



HoloBrain: A Harmonic Holography for Self-organized Brain Function

Huan Liu¹, Tingting Dan¹, Zhuobin Huang¹, Defu Yang², Won Hwa Kim³,
Minjeong Kim⁴, Paul Laurienti⁵, and Guorong Wu^{1,6}✉

¹ Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill,
NC 27599, USA

grwu@med.unc.edu

² Intelligent Information Processing Laboratory and School of Automation, Hangzhou Dianzi
University, Hangzhou 310018, China

³ Computer Science and Engineering/Graduate School of AI, POSTECH,
Pohang 37673, South Korea

⁴ Department of Computer Science, University of North Carolina at Greensboro, Greensboro,
NC 27402, USA

⁵ Department of Radiology, Wake Forest School of Medicine, Winston Salem, NC 27157, USA

⁶ Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill,
NC 27599, USA

Abstract. Functional neuroimaging technology offers a new window to snapshot the transient neural activity *in-vivo*. Although tremendous efforts have been made to characterize spontaneous functional fluctuations, little attention has been paid to the functional mechanisms of neural interactions. Inspired by the notion of holography, we propose an explainable machine learning approach to establishing a novel underpinning of self-organized cross-frequency coupling (CFC) through the lens of brain wave interference on top of the network topology. Specifically, we conceptualize that the interaction between ubiquitous neural activities and a collection of reference harmonic wavelets forms a region-adaptive interference pattern that captures cross-frequency coupling of remarkable neuronal oscillations. In this regard, assembling whole-brain CFC patterns under the constraint of brain wiring mechanisms constitutes a *HoloBrain* mapping that records a wide spectrum of spontaneous neural activities. Since each local interference pattern is a symmetric and positive-definite (SPD) matrix, we tailor a deep model of *HoloBrain* (coined *DeepHoloBrain*) to infer the latent feature representations on the Riemannian manifold of SPD matrices for predicting brain states and recognizing disease connectomes. We have applied *DeepHoloBrain* to the Human Connectome Project and several dementia-related datasets. Compared with current state-of-the-art deep models, our *DeepHoloBrain* not only improves the recognition/prediction accuracy but also sheds new light on understanding the neurobiological mechanisms of brain function and cognition.

Keywords: Cross Frequency Coupling · Harmonic Wavelets · SPD Manifold · Deep Neural Network

H. Liu and T. Dan—These authors contributed equally to this work.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

A. Frangi et al. (Eds.): IPMI 2023, LNCS 13939, pp. 29–40, 2023.

https://doi.org/10.1007/978-3-031-34048-2_3

1 Introduction

Recent developments of *in-vivo* neuroimaging technology, such as functional magnetic resonance imaging (fMRI), allow us to investigate the connectivity between distinct regions in the brain, where we used to study each region separately. Since the research paradigm has been shifted to investigate region-to-region interactions, network neuroscience comes to the stage, which conceptualizes the brain as a connectome—an interactive network map where distinct brain regions synchronize their neuroactivities via myriads of interconnecting nerve fibers [1]. Network neuroscience provides a simple yet elegant system-level approach to understanding how neural circuits support brain function, how they constrain each other, and how they differ across individuals, which sheds new light on gaining insight into cognitive science and uncovering system-level principles of disease mechanisms.

The presumption of functional connectivity (FC) is a statistical dependency between the time series measured neurophysiological signals, which allows us to study whole-brain functional connections as a complex wiring system of the brain network [2]. In this context, graph theory is introduced to describe the global and local topological properties [3]. Recently, the research interest has shifted to dynamic functional connectivities, where the network topology changes over time, even in the resting state [4]. Despite great success in understanding cognition and behavior from a network neuroscience perspective, the neurobiological mechanism of functional co-activation is still largely unknown.

Mounting evidence in neuroscience shows that neuronal oscillations (aka. Brain waves), presenting across a broad frequency spectrum, might emerge remarkable rhythmic changes that support the functional mechanism of the interaction between different frequency bands, a phenomenon termed cross-frequency coupling (CFC) [5]. CFC has been reported in many cortical and subcortical regions in multiple species, yielding distinct signatures in neural dynamics [6]. In computer vision, holography is a stereo-imaging technique that generates a hologram by superimposing a reference beam on the wavefront of interest [7]. The resulting hologram is an interference pattern that can be recorded on a physical medium. Taking together, we present a proof-of-concept approach, called *HoloBrain*, to computationally “record” the cross-frequency couplings of time-evolving interference patterns that are formed by superimposing the harmonic wavelets (with predefined oscillation frequencies) on the subject-specific neural activities.

From the data structure perspective, *HoloBrain* is a graph consisting of nodes (corresponding to brain regions) and edges (weighted by functional co-activations). In contrast to conventional graph embedding vectors, *HoloBrain* uses a symmetric and positive-definite (SPD) matrix to encode all possible CFCs at each brain region (node). In this context, a new machine-learning challenge arises, that is, how to find the most relevant and explainable CFC feature representations throughout the brain network for predicting brain states and recognizing disease brain connectomes. To address this challenge, we formulate the machine-learning problem as a message-exchanging process on a compact subgraph, where we progressively refine the CFC feature representations via the learned random walks on the subgraph. Since the CFC holds the unique data geometry,

we further present a manifold-based deep model (coined *DeepHoloBrain*) to infer the explainable CFC features on the Riemannian manifold of SPD matrices.

We have elucidated the neuroscience insight of *HoloBrain* via a set of group comparison studies. Meanwhile, we have evaluated the clinical value of *DeepHoloBrain* in recognizing brain states on Human Connectome Project (HCP) data and diagnosing Alzheimer’s disease (AD), obsessive-compulsive disorder (OCD), and schizophrenia (SZ). Compared to the current “black-box deep” models, our *DeepHoloBrain* not only improves the prediction/recognition accuracy using functional neuroimaging data but also offers an explainable solution with great biophysics and neuroscience insight.

2 Methods

Suppose we partition a brain into N distinct regions. At each region, we observe the mean time course of BOLD (blood-oxygen-level-dependent) signal $x_i(t) \in \mathcal{R}$ ($i = 1, \dots, N$) where t is a continuous variable of time. Thus, we form a discretized time-varying signal matrix $\mathbf{X}(t) = [x_t]_{t=1}^T \in \mathcal{R}^{N \times T}$, where $x_t \in \mathcal{R}^N$ is a column vector of whole-brain BOLD signal snapshot at time t . In this context, we can represent the brain network as a graph $\mathcal{G} = (\mathbf{V}, \mathbf{W})$ with N nodes (brain regions) $\mathbf{V} = \{v_i | i = 1, \dots, N\}$ and the adjacency matrix $\mathbf{W} = [w_{ij}]_{i,j=1}^N \in \mathcal{R}^{N \times N}$ describing FC strength (measured by Pearson’s correlation). Let $\mathbf{L} = \mathbf{D} - \mathbf{W}$ as the underlying Laplacian matrix, where \mathbf{D} is a diagonal matrix of the total connectivity degree at each node. In the following, we first introduce the neuroscience insight and physics principles of our new model *HoloBrain* for studying brain functions. After that, we present an explainable deep model for *HoloBrain* using graph neural network and Riemannian manifold techniques.

2.1 *HoloBrain*: A Physics-Informed Model for Brain Function Analysis

Our work is inspired by frequency-specific harmonic waves [8], where the oscillation patterns of harmonic waves act as the basis functions to express the neural activities in the graph spectrum domain (constrained by brain network topology). By applying SVD on graph Laplacian matrix \mathbf{L} , the top K eigenvectors $\boldsymbol{\phi} = [\phi_k]_{k=1}^K$ (corresponding to the first K smallest eigenvalues) are used as subject-specific harmonic waves.

Harmonic Wavelets. Furthermore, we extend the harmonic technique to the regime of harmonic wavelets where the region-adaptive oscillation patterns allow us to characterize localized neural activities. To do so, we follow the approximation method in [9] to construct the harmonic wavelets for each node v_i in two steps. First, we generate a subgraph mask $u_i \in \mathcal{R}^N$ where all the nodes are connected to the underlying node v_i within h hops. Thus, u_i is an index vector where $u_i(j) = 1$ denotes region v_j is in the subgraph centered at v_i and $u_i(j) = 0$ otherwise. Second, we estimate a collection of harmonic wavelets $\boldsymbol{\psi}_i = [\psi_i^s]_{s=1}^S$ across frequency s by:

$$\min_{\boldsymbol{\psi}_i} \text{tr}(\boldsymbol{\psi}_i^T \mathbf{L} \boldsymbol{\psi}_i) + \text{tr}(\boldsymbol{\psi}_i^T \text{diag}(1 - u_i) \boldsymbol{\psi}_i) + \text{tr}(\boldsymbol{\psi}_i^T \boldsymbol{\phi}^T \boldsymbol{\phi} \boldsymbol{\psi}_i), \text{ s.t., } \boldsymbol{\psi}_i^T \boldsymbol{\psi}_i = \mathbf{I}. \quad (1)$$

Minimizing the first trace norm with the orthogonal constraint leads deterministic solution of harmonic waves in [8]. Since we sought to localize each harmonic wavelet

within a subgraph, the second trace norm is used to encourage $\|\psi_i\|^2$ to be zero out of the subgraph. In order to achieve complementary basis functions, we introduce the third trace norm to make the harmonic wavelets ψ_i orthogonal to the global harmonic wave ϕ . Due to the linearity of the trace norms, we unify three trace terms in Eq. 1 into one trace term with a matrix $\mathbf{\Pi}_i = \mathbf{L} + \text{diag}(1 - u_i) + \phi^T \phi$. Thus, the optimization of Eq. 1 is boiled down to the eigendecomposition of the matrix $\mathbf{\Pi}_i$.

Examples of harmonic wavelets are shown at the top-left of Fig. 1. It is clear that each harmonic wavelet captures the oscillation patterns (indicated by colored arrows, blue for positive and red for negative values) within a local network topology centered at the underlying brain region, where the waves oscillate slowly in the low-frequency and rapidly in the high-frequency band.

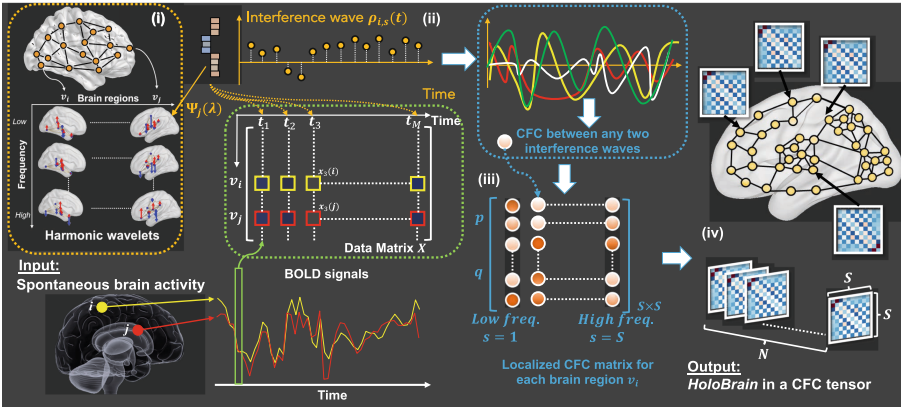


Fig. 1. The workflow of constructing *HoloBrain* for each subject, which consists of four major steps: (i) estimate harmonic wavelets, (ii) calculate interference waves, (iii) construct local CFC matrix, and (iv) assemble localized CFC matrices into a tensor, yielding *HoloBrain*.

Construction of *HoloBrain*. As shown in the middle of Fig. 1, we regard the snapshot of the whole-brain BOLD signal x_t as a time-dependent wave of subject-specific neural activities. Likewise, each harmonic wavelet ψ_i^s is considered as a predefined reference wave corresponding to a neuronal oscillation frequency. In this context, we define an interference wave $\rho_{i,s}(t) = (x_t)^T \cdot \psi_i^s$, which quantifies the dynamic engagement effect of superimposing the reference harmonic wavelets and the subject-specific neural activities.

Following the notion of CFC, we can construct a local CFC matrix for each region v_i by $\mathbf{C}_i = [c_i^{pq}]_{p,q=1}^S$, where each local CFC pattern $c_i^{pq} = (\rho_{i,p})^T \cdot \rho_{i,q}$ captures the synchronization between interface waves $\rho_{i,p}$ at p^{th} frequency and $\rho_{i,q}$ at q^{th} frequency band. To that end, *HoloBrain* is defined as an assembly of CFC matrices throughout the brain as $\mathcal{H} = [\mathbf{C}_i]_{i=1}^N$, where \mathcal{H} is formulated as a $S \times S \times N$ tensor.

Insight of *HoloBrain*. From a *Neuroscience* perspective, we position each harmonic wavelet as a predefined neuronal oscillation. In this regard, a set of harmonic wavelets at the region v_i constitute a spectrum of brain rhythms with varying frequencies along the

brain cortex. Since the neuronal activity associated with stimulus processing might differ depending on its timing relative to the ongoing oscillation [5], the interaction between harmonic wavelets and BOLD signal snapshot at time t indicates the contribution of the underlying harmonic wavelet ψ_i^s to the local cortical excitability. The proposition of considering the time-evolving interaction as the frequency-specific interference pattern sets the stage for understanding the physiological mechanism of CFC in the human brain. Since the human brain is a complex system, our *HoloBrain* “records” the footprint of whole-brain neural activities in the representation of evolving CFC.

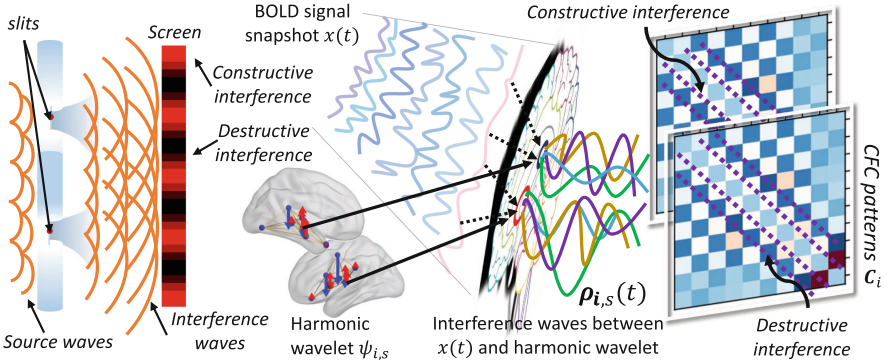


Fig. 2. The physics insight of HoloBrain (right) in analogy to the wave interference principle (left), which both yield constructive and destructive interference patterns on the screen and cross-frequency couplings, respectively.

From a *physics* point of view, the waves from two slits interfere constructively or destructively, as shown in Fig. 2 left. In our *HoloBrain*, each harmonic wavelet acts as the multi-slits where the geometry (e.g., the slit gaps) reflects its oscillation patterns. As shown in the right panel of Fig. 2, we conceptualize the snapshot of BOLD signals $x(t)$ is the source wave of neural activity. In this context, the temporal interaction between a harmonic wavelet and the source wave, i.e., $\rho_{i,s}(t)$, can be regarded as an interference wave. Since each interference wave is associated with a harmonic frequency, we superimpose the interference waves across frequencies and form the $S \times S$ region-specific CFC patterns, which hold the exact nature of interference patterns in the art of 3D hologram [7]. In analogy to the constructive and destructive interferences observed in Young’s double slit [10], we also detect a similar phenomenon along the off-diagonal lines in the $S \times S$ CFC matrix (purple dash lines in Fig. 2). Note, the first off-diagonal line (closest the diagonal line) records the CFC patterns having one shift between one pair of frequency bands. The second off-diagonal line captures two-frequency-shift CFC, etc. As a proof-of-concept approach, we sought to examine the neuroscience insight of how these constructive/destructive interferences in our *HoloBrain* underline the cognition status and the separation between healthy and disease connectomes.

From the *machine learning* perspective, the workflow in Fig. 1 is essentially a data representation process where we use the learned harmonic wavelets (as the basis functions) to express the observed BOLD signal at each time point. To that end, the resulting

CFC tensor of *HoloBrain* \mathcal{H} can be understood as an expression pattern of location and frequency associated with the evolving brain states.

2.2 *DeepHoloBrain*: An Explainable Deep Model for *HoloBrain*

Challenges. Without loss of generality, we sought to find a non-linear mapping from *HoloBrain* \mathcal{H} to an outcome label (e.g., disease connectome). There are three major challenges. *First*, the data dimensionality of CFC pattern at each brain region grows quadratic ($\mathcal{O}(S^2)$) with the number of frequency bands. Since the contribution of each brain region varies in different cognitive statuses or disease progression, a compact representation of \mathcal{H} allows us to find the most relevant features and thus enhance the prediction/recognition accuracy. *Second*, each CFC is an instance of SPD matrix. Since the data structure of CFC holds neuroscience insight, it is important to preserve such intrinsic data geometry in feature representation learning. *Third*, since CFC patterns in the *HoloBrain* are topologically wired, the machine-learning model should take network topology into account.

Problem Formulation. The input is the tensor \mathcal{H} of *HoloBrain* associated with a graph topology \mathcal{G} . We train a deep model \mathcal{M} to learn the latent low-dimension feature \mathbf{F} and a neural circuit $\mathcal{G}_{\mathbf{F}}$ (subgraph of \mathcal{G}) that explains the formation of application-specific feature representation \mathbf{F} . Due to the consideration of model explainability, we expect the output feature representation \mathbf{F} to preserve the geometric structure of $S \times S$ CFC pattern, i.e., $\mathbf{F} \in \text{Sym}_S^+$ is a data instance of co-activation patterns on the Riemannian manifold of SPD matrices Sym_S^+ . Since the conventional graph convolution plus pooling technique [11] has limitations in interpreting the neuroscience insight of learned features, we present a reinforcement learning framework to find the representative CFC pattern \mathbf{F} via a set of random walks on the graph \mathcal{G} , as described next.

Design of our Explainable Deep Model for *HoloBrain*. The overall workflow of our *DeepHoloBrain* is shown in Fig. 3. Following the spirit of the partially observed Markov decision process [12], the agent \mathcal{A} is trained to learn a stochastic policy that, at each step l , maps the history of past interactions with the environment (existing nodes $v^{(1)}, v^{(2)}, \dots, v^{(l-1)}$ in the subgraph $\mathcal{G}_{\mathbf{F}}$) to a probability distribution over actions for the current step l . At each step l , the agent performs two actions: (i) update CFC patterns in the *HoloBrain* through a re-wiring process (steering the feature representation learning), and (ii) add node $v^{(l)}$ to the subgraph $\mathcal{G}_{\mathbf{F}}$ that supports the formation of the representative CFC pattern \mathbf{F} (for model explainability). In this context of reinforcement learning, our *DeepHoloBrain* is made of the following learning components (indicated by color in Fig. 3).

- (1) **Sensor.** As explained in the ‘‘Action’’ module, the actions from the agent \mathcal{A} include the l^{th} candidate node $v^{(l)}$ and a vector of re-wiring message $m^{(l)} = \left[m_j^{(l)} \right]_{j=1}^N \in \mathcal{R}^N$. Here $v^{(l)}$ is an index function that returns the node index in the l^{th} step. Suppose $v^{(l)} = v_i$. Thus, each element $m_j^{(l)}$ indicates the interaction between C_i and other CFC pattern C_j . Furthermore, we use $\mathcal{G}_{\mathbf{F}}^{(l)} = \{v^{(1)}, v^{(2)}, \dots, v^{(l)}\}$ denote the set of

node indexes selected by the agent in the past l steps. By integrating the re-wiring message $m^{(l)}$ with the connectivity profile (i^{th} row in the adjacency matrix \mathbf{W}) at node v_i , we define the external state of the agent $\mathbf{E}^{(l)}$ by: $\mathbf{E}^{(l)} = \sum_{j=1}^N m_j^{(l)} w_{ij} \mathbf{C}_j$.

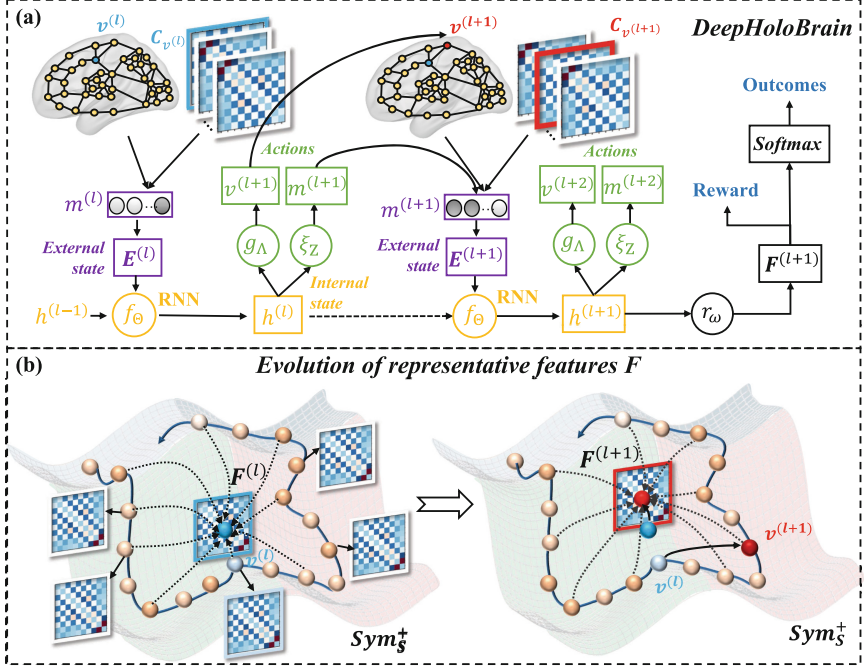


Fig. 3. (a) The overall design of *DeepHoloBrain* falls into a reinforcement learning framework, which consists of a sensor (in purple), an internal state (in brown), a set of actions (in red), and a reward function (in green). (b) The learning process of representative CFC features \mathbf{F} forms a trajectory on the manifold $\text{Sym}_{\mathcal{S}}^+$ as we sequentially walk from $v^{(l)}$ to $v^{(l+1)}$ at each step of reinforcement learning.

- (2) **Internal state.** The agent maintains an internal state \mathbf{h} that summarizes the representative CFC pattern extracted from the history of selected nodes $\mathcal{G}_F^{(l)}$. The agent \mathcal{A} perceives the evolving environment (updated representative feature $\mathbf{F}^{(l)}$) by deciding how to act (for exchanging information between CFC patterns) and where to deploy the new sensor (for steering the random walk). In this context, we train a recurrent neural network (RNN) to update the internal state by $\mathbf{h}^{(l)} = f_{\Theta}(\mathbf{h}^{(l-1)}, \mathbf{E}^{(l)})$, where f_{Θ} denotes the RNN with trainable parameters Θ .
- (3) **Actions.** At each step, the output of RNN is used to predict the node index $v^{(l+1)}$ of the next candidate node in $\mathcal{G}_F^{(l+1)}$ and a re-wiring message $m^{(l+1)}$ that contributes to the change of the environment state. Specifically, the node index $v^{(l+1)}$ is determined stochastically from a distribution parameterized by the random walk network, i.e., $v^{(l+1)} \sim p(v|g_{\Lambda}(\mathbf{h}^{(l)}))$, where g denotes a neural network with trainable parameters Λ . Similarly, the re-wiring vector is drawn from a distribution conditioned on the

output of the attention network $\xi_Z(\mathbf{h}^{(l)})$, i.e., $m^{(l+1)} \sim p(m|\xi_Z(\mathbf{h}^{(l)}))$, where Z is the attention network parameters. Based on the current hidden state $\mathbf{h}^{(l)}$, we further train a prediction network r_ω to learn the representative features $\mathbf{F}^{(l)}$, where the output is a softmax classifier for outcome prediction (e.g., recognizing disease connectomes).

- (4) **Rewards.** After executing the above actions, the agent goes to a new external state $\mathbf{E}^{(l+1)}$ and a reward $\tau^{(l+1)}$. In most applications, $\tau^{(l+1)} = 1$ suggests that the learned representative features $\mathbf{F}^{(l)}$ successfully predict the outcome label for the underlying subject after $l + 1$ steps and 0 otherwise.

Extend Our Deep Model to the Riemannian Manifold. Since multiple lines of evidence show the importance of preserving the geometry of SPD matrices in functional brain network analysis [13, 14], we extend our *DeepHoloBrain* model in Fig. 3 to the manifold-based deep model by replacing the RNN model f_Θ to SPD-SRU [15], where the operations in the SPD-SRU use the manifold algebra for SPD matrices. To that end, the hidden state $\mathbf{h}^{(l)} \in \text{Sym}_S^+$ becomes a $S \times S$ SPD matrix.

Training *DeepHoloBrain*. The driving force of *DeepHoloBrain* consists of two parts. *First*, we use cross entropy to monitor the classification error between the ground truth and predicted outcome label by the current representative CFC feature $\mathbf{F}^{(l)}$ at each step l . *Second*, since the random walk from the current brain region $v^{(l)}$ to the next spot $v^{(l+1)}$ is steered by the re-wiring message $m^{(l)}$, we evaluate the effectiveness of random walk by a composite score of $\tau^{(l)} \cdot m_{v^l}^{(l)}$, where $\tau^{(l)}$ is the reward at the current step and $m_{v^l}^{(l)}$ reflects the reliability of suggesting good candidate regions in reinforcement learning. The parameters of *DeepHoloBrain* $\{\Theta, Z, \Lambda, \omega\}$ are fine-tuned by the Adam optimizer, where the learning rate is set to 0.001.

3 Experiments

In the following experiments, we first investigate the neuroscience insight of *HoloBrain* underlying the relationship between brain function and cognitive status. After that, we evaluate the diagnostic power of *DeepHoloBrain* in recognizing disease connectomes for AD, OCD, and schizophrenia using resting-stage fMRI scans, and recognizing cognitive tasks from task-based fMRI scans in HCP.

Data Description. We evaluate our proposed *HoloBrain* and *DeepHoloBrain* on the task fMRI data from HCP and resting-stage fMRI data from three disease-related datasets. For the task fMRI in HCP, we use Yale’s functional atlas [16] which consists of 268 brain parcellations. The task includes 2-back and 0-back task conditions for body, place, face, and tool stimuli, as well as fixation periods. For the resting-state fMRI, the classic AAL atlas [17] is employed. For each fMRI scan, ICA-AROMA [18] is used to remove motion signal artifacts based on temporal and spatial features in the data related to head motion. A band-pass filter (0.009–0.08 Hz) is then applied to each scan. After spatial normalization, we compute the mean BOLD time course for each parcellated brain region. We perform experiments of cognitive normal (CN) vs. disease connectomes classification on ADNI dataset (102 CN vs. 63 AD), OCD study (63 CN vs. 61 OCD), and schizophrenia study (159 CN vs. 182 SZ). Meanwhile, we perform task recognition experiment on 264 subjects with task-fMRI data from HCP.

3.1 Understanding the Neuroscience Insight of HoloBrain

Since harmonic wavelets are widely used for signal reconstruction, we determine the number of frequency bands S by evaluating the residual error between the original signals and reconstructed signals with limited bandwidth. By doing so, we fix $S = 10$ in all the following experiments. Thus, each CFC pattern is a 10×10 SPD matrix.

We show the population average of global CFC patterns (i.e., averaging throughout brain regions and across individuals) for each clinic cohort in Fig. 4. There is a clear sign that the population-wise CFC average exhibit remarkable off-diagonal striping patterns, which resembles the constructive/destructive interference phenomena in Young’s double slit shown in Fig. 2. In light of this, we hypothesize that the consistency of CFC degree along these striping patterns might offer a new explanation how brain functions emerge diverse cognition and behaviors. Since the striping patterns are visually distinguishable between healthy and disease cohorts, we further speculate that the coherence of maintaining the striping patterns might be an indicator of how brain function is altered as the disease progresses.

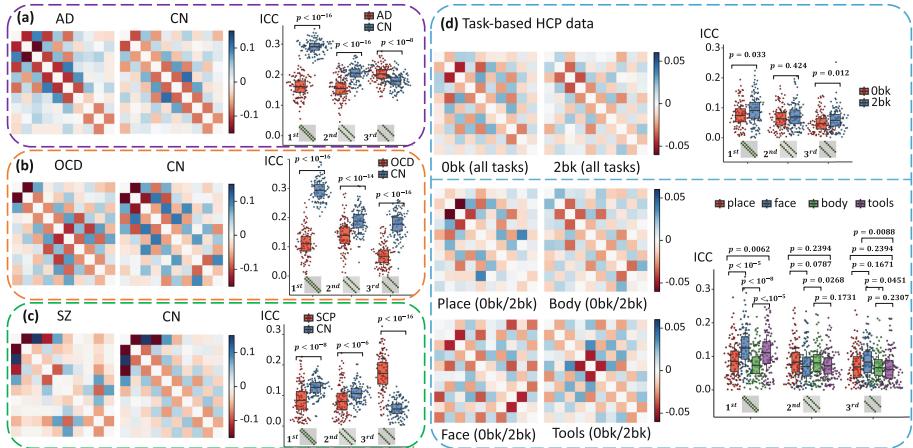


Fig. 4. The statistical power of *HoloBrain* in healthy vs. disease connectome (left) comparison and separating cognitive tasks (right). In each group comparison, we display the average CFC patterns and the quantitative measures that evaluate the discriminative power of *HoloBrain*.

Following this clue, we calculate the interclass correlation coefficient (ICC [19]) for each off-diagonal line, where a higher ICC indicates less variance between CFC degrees under consideration. In Fig. 4, we plot the distribution of ICC for the first three off-diagonal lines for each cohort. It is apparent that the ICC level manifests strong differences between healthy and disease connectomes at the significance level $p < 10^{-6}$. Although the difference between 0-bk and 2-bk tasks is not significant, the CFC patterns show significant differences ($p < 0.005$) in ICC along the first diagonal line among cognitive tasks.

3.2 Evaluating the Clinical Value of *DeepHoloBrain*

In this section, we shift the focus to predicting outcome labels for each functional connectome using the trained *DeepHoloBrain*. The benchmark methods include classic support vector machine (SVM) and graph convolution network (GCN) [11]. For SVM, we vectorize the CFC matrix at each brain region and concatenate them across brain regions. Then we train SVM to predict the outcome from the concatenated vector. For GCN, we consider the vectorized CFC matrix as the graph embedding vector. Then, we train GCN to predict the outcomes based on the topology of the functional connectivities. For all learning-based methods, we use 5-fold cross-validation.

Evaluating the Prediction Accuracy. The prediction results of classifying AD, OCD, and schizophrenia from resting-stage fMRI and recognizing cognitive tasks from task-based fMRI are summarized in Table 1. It is apparent that our *DeepHoloBrain* outperforms the other counterpart methods in terms of accuracy, sensitivity, and specificity.

Table 1. The statistics of machine learning performance in recognizing disease connectomes (top) and cognitive tasks (bottom).

Experiments	Methods	Accuracy	Sensitivity	Specificity
<i>CN vs. AD</i>	SVM	0.713 ± 0.031	0.231 ± 0.065	0.842 ± 0.039
	GCN	0.772 ± 0.026	0.430 ± 0.056	0.743 ± 0.025
	Ours	0.808 ± 0.036	0.650 ± 0.048	0.941 ± 0.036
<i>CN vs. OCD</i>	SVM	0.513 ± 0.020	0.565 ± 0.035	0.452 ± 0.049
	GCN	0.573 ± 0.011	0.560 ± 0.023	0.530 ± 0.055
	Ours	0.611 ± 0.030	0.664 ± 0.026	0.664 ± 0.041
<i>CN vs. SZ</i>	SVM	0.508 ± 0.014	0.535 ± 0.021	0.478 ± 0.021
	GCN	0.578 ± 0.021	0.563 ± 0.016	0.576 ± 0.025
	Ours	0.605 ± 0.016	0.647 ± 0.015	0.690 ± 0.017
<i>HCP 0-bk vs. 2-bk</i>	SVM	0.521 ± 0.035	0.478 ± 0.059	0.570 ± 0.066
	GCN	0.555 ± 0.016	0.244 ± 0.053	0.862 ± 0.041
	Ours	0.595 ± 0.017	0.567 ± 0.045	0.671 ± 0.047
<i>HCP Tools/Place/ Body/Face</i>	SVM	0.251 ± 0.032	0.155 ± 0.068	0.344 ± 0.078
	GCN	0.272 ± 0.041	0.166 ± 0.056	0.422 ± 0.065
	Ours	0.363 ± 0.034	0.213 ± 0.044	0.396 ± 0.059

Evaluating the Model Explainability. *First*, since our *DeepHoloBrain* model is trained to select the best combination of CFC patterns for connectome classification by reward-guided reinforcement learning, we display the most frequently-visited brain regions and frequently-selected random walks in Fig. 5, for CN vs. AD (top-left), CN vs. OCD (top-right), and CN vs. schizophrenia (bottom-left), respectively. Specifically, we use node

size and brightness to reflect the selection frequency, where high frequency indicates the importance in disease classification. Take CN vs. AD classification as an example. Most of the brain regions fall in the default mode network (e.g., precuneus and superior frontal gyrus) and frontoparietal network (e.g., middle frontal gyrus). Meanwhile, we show the population-average subgraph \mathcal{G}_F in each disease classification, which suggests that the CFC patterns along the subgraph reach the best balance between diagnostic power and dimensionality. Furthermore, we display the weighed average of CFC patterns for each clinical cohort in the bottom-right of Fig. 5. Compared to the group comparison results in Fig. 4, it is clear that the constructive and destructive patterns are not only well maintained in the off-diagonal lines in the learned CFD patterns but also greatly enhanced with respect to group-to-group differences.

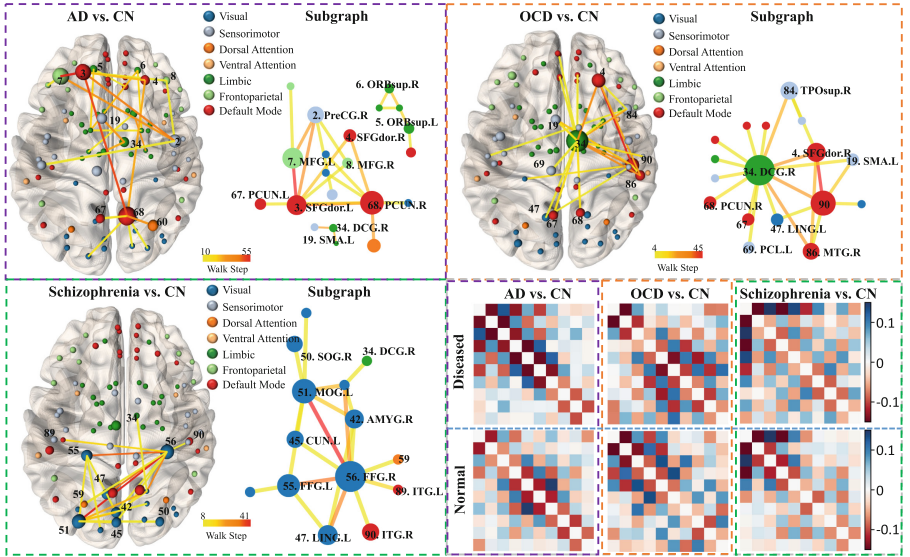


Fig. 5. The most frequently-visited brain regions and subgraph \mathcal{G}_F that are used in CN vs. AD (top-left), CN vs. OCD (top-right), and CN vs. SZ (bottom-left), where the node size and link color indicate the likelihood of being selected in *DeepHoloBrain*. Meanwhile, we display the group average of CFC patterns within the subgraph \mathcal{G}_F for different clinical cohorts in the bottom-right panel.

4 Conclusions

In this work, we propose a principled computational framework to characterize the holography of spontaneous functional connectivities in the human brain. Specifically, our *HoloBrain* technique allows us to elucidate the neuroscience insight of cross-frequency couplings which plays a mechanistic role in neuronal computation, communication, and

learning. Furthermore, we tailor a deep model to translate the *HoloBrain* to various neuroscience and clinical applications, such as predicting cognitive states and recognizing disease connectomes.

References

1. Bassett, D.S., Sporns, O.: Network neuroscience. *Nat. Neurosci.* **20**(3), 353–364 (2017)
2. Airan, R.D., et al.: Factors affecting characterization and localization of interindividual differences in functional connectivity using MRI. *Hum. Brain Mapp.* **37**(5), 1986–1997 (2016)
3. Rubinov, M., Sporns, O.: Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* **52**(3), 1059–1069 (2010)
4. Hutchison, R.M., et al.: Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage* **80**, 360–378 (2013)
5. Canolty, R.T., Knight, R.T.: The functional role of cross-frequency coupling. *Trends Cogn. Sci.* **14**(11), 506–515 (2010)
6. Hyafil, A., et al.: Neural cross-frequency coupling: connecting architectures, mechanisms, and functions. *Trends Neurosci.* **38**(11), 725–740 (2015)
7. Gabor, D.: A new microscopic principle. *Nature* **161**(4098), 777 (1948)
8. Atasoy, S., Donnelly, I., Pearson, J.: Human brain networks function in connectome-specific harmonic waves. *Nat. Commun.* **7**(1), 10340 (2016)
9. Melzi, S., et al.: Localized manifold harmonics for spectral shape analysis. *Comput. Graph. Forum* **37**(6), 20–34 (2018)
10. Young, T.: I. The Bakerian lecture. Experiments and calculations relative to physical optics. *Philos. Trans. Roy. Soc. London* **94**, 1–16 (1804)
11. Defferrard, M., Bresson, X., Vandergheynst, P.: Convolutional neural networks on graphs with fast localized spectral filtering. In: *Proceedings of the 30th International Conference on Neural Information Processing Systems*, pp. 3844–3852. Curran Associates Inc., Barcelona (2016)
12. Mnih, V., et al.: Human-level control through deep reinforcement learning. *Nature* **518**(7540), 529–533 (2015)
13. Dan, T., et al.: Learning brain dynamics of evolving manifold functional MRI data using geometric-attention neural network. *IEEE Trans. Med. Imaging* **41**(10), 2752–2763 (2022)
14. Dan, T., et al.: Uncovering shape signatures of resting-state functional connectivity by geometric deep learning on Riemannian manifold. *Hum. Brain Mapp.* **43**(13), 3970–3986 (2022)
15. Chakraborty, R., et al.: A statistical recurrent model on the manifold of symmetric positive definite matrices. In: *Neural Information Processing Systems* (2018)
16. Shen, X., et al.: Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuroimage* **82**, 403–415 (2013)
17. Tzourio-Mazoyer, N., et al.: Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**(1), 273–289 (2002)
18. Pruim, R.H.R., et al.: ICA-AROMA: a robust ICA-based strategy for removing motion artifacts from fMRI data. *Neuroimage* **112**, 267–277 (2015)
19. Koo, T.K., Li, M.Y.: A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **15**(2), 155–163 (2016)