RESEARCH ARTICLE

EEG Signal-Assisted Algebraic Topological Feature-Enhanced Deep Neural Networks for Gestalt Illusory Contour Perception

FEI NI[1], ZHIFENG ZHAO[1,2], (Member, IEEE), RONGPENG LI[1], (Member, IEEE), AND HONGGANG ZHANG[1,2], (Senior Member, IEEE)

[1]College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China
[2]Zhejiang Laboratory, Hangzhou 311121, China

Corresponding author: Rongpeng Li (lirongpeng@zju.edu.cn)

ABSTRACT Deep neural networks (DNNs) have been incredibly successful at correctly classifying various input types, including images, speech, and data, according to consumer preferences. However, the efficacy of DNNs in Gestalt classification tasks, an important case in manifesting human perceptual capability, remains challenging, as DNNs are generally unable to perceive illusory closure from Gestalt images unless DNNs are carefully calibrated using a significant amount of priori information. However, altering the input, as minute changes typically imperceptible to humans, can confound carefully calibrated DNNs. In this study, EEG signal-empowered deep clustering based on topological data analysis (TDA) is proposed. The adoption of TDA manifests the separability of EEG signals responsive to Gestalt and non-Gestalt images. It also yields new families of features and descriptors for Gestalt illusory contours by extracting topological and geometric information. Furthermore, the combination of analyzed EEG signals and digital images further benefits the recognition of Gestalt illusory contours. Extensive experiments have shown convincing improvements over the state-of-the-art DNNs (e.g., DeepCluster and multi-view clustering [MVC]). In particular, DeepCluster with TDA can perceive illusory contours to some degree, given its 66.5% classification accuracy. Nevertheless, on top of extracted topological features from EEG signals, it produces higher classification accuracy (i.e., 71.9%), indicating the features extracted from the EEG signal contribute to perceive Gestalt illusory contours. On the other hand, the applications of topological features to the latest MVC also bring 4.5% improvement and demonstrate the effectiveness.

INDEX TERMS Deep neural network, topological data analysis, EEG signal, gestalt illusory contour perception.

I. INTRODUCTION

State-of-the-art DNN methods enable computational models consisting of multiple processing layers to mimic how the human brain perceives and understands multi-modal information with different levels of abstraction. In particular, benefitting from advances in Moore's Law, DNNs can explore Gestalt grouping in relation to illusory contours and

The associate editor coordinating the review of this manuscript and approving it for publication was Michele Nappi.

thereby implicitly capture intricate structures of large-scale data within a satisfactory amount of time [1]. For example, Kim et al. train a supervised learning DNN (e.g., InceptionNet) to derive a closure measure metric, which capably computes the similarity between Gestalt images and triangles, and verify the consistency of the proposed metric [2]. Based on a dozen of video frames, Lotter et al. devise a PredNet to capably predict the next frame. Surprisingly, they find that the responses of these DNNs resemble neural activities in brains of monkeys responsive to Gestalt illusory stimuli,

but are quite different from those for non-Gestalt stimuli [3]. In other words, it demonstrates that DNNs are capable to exhibit Gestalt illusory contours. Pang et al. conduct more comprehensive experiments and validate the reproduction of similar results [4]. However, specifically-trained DNNs therein, are intrinsically task-oriented and require a large amount of prior information to train, thus lacking the widely existing deconstruction, decomposition and reasoning capabilities in biological intelligence, [5]. In other words, it is generally difficult for DNNs to break away from the characteristics of data and extract hidden characteristics. In that regard, the aforementioned works [2], [3], [4] do not consider the Gestalt classification capability of DNNs. On the other hand, although DNNs trained on non-illusory fat / thin rectangles successfully classify testing images of real and illusory contours, they do not appear to interpolate values between tangent discontinuities in the same way as human observers do [6]. Thus, it is generally conceived [6] that DNNs do not capably perceive illusory contours. Meanwhile, there are effective means to process the digital images, like denoising [7], [8], while the classification of Gestalt illusory contours is unable to be simply understood as the process of denoising. Thus, instead of directly adopting DNNs, it is meaningful to introduce bio-intelligence patterns to common DNNs, so as to discover effective means to perceive Gestalt illusory contour [9].

On the other hand, the brain is not homogeneous like our current DNN structure, but has different modules [10]. The brain's cognition of a Gestalt image is the result of the collaborative work of various encephalic regions. Nonetheless, it is difficult to simply interpret the relationship between different parts of the brain [11], [12], [13], [14]. Although recent progresses have shown that DNNs can extract some apparent features of EEG signals, such as distinguishing between normal and abnormal EEG signals [15], [16], they place too much emphasis on individual differences between healthy and sick beings. In contrast, TDA is an efficient method for inferring, analyzing and exploiting the complex topological and geometric structures underlying the data [17], [18]. TDA leverages topology and geometry to robustly infer qualitative and quantitative information about data structure. It results in either a full reconstruction, often a triangulation, of the underlying shape of EEG signals, enabling the extraction of topological and geometric features. Alternatively, it produces recapitulative summaries or approximations, from which specific methods like persistent homology (PH) can extract meaningful information [19], [20], [21], [22], [23]. The extracted topological and geometric information from EEG signals, which focuses on the group relevance of different regions of the brain, potentially provides new families of features and descriptors of the data. Inspired by these promising findings, there emerges a strong incentive to investigate the performance of the EEG data into DNNs. To ensure the logic of the experiment, we first should extract effective features from the unlabeled Gestalt training datasets, and study the effective means to integrate with unsupervised DNN

structures. Specifically, in the paper, `DeepCluster` [9] and multi-view clustering (MVC, e.g., `SiMVC`) are used as benchmarks.

In this study, we present a novel systematic framework to form algorithmic integration between DNN structures and intuitive TDA patterns to improve Gestalt illusory contour perception effectiveness. The framework is further evaluated with a combination with algebraic topology-deduced EEG signals. As the block diagrm illustrated in Fig. 1, we argue this new framework performs more effectively and more flexibly than conventional DNNs with a large amount of parameters to calibrate and train. In a nutshell, Table 1 summarizes the key differences with highly related papers, while our contributions are as follows:

- We leverage two TDA-based metrics (i.e., the Euler characteristic and persistent entropy) to verify the separability of EEG responses after observing images with Gestalt and random patterns.
- For perception of Gestalt illusory contour, we propose a framework to combine TDA patterns with DNN structures. Besides, contingent on the separability of EEG data, we supplement it by algebraic topology-deduced features from EEG signals.
- Finally, we confirm the effectiveness of the proposed framework for Gestalt classification task. We demonstrate that general DNN structures cannot always perform this task, and our TDA-based framework obtains significantly improved performance on Gestalt illusory contour perception tasks.

The rest of the paper is organized as follows. In Section II we describe the datasets for analyses and present the preprocessing means. In Section III, after introducing the fundamentals of TDA, we discuss the mapping results of EEG data and present results on the separabilty of EEG data via TDA. In Section IV, we introduce the topological layer and elaborate on the DNN structure with this topological layer. The results of computational experiments are presented in Section V, which are followed by our discussions and conclusion in Section VI and Section VII.

II. DATASET DESCRIPTION AND PREPROCESSING
In this section, we provide a brief description of the leveraged two datasets (i.e., Gestalt image datatset and Gestalt EEG dataset), and give the data preprocessing method.

A. DESCRIPTION OF DATASETS
1) THE GESTALT IMAGE DATASET
The Gestalt image datasets contain four types of contrasting images, which have been illustrated in Fig. 2. Specifically, Fig. 2(a) and Fig. 2(b) show images consisting of 5, 000 different Gestalt (GST) and random sequence diagram (RSD) images but with line and pancake contours respectively, whereas Fig. 2(c) and Fig. 2(d) depict images composed of 5, 000 GST, RSD, and triangle images with different contours. Notably, compared to pancake images, line-broken images with noise addition may bring about larger

TABLE 1. The summary of differences with highly related papers.

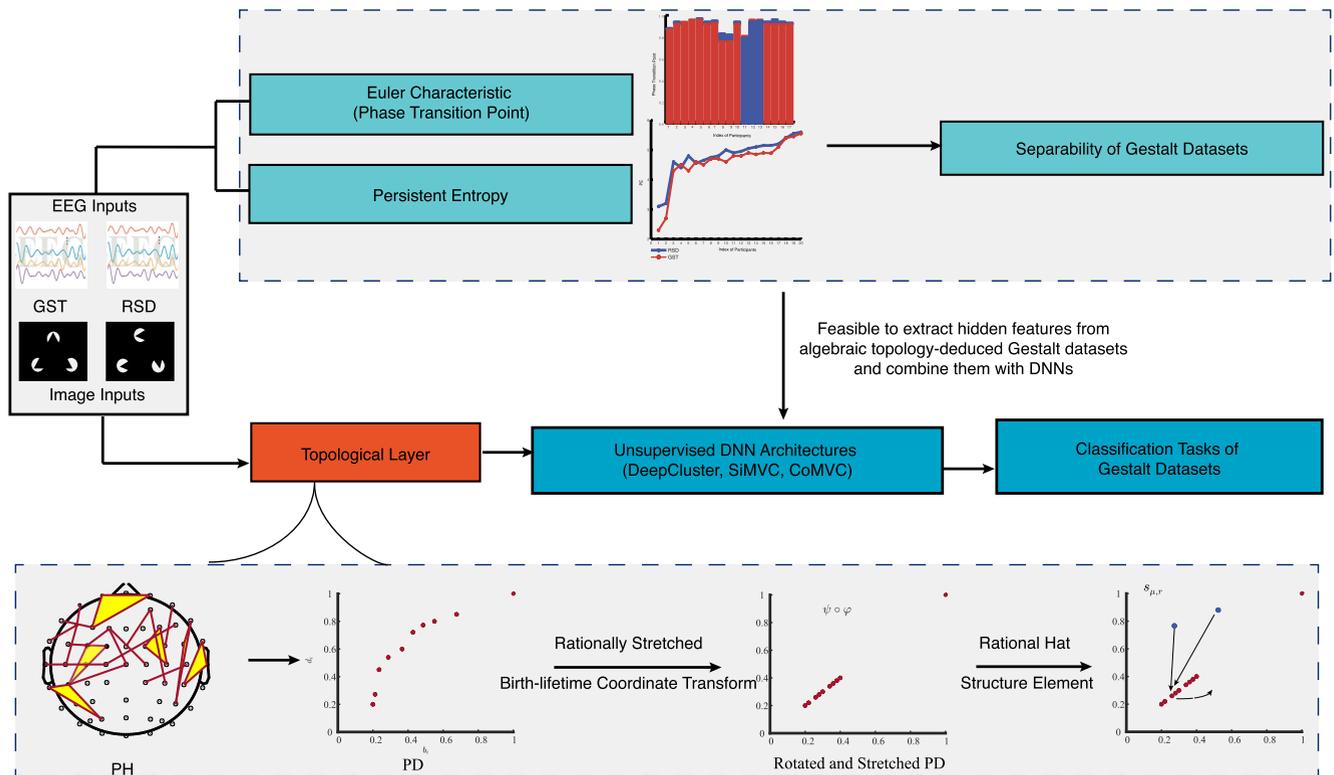| Articles exploring Gestalt illusory contours with DNNs	
[2]: DNN structures (e.g., InceptionNet) pretrained on ImageNet are proposed to testify a closure measure metric, which computes the similarity between Gestalt images and triangles.	**Ours**: Verification of the capability to perceive Gestalt illusory contours by EEG signal-assisted algebraic topological feature-enhanced DNNs.
[3]: Responses of well-trained PredNet resemble the neural activities in brains of monkeys responsive to Gestalt illusory stimuli.	
[4]: Predictive coding networks obtain improvements on the prediction probability of illusory contours.	
[6]: DNNs trained on non-illusory fat / thin rectangles are unable to interpolate values between tangent discontinuities in the same way as human observers do.	**Ours**: DNNs are able to perceive illusory contours in terms of the improvements of Gestalt classification accuracy.
Articles classifying EEG signals with TDA	
[15]: Detecting the emergence of epileptic seizures via persistent entropy.	**Ours**: Analyzing the separability of EEG visual signals responsive to Gestalt and non-Gestalt images via persistent entropy and phase transition point.
[16]: Detecting autism spectrum disorder via persistent entropy.	

FIGURE 1. The diagram of EEG signal-assisted DNNs via TDA.
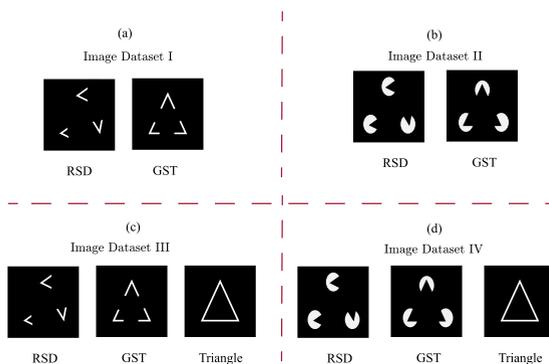

FIGURE 2. An illustration of Gestalt image datasets.

training differences on Gestalt illusory contour perception tasks. Besides, these images are generated by randomly adjusting the edge position, length, or angle.

2) THE GESTALT EEG DATASET

Collected from a brain cognition experiment with informed consent, the Gestalt EEG dataset records the neural activities of the human brain responding to watching and identifying the geometric patterns of GST and RSD images, by using a special cap with sensing electrodes at a sampling rate of $1,000$ Hz. In particular, as depicted in Fig. 3, 20 participants (9 males and 11 females; in an age group of $19-27$) with normal (or corrected to normal) vision were seated in front of a video screen and first watched 30 RSD images and then 10 GST images repeatedly at a fixed interval (i.e., 10 seconds). For each participant, 64 channels of electronic data are measured through EEG caps. Hence, the Gestalt EEG dataset encompasses 20 EEG recordings with 30 and 10 trials of RSD and GST, respectively. The test equipment is Neuracle 64 System (Neuracle product; sensor array: 64-channel adult-sized head cap (i.e., $P = 64$; EEG
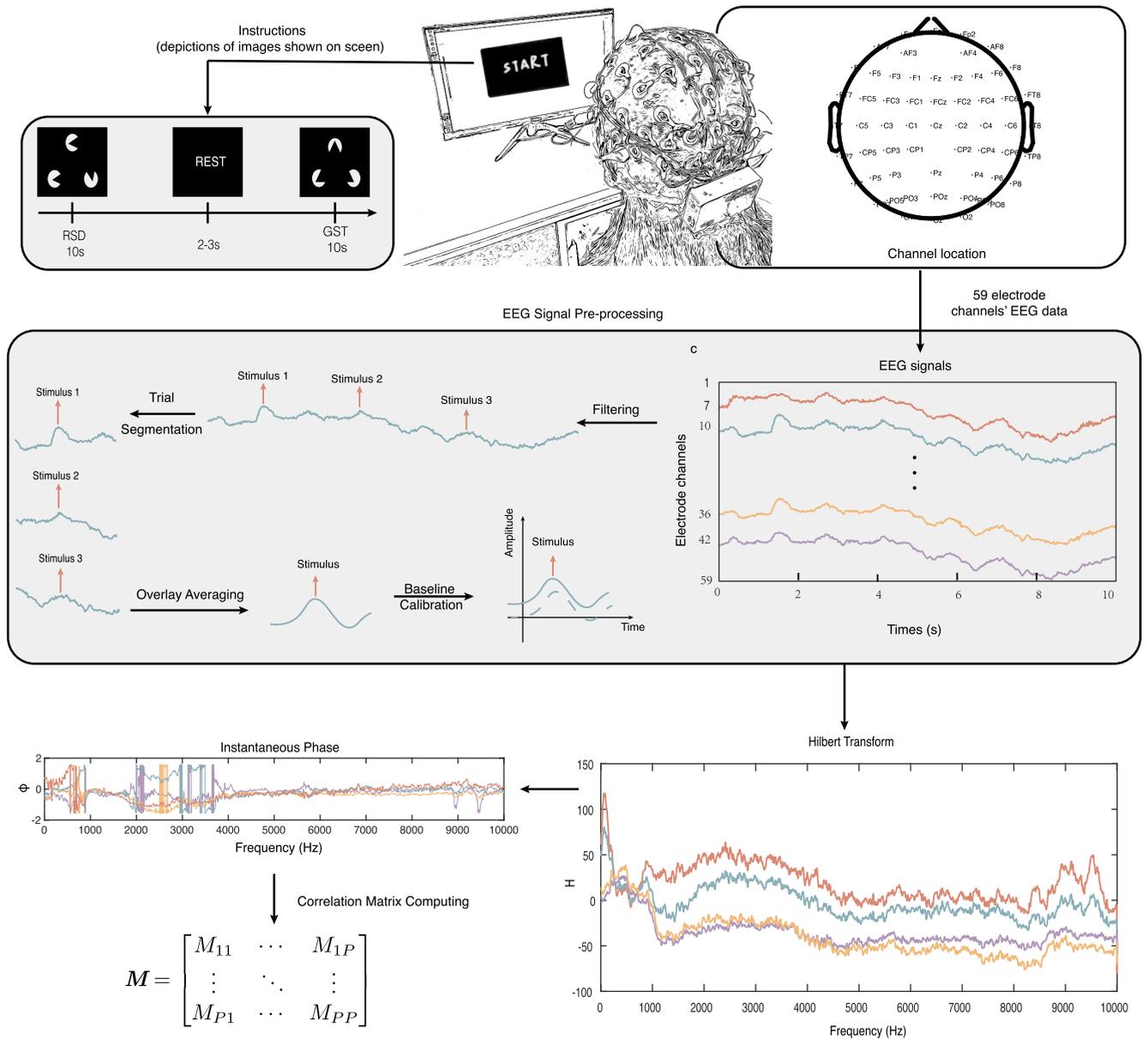
FIGURE 3. Gestalt EEG datasets and their mathematical representation.

acquisition software: EEGRecorder; amplifier: NSW364; reference electrode: middle of Cz and Pz).

B. DATA PREPROCESSING

1) PREPROCESSING OF GESTALT IMAGE DATASET

Before mining the topological features, we preprocess the images to filter and highlight important features. Specifically, we binarize the image, and map all pixel values to the interval [0, 1] according to the set threshold, thus highlighting the outline. Afterwards, we perform the radical filtration [24] to construct the cubical complexes, by following the l_2-norm

distance between a pixel point and the preset center point. Besides, as discussed lately, the PH can be conducted according to the radical distance.

2) PREPROCESSING OF GESTALT EEG DATASET

As for the EEG dataset, the preprocessing includes the following steps:

1) *Filtering*. The DC component and high-frequency interference are filtered out by a bandpass filter, thus removing redundant EEG frequency bands. The EEG signals of all leads are filtered by a

digital passband filter and the filtering window is $1 - 45$ Hz [25].

2) *Trial segmentation.* The EEG waveforms recorded in each trial are segmented according to the starting point and duration of each stimulus. Besides, a small segment of the signal before the start of the stimulation time is generally retained, so as to facilitate the subsequent baseline determination and calibration of spontaneous EEG.

3) *Artifact removal.* In order to combat various artifact noises such as eye blinks, eye movements, and electromyograms, trials with amplitude spikes exceeding a predefined noise threshold are discarded. Besides, taking account of regional activities of the cognition process in the human brain, we voluntarily delete the data of five electrode channels (ECG, HEOR, HEOL, VEOU, VEOL), and the locations of the remaining channels are visualized in Fig. 3.

4) *Overlay average.* In terms of stimulus categories, all trials within the same category are superimposed and averaged.

5) *Baseline calibration.* We take the average of a small segment of signals before the start of the simulation time as the baseline, and obtain the corresponding event-related potential by subtracting the baseline from the EEG data.

After the preprocessing of EEG signals in Fig. 3, we could construct an EEG signal matrix by augmenting and normalizing singals from 59 electrode channels as

$$\boldsymbol{F}_{\text{EEG}} = \begin{bmatrix} f_{11} & \cdots & f_{1Q} \\ \vdots & \ddots & \vdots \\ f_{P1} & \cdots & f_{PQ} \end{bmatrix}, \qquad (1)$$

where $Q = 10,000$ denotes the total data length, equal to the sampling rate multiplied by the measurement time, and the notation of P is slightly abused to denote the number of effectively used electrodes (i.e., $P = 59$).

Next, for topological feature mining, we attempt to set up a corresponding relationship between the 59 electrode channels. Specifically, for each row vector $\boldsymbol{F}_{\text{EEG}}(p, :)$, $\forall p = \{1, \cdots, P\}$, we apply a Hilbert transform to obtain $\boldsymbol{h}(\boldsymbol{F}_{\text{EEG}}(p, :))$. The instantaneous phase of each electrode $\boldsymbol{\phi}(p, :)$ is calculated as

$$\boldsymbol{\phi}(p, :) = \arctan\left(\frac{\boldsymbol{h}(\boldsymbol{F}_{\text{EEG}}(p, :))}{\boldsymbol{F}_{\text{EEG}}(p, :)}\right). \qquad (2)$$

Subsequently, a correlation matrix $\boldsymbol{M} = [M_{p_1,p_2}] \in \mathbb{R}^{P \times P}, \forall p_1, p_2 = \{1, \cdots, P\}$ is obtained by phase-locking analysis [26], [27]. In particular, M_{p_1,p_2} is computed as

$$M_{p_1,p_2} = \begin{cases} \frac{1}{Q} \left| \sum_{q=1}^{Q} \exp(j\{\phi(p_1, q) - \phi(p_2, q)\}) \right|, & p_1 \neq p_2; \\ 0, & p_1 = p_2, \end{cases} \qquad (3)$$

where j represents an imaginary unit. Assuming that if M_{p_1,p_2} is greater than the threshold $1 - \varepsilon$ (i.e., the filtration value discussed below), $\forall \varepsilon \in [0, 1]$, the electrodes p_1 and p_2 are "logically" connected; otherwise, they are unconnected. Thus, a modified correlation matrix $\hat{\boldsymbol{M}} = [\hat{M}_{p_1,p_2}(\varepsilon)]$ can be achieved with

$$\hat{M}_{p_1,p_2}(\varepsilon) = \begin{cases} |M_{p_1,p_2}|, & |M_{p_1,p_2}| \geq 1 - \varepsilon; \\ 0, & |M_{p_1,p_2}| < 1 - \varepsilon. \end{cases} \qquad (4)$$

III. TDA-BASED SEPARABILITY ANALYSIS OF EEG DATA
A. FUNDAMENTALS OF TDA
1) VIETORIS-RIPS COMPLEX
TDA is an emerging research field that studies topological approaches to explore and provide tight representations of complex, high-dimensional data [28], [29]. By leveraging the idea of topology to recognize patterns within data, TDA transforms data into compressed, useful knowledge. Specifically, TDA uses PH to capture topological features of a space at different spatial resolutions and enables a fast comparison of functions defined on the data [30]. To find the PH of a space, all point cloud data (PCD) in that space could be filtrated by circles with different radii ε to form a certain geometric structure called a complex. For example, on top of the aforementioned modified correlation matrix $\hat{\boldsymbol{M}}$, the functional brain network can be filtrated for each value $\varepsilon \in [0, 1]$ by assigning an edge between connected electrode channels p_1 and p_2 with a non-zero $\hat{M}_{p_1,p_2}$. In other words, as illustrated in Fig. 4, an undirected graph $\mathcal{G}(\varepsilon) = (\mathcal{V}, \mathcal{E}(\varepsilon))$ can be attained, where $\mathcal{V}$ represents the vertex set of 59 correlated points and $\mathcal{E}$ represents the edge set under ε. Furthermore, a Vietoris-Rips (VR) complex, which belongs to a widely used simplicial complex to characterize the properties of PCD [31], can be represented as

$$R(\mathcal{G}(\varepsilon)) = \mathcal{V} \cup \mathcal{E}(\varepsilon) \cup \left\{ \sigma \middle| \binom{\sigma}{2} \subseteq \mathcal{E}(\varepsilon) \right\}, \qquad (5)$$

where σ represents any simplex in $\mathcal{E}(\varepsilon)$ [31]. As the threshold ε increases, new edges are gradually attached, thereby possibly changing the topology of the constructed networks.

2) HOMOLOGY AND BETTI NUMBERS
Considering the aforementioned VR complex $R(\mathcal{G}(\varepsilon))$, a k-chain c_k is a formal sum of the k-simplices in $R(\mathcal{G}(\varepsilon))$ (i.e., $c_k = \sum_{i=1}^{n_k} a_i \sigma_k^{(i)}$, where $a_i \in \{0, 1\}$). Besides, there are n_k k-simplices in $R(\mathcal{G}(\varepsilon))$, one of which $\sigma_k^{(i)}$ belongs to. Furthermore, the boundary of k-simplex can be defined as the summation of its $k-1$-dimensional faces. In other words, for a k-simplex σ_k spanned by the vertices $v_0, \cdots, v_k \in \mathcal{V}$, its boundary is defined as

$$\partial_k \sigma_k = \sum_{i=1}^{k} [v_0, \cdots, \hat{v}_i, \cdots, v_k], \qquad (6)$$

where $\hat{v}_i$ indicates the drop of v_i in the summation. For the k-chain c_k, the boundary is the sum of boundaries of its simplices $\partial_k c_k = \sum a_i \partial_k \sigma_k^{(i)}$. Therefore, ∂_k can be regarded as a homomorphism from the k-chain group $\mathcal{C}_k$ to the $k-1$-chain
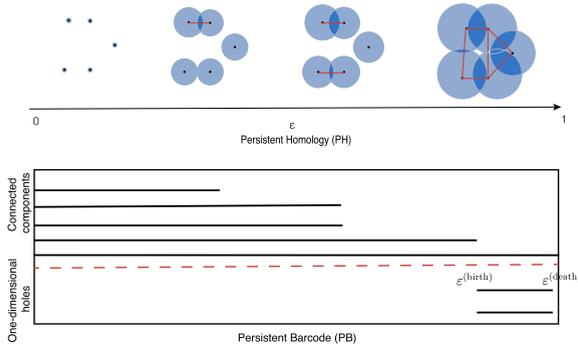
FIGURE 4. PH and its corresponding PB.

group $\mathcal{C}_{k-1}$. Meanwhile, a chain complex is the sequence of chain groups connected by boundary homomorphisms.

After defining $\mathcal{C}_k$, we can present two special types of chain groups (i.e., the boundary group $\mathcal{B}_k = \text{im } \partial_{k+1} \subseteq \mathcal{C}_k$ and the cycle group $\mathcal{Z}_k = \text{ker } \partial_k \subseteq \mathcal{C}_k$), which denote the image set of $\mathcal{C}_{k+1}$ and the kernel set to a zero vector (i.e., $\partial_k \mathcal{Z}_k = 0$), respectively [17], [32]. Notably, as $\partial_k \partial_{k+1} c = 0$, the defined groups are nested as $\mathcal{B}_k \subseteq \mathcal{Z}_k \subseteq \mathcal{C}_k$. Afterwards, another group (i.e., homology group) can be defined by the quotient group $\mathcal{H}_k = \text{ker } \partial_k / \text{im } \partial_{k+1}$ [33]. The quotient operation can be regarded as calculating a special subset (i.e., the kernel of the boundary operator) and then removing another subset (i.e., the image of the boundary operator with an increased dimension) [34]. Equivalently, Betti numbers can be derived from the rank of the difference between $\mathcal{Z}_k$ and $\mathcal{B}_k$ as

$$\beta_k = \text{rank} (\mathcal{Z}_k) - \text{rank} (\mathcal{B}_k). \tag{7}$$

Intuitively, the Betti numbers $\beta_k, k \in \{0, 1, \cdots, K\}$ indicate the number of k-dimensional holes. More specifically, β_0 and β_1 denote the number of connected components and one-dimensional or "circular" holes, respectively.

Recalling the definition of boundary in (6), we can construct a k-th boundary matrix as

$$\partial_k = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n_k} \\ \vdots & \ddots & \vdots \\ a_{n_{k-1},1} & \cdots & a_{n_{k-1},n_k} \end{bmatrix}, \tag{8}$$

wherein for an ordering pair (i, j) ($\forall i \in \{1, \cdots, n_{k-1}\}$ and $j \in \{1, \cdots, n_k\}$), $a_{i,j} = 1$ if and only if the $k-1$-simplex σ_{k-1} is the face of the k-simplex σ_k [35]. By elementary row and column operations, we can get the boundary matrix's Smith normal form $\hat{\partial}_k$ [36], and have rank$(\mathcal{B}_{k-1})$ = rank $\hat{\partial}_k$ while rank$(\mathcal{Z}_k)$ = rank$(\mathcal{C}_k)$ − rank$(\mathcal{B}_{k-1})$ = n_k − rank$(\mathcal{B}_{k-1})$. Correspondingly, for each ε, the Betti numbers can be easily achieved.

Notably, the VR complex obtained from the filtration [37] implies the topological structures of data. In mathematical words, considering the disappearance of old simplices and emergence of new ones, the filtration of a VR complex potentially demonstrates the persistence or vitality of the Betti numbers.

3) PERSISTENT BARCODE AND PERSISTENT DIAGRAM
As mentioned above, the filtration of ε results into different VR complexes, and correspondingly a simplex (e.g., a hole) might appear and vanish. As illustrated in Fig. 4, persistent barcode (PB) is leveraged to characterize this procedure in terms of bars. Specifically, for a simplex σ persisting from the birth time $\varepsilon^{(\text{birth})}$ to the death time $\varepsilon^{(\text{death})}$, the PB is defined as the multiset of Betti intervals $[\varepsilon^{(\text{birth})}, \varepsilon^{(\text{death})}]$ and thus capably characterizes the topological patterns of the functional brain network.

Different from the calculation of Betti numbers from VR complexes, we start with a complex corresponding to the largest correlation threshold ε to calculate the PB. In other words, all the potential simplicies hidden in all ε values are merged to form a coherent filter stream (or rather, a series of nested complexes). Furthermore, we order these simplicies as follows: For a k_1-simplex σ_{k_1} and a k_2-simplex σ_{k_2}, the simplicies are first sorted according to the ascending order of k_1 and k_2 and the 0-simplices (i.e., the nodes) are randomly sorted at first; for equal k_1 and k_2, the order of simplicies is determined by its longest (most specific) 1-simplex (i.e., the edge); for equal longest edges (i.e., their longest edges emerge at the same value ε), the simplicies are sorted in an ascending order between the maximum number of nodes. Subsequently, during the whole filtration range, we can achieve a boundary matrix $\partial^{(\text{whole})}$ in which each row and column represents all emerging ordered simplicies. Similarly, if the $k − 1$-simplex σ_{k-1} is the face of the k-simplex σ_k, the corresponding entry in $\partial^{(\text{whole})}$ refers to 1. Besides, we introduce a weight matrix $\mathbf{W} = [\varepsilon_\sigma]$, where ε_σ records the corresponding value of ε when σ emerges. Thus, to dig out the birth time $\varepsilon_\sigma^{(\text{birth})}$ and death time $\varepsilon_\sigma^{(\text{death})}$ of topological features, we reduce the matrix $\partial^{(\text{whole})}$ into a column-echelon form $\hat{\partial}^{(\text{whole})}$ [38]. Meanwhile, the transformation steps are stored in an auxiliary memory matrix.

Next, it becomes ready to read and plot the related Betti intervals $[\varepsilon_\sigma^{(\text{birth})}, \varepsilon_\sigma^{(\text{death})}]$ of a feature into a PB, by simply scanning each column from left to right of the reduced matrix $\hat{\partial}^{(\text{whole})}$. If the values of a column u (i.e. the index of the ordered simplices) are all 0, we resort to lookup the weight matrix $\mathbf{W}$, and record $[\varepsilon_u, \inf]$ as the Betti interval of the simplex. Otherwise, if a column v contains non-zero entries (i.e., 1) and the row index corresponding to the last 1 equals u, it means that the simplex generated at ε_u disappears at ε_v. Correspondingly, its barcode can be represented as $[\varepsilon_u, \varepsilon_v]$ (i.e. $\varepsilon^{(\text{birth})} = \varepsilon_u$ and $\varepsilon_\sigma^{(\text{death})} = \varepsilon_v$). Meanwhile, a PB can be transformed into a persistent diagram (PD), which represents topological features as multisets of points $(\varepsilon_u, \varepsilon_v)$.

B. RESULTS OF SEPARABILITY ANALYSIS
In this section, we boldly conjecture that the EEG data responding to watching images with Gestalt and random pattern could be distinguishable and try to testify this. Our conjecture are significantly motivated by the following direct findings from Fig. 5:
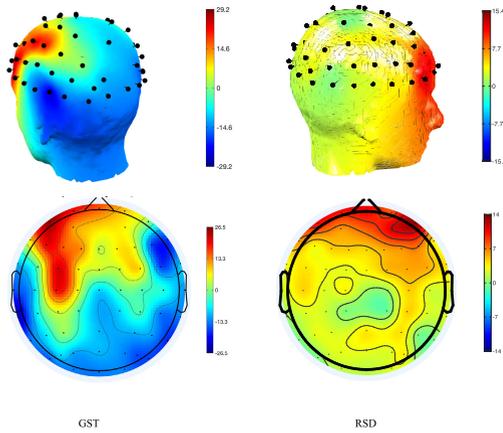
GST   RSD

FIGURE 5. EEG mapping.

- The brain reacts differently to the perception of images with different shape and contour, and perceiving images yields intense brain activities mostly in the frontal lobe, which is the physiological basis of advanced mental activities such as planning, regulating, and controlling other advanced and purposeful behaviors [39]. Nonetheless, when the brain perceives irregularly distributed images (i.e., RSD images), more brain regions are involved, but with nonprominent values. However, when it perceives implicitly structured images (i.e., GST images), there emerge clear reaction areas with more prominent features. Thus, this implies a potential separability of the two types of EEG signals (RSD & GST).
- The aforementioned observation and the Gestalt law of closure are also anastomosed, which identifies that, faced with implicit Gestalt closure, our brain draws contours by filling in the blanks and creating a unified ensemble to abate visual noise and convey information with reduced dispensable activities and information exchange [40]. In other words, it is feasible to further extract hidden features from algebraic topology-deduced EEG signals and combine them with DNNs.

Based on the aforementioned findings, we try to prove the separability of two types of EEG signals (RSD & GST) in terms of both Euler characteristic and persistent entropy analysis.

1) EULER CHARACTERISTIC ANALYSIS

Classically, in Hamiltonian function-described exactly solvable systems [41], the singularities of the Euler entropy occurs exactly at the transitions. This result is also consistent with the statement of the Yang-Lee theorem [42], which states that the singular behavior exhibited by thermodynamic quantities in equilibrium phase transitions coincides precisely with the zeros of the system's partition function. Unfortunately, as for the topological phase transition (TPT) in complex networks, the Hamiltonian function [41] is usually missing (or nonexistent). Therefore, instead of studying the system behavior, intrinsic correlations between the system

components determined from the empirical data are frequently leveraged to define the network topology. In particular, the Euler characteristic and the Betti numbers emerge in this scenario as natural quantities to investigate the TPT point [43], which may further signal a major change in a brain network.

Generally, each simplicial complex is constituted by its nodes ($k = 0$), edges ($k = 1$), triangles ($k = 2$), tetrahedrons ($k = 3$), and higher k-dimensional parts [43]. Basically, Euler characteristic numbers $\langle \chi \rangle$ could be calculated as the alternate sum of the number of k-simplices (or equivalently the alternate sum of Betti numbers) [44] as

$$\langle \chi \rangle (\varepsilon) = \sum_{k=0}^{K} (-1)^k n_k(\varepsilon) = \sum_{k=0}^{K-1} (-1)^k \beta_k(\varepsilon). \quad (9)$$

For example, Fig. 6 illustrates Euler characteristic numbers under different filtration numbers when $K = 3$. Correspondingly, the Euler entropy S_χ can be computed as

$$S_\chi(\varepsilon) = \ln |\langle \chi \rangle (\varepsilon)|. \quad (10)$$

Afterwards, the TPT value can be obtained as [43]

$$\varepsilon_{\text{TPT}} = \arg \min_\varepsilon (S_\chi(\varepsilon)). \quad (11)$$

Therefore, based on the filtrated EEG signals in $R(\mathcal{G}(\varepsilon))$, we apply the aforementioned persistent homology theory and track the TPT point of every participant. Fig. 7 provides the corresponding result. It can be observed from Fig. 7 that the TPT points of GST trials are generally smaller than that of the RSD, which intuitively shows the separability of the overall topological characteristics of the EEG signal.

2) PERSISTENT ENTROPY ANALYSIS

Besides Euler characteristic analysis, the persistent entropy (PE) is also adopted for evaluating whether a filtered simplicial complex is "ordered" based on PBs from an entropy perspective [45].

Without loss of generality, for any bar i with $[\varepsilon_i^{(\text{birth})}, \varepsilon_i^{(\text{death})}]$ in the PB, the PE H_{PE} of the simplicial complex filtration is calculated by the following equation:

$$H_{\text{PE}} = -\sum_i p_i \log p_i, \quad (12)$$

where $p_i = \frac{\varepsilon_i^{(\text{death})} - \varepsilon_i^{(\text{birth})}}{L}$ and $L = \sum_i \left(\varepsilon_i^{(\text{death})} - \varepsilon_i^{(\text{birth})} \right)$.

For cases with a unbounded interval $[\varepsilon_i^{(\text{birth})}, \inf)$, an interval $[\varepsilon_i^{(\text{birth})}, m)$ is used instead, where $m = \max \boldsymbol{F}_{\text{EEG}} + 1 = 2$. Moreover, in order to facilitate the comparison among different barcodes, we adopt the stability theorem and normalize H_{PE} into the interval $[0, 1]$, denoted by $\hat{H}$, and defined as

$$\hat{H}_{\text{PE}} = \frac{H_{\text{PE}}}{\log l_{\max}}, \quad (13)$$

where $l_{\max}$ denotes the maximum interval in the considered PB group.

Fig. 8 depicts the respective results and demonstrates an nearly unanimous difference in the EEG signals responding
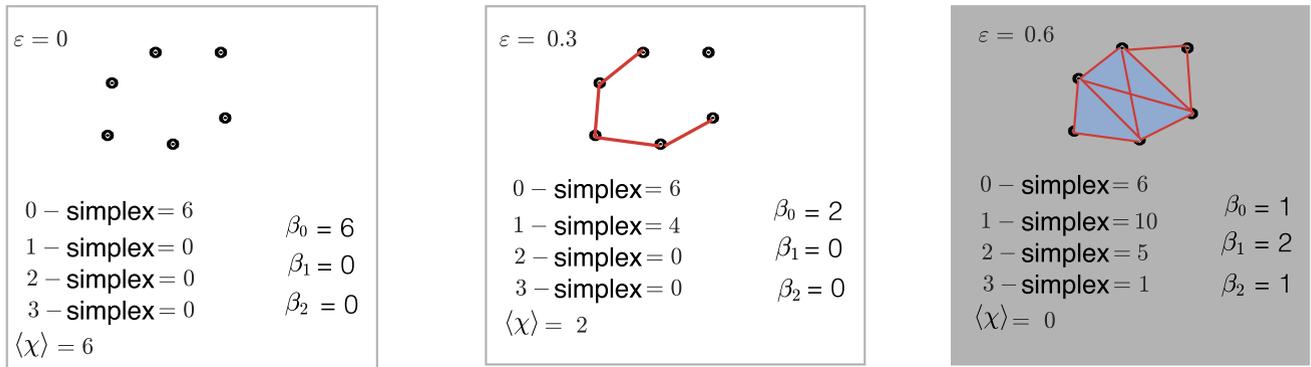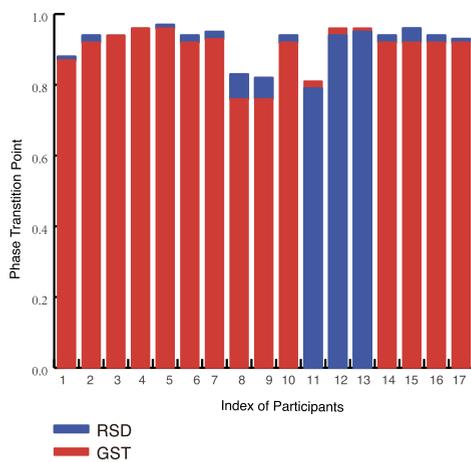
FIGURE 6. Euler characteristic numbers.

FIGURE 7. The topological phase transition point of every participant in terms of Euler characteristic analysis.
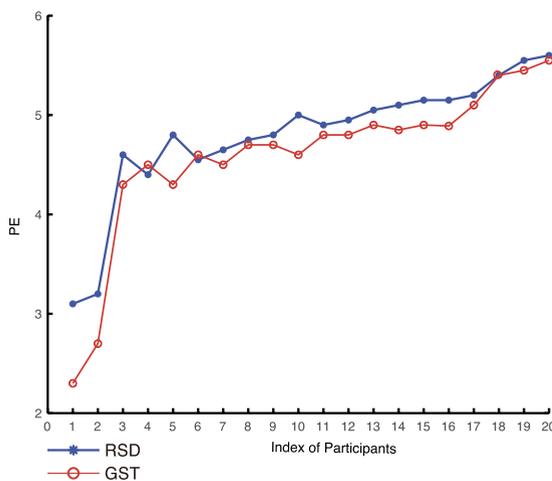

FIGURE 8. The PE of every participant by PH analysis.

to RSD and GST patterns for all 20 participants. Specifically, almost all the PE values (i.e., 85%) for the topological structure of the EEG signals induced by RSD are greater than those induced by GST.

The outcome of Euler characteristic and the persistent entropy analysis lays the very foundation for the separability of EEG signals. Meanwhile, it becomes preferential to select participants whose EEG data exhibit consistency and leverage their TDA-processed datasets as the training data to perform Gestalt classification tasks using DNNs.

IV. THE DNN STRUCTURE WITH THE TOPOLOGICAL LAYER FOR GESTALT IMAGE CLASSIFICATION

In this part, we give the means to extract topological and geometric features via TDA, and then present a comprehensive DNN framework with this topological layer.

A. TOPOLOGICAL LAYER

The summary representations (e.g., PBs or PDs) of topological features of data, come in an unusual format of multisets, equipped with computationally expensive metrics. Therefore, they can not readily be processed with conventional DNNs. Accordingly, suitable vectorization techniques based on algebraic topology ideas [46] are proposed to bridge the gap between DNNs and TDA, which leverage advances in representation learning via DNNs. The core idea is to project points in a barcode by a collection of parametrized functionals, so called "structure elements" [47].

In particular, we start by rotating the points of PD by ψ, that is

$$\psi : (\varepsilon_i^{(\text{birth})}, \varepsilon_i^{(\text{death})}) \mapsto (\varepsilon_i^{(\text{birth})}, \varepsilon_i^{(\text{death})} - \varepsilon_i^{(\text{birth})}), \quad (14)$$

so as to ensure that the x-axis still represents the birth time of topological features, whereas the y-axis can then be interpreted as the persistence of features [48].

Afterwards, on top of ψ, we introduce another transformation

$$\varphi_\xi(x_0, x_1) = \begin{cases} (x_0, x_1), & x_1 \in [\xi, \infty) \\ (x_0, 2\xi - \dfrac{\xi^2}{x_1}), & x_1 \in (0, \xi) \\ 0, & x_1 = 0 \end{cases} \quad (15)$$

towards the points $\boldsymbol{x} = (x_0, x_1) = (\varepsilon_i^{(\text{birth})}, \varepsilon_i^{(\text{death})} - \varepsilon_i^{(\text{birth})})$, where ξ represents the threshold at which the transformation starts to operate. Since each PD contains points at the
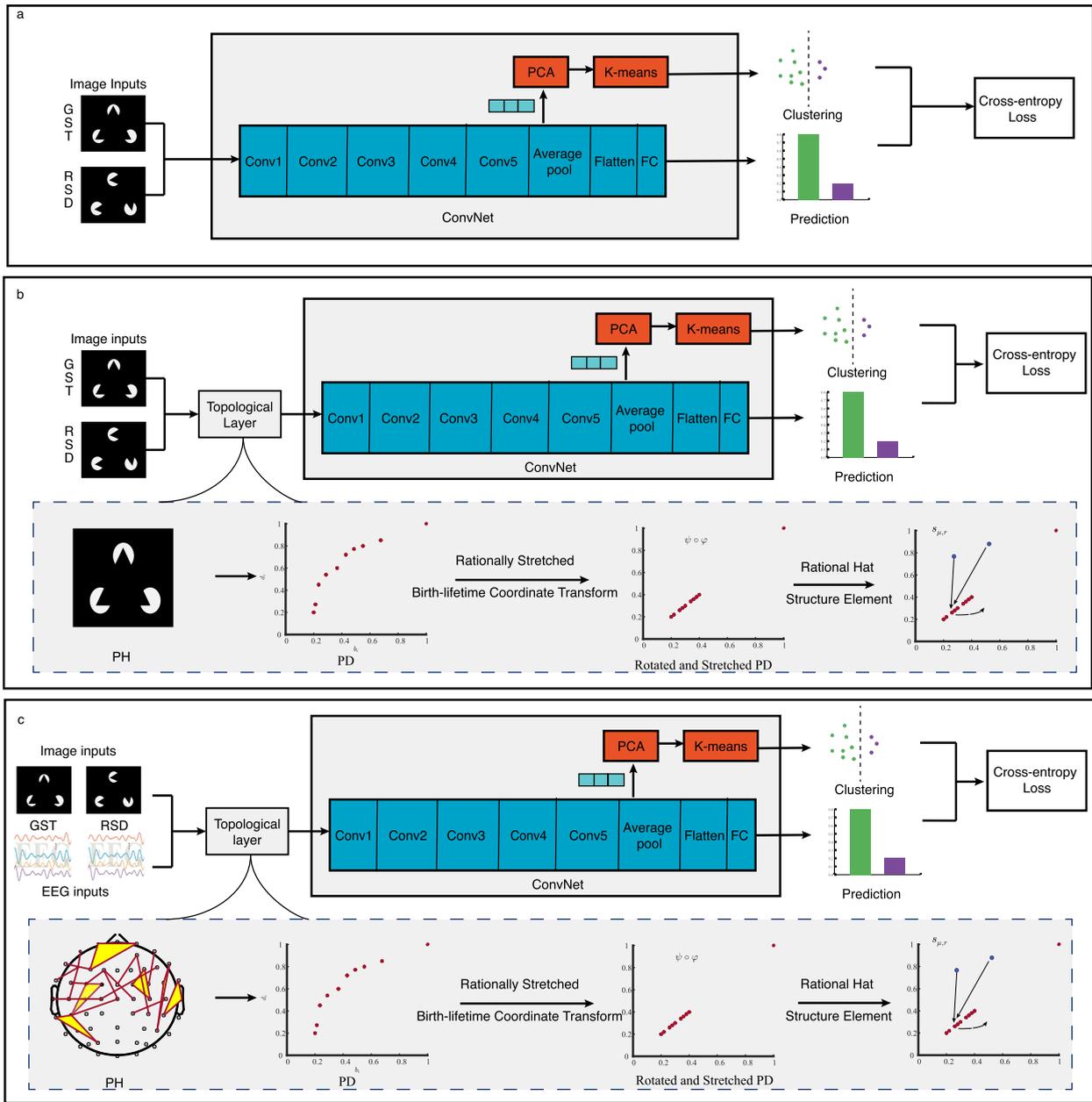
FIGURE 9. Experiment framework. (a) The architecture of `DeepCluster`. (b), (c) Topological layer with `DeepCluster`.

diagonal with infinite multiplicity [47], the transformation stretches the space between the x-axis and the line drawn at $x+\xi$ to infinite length, so as to guarantee Lipschitz continuity of all proposed structure elements. In this study, $\xi = 0.1$.

Following the transformation, we further add a rational hat structure element [47]

$$t_{\boldsymbol{\mu},r}(\boldsymbol{x}) = \frac{1}{1 + ||\boldsymbol{x} - \boldsymbol{\mu}||_2} - \frac{1}{1 + ||r| - ||\boldsymbol{x} - \boldsymbol{\mu}||_2|}. \quad (16)$$

In a nutshell, the rational hat structure element, whose maximum is reached when the distance of a point to $\boldsymbol{\mu}$ is r, can be regarded as a collection of parameterized function.

By implementing the vectorization of a PD, it yields a more balanced gradient scaling behavior [47], [48]. These vectorizations are then fed to a DNN, e.g., a discriminant classifier. Notably, the parametrization $t_{\boldsymbol{\mu},r}$ is learned during the training and allows to obtain a task-specific vectorization of barcodes.

B. DNN STRUCTURE

As for the DNN structure, we use the classical unsupervised classification architecture `DeepCluster` [9], shown in Fig. 9(a), as a benchmark. In order to generate the initial labels for training, we can leverage the available pre-trained

models such as AlexNet [49] and VGG-16 [50].[1] Specifically, we first take images as input and feed forward the DNN model. Correspondingly we construct a feature matrix of images by exploiting the final average pool layer. Then, we reduce the dimension of the feature matrix by the principal component analysis (PCA). In order to perform the classification, K-means clustering with $K = 2$ is applied to produce the clustering results.

In terms of complexity and completeness, we add the topological layer to `DeepCluster` incrementally.

- We introduce TDA-processed data into `DeepCluster` (`TFD-DC` for short). Specifically, the DNN structure remains the same. However, as depicted in Fig. 9(b), the Gestalt image datasets will be further processed by filtration to obtain the corresponding VR complex and the PD. The PD would become the inputs of the DNN.
- TDA-processed EEG data are introduced into `DeepCluster` (`ETFD-DC` for short). As shown in Fig. 9(c), the EEG patterns chosen in terms of PE are supplemented to Gestalt image datasets as a whole, while the processing method remains the same. The pseudocode of ETFD is shown in Algorithm 1.

V. EXPERIMENTAL RESULTS

In order to reflect the effectiveness of TDA, we perform the Gestalt classification task and evaluate the performance on the basis of `DeepCluster` (with either AlexNet [49] or VGG-16 [50]) and MVC [51].
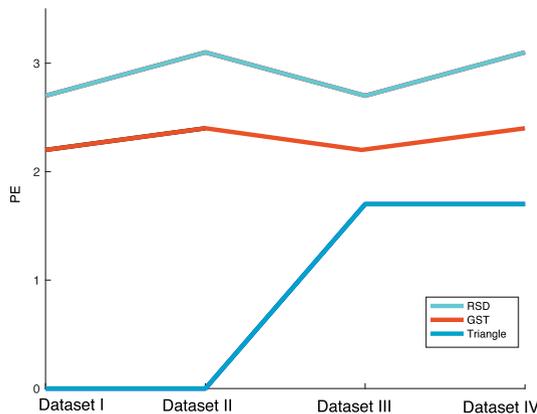
FIGURE 10. The persistent entropy of Gestalt image datasets.

During the training, the consistency between the clustering results and the original labels (with distributions δ_1 and δ_2 repectively) are measured by the evolution of the normalized mutual information (NMI). Specifically, NMI is computed as $\mathrm{NMI}(\delta_1; \delta_2) = \frac{\mathrm{MI}(\delta_1; \delta_2)}{\sqrt{H_{\mathrm{SE}}(\delta_1)H_{\mathrm{SE}}(\delta_2)}}$ where $\mathrm{MI}(\delta_1; \delta_2)$ denotes the mutual information and $H_{\mathrm{SE}}(\cdot)$ denotes the Shannon entropy. If δ_1 and δ_2 are independent, the NMI is equal to 0; while if one can be deterministically predictable from the other, the NMI is equal to 1 [9], [52]. The DNN is trained with dropout, a momentum of 0.9 and an l_2 penalization of

---

[1]Notably, other DNN structures can be leveraged as well and shall be further detailed in Section V.

Algorithm 1 ETFD for Gestalt Illusory Contour Classification

Input: Gestalt EEG datasets, filtration dimension k.
Output: Clustering results.
1 Initialize nodes $\mathcal{V}$, edges $\mathcal{E}$, weights $\boldsymbol{W}$;
2 Compute corresponding correlation matrix M of Gestalt EEG datasets as in Eq. (3);
3 Update $\mathcal{V}$, $\mathcal{E}$ with correlation Matrix $\boldsymbol{M}$;
4 **for** $k \in [0, 1]$ **do**
5   **if** $k = 0$ **then**
6   Initialize simplex σ_k with $\mathcal{V}$;
7   Add σ_k to σ;
8   Update corresponding value in $\boldsymbol{W}$ with 0;
9   **else if** $k = 1$ **then**
10   Initialize simplex σ_k with $\mathcal{E}$;
11   Add σ_k to σ;
12   **for** $\varepsilon \in (0, 1)$ **do**
13   Update modified correlation matrix $\hat{\boldsymbol{M}}$ with Eq. (4);
14   Update corresponding value in $\boldsymbol{W}$ with $\hat{\boldsymbol{M}}$;
15   **end**
16   **end**
17 **end**
18 Construct complex R with Eq. (5);
19 Update the boundary matrix $\partial^{(\mathrm{whole})}$ by extending Eq. (8) to characterize the whole filtration range;
20 Reduce the matrix $\partial^{(\mathrm{whole})}$ into a column-echelon form $\hat{\partial}^{(\mathrm{whole})}$;
21 Update birth time $\varepsilon_\sigma^{(\mathrm{birth})}$ and death time $\varepsilon_\sigma^{(\mathrm{death})}$ of all simplices with $\hat{\partial}^{(\mathrm{whole})}$;
22 Update (x_0, x_1) with persistence $(\varepsilon_\sigma^{(\mathrm{birth})}, \varepsilon_\sigma^{(\mathrm{death})})$;
23 Vectorize (x_0, x_1) with composite mapping consists of rotation ψ (i.e., Eq. (14)), transformation φ (i.e., Eq. (15)) and structure element t (i.e., Eq. (16));

24 Input the extracted feature (x_0, x_1) to the DNNs (e.g., DeepCluster, MVC);
25 Reduce the feature dimension from the average pool layer of the model by PCA;
26 Classify the features by K-means clustering.

---

the weights θ. Each mini-batch contains 100 images. We train the models for 100 epochs. We choose 5-fold cross validation and summarize the main hyperparameter in Table 2. All experiments have been implemented in PyTorch and run on NVIDIA GeForce RTX 3090. The final performance evaluation is chosen as the average value of five tests.

TABLE 2. The default hyperparameter settings.

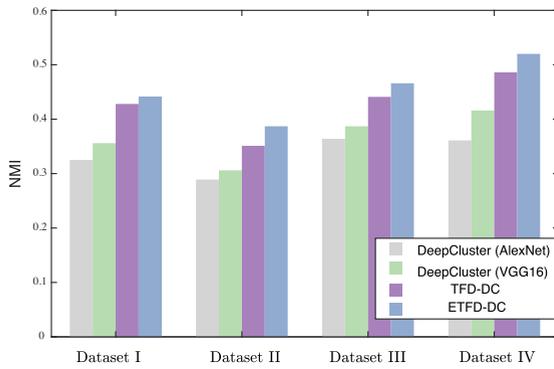| Hyperparameter	K-means	Weight deacy	Learning rate	Momentum
Setting	2	1e-4	0.05	0.9
Hyperparameter	Number of the structure element		Batch size	Epoch
Setting	100		100	100

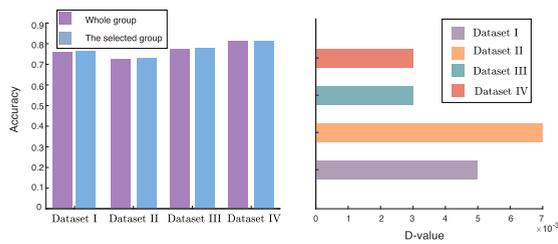**FIGURE 11.** The comparison of the clustering quality.

FIGURE 12. The comparison of the classification accuracy of two categories of EEG data (i.e.,the entire group and participants whose EEG data exhibit consistency in the entire group).

Beforehand, the PE results of Gestalt image datasets are given in Fig. 10. Specifically, the PE values of the RSD images are greater than those of GST and triangle images, thus verifying the separability of TDA-extracted features. Hence, it re-confirms the feasibility to combine TDA-processed datasets with DNN structures.

Fig. 11 presents perception capability in terms of NMI. Besides, Table 3 summarizes the evaluation accuracy of the aforementioned methods and validate this observation. Consistent with works which have shown DNNs cannot always perceive illusory contours [2], [6], the results of `DeepCluster` in Fig. 11 and Table 3 illustrate relatively worse perception performance in terms of the NMI and the classification accuracy. Meanwhile, the results also suggest that `DeepCluster` with VGG-16 outperforms that with AlexNet. Therefore, the following experiments are primarily based on `DeepCluster` (VGG-16).

TABLE 3. The comparisons of classification accuracy of different methods based on `DeepCluster`.

	DeepCluster (AlexNet)	DeepCluster (VGG16)	TFD-DC	ETFD-DC
Dataset I	64.75%	67.25%	72.75%	75.05%
Dataset II	56.25%	59.3%	66.5%	71.9%
Dataset III	69.83%	71.73%	75.47%	77.13%
Dataset IV	68.5%	73.43%	78.37%	80.57%

The contribution of TDA to `DeepCluster` can be further verified in terms of the classification efficacy and false alarm ratio. From Table 3, we can conclude that `TFD-DC` significantly outperforms `DeepCluster` because `TFD-DC`

TABLE 4. Comparisons of false alarming ratio of different methods.

	DeepCluster (AlexNet)	DeepCluster (VGG16)	TFD-DC	ETFD-DC
Dataset I	55.4%	50.6%	41.4%	36.6%
Dataset II	54.9%	49.4%	35.7%	25.6%
Dataset III	49.2%	44.9%	34.9%	32.2%
Dataset IV	51.3%	39.3%	25.1%	20.6%

TABLE 5. Comparisons of classification accuracy under different dimension no. of PCA.

	DeepCluster (AlexNet)		DeepCluster (VGG16)		TFD-DC		ETFD-DC	
	100	200	100	200	100	200	100	200
Dataset I	60.1%	66.5%	62.7%	68.2%	71.1%	73.9%	72.55%	76.6%
Dataset II	56.75%	58.2%	59.7%	61.15%	65.4%	67.3%	70.9%	73.25%
Dataset III	66.8%	71.73%	69.03%	72.8%	74.9%	77.6%	76.4%	78.93%
Dataset IV	66.63%	70.8%	71.3%	75.1%	76.93%	79.73%	79.7%	81.83%

TABLE 6. The comparisons of classification accuracy of different methods based on MVC.

	SiMVC	TFD-MVC	ETFD-MVC
Dataset I	78.6%	75.1%	77.05%
Dataset II	66.7%	72.8%	75.6%
Dataset III	77.8%	76.8%	78.9%
Dataset IV	78.9%	81.2%	83.4%

incorporates additional topological and geometric features. Specifically, for the Dataset IV, the maximum improvement in classification accuracy (i.e., 4.93%) can be expected. Therefore, the integration of TDA and DNNs can improve the perception of the illusory contour to a certain extent. Furthermore, adding the EEG signal extracted from the visual experiment further improves the accuracy. For example, for Dataset II, the accuracy boosts from 66.5% to 71.9%, indicating that the features extracted from EEG signals more accurately reflect the topological differences. Besides, as depicted in Fig. 11, as the iterations evolve, ETFD-DC with more sufficient topological features prominently capture the difference between GST and RSD in a more accurate manner. On the other hand, it can be observed from Table 4 that our method can lead to least false alarming ratio. The comparisons of classification accuracy under different settings of PCA (i.e., the dimension to whcih a feature matrix of datasets reduces) is also discussed in Table 5, which shows that it's more reliable to set the value as 200.

As for latest unsupervised learning structures, `SiMVC` [51] shows remarkable perfomance on many datasets [53]. Hence, Table 6 gives the comparison with the SOTA method (i.e., SiMVC) on top of MVC. From Table 6, it's evident that EEG signal-assisted algebraic topological feature-enhanced MVC (ETFD-MVC) outperforms SiMVC, especially on Dataset IV. This highlights the beneficial impact of algebraic topology-deduced EEG signals on the classification of Gestalt illusory contours.

We also explore the impact of the quality of EEG signals by selecting a subgroup of participants with consistent EEG data.

TABLE 7. Comparisons of running time of different methods.

	DeepCluster (AlexNet)	DeepCluster (VGG16)	ETFD-DC
Dataset I	1.5	1.9	2.4
Dataset II	1.7	2.4	3
Dataset III	1.6	2.1	2.6
Dataset IV	1.9	2.6	3.2

Fig. 12 presents the related accuracy. It can be observed that given the difference (D-value) in Fig. 12(b), the classification accuracy can be further improved when the selected consistent EEG data are used to train the `ETFD-DC`. In other words, the higher quality of EEG data promotes the improvement of performance.

Fig. 13 depicts the accuracy for all four datasets with respect to the number of the structure element and demonstrates that a larger number of structure elements produces superior performance. Together with Fig. 14, an illustration of structure elements centered at $\mu = (1, 1)$, it provides the guidance to set the number to 100.
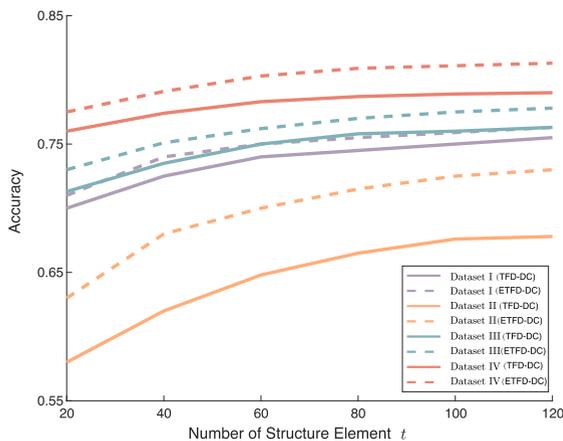

FIGURE 13. The accuracy on four datasets in Fig. 2 as a function of the number of the structure element.
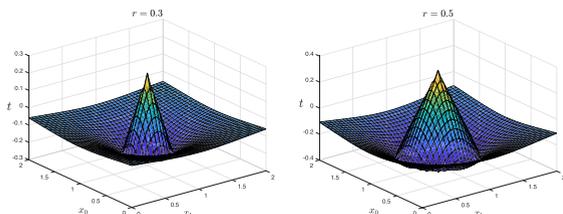

FIGURE 14. An illustration of structure elements centered at $\mu = (1, 1)$.

Finally, the comparison of running time (*h*) on NVIDIA GeForce RTX 3090 is provided in Table 7. It can be observed that introduction of the topological layer adds trivial time for training.

VI. DISCUSSIONS
DNNs exhibit remarkable performance on many types of classification problems, especially in computer vision and natural language processing [49]. The basic architecture of DNNs resembles the human brain networks on many performance-based measures, and many studies have been conducted to evaluate the similarity between DNNs and human perception [2]. The Gestalt principles describe how visual elements are grouped and interpreted [54], [55], [56]. For example, the Gestalt principle of closure asserts that human visual perception tends to "close the gap" by grouping elements that can jointly be interpreted as a complete figure or object [2]. Thus, this principle provides a basis for predicting how humans will parse, interpret, and attempt to display fragments. In the case of illusory contours like Gestalt images, the representations of humans and artificial DNNs behave very differently, as DNNs do not appear to interpolate between tangent discontinuities in the same way as human observers do [57].

In this work, we introduce EEG signals as additional input for DNNs. EEG signals capture changes in electric waves during brain activity and are an overall reflection of the electro-physiological activities of brain neural cells on the surface of the cerebral cortex or scalp [58]. In essence, we first analyze the EEG signals via algebraic topology. The core idea is to project points into a series of PBs by filtrating the corresponding correlation matrixes, and adopting a collection of parameterized functional, or the so-called structure elements. The parametrization is learned during the training and allows to obtain a task-specific vectorization of PBs. Besides, the results verify its effectiveness in improving the performance of DNNs to perceive Gestalt illusory contours. Although it is limited to the performance of DNNs with the EEG signals, TDA-based processing methods that filter images by algebraic topology seem to be superior than DNNs in terms of the accuracy in classifying Gestalt and random images. It can be also observed that supplementing images by EEG signals with consistent and clear differences in terms of topological phase transition point or PE from the field of algebraic topology can further improve the performance of the DNN.

VII. CONCLUSION AND FUTURE WORKS
In the study, towards the perception of illusory contour, we have made comprehensive analyses toward the topological features of EEG signals responding to watching images with Gestalt and random patterns. Contingent on a series of TDA-based analyses including Euler characteristic analysis and persistent entropy, we have confirmed the separability of these EEG signals. Furthermore, considering general DNN structures' incapability to possess the perceptual ability, we have combined TDA patterns with popular DNN structures (i.e., `DeepCluster`, `SiMVC` and `CoMVC`). Extensive simulation results unveil that feeding algebraic topology-deduced features from images and EEG signals as inputs to the TDA-based DNN significantly enhances its

ability to perceive illusory contours, resulting in improved classification accuracy. This indicates that TDA can be beneficial for understanding human perceptual capability in a solid manner.

In addition, EEG signal-assisted algebraic topological feature-enhanced DNNs' performance is limited to the quantity and quality of involved EEG signal datasets. Nevertheless, it is very cost-ineffective to record human brain activities for various visual categories. Besides, the human ability to detect, discriminate, and recognize perceptual visual stimuli is influenced by both visual features and people's prior experiences. In this regard, in order to obtain more improvement, our ongoing research leverages a mixture-of-product-of-experts formulation, which helps to infer a latent embedding for a coherent joint generation of two modalities and learn a more consistent joint representation. Therefore, our method promises to benefit from the relationships between brain activities and visual features via multimodal deep generative models, so as to improve the data efficiency in the case of limited brain activity data.

REFERENCES

[1] D. Heaven, "Why deep-learning AIs are so easy to fool," *Nature*, vol. 574, no. 7777, pp. 163–166, Oct. 2019.

[2] B. Kim, E. Reif, M. Wattenberg, S. Bengio, and M. C. Mozer, "Neural networks trained on natural scenes exhibit gestalt closure," *Comput. Brain Behav.*, vol. 4, no. 3, pp. 251–263, Sep. 2021.

[3] W. Lotter, G. Kreiman, and D. Cox, "A neural network trained for prediction mimics diverse features of biological neurons and perception," *Nature Mach. Intell.*, vol. 2, no. 4, pp. 210–219, Apr. 2020.

[4] Z. Pang, C. B. O'May, B. Choksi, and R. VanRullen, "Predictive coding feedback results in perceived illusory contours in a recurrent neural network," *Neural Netw.*, vol. 144, pp. 164–175, Dec. 2021.

[5] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Comput. Intell. Neurosci.*, vol. 2018, Feb. 2018, Art. no. 7068349.

[6] N. Baker, G. Erlikhman, P. J. Kellman, and H. Lu, "Deep convolutional networks do not perceive illusory contours," in *Proc. 40th Annu. Meeting Cognition Sci. Soc.*, Jul. 2018, pp. 1310–1315.

[7] A. Khmag, A. R. Ramli, and N. Kamarudin, "Clustering-based natural image denoising using dictionary learning approach in wavelet domain," *Soft Comput.*, vol. 23, no. 17, pp. 8013–8027, Sep. 2019.

[8] A. Khmag, "Additive Gaussian noise removal based on generative adversarial network model and semi-soft thresholding approach," *Multimedia Tools Appl.*, vol. 82, no. 5, pp. 7757–7777, Feb. 2023.

[9] M. Caron, P. Bojanowski, A. Joulin, and M. Douze, "Deep clustering for unsupervised learning of visual features," in *Proc. Eur. Conf. Comput. Vis.*, Munich, Germany, Sep. 2018, pp. 132–149.

[10] D. George, M. Lázaro-Gredilla, and J. S. Guntupalli, "From CAPTCHA to commonsense: How brain can teach us about artificial intelligence," *Frontiers Comput. Neurosci.*, vol. 14, Oct. 2020, Art. no. 554097.

[11] J. A. Gray, "Brain systems that mediate both emotion and cognition," *Cognition Emotion*, vol. 4, no. 3, pp. 269–288, Jul. 1990.

[12] A. Bubic, D. Y. von Cramon, and R. I. Schubotz, "Prediction, cognition and the brain," *Frontiers Hum. Neurosci.*, vol. 4, p. 1094, Mar. 2010.

[13] S. L. Bressler and V. Menon, "Large-scale brain networks in cognition: Emerging methods and principles," *Trends Cogn. Sci.*, vol. 14, no. 6, pp. 277–290, Jun. 2010.

[14] R. E. Beaty, M. Benedek, P. J. Silvia, and D. L. Schacter, "Creative cognition and brain network dynamics," *Trends Cogn. Sci.*, vol. 20, no. 2, pp. 87–95, Feb. 2016.

[15] M. Piangerelli, M. Rucco, L. Tesei, and E. Merelli, "Topological classifier for detecting the emergence of epileptic seizures," *BMC Res. Notes*, vol. 11, no. 1, pp. 1–7, Dec. 2018.

[16] S. Majumder, F. Apicella, F. Muratori, and K. Das, "Detecting autism spectrum disorder using topological data analysis," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 1210–1214.

[17] F. Chazal and B. Michel, "An introduction to topological data analysis: Fundamental and practical aspects for data scientists," *Frontiers Artif. Intell.*, vol. 4, p. 108, Sep. 2021.

[18] P. Bubenik, "Statistical topological data analysis using persistence landscapes," *J. Mach. Learn. Res.*, vol. 16, no. 1, pp. 77–102, 2015.

[19] H. Adams, T. Emerson, M. Kirby, R. Neville, C. Peterson, P. Shipman, S. Chepushtanova, E. Hanson, F. Motta, and L. Ziegelmeier, "Persistence images: A stable vector representation of persistent homology," *J. Mach. Learn. Res.*, vol. 18, no. 8, pp. 1–35, Jan. 2017.

[20] S. Balakrishnan, A. Rinaldo, D. Sheehy, A. Singh, and L. Wasserman, "Minimax rates for homology inference," in *Proc. 15th Int. Conf. Artif. Intell. Stat.*, Apr. 2012, pp. 64–72.

[21] T. Bonis, M. Ovsjanikov, S. Oudot, and F. Chazal, "Persistence-based pooling for shape pose recognition," in *Proc. Int. Conf. Comput. Topol. Image Context Workshop*, Marseille, France, Jun. 2016, pp. 19–29.

[22] Z. Cang and G.-W. Wei, "TopologyNet: Topology based deep convolutional and multi-task neural networks for biomolecular property predictions," *PLoS Comput. Biol.*, vol. 13, no. 7, Jul. 2017, Art. no. e1005690.

[23] F. Chazal, M. Glisse, C. Labruère, and B. Michel, "Convergence rates for persistence diagram estimation in topological data analysis," in *Proc. Int. Conf. Mach. Learn.*, Beijing, China, Jun. 2014, p. 3604.

[24] A. Garin and G. Tauzin, "A topological, 'reading' lesson: Classification of MNIST using TDA," in *Proc. IEEE 18th Int. Conf. Mach. Learn. Appl.*, Boca Raton, FL, USA, Dec. 2019, pp. 1551–1556.

[25] F. Deligianni, M. Centeno, D. W. Carmichael, and J. D. Clayden, "Relating resting-state fMRI and EEG whole-brain connectomes across frequency bands," *Frontiers Neurosci.*, vol. 8, p. 258, Aug. 2014.

[26] B. He, L. Astolfi, P. A. Valdés-Sosa, D. Marinazzo, S. O. Palva, C.-G. Bénar, C. M. Michel, and T. Koenig, "Electrophysiological brain connectivity: Theory and implementation," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 7, pp. 2115–2137, Jul. 2019.

[27] J.-P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela, "Measuring phase synchrony in brain signals," *Hum. Brain Mapping*, vol. 8, no. 4, pp. 194–208, 1999.

[28] G. Carlsson, "Topology and data," *Bull. Amer. Math. Soc.*, vol. 46, no. 2, pp. 255–308, Apr. 2009.

[29] P. Y. Lum, G. Singh, A. Lehman, T. Ishkanov, M. Vejdemo-Johansson, M. Alagappan, J. Carlsson, and G. Carlsson, "Extracting insights from the shape of complex data using topology," *Sci. Rep.*, vol. 3, no. 1, pp. 1–8, Feb. 2013.

[30] M. G. Bergomi, P. Frosini, D. Giorgi, and N. Quercioli, "Towards a topological–geometrical theory of group equivariant non-expansive operators for data analysis and machine learning," *Nature Mach. Intell.*, vol. 1, no. 9, pp. 423–433, Sep. 2019.

[31] A. Zomorodian, "Fast construction of the Vietoris–Rips complex," *Comput. Graph.*, vol. 34, no. 3, pp. 263–271, Jun. 2010.

[32] H. Edelsbrunner and J. Harer, *Computational Topology: An Introduction*. Providence, RI, USA: American Mathematical Society, 2010.

[33] P. J. Hilton and S. Wylie, *Homology Theory: An Introduction to Algebraic Topology*. Cambridge, U.K.: CUP Arch., 1967.

[34] M. Horn, E. De Brouwer, M. Moor, Y. Moreau, B. Rieck, and K. Borgwardt, "Topological graph neural networks," 2021, *arXiv:2102.07835*.

[35] A. Zomorodian and G. Carlsson, "Computing persistent homology," in *Proc. 20th Annu. Symp. Comput. Geometry*, Jun. 2004, pp. 347–356.

[36] J.-G. Dumas, F. Heckenbach, D. Saunders, and V. Welker, "Computing simplicial homology based on efficient Smith normal form algorithms," in *Algebra, Geometry and Software Systems*, Berlin, Germany: Springer, Mar. 2003.

[37] C. Seng Pun, K. Xia, and S. Xian Lee, "Persistent-homology-based machine learning and its applications—A survey," 2018, *arXiv:1811.00252*.

[38] N. Otter, M. A. Porter, U. Tillmann, P. Grindrod, and H. A. Harrington, "A roadmap for the computation of persistent homology," *EPJ Data Sci.*, vol. 6, no. 1, pp. 1–38, Dec. 2017.

[39] L. Pessoa, "Understanding brain networks and brain organization," *Phys. Life Rev.*, vol. 11, no. 3, pp. 400–435, Sep. 2014.

[40] D. Chang, L. Dooley, and J. E. Tuovinen, "Gestalt theory in visual screen design—A new look at an old subject," in *Proc. 7th World Conf. Comput. Educ.*, Jul./Aug. 2002, pp. 1–9.

[41] M. Pettini, *Geometry Topology Hamiltonian Dynamics and Statistical Mechanics*. Cham, Switzerland: Springer, 2007.

[42] C. N. Yang and T. D. Lee, "Statistical theory of equations of state and phase transitions. I. Theory of condensation," *Phys. Rev.*, vol. 87, no. 3, pp. 404–409, Aug. 1952.

[43] F. A. N. Santos, E. P. Raposo, M. D. Coutinho-Filho, M. Copelli, C. J. Stam, and L. Douw, "Topological phase transitions in functional brain networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 100, no. 3, Sep. 2019, Art. no. 032414.

[44] J. Harer and D. Zagier, "The Euler characteristic of the moduli space of curves," *Inventiones Mathematicae*, vol. 85, no. 3, pp. 457–485, Oct. 1986.

[45] M. Rucco, F. Castiglione, E. Merelli, and M. Pettini, "Characterisation of the idiotypic immune network through persistent entropy," in *Proc. Eur. Conf. Complex Syst.* Cham, Switzerland: Springer, Sep. 2016, pp. 117–128.

[46] J. Reininghaus, S. Huber, U. Bauer, and R. Kwitt, "A stable multi-scale kernel for topological machine learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Boston, MA, USA, Jun. 2015, pp. 4741–4748.

[47] C. D. Hofer, R. Kwitt, and M. Niethammer, "Learning representations of persistence barcodes," *J. Mach. Learn. Res.*, vol. 20, no. 126, pp. 1–45, Jan. 2019.

[48] C. Hofer, R. Kwitt, M. Niethammer, and A. Uhl, "Deep learning with topological signatures," in *Proc. Conf. Adv. Neural Inf. Process. Syst.*, Long Beach, CA, USA, Dec. 2017, pp. 1–11.

[49] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[50] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[51] D. J. Trosten, S. Lokse, R. Jenssen, and M. Kampffmeyer, "Reconsidering representation alignment for multi-view clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Mar. 2021, pp. 1255–1265.

[52] P. A. Estevez, M. Tesmer, C. A. Perez, and J. M. Zurada, "Normalized mutual information feature selection," *IEEE Trans. Neural Netw.*, vol. 20, no. 2, pp. 189–201, Feb. 2009.

[53] D. J. Trosten, S. Løkse, R. Jenssen, and M. C. Kampffmeyer, "On the effects of self-supervision and contrastive alignment in deep multi-view clustering," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 23976–23985.

[54] A. Amanatiadis, V. G. Kaburlasos, and E. B. Kosmatopoulos, "Understanding deep convolutional networks through gestalt theory," in *Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST)*, Oct. 2018, pp. 1–6.

[55] D. Bau, B. Zhou, A. Khosla, A. Oliva, and A. Torralba, "Network dissection: Quantifying interpretability of deep visual representations," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3319–3327.

[56] E. Brunswik and J. Kamiya, "Ecological cue-validity of 'proximity' and of other Gestalt factors," *Amer. J. Psychol.*, vol. 66, no. 1, pp. 20–32, Jan. 1953.

[57] M. Brodeur, F. Lepore, and J. B. Debruille, "The effect of interpolation and perceptual difficulty on the visual potentials evoked by illusory figures," *Brain Res.*, vol. 1068, no. 1, pp. 143–150, Jan. 2006.

[58] S. Siuly, S. K. Khare, V. Bajaj, H. Wang, and Y. Zhang, "A computerized method for automatic detection of schizophrenia using EEG signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 11, pp. 2390–2400, Nov. 2020.

ZHIFENG ZHAO (Member, IEEE) received the B.E. degree in computer science, the M.E. degree in communication and information systems, and the Ph.D. degree in communication and information systems from the PLA University of Science and Technology, Nanjing, China, in 1996, 1999, and 2002, respectively. From 2002 to 2004, he was a Postdoctoral Researcher with Zhejiang University, Hangzhou, China, where his researches were focused on multimedia next-generation networks (NGNs) and softswitch technology for energy efficiency. From 2005 to 2006, he was a Senior Researcher with the PLA University of Science and Technology, where he performed research and development on advanced energy-efficient wireless router, ad-hoc network simulator, and cognitive mesh networking test-bed. From 2006 to 2019, he was an Associate Professor with the College of Information Science and Electronic Engineering, Zhejiang University. Currently, he is with the Zhejiang Laboratory, Hangzhou, as the Chief Engineering Officer. His research interests include software defined networks (SDNs), wireless network in 6G, computing networks, and collective intelligence. He is also the Symposium Co-Chair of ChinaCom, in 2009 and 2010. He is also the Technical Program Committee (TPC) Co-Chair of the Tenth IEEE International Symposium on Communication and Information Technology (ISCIT 2010).

RONGPENG LI (Member, IEEE) was a Research Engineer with the Wireless Communication Laboratory, Huawei Technologies Company Ltd., Shanghai, China, from August 2015 to September 2016. He was a Visiting Scholar with the Department of Computer Science and Technology, University of Cambridge, Cambridge, U.K., from February 2020 to August 2020. He is currently an Associate Professor with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China. His current research interests include networked intelligence for communications evolving (NICE). He received the sponsorship "The National Postdoctoral Program for Innovative Talents," in 2016, and the Wu Wenjun Artificial Intelligence Excellent Youth Award, in 2021, respectively. He serves as an Editor for *China Communications*.

HONGGANG ZHANG (Senior Member, IEEE) was a Honorary Visiting Professor with the University of York, York, U.K., and an International Chair Professor of Excellence with Université Européenne de Bretagne and Supélec, France. He is also the Chief Managing Editor of *Journal of Intelligent Computing* and *Science Partner Journals* and a Professor with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China. He has coauthored and edited two books: *Cognitive Communications: Distributed Artificial Intelligence (DAI), Regulatory Policy & Economics, Implementation* (John Wiley & Sons) and *Green Communications: Theoretical Fundamentals, Algorithms and Applications* (CRC Press), respectively. His research interests include cognitive radio and networks, green communications, mobile computing, machine learning, artificial intelligence, and the Internet of Intelligence (IoI). He was a co-recipient of the 2021 IEEE Communications Society Outstanding Paper Award and the 2021 IEEE INTERNET OF THINGS JOURNAL (IoT-J) Best Paper Award. He was a leading Guest Editor of the Special Issues on Green Communications of the *IEEE Communications Magazine*. He served as a Series Editor for the *IEEE Communications Magazine* (Green Communications and Computing Networks Series), from 2015 to 2018, and the Chair of the Technical Committee on Cognitive Networks for the IEEE Communications Society, from 2011 to 2012. He is the Associate Editor-in-Chief of *China Communications*.

• • •

FEI NI is currently pursuing the Ph.D. degree with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China. Her research interests include algebraic topology and deep learning.