhance its utility in clinical contexts, where confidence metrics are essential.

In summary, BCNE offers a principled framework for visualizing and interpreting dynamic brain data by combining temporospatial encoding with recursive deep manifold learning. Its ability to reveal multiscale neural dynamics, coupled with its computational efficiency and adaptability, positions it as a powerful tool for both basic neuroscience and clinical research. Although current limitations highlight areas for improvement—particularly in handling long-range dependencies,

optimizing hyperparameters for diverse datasets and exploring higher-dimensional embeddings—BCNE lays a strong foundation for future advancements. By extending its applicability to multimodal and pathological datasets, BCNE could become a key enabler of precision neuroscience, driving discoveries that span from fundamental brain mechanisms to clinical translation.

Methods

Technical framework of BCNE

Temporospatial correlation-based reconfiguration. *Temporal projection*. To effectively visualize the dynamic brain signal with informative temporal trajectory, we used a temporal autocorrelation-based matrix[13] to capture the temporal dynamics of brain activity and reveal the evolving relationships during the dynamic recognition process of a living participant. Let $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T\}$, where $\mathbf{x}_t \in \mathbb{R}^N$ is the N-dimensional brain-signal vector at time t. We first compute the per-channel autocorrelation matrix $A \in \mathbb{R}^{N\times(T-1)}$:

$$A(i,\tau) = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}(X_{t,i}-\bar{X}_i)\left(X_{t+\tau,i}-\bar{X}_i\right), i = 1,\ldots,N, \tau = 1,\ldots,T-1, \quad (1)$$

where τ represents the time lag and i denotes the recording channel. $\bar{X}_i = \frac{1}{T}\sum_{t=1}^{T} X_{t,i}$ represents the mean signal of channel i. The autocorrelation matrix is then smoothed using a centered moving average filter:

$$\bar{A}(i,\tau) = \frac{1}{\omega}\sum_{p=\tau-\lfloor(\omega-1)/2\rfloor}^{\tau+\lfloor(\omega-1)/2\rfloor} A(i,p), \quad (2)$$

where ω represents the window size for smoothing. In this process, the convolution operation applies a rolling average, ensuring a centered smoothing effect. If the window exceeds the boundaries of the time series, the effective window size is automatically reduced near the edges, avoiding the need for explicit padding.

We define drop-off point as the first index where the smoothed autocorrelation function becomes negative, serving as a threshold beyond which correlations are considered negligible. Based on the drop-off point τ_{drop}, a time-dynamic correlation matrix $M \in \mathbb{R}^{T\times T}$ is constructed as follows:

$$M(i,j) = \begin{cases} \frac{1}{N}\sum_{f=1}^{N}\bar{A}(f,|i-j|), & 0 < |i-j| < \tau_{\text{drop}}, \\ 0, & \text{otherwise.} \end{cases}, \quad (3)$$

where $M(i,j)$ represents the strength of correlation between time points i and j, incorporating the temporal context captured by the autocorrelation function.

Once the time-dynamic correlation matrix M is obtained, the temporally reconfigured data matrix $\tilde{X} \in \mathbb{R}^{N\times T}$ is computed as

$$\tilde{X} = XM \oslash (\mathbf{1}_T^{\top} M), \quad (4)$$

where $1_T \in \mathbb{R}^T$ is a vector of ones, so that $1_T^{\top} M$ produces the row-sums of M as a $1\times T$ vector, and $\oslash$ denotes elementwise division broadcast across each row of XM. Equivalently, the tth column is

$$\tilde{X}_{,t} = \frac{\sum_{t'=1}^{T} M(t,t')X_{,t'}}{\sum_{t'=1}^{T} M(t,t')}. \quad (5)$$

Here, the new column $\tilde{X}_{,t}$ is a weighted average of the original time-point vectors $X_{,1},\ldots,X_{,T}$. A large weight $M(t,t')$ indicates that time points t and t' are strongly correlated, so $X_{,t'}$ contributes more to $\tilde{X}_{,t}$. Dividing by $\sum_{t'=1}^{T} M(t,t')$ ensures that each reconfigured column has a consistent scale, preventing any single time-point cluster from dominating. In essence, equation (5) smooths and aggregates each time point t according to the nearest correlated neighbors in time.

Image representation of the brain-dynamic data.

(1) Motivation

We represent the responses of the spatial channels, the HD brain signals, at each time point as a schematically meaningful 2D image to effectively capture the interactions among recording channels. The transformation maps the responses of the spatial channels onto image pixels in such a way that the interchannel relationships are encoded within the contextual pattern of the image. Specifically, pairwise interchannel correlations are computed from the inverse covariance of temporally processed signals, obtained by autocorrelation-based lag-weighted averaging (Fig. 1b). These interactions are encoded contextually into images by minimizing the Gromov–Wasserstein (GW) discrepancy between the interaction matrix and pixelated grid-distance matrix. This process adapts the GenoMap[19] concept, where gene–gene interactions are encoded as images for high-performance analysis of HD single-cell RNA sequence data, to dynamic brain signals. By representing interchannel dependencies in a structured 2D image, standard CNNs can directly exploit these correlations for more effective pattern discernment and downstream analyses.

(2) Quantification of feature–feature interactions

At each time point t, the temporally reweighted data matrix is denoted $V = \tilde{X} \in \mathbb{R}^{N\times T}$. Let

$$\Omega_{ij} = \frac{1}{T}\sum_{t=1}^{T}(V_{i,t}-\bar{V}_i)(V_{j,t}-\bar{V}_j), \qquad \bar{V}_i = \frac{1}{T}\sum_{t=1}^{T} V_{i,t}, \quad (6)$$

so that Ω is the sample covariance and Ω^{-1} its precision matrix. We defined the channel interaction matrix

$$C_{ij} = \begin{cases} -\dfrac{(\Omega^{-1})_{ij}}{\sqrt{(\Omega^{-1})_{ii}(\Omega^{-1})_{jj}}}, & i \neq j, \\ 1, & i = j, \end{cases} \quad (7)$$

which measures standardized pairwise dependencies between channels.

(3) 2D grid template

We create a square lattice with side $w = \lceil\sqrt{N}\rceil$, assigning each pixel integer coordinates $(x_k, y_k) \in \mathbb{Z}^2$ centered at the origin. The Euclidean distance matrix of the grid $\bar{C} \in \mathbb{R}^{N\times N}$ is

$$\bar{C}_{kl} = \sqrt{(x_k-x_l)^2 + (y_k-y_l)^2} \quad (8)$$

Both C and $\bar{C}$ are divided by their means so that $\sum_{i,j} C_{ij} = \sum_{i,j}\bar{C}_{ij} = N^2$.

(4) Optimal assignment via GW alignment

Following optimal-transport theory, we seek a coupling S that aligns the intra-feature geometry of C with the intrapixel geometry of $\bar{C}$. Define uniform marginals $u = v = \frac{1}{N}\mathbf{1}_N$, where $\mathbf{1}_N$ is the N-dimensional all-ones vector. The GW discrepancy is

$$d_{\text{GW}}(C,\bar{C},u,v) = \min_{S\in\Pi(u,v)}\sum_{i,j,k,l} L(C_{ik},\bar{C}_{jl})S_{ij}S_{kl}, \quad (9)$$

with the admissible coupling set $\Pi(u,v) = \{S \geq 0 \mid S\mathbf{1}_N = u,\ S^{\top}\mathbf{1}_N = v\}$, and cost $L(a,b) = a\log\frac{a}{b} - a + b$ (the KL divergence). We solve equation (9) via an entropically regularized Sinkhorn algorithm[19,27]. The resulting $S\epsilon\mathbb{R}_+^{N\times N}$ is near-bistochastic (rows and columns sum to $1/N$); scaling by N enforces exact bistochasticity. A hard, one-to-one assignment is obtained by selecting, for each channel i, the pixel j with maximum S_{ij}. Here, S_{ij} represents the probability mass transported from feature i to pixel j. Large S_{ij} values indicate that feature i is best

placed at grid position j, so channels with strong mutual correlations (large C_{ik}) are steered toward spatially proximate pixels (small $\tilde{C}_{jl}$).

(5) Generating 2D image representation

For each time point t, let $\mathbf{v}_t \in \mathbb{R}^N$ be the tth column of the temporally reweighted data. We then form

$$\tilde{V}_t = \text{reshape}(S^\top \mathbf{v}_t, w, w) \in \mathbb{R}^{w\times w}, \tag{10}$$

yielding an image sequence $\tilde{V} \in \mathbb{R}^{T\times w\times w}$. In each image representation V_t, interactions between spatial channels are uniquely encoded within the image's contextual pattern, making it well suited for analysis by the subsequent CNN. It is important to note that, in our spatial transformation, the generation of a contextual image representation of the dynamic brain data is directly based on the mathematical correlations among spatial channels, rather than on explicit functional connectivity analyses.

Convolutional-based embedding and recursive optimization. *Convolutional-based BCNE architecture*. Following the construction of temporospatial correlations, the signal from each participant at each time point is reformatted into an image. This allows the temporospatial correlation-induced reconfigured data to be input into the BCNE. The network is established on the basis of convolutional architecture, as illustrated in Fig. 1b. The architecture initiates with a module that includes four convolutional layers to extract shallow features from the input data in image format. The temporospatial structure of the reconfigured data ensures that points of high correlation are centrally relocated and positioned close to each other, making convolutional layers an apt choice for processing this type of data. They focus on local connections, thus enhancing the depth of the latent space while incurring a substantially lower computational cost compared with that of a purely MLP framework. Following this, a flattening layer converts the 2D latent features into a one-dimensional format. This conversion is succeeded by a fully connected module, consisting of several dense layers. This module projects the data from a global perspective to generate the final output embedding. After the network is trained, the brain signals can directly produce participant-specific visualizations. These visualizations are applicable to the external data embeddings of the same participant's analogous dynamic cognitive and behavioral tasks.

Recursive optimization strategy.

(1) Motivation

An important function of the multiple dense layers in our BCNE is to sequentially reduce the dimensionality of the feature vector representation of the input HD data, from 1,024 in layer 1, 512 in layer 2, 256 in layer 3, 8 in layer 4 and finally to 2 or 3 dimensions. Broadly, each step of the dimensionality reduction here compresses the feature vector to a lower dimension while injecting the vector components with more distant and global relational information of the data as a result of MLP optimization, which captures complex relationships across the data space as information flows through the dense layers[28–31]. Our recursive derivation of the brain state trajectory capitalizes on this unique characteristic of sequential dimensionality reduction to incorporate the information of brain data at different scales. Specifically, in stage 1 (Fig. 1c), we first optimize the BCNE embedding by minimizing its KL divergence with the original HD data. Similar to traditional parametric t-SNE, this embedding primarily takes into consideration the local interactions of the data. However, it benefits from the introduction of image representation and the use of convolutional operations in embedding. In stage 2 (Fig. 1c), the BCNE embedding network is fine-tuned by replacing the original HD input with the feature representation vector from the first dense layer, which helps to integrate the next level of nonlocal data relationships into the embedding process. The recursion continues to leverage the rich structural information of the subsequent dense layers until the quality of embedding saturates. Empirically, two to three refinement iterations suffice to expose long-range temporal transitions and subtle spatial interactions that a single embedding pass cannot easily uncover. As shown in Fig. 3a–c and Fig. 4b,g, this strategy markedly improves the identification of complex stage transitions and inter-region interactions.

(2) Stage 1: initial embedding

BCNE initially trains a deep network using a KL divergence objective similar to parametric t-SNE. This process yields a mapping from the temporospatial preprocessed HD data to an LD space, primarily preserving local neighborhoods. Analogous to standard t-SNE, the pairwise similarities (p_{ij}, q_{ij}) emphasize near neighbors, potentially overlooking broader 'global' structures—a particular concern in rich neural data where long-range interactions or transitions can be critical for understanding distinct cognitive or behavioral stages. Specifically, this process involves minimizing the KL divergence between two probability distributions, P and Q, where P corresponds to x in the HD space, and Q represents the LD space counterpart, η. Define the BCNE loss

$$L_{\text{BCNE}} = \text{KL}(P||Q) = \sum_{i\neq j} p_{ij} \log \frac{p_{ij}}{q_{ji}}, \tag{11}$$

where p_{ij} and q_{ij} represent normalized pairwise similarities: p_{ij} is the likelihood that a data point x_i in the HD space would select x_j as its neighbor based on the Gaussian distribution. Conversely, q_{ij} is the likelihood that a data point η_i in the LD space would choose η_j as its neighbor based on the t distribution. The calculation of p_{ij} and q_{ij} is executed as follows:

$$p_{j|i} = \frac{\exp\left(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/(2\sigma_i^2)\right)}{\sum_{k\neq i}\exp\left(-\|\mathbf{x}_i - \mathbf{x}_k\|^2/(2\sigma_i^2)\right)} \tag{12}$$

$$p_{ij} = \frac{P_{j|i} + P_{i|j}}{2n} \tag{13}$$

$$q_{ij} = \frac{\left(1 + \|\boldsymbol{\eta}_i - \boldsymbol{\eta}_j\|^2/\sigma\right)^{-(\alpha+1)/2}}{\sum_{k\neq i}\left(1 + \|\boldsymbol{\eta}_i - \boldsymbol{\eta}_k\|^2/\sigma\right)^{-(\alpha+1)/2}}, \tag{14}$$

where σ represents the Gaussian kernel and α is the degrees of freedom for the t distribution, defined as $\alpha = O - 1$, where O represents the desired output dimensionality.

(3) Stage $r \geq 1$: progressive global refinement

(i) Deeper latent representations

After the network is trained, each input $\mathbf{x}_i$ is passed through the intermediate layers—the deeper part of the BCNE architecture—to extract latent features, denoted by $f^r(\mathbf{x}_i) \in \mathbb{R}^{d_r}$, where d_r is the dimensionality of the selected layer under recursion stage r. These latent representations capture abstract patterns, transforming local elements into more cohesive, higher-level structures.

(ii) Constructing an updated HD similarity

We treat set $\{f^r(\mathbf{x}_i)\}$ as a new HD space and recompute conditional neighbor affinities using a Gaussian kernel. Specifically, for each point i and candidate neighbor j, we define

$$\tilde{p}^{\,r}_{j|i} = \frac{\exp\left(-\|f^r(\mathbf{x}_i) - f^r(\mathbf{x}_j)\|^2/(2\sigma_i^2)\right)}{\sum_{k\neq i}\exp\left(-\|f^r(\mathbf{x}_i) - f^r(\mathbf{x}_k)\|^2/(2\sigma_i^2)\right)}. \tag{15}$$

where σ_i is the bandwidth chosen (as in the initial stage) to match a specified perplexity for point i. The joint affinities are then symmetrized by $\tilde{p}^r_{ij} = \frac{\tilde{p}^r_{j|i} + \tilde{p}^r_{i|j}}{2n}$.

(iii) Refining the embedding with updated similarities

Let $\tilde{P}^r = [\tilde{p}^r_{ij}]$ denote the updated HD similarity matrix. We keep $Q = [q_{ij}]$ as the LD Student's t-distribution affinities on the current embedding coordinates $\boldsymbol{\eta}_i$. We then fine-tune the same BCNE network by minimazing L'_{BCNE}:

$$L'_{\mathrm{BCNE}} = \mathrm{KL}(\tilde{P}^r || Q) = \sum_{i \neq j} \tilde{p}^r_{ij} \log \frac{\tilde{p}^r_{ij}}{q_{ji}}. \tag{16}$$

At every refinement stage r, the network still ingests the original image input x_i at layer 1. Layers $[1, r]$ remain fully trainable and receive gradient updates from L'_{BCNE}. Empirically, freezing layers 1 through r degraded performance. Allowing these layers to remain adaptive enables the similarity constraints to redistribute representational capacity toward features identified as informative by deeper layers, resulting in improved embedding quality and more stable convergence. Conceptually, each recursive iteration leverages increasingly abstract latent relationships learned by the network. Consequently, the final embedding seamlessly integrates detailed local neighborhoods with overarching manifold topologies.

These three steps are repeated for two to three recursions (Dense Layer 1 at Recur 1 → Dense Layer 3 at Recur 3 in our implementation). Each pass 'zooms out', weaving the freshly discovered global relationships into the existing manifold while retaining local fidelity. The final embedding unites high-resolution local neighborhoods with coherent long-range topology, exposing subtle state transitions and distributed neural interactions that a single embedding pass fails to reveal. This property is essential for dynamic brain data, where behavioral or cognitive shifts are often encoded in extended, cross-regional patterns.

Dataset

Sherlock dataset (human fMRI BOLD imaging). The Sherlock dataset[20] includes functional MRI data from 16 participants who viewed a 48-min segment from the BBC television series Sherlock. MRI data were collected on a Siemens Skyra 3-T scanner equipped with a 20-channel head coil. Functional images were acquired using a T2*-weighted echo-planar imaging sequence, and anatomical images were obtained using a T1-weighted MPRAGE sequence. The stimulus was projected onto a rear-projection screen within the magnet bore and viewed through an angled mirror. Data preprocessing followed established procedures[13] and included slice-time correction, motion correction, linear detrending, high-pass filtering with a cutoff of 140 s and spatial normalization to the Montreal Neurological Institute template using fMRI Software Library (FSL). Images were resampled to isotropic voxels (3 mm^3), z-scored and spatially smoothed with a 6-mm Gaussian kernel. ROIs were selected on the basis of previous studies[13,20], including EV, EA, HV and PMC regions, chosen for their consistent activation during audiovisual movie viewing. Voxel counts within each ROI were 307 for EV, 571 for HV, 1,018 for EA and 481 for PMC.

Rat dataset (hippocampal electrophysiology). The rat hippocampus dataset[32] detailed electrophysiological recordings from area 1 of the cornu ammonis (CA1) hippocampal region of four male Long–Evans rats, obtained using bilaterally implanted silicon probes. Each rat, during the sessions, navigated a 1.6-m linear track to receive water rewards at both ends. The count of identified putative pyramidal neurons varied from 48 to 120 per rat. We followed the data preprocessing procedure[14], where neuronal spikes were segmented into 25-ms bins. The spatial position and movement direction (either left or right) of each rat were represented by a 3D vector, incorporating a continuous position value alongside two binary indicators for the direction.

Macaque dataset (S1 electrophysiology). The macaque dataset[24] consists of electrophysiological recordings from area 2 of the somatosensory cortex (S1) in a rhesus macaque during a center-out reaching task using a manipulandum. In this task, the macaque was required to perform reaches in eight different directions. On half of the trials, the macaque actively executed center-out movements toward a specified target. The other half were 'passive' trials, during which an unexpected 2 Newton force was applied to the manipulandum in one of the eight target directions while the macaque was holding it. The data preprocessing follows reference14, which followed the procedure outlined by Pei et al.[33], covering a timeframe from −100 ms to 500 ms relative to the onset of movement. The analysis involved segmenting the data into 1-ms time bins and smoothing with a Gaussian kernel (standard deviation 40 ms).

Evaluation metrics

We applied comprehensive evaluation metrics across three datasets—Sherlock fMRI dataset, rat hippocampal electrophysiological recordings and macaque S1 electrophysiological data—to assess how effectively our embeddings capture meaningful neurodynamic structures. Formally, let $\mathbf{X} \in \mathbb{R}^{N \times T}$ denote the original brain-data matrix, where N represents the number of input features (voxels or electrodes) and T is the number of time points. Our method produces an LD embedding $\eta(t) \in \mathbb{R}^O$, where O (typically 2 or 3) is the target embedding dimension for brain-state trajectory visualization.

Sherlock dataset (human fMRI BOLD imaging).

(1) KNN scene classification:
We used a KNN classifier to evaluate how well the embeddings η distinguish brain states corresponding to different movie scenes. Let $L(t) \in \{1, \ldots, N_{sc}\}$ be the ground-truth movie-scene label at time t, where N_{sc} is the total number of predefined scenes. We use a KNN classifier, where $k \in \{1, 3, 5, 8, 10, 30\}$ denotes the number of neighbors. Let $\hat{L}(t)$ be the label predicted by the KNN classifier for time point t. Define classification accuracy (Acc) as

$$\mathrm{Acc} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{1}(\hat{L}(t) = L(t)), \tag{17}$$

where $\mathbf{1}(\bullet)$ is the indicator function, returning 1 if its argument is true and 0 otherwise

$$\mathbf{1}(\hat{L}(t) = L(t)) = \begin{cases} 1, & if \hat{L}(t) = L(t) \\ 0, & \text{otherwise.} \end{cases} \tag{18}$$

This scene-category KNN provides discrete yet comprehensive assessment, effectively capturing rapid transitions and long-range dependencies in neural activity patterns. Accuracy was determined via tenfold cross-validation; higher values indicate better separation of scene-specific brain states.

(2) Behaviorally rated event boundaries:

To evaluate alignment with behaviorally annotated events, we assessed embedding similarities relative to externally defined event boundaries. Each time point t has a behavioral event label $L(t)$, partitioning the series into discrete events. Boundaries occur at label transitions ($L(t) \neq L(t-1)$), with series endpoints included. Following refs. 13,22,23, we constructed paired time points at fixed lag d, retaining pairs straddling event boundaries. Let $\rho_{\mathrm{event}}(t, t')$ denote the embedding similarity for a pair of time points t, t' across or within event boundaries.

$$\rho_{\text{event}}(t,t') = \frac{\sum_{i=1}^{O}(\eta_i(t)-\bar{\eta}(t))(\eta_i(t')-\bar{\eta}(t'))}{\sqrt{\sum_{i=1}^{O}(\eta_i(t)-\bar{\eta}(t))^2}\sqrt{\sum_{i=1}^{O}(\eta_i(t')-\bar{\eta}(t'))^2}}, \quad (19)$$

where $\eta_i(t)$ is the ith component of the embedding vector $\boldsymbol{\eta}(t)$. Define $\bar{\eta}(t) = \frac{1}{O}\sum_i \eta_i(t)$ as the mean of the embedding components at time t. Let $\mathcal{W}$ denote the set of all within-event time point pairs and $\mathcal{B}$ the set of all across-boundary pairs; $|\cdot|$ denotes set cardinality. We then define

$$W_{\text{avg}} = \frac{1}{|\mathcal{W}|}\sum_{(t,t')\in\mathcal{W}}\rho_{\text{event}}(t,t'), B_{\text{avg}} = \frac{1}{|\mathcal{B}|}\sum_{(t,t')\in\mathcal{B}}\rho_{\text{event}}(t,t'). \quad (20)$$

The behavioral difference $D_{\text{event}} = W_{\text{avg}} - B_{\text{avg}}$ quantifies how distinctly the embedding reflects behaviorally defined transitions, with larger values indicating clearer alignment. It is important to note that our reported event-segmentation results serve as a sanity check to evaluate whether the visualizable trajectories naturally align with behavioral boundaries, rather than to maximize downstream scores. By contrast, T-PHATE[13] uses higher-dimensional embeddings for this task, which probably explains the differences in observed outcomes.

Rat dataset (hippocampal electrophysiology).

(1) Learning stage classification:
Each time point t was labeled with learning stage $L_{\text{stage}}(t)$ based on water-reward thresholds (6, 10, 12, 13, 14 or 18 rewards). To evaluate the robustness of the embeddings in capturing the specific behavioral transitions, we evaluated embedding quality by KNN classification accuracy (as defined previously) for $k \in \{1,3,5,8,10,30\}$), using tenfold cross-validation under each threshold.

(2) Moving direction classification:
We labeled each time point by movement direction and computed KNN accuracy
on $\boldsymbol{\eta}(t)$ similarly to assess directional encoding in the embedding.

(3) Position–embedding representational similarity:
This metric evaluates how well the spatial relationships among neural positions are preserved in the embedding. Let t_i and t_j be time points with original spatial coordinates $\mathbf{p}(t)$ and embeddings $\boldsymbol{\eta}(t)$. Define distance matrices

$$D_p(t_i,t_j) = \left\|\mathbf{p}(t_i) - \mathbf{p}(t_j)\right\| \quad (21)$$

$$D_\eta(t_i,t_j) = \left\|\boldsymbol{\eta}(t_i) - \boldsymbol{\eta}(t_j)\right\|. \quad (22)$$

Extracting the upper-triangle entries ($i<j$), the ρ_{RSA} is defined as

$$\rho_{\text{RSA}} = \frac{\sum_{i<j}(D_p(t_i,t_j)-\bar{D}_p)(D_\eta(t_i,t_j)-\bar{D}_\eta)}{\sqrt{\sum_{i<j}(D_p(t_i,t_j)-\bar{D}_p)^2}\sqrt{\sum_{i<j}(D_\eta(t_i,t_j)-\bar{D}_\eta)^2}}, \quad (23)$$

where $\bar{D}_p$ and $\bar{D}_\eta$ are the mean values of the all upper-triangle entries of D_p and D_η. A higher ρ_{RSA} indicates faithful spatial preservation.

(4) Trustworthiness:
This metric measures preservation of local neighborhoods in the embedding. For neighborhood size k, let $N^{(k)}(t_i)$ be the set of the k nearest neighbors of time point t_i in the original space, and let $M^{(k)}(t_i)$ be the analogous set in the embedding. We define the set of spurious neighbors

$$U_i^{(k)} = M^{(k)}(t_i)\backslash N^{(k)}(t_i). \quad (24)$$

Let $r(t_i,t_j)$ be the rank of time point t_j sorted by distance from t_i in the original space. Trustworthiness is defined as

$$T(k) = 1 - \frac{2}{Nk(2N-3k-1)}\sum_{i=1}^{N}\sum_{t_j\in U_i^{(k)}}(r(t_i,t_j)-k). \quad (25)$$

Here, the trustworthiness $T(k)$ penalizes points that appear close in the embedding but are not truly neighbors in the original space. Higher $T(k)$ values indicate better local structure preservation.

(5) Continuity:

This metric quantifies how well the original temporal ordering is preserved. The set of missing neighbors (neighbors present in the original space but absent in the embedding) is

$$U'^{(k)}_i = N^{(k)}(t_i)\backslash M^{(k)}(t_i). \quad (26)$$

Let $r'(t_i,t_j)$ be the rank of time point t_j sorted by distance from t_i in the embedding space. Continuity is defined as

$$C(k) = 1 - \frac{2}{Nk(2N-3k-1)}\sum_{i=1}^{N}\sum_{t_j\in U'^{(k)}_i}(r'(t_i,t_j)-k). \quad (27)$$

Here, $C(k)\in[0,1]$, with larger values indicating stronger temporal-sequence fidelity.

Macaque dataset (S1 electrophysiology).

(1) Arm-angle classification:
Each t is labeled $L_{\text{angle}}(t) \in \{0^\circ, 45^\circ, \ldots, 315^\circ\}$. KNN accuracy (as defined previously) on $\boldsymbol{\eta}(t)$ measures angle encoding.

(2) Position–embedding representational similarity:

For each angle i, compute Pearson correlation ρ_i between D_p and D_η restricted to time points at angle i, analogous to the RSA defined above.

Roll-shift test. To assess temporal sensitivity, let δ denote the circular time-shift offset. We circularly shift the spatial-coordinate series $\mathbf{p}(t)$ by δ, then recompute $\rho_{\text{RSA}}(\delta)$ (the RSA correlation after shifting by δ) and ρ_i (the RSA correlation restricted to time points at reach-direction angle i). A peak in $\rho_{\text{RSA}}(\delta)$ at $\delta = 0$ with a monotonic decrease for $|\delta| > 0$ indicates correct temporal encoding. Applied to both rat and macaque datasets, this roll-shift test complements standard RSA to verify spatiotemporal fidelity.

Compared methods

In our evaluation, we considered six methods for comparison: PCA, t-SNE, UMAP, PHATE, T-PHATE and CEBRA. PCA is a widely adopted dimensionality reduction technique frequently used as a preprocessing step. t-SNE and UMAP are established manifold learning techniques, commonly applied for visualizing HD data. PHATE, a more recent method, leverages Markov diffusion to capture and visualize complex trajectories, particularly in genomic cell differentiation. T-PHATE builds upon PHATE by incorporating atemporal correlations, rendering it particularly well suited for brain imaging data such as fMRI. CEBRA is a deep learning-based embedding framework specifically designed for brain data analysis, offering multiple operational modes. For the comparison, we used the CEBRA-Time mode, an unsupervised, data-driven approach that aligns with the goal of BCNE in uncovering intrinsic neural structures without relying on predefined labels. All reported results for CEBRA in this study were obtained using the CEBRA-Time mode.

Parameter tuning. In the experiments, we used the default settings for the PCA reimplementation, as the results were not notably affected.

For the other methods, we conducted grid searches to optimize key hyperparameters for each approach.

(1) t-SNE: Perplexity∈{5, 10, 20, 30, 40, 50}, early exaggeration∈{12, 18, 24, 32}.
(2) UMAP: Number of neighbors∈{5, 12, 24, 48, 100, 200}, min_dist∈{0.0001, 0.001, 0.01, 0.1, 0.3, 0.5, 0.99}.
(3) PHATE: n_landmark∈{500, 1,000, 2,000, $N_{\text{temporal\_dim}}$}, t∈{1, 3, 5, auto}, knn∈{5, 10}, decay∈{20, 40}.
(4) T-PHATE: n_landmark∈{500, 1,000, 2,000, $N_{\text{temporal\_dim}}$}, t∈{1, 3, 5, auto}, knn∈{5, 10}, decay∈{20, 40}.
(5) CEBRA: Batch size∈{512, 1,024, 2,048, 4,096}, learning rate∈{1×10^{-4}, 3×10^{-4}}, temperature∈{0.5, 1, 1.5, 2, 3}, distance∈{Cosine, Euclidean}.

Here, $N_{\text{temporal\_dim}}$ denotes the temporal dimensionality of the input brain data. We performed these grid searches on a randomly selected participant from each dataset and then applied the selected hyperparameters to all individual participants within the same dataset, following the same procedure as the BCNE implementation.

Final parameter selection. We selected the optimal hyperparameters for each method and dataset by maximizing average scores across multiple runs. Below are the final configurations:

(1) Sherlock dataset
t-SNE: perplexity = 30, early_exaggeration = 12. UMAP: min_dist = 0.1, n_neighbors = 5. PHATE: n_landmark = $N_{\text{temporal\_dim}}$, t = 5, knn = 5, decay = 20. TPHATE: n_landmark = $N_{\text{temporal\_dim}}$, t = 5, knn = 5, decay = 20. CEBRA: batch_size = 512, distance = Euclidean, learning_rate = 3×10^{-4}, temperature = 1.0.
(2) Rat hippocampus dataset
t-SNE: perplexity = 10, early_exaggeration = 16.
UMAP: min_dist = 0.0001, n_neighbors = 24.
PHATE: n_landmark = 2,000, t = auto, knn = 10, decay = 40.
TPHATE: n_landmark = 2,000, t = 1, knn = 10, decay = 40.
CEBRA: batch_size = 512, distance = Euclidean, learning_rate = 1×10^{-4}, temperature = 2.0.
(3) Macaque dataset
t-SNE: perplexity = 30, early_exaggeration = 12.
UMAP: min_dist = 0.1, n_neighbors = 15.
PHATE: n_landmark = 1,000, t = 10, knn = 10, decay = 40.
TPHATE: n_landmark = 500, t = 10, knn = 5, decay = 20.
CEBRA: batch_size = 2,048, distance = Cosine, learning_rate = 3×10^{-4}, temperature = 1.0.

We observed that PHATE and T-PHATE were the most sensitive to hyperparameter settings across the evaluated datasets, underscoring the importance of meticulous parameter tuning for these methods. By contrast, PCA, t-SNE, UMAP, CEBRA and BCNE demonstrated relatively stable performance across different parameter configurations.

Inclusion and ethics statement. The research used deidentified, publicly available data, with no direct involvement of human participants or animal subjects. The study design and analyses aimed to avoid bias and ensure equitable representation across datasets.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Source data are provided with this paper. This study utilized three publicly available datasets: the Sherlock fMRI dataset[34] (preprocessing pipeline available via GitHub at https://github.com/KrishnaswamyLab/TPHATE), the hippocampus dataset[35] and the macaque dataset[24] (preprocessing code available via GitHub at https://github.com/AdaptiveMotorControlLab/CEBRA). All datasets are publicly accessible without restriction, and no clinical or proprietary third-party data were used in this work.

Code availability

The BCNE code, including the complete pipeline to reproduce all analyses and the preprocessed datasets used in this study, is available via Code Ocean at https://codeocean.com/capsule/3710904/tree (ref. 36). All custom scripts and implementation details supporting the findings of this work are available via GitHub at https://github.com/ZixiaZ/BCNE (ref. 37).

References

1. Logothetis, N. K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).
2. Thakor, N. V. & Tong, S. Advances in quantitative electroencephalogram analysis methods. *Annu. Rev. Biomed. Eng.* **6**, 453–495 (2004).
3. Bassett, D. S. et al. Dynamic reconfiguration of human brain networks during learning. *Proc. Natl Acad. Sci. USA* **108**, 7641–7646 (2011).
4. Poldrack, R. A. et al. Interactive memory systems in the human brain. *Nature* **414**, 546–550 (2001).
5. LeDoux, J. E. Emotion circuits in the brain. *Annu. Rev. Neurosci.* **23**, 155–184 (2000).
6. Mitra, P. P. & Pesaran, B. Analysis of dynamic brain imaging data. *Biophys. J.* **76**, 691–708 (1999).
7. Rao, S. M., Mayer, A. R. & Harrington, D. L. The evolution of brain activation during temporal processing. *Nat. Neurosci.* **4**, 317–323 (2001).
8. Makeig, S. et al. Dynamic brain sources of visual evoked responses. *Science* **295**, 690–694 (2002).
9. Breakspear, M. Dynamic models of large-scale brain activity. *Nat. Neurosci.* **20**, 340–352 (2017).
10. Healy, J. & McInnes, L. Uniform manifold approximation and projection. *Nat. Rev. Methods Primers* **4**, 83 (2024).
11. van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
12. Moon, K. R. et al. Visualizing structure and transitions in high-dimensional biological data. *Nat. Biotechnol.* **37**, 1482–1492 (2019).
13. Busch, E. L. et al. Multi-view manifold learning of human brain-state trajectories. *Nat. Comput. Sci.* **3**, 240–253 (2023).
14. Schneider, S., Lee, J. H. & Mathis, M. W. Learnable latent embeddings for joint behavioural and neural analysis. *Nature* **617**, 360–368 (2023).
15. Shin, H.-C. et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imaging* **35**, 1285–1298 (2016).
16. Zhou, Z., Wang, Y., Guo, Y., Qi, Y. & Yu, J. Image quality improvement of hand-held ultrasound devices with a two-stage generative adversarial network. *IEEE Trans. Biomed. Eng.* **67**, 298–311 (2020).
17. Zhou, Z. et al. Virtual multiplexed immunofluorescence staining from non-antibody-stained fluorescence imaging for gastric cancer prognosis. *EBioMedicine* **107**, 105287 (2024).
18. Islam, M. T. et al. Revealing hidden patterns in deep neural network feature space continuum via manifold learning. *Nat. Commun.* **14**, 8506 (2023).
19. Islam, M. T. & Xing, L. Cartography of genomic interactions enables deep analysis of single-cell expression data. *Nat. Commun.* **14**, 679 (2023).

20. Chen, J. et al. Shared memories reveal shared structure in neural activity across individuals. *Nat. Neurosci.* **20**, 115–125 (2017).
21. Huang, J. et al. Learning shared neural manifolds from multi-subject fMRI data. In *2022 IEEE 32nd International Workshop on Machine Learning for Signal Processing* 01–06 (IEEE, 2022); https://doi.org/10.1109/MLSP55214.2022.9943383.
22. Yates, T. S. et al. Neural event segmentation of continuous experience in human infants. *Proc. Natl Acad. Sci. USA* **119**, e2200257119 (2022).
23. Baldassano, C. et al. Discovering event structure in continuous narrative perception and memory. *Neuron* **95**, 709–721.e5 (2017).
24. Chowdhury, R. H., Glaser, J. I. & Miller, L. E. Area 2 of primary somatosensory cortex encodes kinematics of the whole arm. *Elife* **9**, e48198 (2020).
25. Alonso, I. et al. Peripersonal encoding of forelimb proprioception in the mouse somatosensory cortex. *Nat. Commun.* **14**, 1866 (2023).
26. Hanin, B. Which neural net architectures give rise to exploding and vanishing gradients? In *Advances in Neural Information Processing Systems* 582–591 (Curran Associates, Inc., 2018).
27. Yan, R., Islam, M. T. & Xing, L. Interpretable discovery of patterns in tabular data via spatially semantic topographic maps. *Nat. Biomed. Eng.* **9**, 471–482 (2025).
28. Vincent, P., Larochelle, H., Bengio, Y. & Manzagol, P.-A. Extracting and composing robust features with denoising autoencoders. In *Proc. 25th International Conference on Machine Learning* (eds McCallum, A. & Roweis, S. T.) 1096–1103 (Association for Computing Machinery, 2008); https://doi.org/10.1145/1390156.1390294
29. Asano, Y. M., Rupprecht, C. & Vedaldi, A. Self-labelling via simultaneous clustering and representation learning. In *Proc. 8th International Conference on Learning Representations (ICLR 2020)* https://openreview.net/pdf?id=Hyx-jyBFPr (2020).
30. Caron, M., Bojanowski, P., Joulin, A. & Douze, M. Deep clustering for unsupervised learning of visual features. In *Computer Vision – ECCV 2018: 15th European Conference, Proceedings, Part XIV* (eds V. Ferrari, M. Hebert, C. Sminchisescu & Y. Weiss) 139–156 (Springer, 2018); https://doi.org/10.1007/978-3-030-01264-9_9
31. Zhou, Z., Zu, X., Wang, Y., Lelieveldt, B. P. F. & Tao, Q. Deep recursive embedding for high-dimensional data. *IEEE Trans. Vis. Comput. Graph.* **28**, 1237–1248 (2022).
32. Grosmark, A. D. & Buzsáki, G. Diversity in neural firing dynamics supports both rigid and learned hippocampal sequences. *Science* **351**, 1440–1443 (2016).
33. Pei, F. et al. Neural Latents Benchmark '21: evaluating latent variable models of neural population activity. In *Advances in Neural Information Processing Systems (NeurIPS), Track on Datasets and Benchmarks* 34 (Curran Associates, Inc., 2021).
34. Chen, J. Sherlock movie watching dataset. *Princeton University* https://doi.org/10.34770/9ndy-8c50 (2016).
35. Grosmark, A. D., Long, J. & Buzsáki, G. Recordings from hippocampal area CA1, PRE, during and POST novel spatial learning. *CRCNS* https://doi.org/10.6080/K0862DC5 (2016).
36. Zhou, Z. BCNE analysis code. *Code Ocean* https://doi.org/10.24433/CO.3710904.v1 (2025).
37. Zhou, Z. BCNE analysis code. *Zenodo* https://doi.org/10.5281/zenodo.16741228 (2025).

Acknowledgements

This work was partially supported by the 2024 Stanford Human-Centered Artificial Intelligence (HAI) seed grant (L.X. and Z.Z.).

Author contributions

Z.Z., M.T.I. and L.X. conceived and designed the study; Z.Z. established the methodology pipeline and did the statistical analyses; Y.G., Q.T. and Y.W. contributed to refining the methodology; L.X. implemented quality control of data and the algorithms; Z.Z. prepared the first draft of the manuscript; Z.Z., J.L., W.E.W., R.F., S.L., Q.W., R.Y., M.T.I. and L.X. revised the manuscript; All authors contributed to manuscript preparation.

Competing interests

The authors declare no competing interests.

Additional information

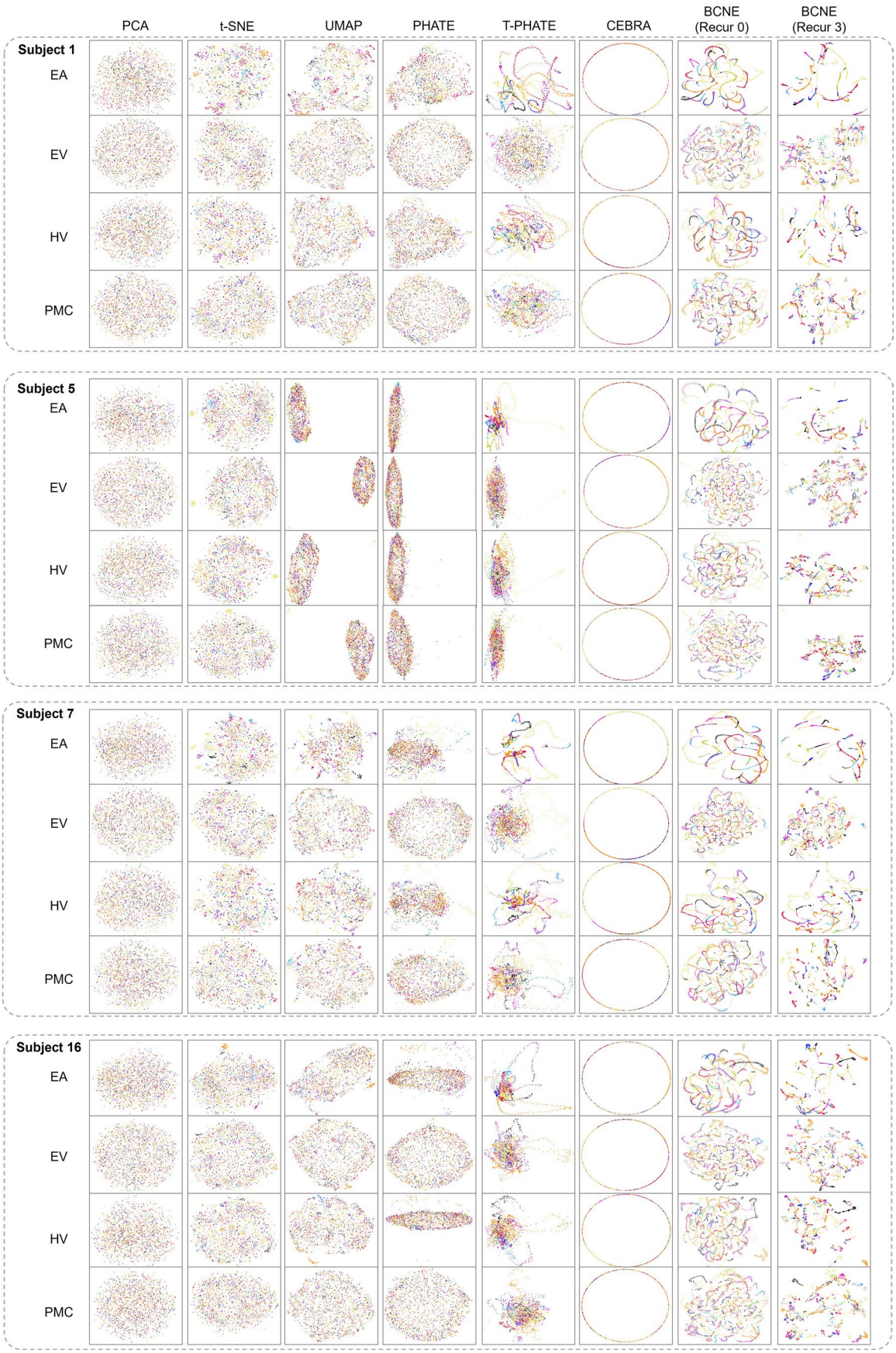
Extended data is available for this paper at https://doi.org/10.1038/s43588-025-00911-9.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43588-025-00911-9.

Correspondence and requests for materials should be addressed to Md Tauhidul Islam or Lei Xing.

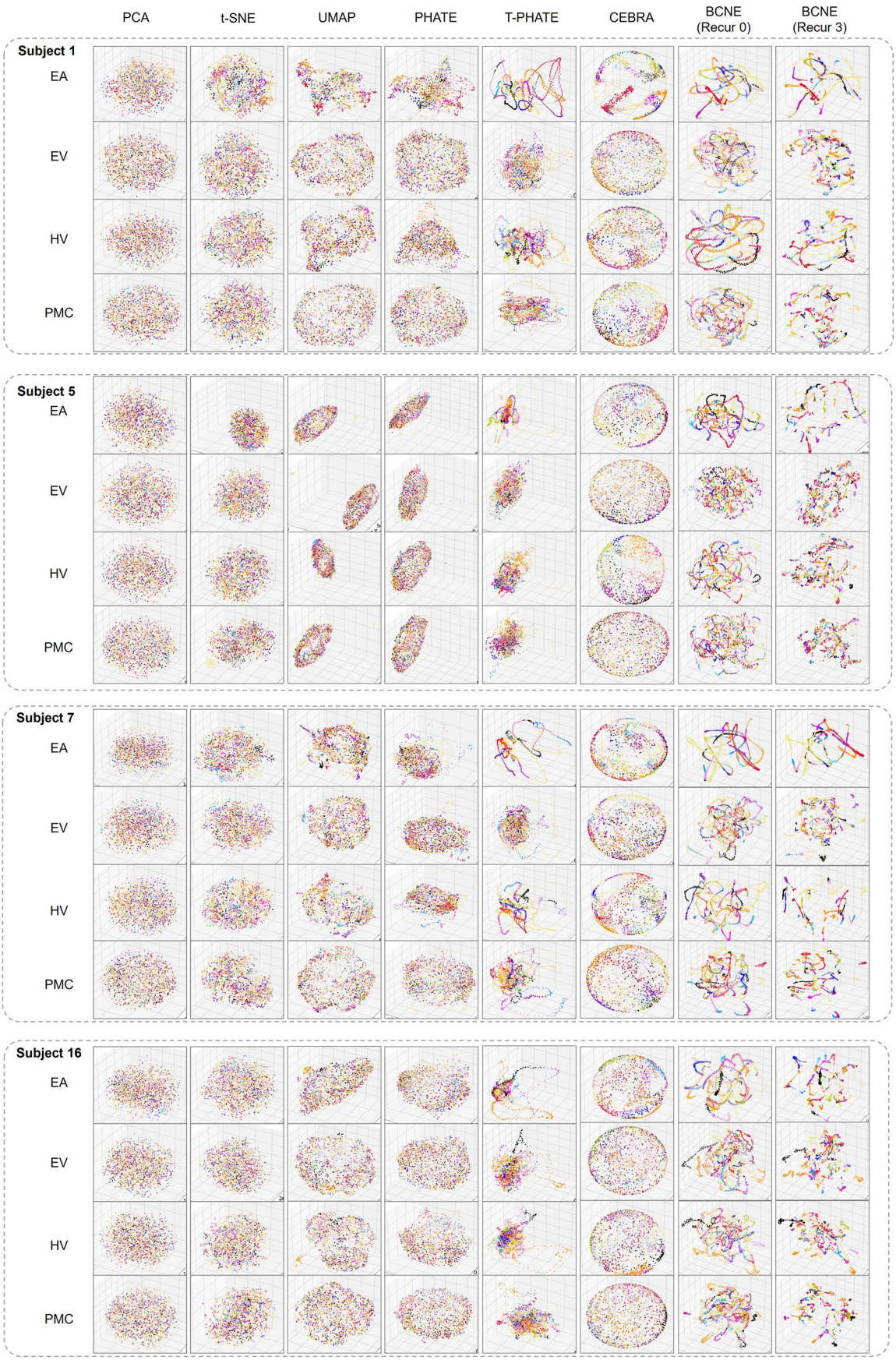
Peer review information *Nature Computational Science* thanks Erica Busch, Issam El Naqa, Mohammad Arafat Hussain and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Ananya Rastogi, in collaboration with the *Nature Computational Science* team.

https://doi.org/10.1038/s43588-025-00911-9



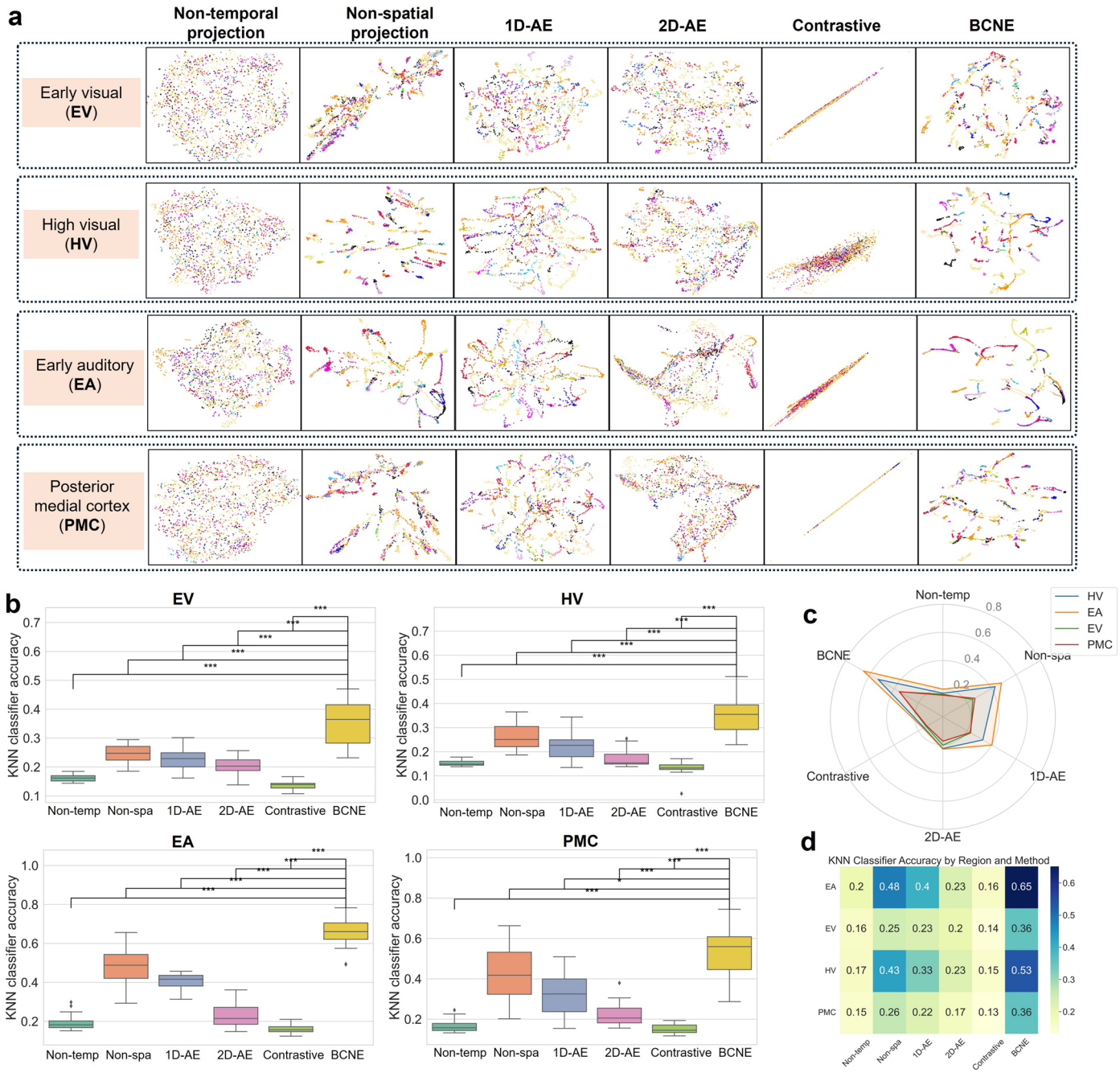
Extended Data Fig. 1 | 2D Visualizations of the Sherlock fMRI Dataset. Additional 2D visualizations generated using PCA, t-SNE, UMAP, PHATE, T-PHATE, CEBRA, and BCNE (Recur 0 and Recur 3) across four ROIs. The colormap is the same as in Fig. 2.

https://doi.org/10.1038/s43588-025-00911-9



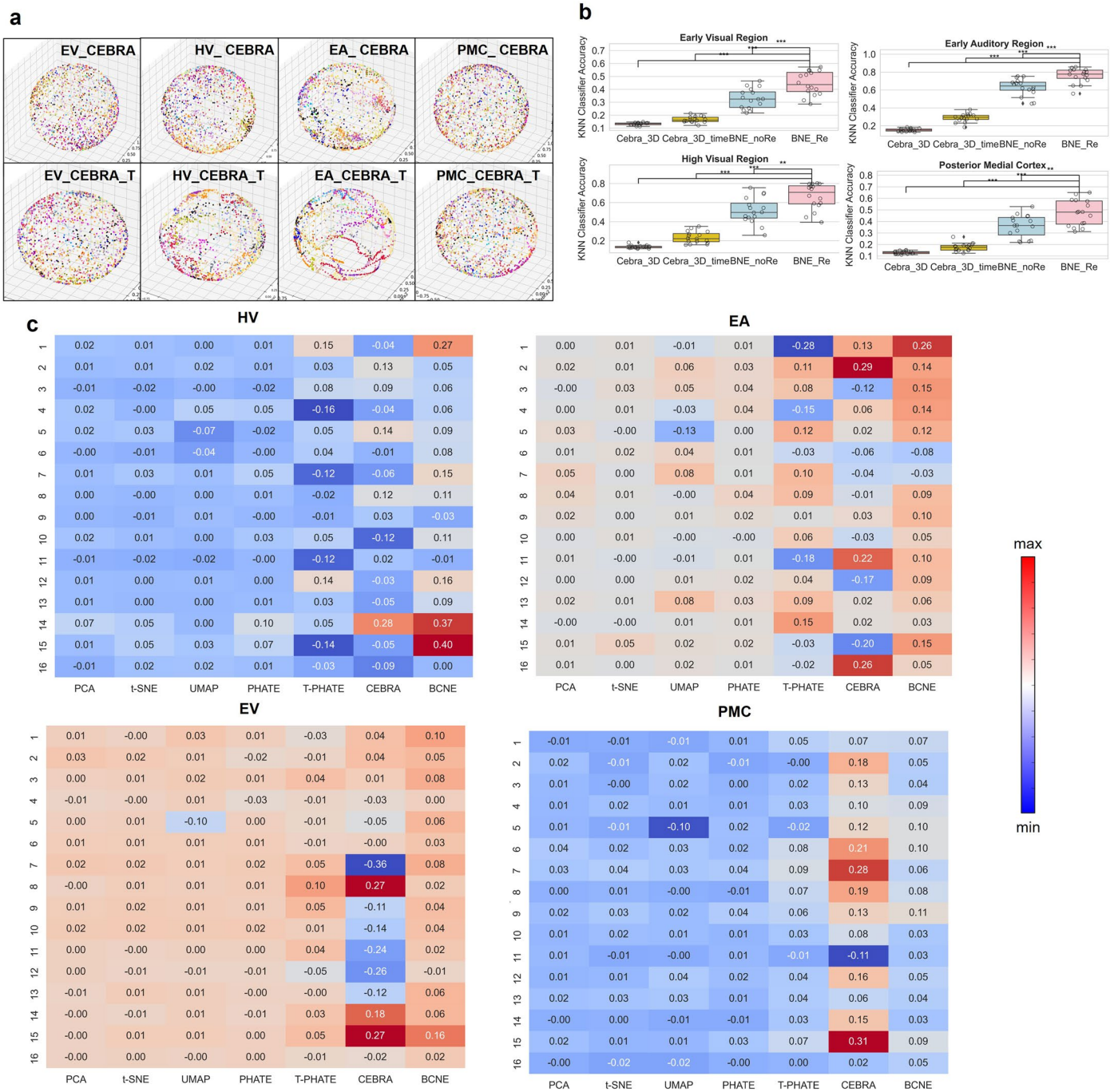
Extended Data Fig. 2 | 3D Visualizations of the Sherlock fMRI Dataset. Additional 3D visualizations generated using PCA, t-SNE, UMAP, PHATE, T-PHATE, CEBRA, and BCNE (Recur 0 and Recur 3) across four ROIs. The colormap is the same as in Fig. 2.

https://doi.org/10.1038/s43588-025-00911-9



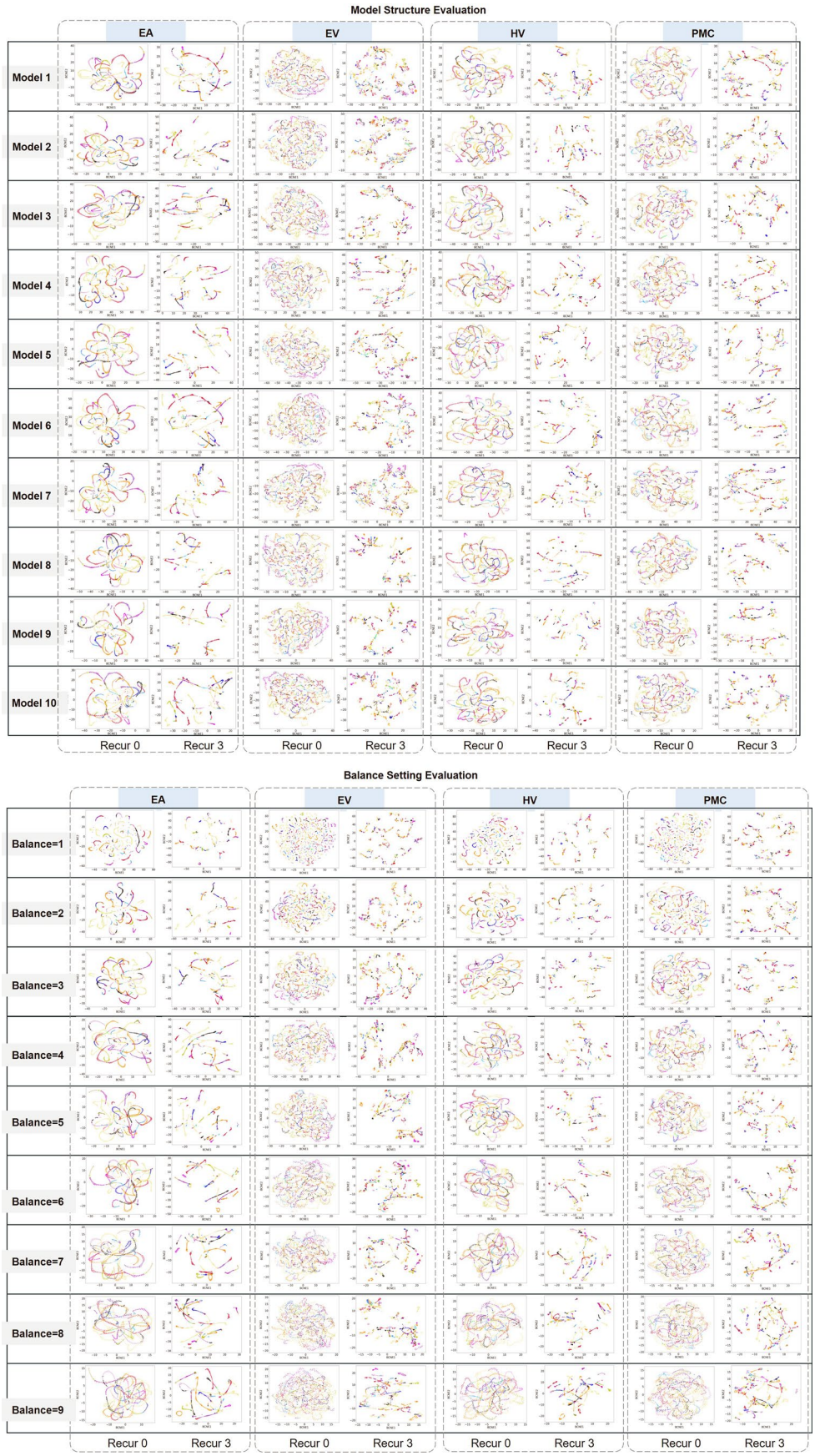
Extended Data Fig. 3 | Ablation Experiments on the Sherlock fMRI Dataset. (**a**) 2D visualizations for multiple ablation experiments, including BCNE without temporal projection, without spatial projection, with a 1D autoencoder as spatial projection, with a 2D autoencoder as spatial projection, with contrastive loss, and the proposed BCNE, across the EV, HV, EA, and PMV regions. The colormap is the same as in Fig. 2. (**b**–**d**) Boxplots, radar maps, and heatmaps displaying KNN classifier accuracy for the 39-event classification task, computed from embeddings produced by the ablation experiments across different regions. KNN accuracies were calculated using a single, randomly selected seed. The same statistical tests and box plot conventions are applied as in Fig. 2.

https://doi.org/10.1038/s43588-025-00911-9



Extended Data Fig. 4 | Comparative Visualizations and Classification Results. (**a**) 3D visualizations of embeddings generated by CEBRA (3D), CEBRA (3D with temporal correlation projection), BCNE (Recur 0), and BCNE (Recur 3). The colormap is the same as in Fig. 2. (**b**) Boxplots of KNN classifier accuracy calculated from these embeddings. The same statistical tests and box plot conventions are applied as in Fig. 2. (**c**) Heatmap displaying results of behaviorally-rated event analyses across PCA, t-SNE, UMAP, PHATE, T-PHATE, CEBRA, and BCNE for the four ROIs.

https://doi.org/10.1038/s43588-025-00911-9

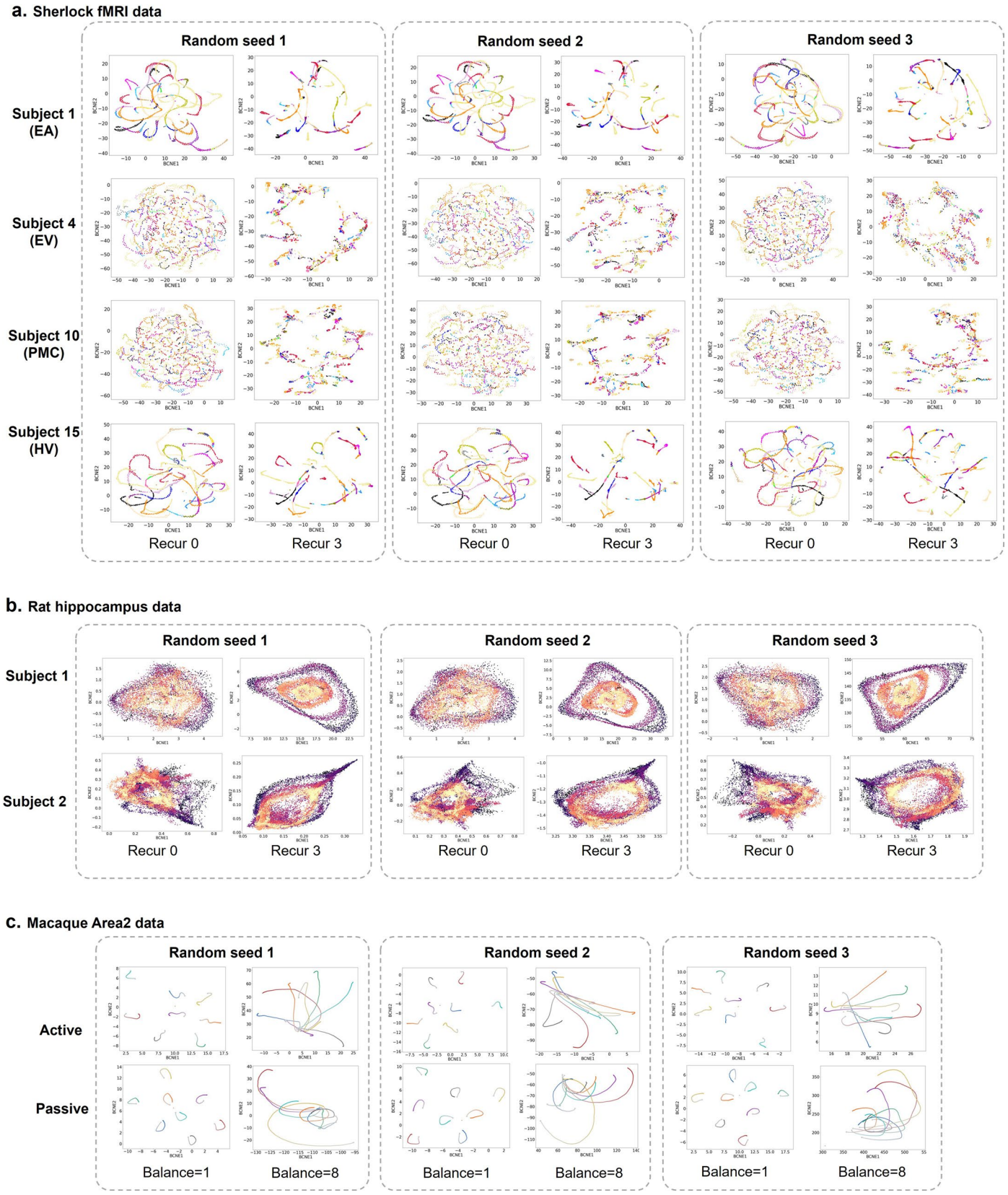


Extended Data Fig. 5 | See next page for caption.

https://doi.org/10.1038/s43588-025-00911-9

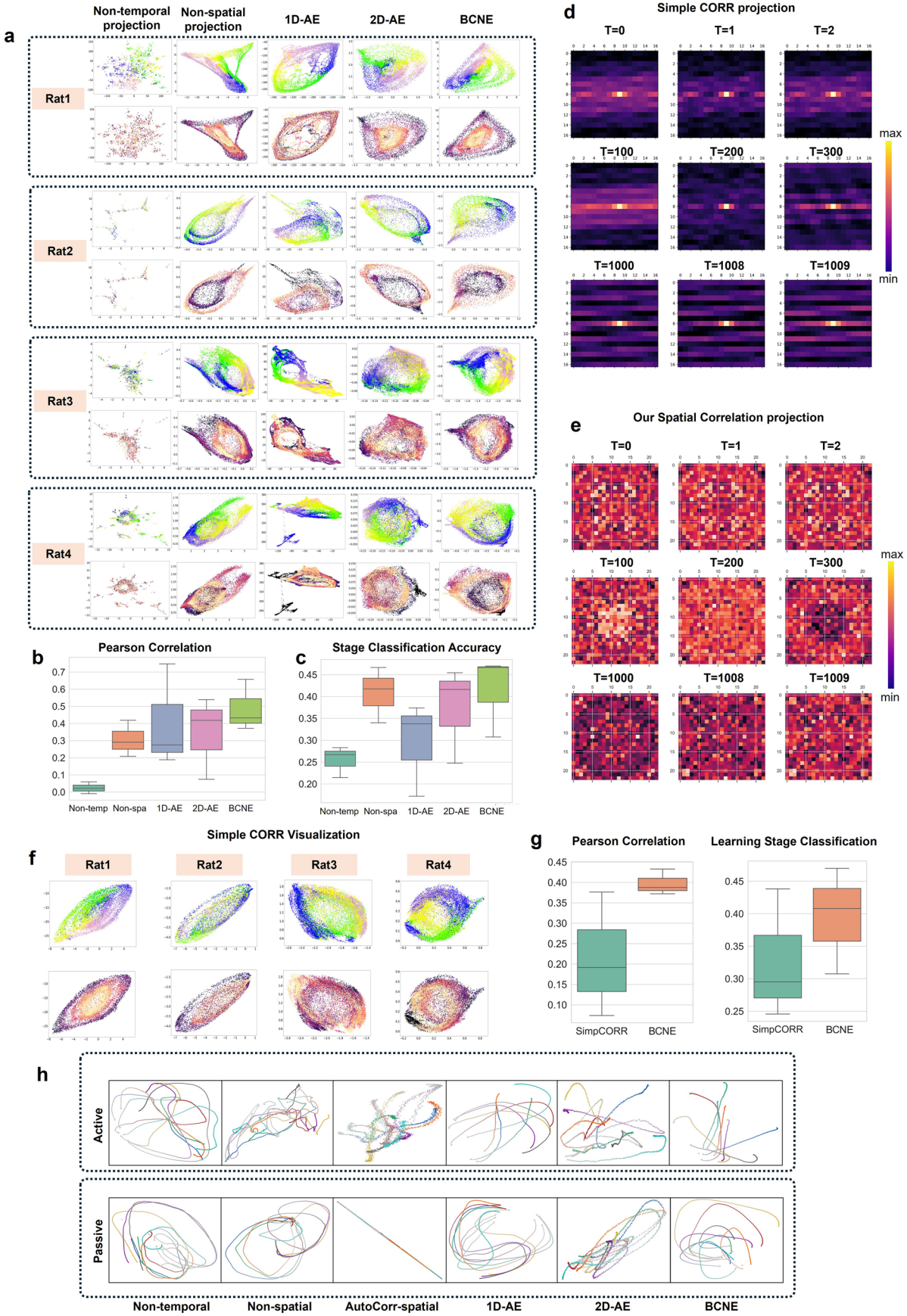
Extended Data Fig. 5 | 2D visualizations under varying BCNE configurations and balance parameters. This figure compares BCNE-generated 2D embeddings of the Sherlock dataset under different architectural configurations (top) and a range of balance-parameter settings (bottom). Each panel shows the low-dimensional trajectory for one of the four ROIs (EA, EV, HV and PMC) at recursion stages 0 and 3. Model-structure experiments evaluate the influence of alternative convolutional and dense-layer designs, while balance-parameter experiments assess the effect of varying the allocation ratio between HD- and LD-manifold components during training. Colormaps follow the same scene-label scheme as in Fig. 2.

https://doi.org/10.1038/s43588-025-00911-9



Extended Data Fig. 6 | 2D Visualizations from BCNE with Different Random Seeds. Comparison of 2D visualizations generated by the BCNE method with different random seeds across (a) the Sherlock dataset (colormap as in Fig. 2), (b) the rat dataset (colormap as in Fig. 4), and (c) the macaque dataset (colormap as in Fig. 5).

https://doi.org/10.1038/s43588-025-00911-9

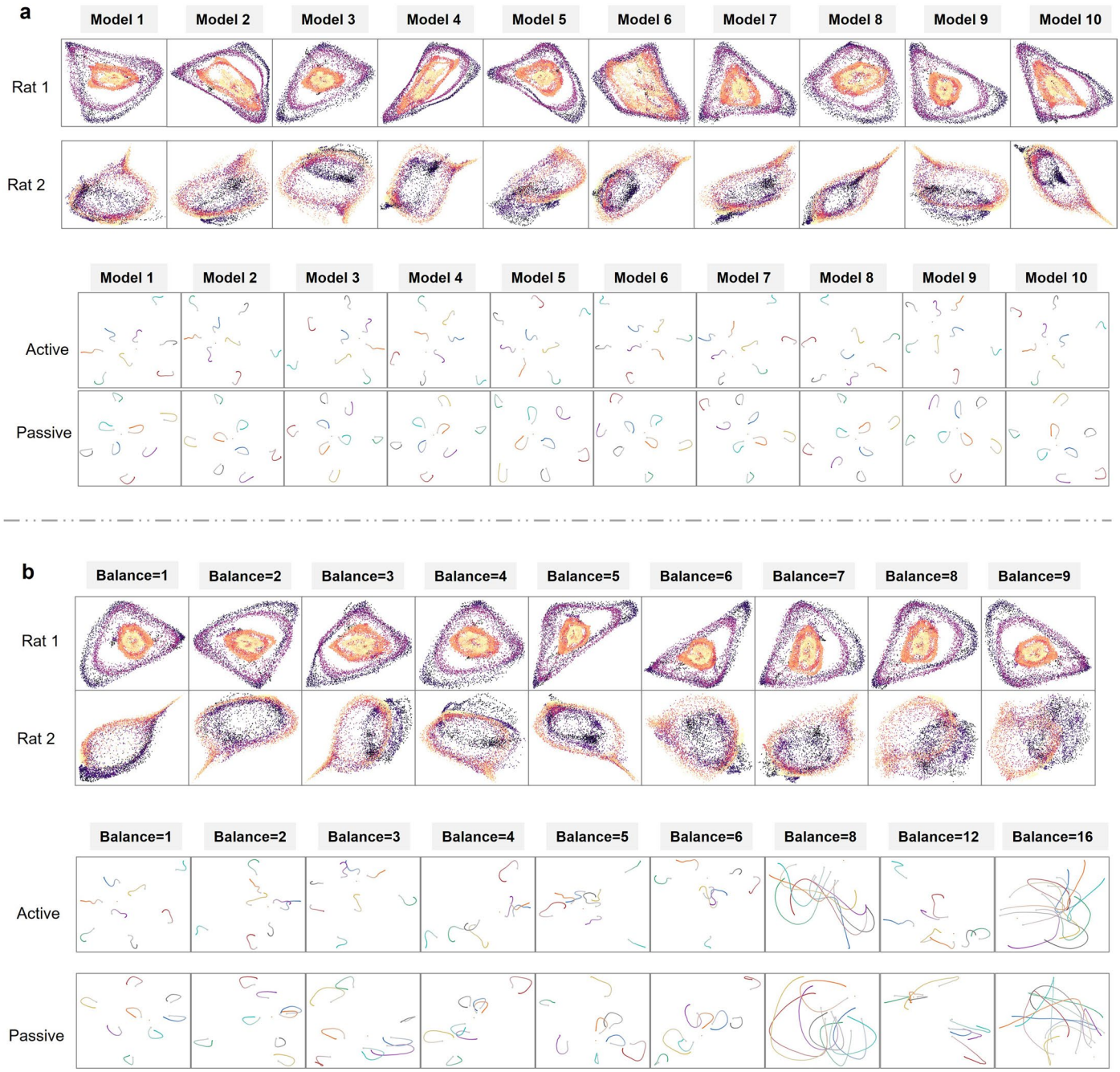


Extended Data Fig. 7 | See next page for caption.

https://doi.org/10.1038/s43588-025-00911-9

Extended Data Fig. 7 | Ablation and Spatial Projection Analyses for the Rat Hippocampus and Macaque Datasets. (**a**) 2D visualizations from ablation experiments on the Rat Hippocampus dataset, including BCNE without temporal projection, without spatial projection, with a 1D autoencoder as spatial projection, with a 2D autoencoder as spatial projection, and the proposed BCNE, shown for four rats. (**b**) Boxplot of Pearson correlation values between real rat positions and embeddings generated by ablation experiments. (**c**) Boxplot of KNN classifier accuracy for the learning stage classification task, calculated from ablation embeddings. (**d**) 2D input data after rearrangement using the compared autocorrelation-based spatial projection, displayed for selected time points (T = 0, 1, 2, 100, 200, 300, 1000, 1008, 1009). (**e**) 2D input data after rearrangement using the proposed spatial projection of BCNE, shown for the same time points. (**f**) 2D visualizations generated by BCNE with the compared spatial correlation projection for four rats. (**g**) Boxplots for Pearson correlation values and KNN classifier accuracy calculated from embeddings generated by the ablation experiments. (**h**) 2D visualizations from ablation experiments on the Macaque dataset, including BCNE without temporal projection, without spatial projection, with simple spatial correlation projection, with a 1D autoencoder as spatial projection, with a 2D autoencoder as spatial projection, and the proposed BCNE under active or passive modes. The colormap, statistical tests, and box plot conventions are the same as those used in Fig. 4.

https://doi.org/10.1038/s43588-025-00911-9



Extended Data Fig. 8 | 2D Visualizations with Varying Configurations and Balance Parameters on the Rat and Macaque Datasets. (**a**) 2D visualizations generated by the BCNE method using the default model and various alternative model configurations on the Rat and Macaque datasets. (**b**) 2D visualizations generated by the BCNE method with different "balance" parameter settings on the Rat and Macaque datasets (colormap as in Figs. 4 and 5).

nature portfolio

Corresponding author(s): Md Tauhidul Islam, Lei Xing

Last updated by author(s): Aug 5, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All public datasets used are reported in the manuscript. No special software was used to collect the data.
Data analysis	The analysis was performed using the following software and packages: Python (v3.8.2), Keras (v2.6.0), TensorFlow (v2.6.2), NumPy (v1.19.5), Pandas (v1.3.5), Scikit-learn (v1.3.2), Matplotlib (v3.5.3), scipy(v1.5.4). Open-source and custom analysis tools:TPHATE preprocessing pipeline for the Sherlock dataset: https://github.com/KrishnaswamyLab/TPHATE (commit/DOI if possible)CEBRA preprocessing for the hippocampus and macaque datasets: https://github.com/AdaptiveMotorControlLab/CEBRA (commit/DOI if possible)Custom BCNE analysis code:Code Ocean: https://doi.org/10.24433/CO.3710904.v1GitHub: https://doi.org/10.5281/zenodo.16741228All software versions and repository links are provided in the Data Availability and Code Availability sections.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Source data for Figs.2-5 is available with this manuscript.
The following publicly available datasets were used in this study:
1) Sherlock fMRI dataset34: https://doi.org/10.34770/9ndy-8c50
Preprocessing pipeline: https://github.com/KrishnaswamyLab/TPHATE.
2) Hippocampus dataset35: http://dx.doi.org/10.6080/K0862DC5.
3) Macaque dataset24: https://datadryad.org/stash/dataset/doi:10.5061/dryad.nk98sf7q7
Preprocessing code for hippocampus and macaque datasets: https://github.com/AdaptiveMotorControlLab/CEBRA.
All datasets are publicly available without restriction. No clinical or proprietary third-party data were used.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Sex and gender information was not specifically analyzed or reported in this study, as all analyses used publicly available datasets without accompanying demographic metadata.
Reporting on race, ethnicity, or other socially relevant groupings	Race, ethnicity, and other socially relevant groupings were not analyzed or reported, as such information was not available in the datasets used.
Population characteristics	All analyses were performed on publicly available datasets. Detailed population characteristics (if available) are described in the original publications and dataset documentation.
Recruitment	No new participants were recruited for this study. All data were obtained from previously published, publicly available datasets.
Ethics oversight	All original data collection procedures were approved by the corresponding institutional review boards as described in the dataset references. No new data collection or experimental procedures involving human or animal subjects were conducted for this study.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	All analyses were performed on publicly available datasets with established sample sizes: Sherlock fMRI dataset (n = 16 human subjects) Rat hippocampus dataset (n = 4 animals) Macaque S1 dataset (n = 3 animals) The full available sample in each dataset was used. These sample sizes are standard in prior studies employing the same datasets and are sufficient to demonstrate method validity and comparative performance.
Data exclusions	No data were excluded from any dataset; all provided samples/trials were analyzed.
Replication	Our analyses replicate previously published results on these benchmark datasets. All key findings were reproduced across multiple random seeds and/or cross-validation folds.
Randomization	As our study uses complete public datasets without experimental interventions, no random assignment was performed. Randomization was applied in the partitioning of data for cross-validation.

Blinding	No experimental groups or interventions were involved; analyses were performed on pre-existing data, so blinding was not applicable.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A