Detecting High Frequency Oscillations for Stereoelectroencephalography in Epilepsy via Hypergraph Learning

Jiayang Guo, Hailong Li, Xiaoshuai Sun, Lei Qi, Hui Qiao, Yijie Pan, Jing Xiang, and Rongrong Ji, *Senior Member, IEEE*

***Abstract*— Successful epilepsy surgeries depend highly on pre-operative localization of epileptogenic zones. Stereoelectroencephalography (SEEG) records interictal and ictal activities of the epilepsy in order to precisely find and localize epileptogenic zones in clinical practice. While it is difficult to find distinct ictal onset patterns generated the seizure onset zone from SEEG recordings in a confined region, high frequency oscillations are commonly considered as putative biomarkers for the identification of epileptogenic zones. Therefore, automatic and accurate detection of high frequency oscillations in SEEG signals is crucial for timely clinical evaluation. This work formulates the detection of high frequency oscillations as a signal segment classification problem and develops a hypergraph-based detector to automatically detect high frequency oscillations such that human experts can visually review SEEG signals. We evaluated our method on 4,000 signal segments from clinical SEEG recordings that contain both ictal and interictal data obtained from 19 patients who suffer from refractory focal epilepsy. The experimental results demonstrate the effectiveness of the proposed detector that can successfully localize interictal high frequency oscillations and outperforms multiple peer machine learning methods. In particular, the proposed detector achieved 90.7% in accuracy, 80.9% in sensitivity, and 96.9% in specificity.**

***Index Terms*— High frequency oscillations, epilepsy, hypergraph learning, deep learning, stereoelectroencephalography.**

I. Introduction

EPILEPSY is the fourth most common neurological disorder, which is characterized by unpredictable seizures. It impacts people of all age, and many patients with epilepsy are medically intractable and require respective neurosurgery to gain seizure freedom [1]. Epilepsy surgery relies critically on pre-operatively localizing epileptogenic zones [1]. Thus, presurgical biomarkers play an essential role in identifying the epileptogenic zones that result in epileptic seizures. While it is difficult to directly measure and localize the epileptogenic zone by insufficiently concordant or inconclusive data from a multitude of tests, intracranial electroencephalography (iEEG) technologies [2], [3] is widely adopted. Among existing iEEG technologies, stereoelectroencephalography (SEEG), which records interictal and ictal activities of the epilepsy is increasingly used to define the epileptogenic zone in complex cases, providing a lower rate of complications than subdural grids [4]. Using the conventional spikes in SEEG to localize epileptogenic zones results in seizure freedom for 40-70% of surgical patients during epilepsy surgery [5]–[7]. Therefore, there is an overriding need for new reliable biomarkers to robustly localize epileptogenic zones.

Recent studies have demonstrated that high frequency oscillations (HFOs) can be used as a putative biomarker to indicate epileptogenic zones [8]–[15]. Thus, it is desirable to correctly identify HFOs in SEEG signals for evaluating epileptic patients. However, it remains challenging to detect high-frequency brain signals. First, HFOs require a high sampling rate to digitize signal data. The high sampling rate generate a substantial amount of data [16], [17], posing a challenge for SEEG data analysis. Second, HFOs are typically much weaker than the conventional low frequency brains signals (e.g., spikes), making it hard to be identified by human experts easily. Third, manual analysis of HFOs is a subjective review, resulting in poor inter-reviewer agreement. These challenges lead to subjective, time consuming, and error

Manuscript received September 6, 2020; revised November 24, 2020; accepted November 28, 2020. Date of publication February 3, 2021; date of current version March 8, 2021. This work was supported in part by the National Science Fund for Distinguished Young Scholars under Grant 62025603; in part by the National Natural Science Foundation of China under Grant U1705262, Grant 62072386, Grant 62072387, Grant 62072389, Grant 62002305, Grant 61772443, Grant 61802324, and Grant 61702136; and in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2019B1515120049. *(Corresponding authors: Hailong Li; Rongrong Ji.)*

Jiayang Guo and Rongrong Ji are with the Media Analytics and Computing Laboratory, Department of Artificial Intelligence, School of Informatics, Xiamen University, Xiamen 361005, China, also with the Peng Cheng Laboratory, Shenzhen 518066, China, and also with the Institute of Artificial Intelligence, Xiamen University, Xiamen 361005, China (e-mail: rrji@xmu.edu.cn).

Hailong Li is with the Perinatal Institute, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229 USA (e-mail: hailong.li@cchmc.org).

Xiaoshuai Sun is with the Media Analytics and Computing Laboratory, Department of Artificial Intelligence, School of Informatics, Xiamen University, Xiamen 361005, China.

Lei Qi and Hui Qiao are with the Beijing Neurosurgical Institute, Capital Medical University, Beijing 100050, China.

Yijie Pan is with the Department of Technology Development, Ningbo Institute of Information Technology Application, Chinese Academy of Sciences (CAS), Ningbo 315040, China.

Jing Xiang is with the MEG Center, Department of Neurology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229 USA.

Digital Object Identifier 10.1109/TNSRE.2021.3056685

prone identification of HFOs in the current clinical setting. Therefore, we aim to develop an automatic and objective detection approach to accurately identify HFOs in complex SEEG data.

Although HFOs detection is left unexploited for SEEG data analysis, a number of automatic approaches [15], [18], [19], have been reported to detect HFOs from electroencephalography (EEG) and magnetoencephalography (MEG) signals. Most of these methods share a similar framework, which extracts discriminative features for decision making, given a signal segment that is retrieved from the whole data. Typically, these approaches employ handcrafted features that were manually obtained from observation or statistical analysis. These features were commonly referred as handcrafted features. For example, Klink *et al.* [18] proposed an automatic HFOs detection and visualization approach in MEG, while Quitadamo *et al.* [20] proposed a Kurtosis-based HFOs detector to exploit statistical properties of the pre-ictal and ictal iEEG time series. Burnos *et al.* [21] used handcrafted features (e.g., high frequency peak and low frequency peak) to automatically distinguish HFOs in EEG. Unfortunately, these handcrafted feature-based approaches require a threshold that must be adjusted or re-optimized to recognize HFOs in EEG and MEG data. Such a circumstance hinders the generalizability of HFOs detection models. Additionally, there is no theoretical evidence derived from these handcrafted features to guarantee an optimal performance of HFOs detection.

Recently, machine learning techniques have been proposed to identify HFOs and reduce human interference. Our recent study discussed a stacked sparse autoencoder-based method for HFOs detection in complex MEG signals [22]. The prior detection technique required minimal human interference by using a golden standard dataset to train the detector and estimate the pairwise distance between the objects of the interest (e.g., SEEG segments), However, it ignored the higher order relationship among samples. In many real-world tasks, representing a set of objects only based on pairwise relationship possibly causes information loss [23]. Hypergraph learning [24] has been proposed to represent the complex relationship between objects. Nevertheless, it remains unknown if hypergraph learning can improve the performance of HFOs detection by considering complex relationships between SEEG segments. Besides, distance measurement between objects are essential in hypergraph construction process. While euclidean distance is one of the most common distance metrics for this purpose, it remains unclear what suits best for an optimal HFOs detection, considering the unique characteristic of biological signals (i.e., SEEG signals). Meanwhile, regularization terms [25], [26] have been introduced into hypergraph learning to control weight matrix complexity. However, it retains as an open problem whether these regularization terms could achieve desirable performance for the HFOs detection task.

This study is to develop a hypergraph-based SEEG HFOs (HSO) detector, aiming to automatically detect HFOs signals to assist human experts for the visual review of SEEG signals. Our key innovation is proposing a hypergraph learning framework to identify HFOs in SEEG signals. First, our method separates SEEG data into a series of signal segments. Second, we evaluate the distance measurement between individual SEEG segments. By utilizing the hyper-relationship of signal segments, our HSO detector classifies each signal segment into either normal-control signals or epileptic HFOs in the time domain. Finally, the identified HFOs are delineated at their original position with SEEG visualization software (e.g., EEG processor [27]). Our proposed HSO detector is not dependent on handcrafted signal features, enabling an objective and automatic detection of HFOs for the pre-operative evaluation. The contribution of this study is three folds: First, we introduced a hypergraph learning framework to detect HFOs from SEEG data. Second, a robust adaptive signed correlation index (ASCI)-based distance measurement approach was utilized to quantify the morphological similarity between SEEG signals. Third, a new regularization term was applied to improve performance of HFOs detection.

The rest of this paper is organized as follows: Firstly, the detailed hypergraph learning framework is presented in the method section. Secondly, we describe the patients and their associated SEEG data in this work. Experiment configurations, such as model evaluation and peer models, are described. Thirdly, we present the HFOs detection performance of the proposed HSO detector as well as related machine learning models. Comprehensive ablation studies are presented to test our model. Finally, we discuss the discoveries and limitation of our work.

II. Hypergraph-Based SEEG HFOs (HSO) Detector

A. Hypergraph Learning

This section briefly reviews the method of the hypergraph learning. A simple graph, in which an edge only connects two related vertices, can be considered as a special case of the hypergraph. The simple graph can be undirected or directed, depending on whether the pairwise relationships between objects are symmetric [23]. Since the graph can represent data distribution precisely, the learning assignment can be performed on the graph. For instance, when classifying objects that are represented by vectors in a feature space, an undirected graph can be constructed based on the pairwise distances, upon which various graph-based learning methods can be applied [23]–[25], [28]. However, in many real-world applications, such a pairwise relationship of a simple graph could not capture the high-order information between objects. Simply representing the complex relationships by pairwise links may inevitably lead to the loss of information that could be valuable in many learning tasks [23], [29], [30]–[32]. Compared with a simple graph, the hypergraph uses a hyperedge to connect any number of vertices (two or more). Table I summarizes the notations and the definitions for hypergraph learning.

A hypergraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ consists of three parts: the vertex set $\mathcal{V}$, the hyperedge set $\mathcal{E}$, and the hyperedge weight set w. In the hypergraph $\mathcal{G}$, a weight value $w(e)$ is assigned to each hyperedge e. A $|\mathcal{V}| \times |\mathcal{E}|$ incidence matrix H which

TABLE I
NOTATIONS AND DEFINITIONS USED IN THE PAPER

Notation	Definition
$\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$	$\mathcal{G}$ is the representation of a hypergraph, where $\mathcal{V}$ and $\mathcal{E}$ denote the sets of vertices and hyperedges, respectly,and w denotes hyperedge weight vector.
$\mathcal{V}$	The set of vertices located in the hypergraph, which contains n objects.
$\mathcal{E}$	The set of hyperedges, which contains m elements.
n	The number of objects in the data, i.e.,$\lvert\mathcal{V}\rvert$.
m	The number of hyperedges, i.e.,$\lvert\mathcal{E}\rvert$.
$w(e)$	The weight of hyperedge e.
$\delta(e)$	The degree of hyperedge e.
$d(v)$	The degree of vertex v.
H	The incidence matrix of the hypergraph.
L	The Laplacian matrix of the hypergraph.
D_v	The diagonal matrix of the vertex degrees.
D_e	The diagonal matrix of the edge degrees.
D_w	The diagonal matrix of the hyperedge weights.
$Y = \begin{bmatrix} y_1, & y_2, & \ldots, & y_n \end{bmatrix}^T$	The $n \times 1$ label vector for hypergraph learning.
$F = \begin{bmatrix} f_1, & f_2, & \ldots, & f_n \end{bmatrix}^T$	The $n \times 1$ to-be-learned label vector for hypergraph learning.

represents the hypergraph $\mathcal{G}$ is defined by [24]:

$$H(v,e) = \begin{cases} 1 & \text{if } v \in e \\ 0 & \text{if } v \notin e \end{cases} \tag{1}$$

According to the definition of H, for any vertex $v \in \mathcal{V}$, its degree is defined by [24]:

$$d(v) = \sum_{e \in \mathcal{E}} w(e) H(v,e), \tag{2}$$

Similarly, for any hyperedge $e \in \mathcal{E}$, its degree is defined by [24]:

$$\delta(e) = \sum_{v \in \mathcal{V}} H(v,e), \tag{3}$$

In hypergraph learning, D_v and D_e denote the diagonal matrix of the vertex degrees and the diagonal matrix of the edge degrees, respectively. D_w is used to denote the $|\mathcal{E}| \times |\mathcal{E}|$ diagonal matrix of the hyperedge weights.

Due to the property of representing high-order relationship, the hypergraph can be applied to different machine learning tasks, such as classification, clustering, ranking, and embedding. Here, we use the binary classification as an example. The hypergraph learning framework with regularization is defined as [23]:

$$\arg\min_f \{\lambda R_{emp}(f) + \Omega(f)\}, \tag{4}$$

where $R_{emp}(f)$ is an empirical loss, $\lambda > 0$ is a tradeoff parameter to balance the two terms in equation (4), $\Omega(f)$ is a regularizer on the hypergraph, f is the classification function, which generates predicted labels. Specifically, the regularization term in the hypergraph is defined as [24]:

$$\Omega(f) = \frac{1}{2} \sum_{e \in \mathcal{E}} \sum_{u,v \in \mathcal{V}} \frac{w(e) H(u,e) H(v,e)}{\delta(e)} \left(\frac{f(u)}{\sqrt{d(u)}} - \frac{f(v)}{\sqrt{d(v)}} \right)^2, \tag{5}$$

By letting $\Delta = D_v^{-\frac{1}{2}} H D_w D_e^{-1} H^T D_v^{-\frac{1}{2}}$ and $L = I - \Delta$, the normalized regularization term can be defined as [24]:

$$\Omega(f) = f^T L f, \tag{6}$$

where L is the hypergraph Laplacian matrix and a positive semi-definite matrix. Meanwhile, the difference between the vector of predicted labels f and the vector of original labels y is measured by empirical loss $R_{emp}(f)$, which is defined as [24]:

$$R_{emp}(f) = \|f - y\|^2. \tag{7}$$

B. The Framework of HSO Detector

This section introduces the proposed approach for HFOs detection. Fig. 1 shows the framework of our HSO detector. Our model consists of four steps (1) SEEG signal segmentation; (2) distance measurement; (3) hypergraph construction; and (4) hypergraph learning. In step (1), SEEG signals are separated into a series of small signal segments. The label of each sample in the gold standard dataset is prepared by clinicians. In step (2), the distance measurement is calculated between every pair of the signal segments based on ASCI similarity measurement. In step (3), the hypergraph is constructed by generating a set of hyperedges from each signal segment and its neighbors. In step (4), the hypergraph Laplacian is applied to the classification framework, and the label probability of testing data and the weights of the hyperedges are simultaneously optimized through hypergraph learning.

1) Signal Segmentation: Using a moving-window technique [33], the SEEG signal from multiple channels could be segmented into a series of signal segments. For the hypergraph learning purpose, the clinical epileptologists pre-selected a number of HFOs and baseline control (BC) signal segments based on the invasive recordings and surgical outcomes. A total of n signal segments (HFOs and BC samples) were composed of a golden standard data set for model development and evaluation. The duration (i.e., window size) of each signal segment is w milliseconds. Namely, both HFOs and BC signal segments

are represented by time series vectors in time domain. Different window sizes and overlaps could be adjusted. In the current work, we applied a time moving window without overlap to segment the processed SEEG signal. The selection of the moving window size is very challenging. If the window is too small, it may not cover visible HFOs signals. If the window is too big, the sample signals contain too much redundant information which may negatively impact machine learning models to understand HFOs. After reviewing multiple previous studies [34]–[36], as well as our own data, we set 1,000 milliseconds as a pre-selected size of moving window. In this way, in spite of various length of SEEG recording for different subjects, the output from the signal segmentation component of HSO detector are all unified to a signal segments with 1,000 milliseconds length. Details of the data are described in section III-A.

2) Distance Measurement: In this work, we used a new distance measurement inspired by ASCI, which is demonstrated to be more suitable for sequential SEEG data compared to the traditional L2 measurement used in classic hypergraph learning framework. ASCI [37] is a normalized similarity index ranging from -1 to 1, accounting for the amplitude difference between signal segments. By taking the advantage of this amplitude dependency, we employ ASCI to measure the morphological similarity between two SEEG signal segments in hypergraph learning. Assume v_a and v_b are the Z-score transformation of two signal segments in the same hyperedge, and let $ASCI(v_a, v_b)$ denote the ASCI between these two signal segments. Signal trichotomization is the first step of ASCI measurement between two samples signals. Let R be the signal amplitude space. We divide R into upper space R_u, middle space R_m, and lower space R_l. $U = u(i)$ and $L = l(i)$ are pre-defined threshold vectors at time instant i, (e.g., $u(i) = 0.25$ and $l(i) = -0.25$). The three sub-spaces are defined as: upper sub-space $R_u = \{V > U\}$, middle sub-space $R_m = \{L \leq V \leq U\}$, and lower sub-space $R_l = \{V < L\}$. The trichotomization $T_a = \{t_a(i)\}$ of signal $v_a = \{v_a(i), 1 \leq i \leq N\}$ is defined by [37]:

$$t_a(i) = \begin{cases} 1 & \text{if } v_a(i) \in R_u \\ 0 & \text{if } v_a(i) \in R_m \\ -1 & \text{if } v_a(i) \in R_l \end{cases} \tag{8}$$

Then, the trichotomization T_b of $v_b = \{v_b(i), 1 \leq i \leq N\}$ can be computed in a similar way. The ASCI between SEEG signal segments v_a and v_b is calculated by [37]:

$$ASCI(v_a, v_b) = \frac{T_a \; o \; T_b}{\sqrt{T_a \; o \; T_a} \times \sqrt{T_b \; o \; T_b}}, \tag{9}$$

where $\otimes$ is the signed correlation between two trichotomized vectors T_a and T_b

$$T_a \; o \; T_b = \sum_{i=1}^{N} t_a(i) \; \otimes \; t_b(i), \tag{10}$$

where $\otimes$ is the signed product of T_a and T_b and defined by [37]:

$$t_a(i) \; \otimes \; t_b(i) = \begin{cases} 1 & \text{if } t_a(i) = t_b(i) \\ -1 & \text{if } t_a(i) \cdot t_b(i) = -1 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

At any time instant, if $v_a(i)$ and $v_b(i)$ are in the same subspace, they are called a concordant pair. Two signal samples are termed discordant if one of them is in upper sub-space and the other one is in lower sub-space. The sample pair is, otherwise, called nilcordant. Apparently, ASCI is increased by more concordant pairs, the unaltered by nilcordant pairs, and decreased by discordant pairs. The ASCI is closer to 1 when two SEEG signal segments demonstrate high morphological similarity, and it approaches -1 when two SEEG signal segments have low similarity. In this way, the morphological similarity between two segments can be effectively measured. Based on the morphological similarity between two signal segments, the correlation distance between v_a and v_b is defined as follow:

$$d_{ASCI}(v_a, v_b) = 1 - ASCI(v_a, v_b). \tag{12}$$

When v_a and v_b have high morphological similarity, the correlation distance $d_{ASCI}(v_a, v_b)$ is closer to 0. Otherwise, when the two signal segments demonstrate low morphological similarity, the correlation distance $d_{ASCI}(v_a, v_b)$ approaches 2.

3) Hypergraph Construction: Hypergraph construction is an important step during the hypergraph learning process. When constructing hypergraph, we consider each signal segment in the data set as a vertex in the hypergraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$. For instance, if there are n signal segments in the data set, then the generated hypergraph consists of n vertices. Let $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ denotes n signal segments with known clinical labels. In our approach, a star-expansion strategy [38] was applied to build a set of hyperedges by exhaustively visiting each vertex v_n in $\mathcal{V}$. Specifically, the vertex v_n in $\mathcal{V}$ was chosen as a centroid vertex and each hyperedge e_n was generated by connecting v_n and its nearest k neighbors. In this way, the hyperedges set $\mathcal{E} = \{e_1, e_2, \ldots, e_n\}$ was formed and each hyperedge e_n connected $k + 1$ vertices.

After the hypergraph was constructed, each hyperedge e_n associates with a positive weight value $w(e)$. This value was estimated by the sum of the similarities between two signal segments, defined by:

$$w(e) = \sum_{v_a, v_b \in e} \exp\left(-\frac{d_{ASCI}(v_a, v_b)^2}{\varphi^2}\right). \tag{13}$$

where v_a and v_b are two signal segments in the same hyperedge. $d_{ASCI}(v_a, v_b)$ is the correlation distance between these two signal segments. In traditional hypergraph learning [24], [26], Euclidean distance is a common metric to measure the distance $d(v_a, v_b)$ between two samples. However, in our study, due to the unique characteristic of biological signals, Euclidean distance may not be an optimal metric to measure the similarity between two signal segments. Thus,

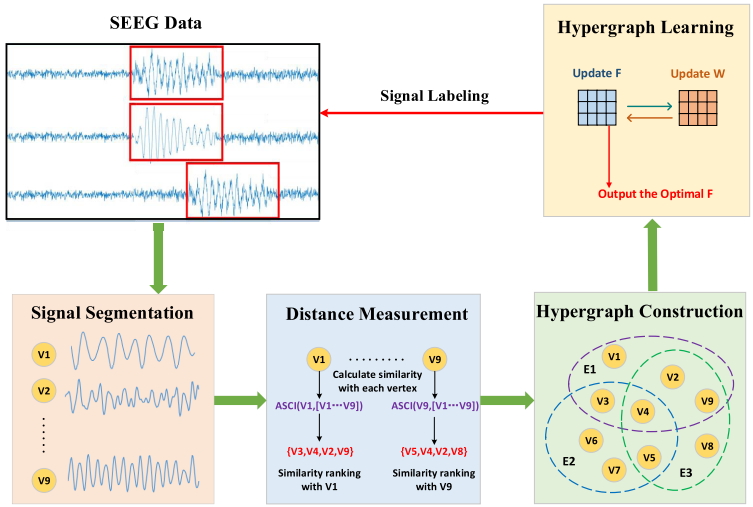

Fig. 1. The overview of hypergraph based learning method for detecting HFOs in SEEG data.

we introduce $d_{ASCI}(v_a, v_b)$ to measure the similarity between SEEG signal segments. In the following sections, we would demonstrate that $d_{ASCI}(v_a, v_b)$ is an effective way to improve the detection of HFOs in the hypergraph-based model. The parameter φ is empirically set to the median value of the distances of all signal segment pairs. The aforementioned process shows that a higher weights would be assigned to the hyperedges if the signal segments in the hyperedge have closer relationships.

4) Hypergraph Learning: In our paper, we consider the HFOs detection task as a binary classification problem [23]. Therefore, the HFOs detection can be formulated as a regularized optimization problem $\arg\min_f \{\lambda R_{emp}(f) + \Omega(f)\}$, where $\Omega(f)$ is the regularization term and $R_{emp}(f)$ is a loss function term. According to Equation 5, we find that the regularization term $\Omega(f)$ is a function which contains with the parameter F and the parameter $w(e)$. In the the regularization term $\Omega(f)$, the term $\sum_{u,v\in\mathcal{V}} \frac{1}{2}\frac{w(e)}{\delta(e)}\left(\frac{f(u)}{\sqrt{d(u)}} - \frac{f(v)}{\sqrt{d(v)}}\right)^2$ measures the quality of the hyperedges e. When vertex u and vertex v belong to the same hyperedge e. A hyperedge with good quality means vertex u and vertex v are in the same class, which makes the term $\left(\frac{f(u)}{\sqrt{d(u)}} - \frac{f(v)}{\sqrt{d(v)}}\right)$ approach close to zero. During the hypergraph learning process, our aim is to minimize the objective function, so a straightforward method is to optimize the parameter F and the parameter w simultaneously. In the previous hypergraph learning study, yu *et al.* [25] proposed a 2-norm regularizer to enforce the sparsity on hyperedge weight matrix. The 2-norm regularizer controls the model complexity motivated by the success of sparse learning. Gao *et al.* [26] proposed a L_2 regularizer to learn the optimal hyperedge weight. The L_2 regularizer enables the model to optimize the parameter F and the parameter w simultaneously. Combining with the advantages of the above two approaches, a L_2 hypergraph weight regularizer $\phi(w)$ can be added in the objective function as follows:

$$\arg\min_{F,w}\{\lambda R_{emp}(f) + \Omega(f) + \mu\phi(w)\}$$

$$\phi(w) = (1-\rho)\|D_w\|^2 + \rho\sum_{i=1}^{m} w_i(e)^2$$
$$s.t. \sum_{i=1}^{m} w_i(e) = 1, \quad 0 \le \rho \le 1, \quad \lambda > 0, \ \mu > 0 \qquad (14)$$

where $0 \le \rho \le 1$ is a positive control parameter. When $\rho = 0$, the hypergraph weight regularizer becomes $\|D_w\|^2$, which induces sparse solutions for hypergraph weight matrix. When $\rho = 1$, the hypergraph weight regularizer becomes $\sum_{i=1}^{m} w_i(e)^{\Theta}$, which improves the performance of classification by learning the optimal hyperedge weight. The loss function term can be defined as:

$$R_{emp}(f) = \|F - Y\|^2 = \sum_{v\in\mathcal{V}} (f(v) - y(v))^2 . \qquad (15)$$

where F is the vector of predicted labels and Y is the vector of ground-truth labels. For each vertex $v \in \mathcal{V}$, $y(v)$ returns its clinical label, where $y(v) = +1$ and $y(v) = -1$ denote HFO and baseline control signal, respectively. Meanwhile, to measure the efficiency of our hypergraph learning method, we need to keep some data for testing. Since the labels on these data are unknown, we use '0' as the labels to represent these unknown data. Let n denote the number of the training set, and $Y = \left[y_1, y_2, \ldots, y_n\right]^T$ is a $n \times 1$ vector.

Our hypergraph learning based method detects HFOs by inspecting the performance of each hyperedge in separating the signal segments into two classes. To achieve this goal, we assumed the to-be-learned label on each signal was unknown, and then we used our hypergraph learning method to estimate the likelihood f_n for each vertex v_n in the hypergraph. Since the likelihood f_n is a probability index, then the value of f_n is between -1 and $+1$. To classify HFOs and BC signals, there are two cases here. If $0 < f_n \le 1$, the vertex v_n will be assigned to HFO. And if $-1 \le f_n \le 0$, the vertex v_n will be assigned to baseline control signal.

According to Equation 15 and Equation 6, the learning task for detecting HFOs is to minimize the sum of the three loss

Algorithm 1 The Iterative Algorithm of Equation (9)

Input: The vertex set $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ and the clinical label vector $Y = [y_1, y_2, \ldots, y_n]^T$.

Output: The optimized value of F and w.

1: Applying star-expansion strategy to the vertex set $\mathcal{V}$ for constructing hypergraph $\mathcal{G}$ and calculating H, D_v, D_e, D_w, L.
2: Initialize $F^{(k)}$ and $w^{(k)}$, when $k = 0$.
3: **while** not convergence **do**
4: $F^{(k+1)}) = \frac{1}{1+\lambda}(I - L)F^{(k)} + \frac{\lambda}{1+\lambda}Y$.
5: Update the $w^{(k+1)})$ based on equation (15).
6: $k = k + 1$.
7: **end while**

function terms as follow:

$$\Psi_w(F) = \lambda \|F - Y\|^2 + F^T LF + \mu\{(1-\rho)\|D_w\|^2 + \rho \sum_{i=1}^{m} w_i(e)^2\}.$$
$$s.t. \sum_{i=1}^{m} w_i(e) = 1, \quad 0 \le \rho \le 1, \quad \lambda > 0, \ \mu > 0 \tag{16}$$

where $\lambda > 0$ and $\mu > 0$ are two positive weight parameters. By differentiating $\Psi_w(F)$ with respect to F, we obtain the following equation:

$$\frac{\partial \Psi}{\partial F}\left(\lambda \|F - Y\|^2 + F^T\left(I - D_v^{-\frac{1}{2}} H D_w D_e^{-1} H^T D_v^{-\frac{1}{2}}\right)F\right) = 0 \Rightarrow F = \left(I + \frac{1}{\lambda}L\right)^{-1} Y \tag{17}$$

Meanwhile, by differentiating $\Psi_w(F)$ with respect to w, we obtain the Equation:

$$\frac{\partial \Psi}{\partial w}\left(\mu\{(1-\rho)\|D_w\|^2 + \rho \sum_{i=1}^{m} w_i(e)^2\} + F^T LF\right) = 0 \tag{18}$$

Here, let $\Theta = F^T D_v^{-\frac{1}{2}} H$ and $\Upsilon = \left(\sum_{i=1}^{m} w_i(e)^{\Theta}\right)^{-\frac{1}{2}}$, the optimization problem can be derived as:

$$-D_{e(i,i)}^{-1}\Theta_i^2 + \mu\left(2\rho + (1-\rho)\Upsilon\right) w_i(e) = 0 \tag{19}$$

Now, We can obtain the following Equation:

$$w_i(e) = \frac{D_{e(i,i)}^{-1}\Theta_i^2}{\mu\left(2\rho + (1-\rho)\Upsilon\right)}. \tag{20}$$

where Θ_i denotes the i-th column of the $m \times m$ matrix Θ. When we get F and w, similar to many existing approaches [39], [40], we can optimize F and w simultaneously by an iterative algorithm. The optimizing process is described in Algorithm 1.

III. Experimental Setups

This section presents the experimental setups, including our SEEG data set, model evaluation, and peer methods. We implement the proposed hypergraph model, as well as peer models with the optimized model parameters. By using the same data set, we provide a fair comparison on the performance of HFOs detection.

A. Data

SEEG data were acquired from 19 clinical patients (3-33 years old, with a mean age of 21.2 years) who suffer from refractory focal epilepsy for clinical purposes. All patients had at least one seizure during the SEEG recording. Consequently, the SEEG data included both ictal and interictal data. The present study focused on interictal SEEG data and was approved by the institutional review board at Beijing Neurosurgical Institute. A presurgical evaluation which includes symptom analysis, surface EEG, anatomic magnetic resonance imaging (MRI), positron emission computed tomography (PET-CT), and other evaluations was conducted to all the patients. For the patients who showed no clear epileptogenic zone in noninvasive evaluations, depth electrodes (8–16 contacts, 0.8 mm diameter, 3.5 mm intercontact distance) with the Integra CRW System (NeuroSight Arc 2.7.1, Integra LifeSciences, SM USA) were implanted in their brain during SEEG recording. The target and trajectory of the electrodes were determined by non-invasive evaluation and hypotheses about the localization of epileptogenic foci and network. SEEG data were acquired by a 256-channel NK EEG 2100 system (Nihon-Kohden, Japan) with a sampling rate of 2,000 Hz. All the patients experienced at least 2 habitual seizures (from 2 to 18 times) during video-SEEG monitoring. Based on the timing of habitual seizures, the monitoring time varied from 2 to 30 days.

All data were de-identified before data analysis. We use different analytical methods in ictal and interictal SEEG data,respectively. For analyzing interictal activities of the epilepsy, 2-3 epochs within 2 minutes long are selected in slow-wave sleep stage, which were 2 hours after a seizure to eliminate the influence of seizure. For analyzing ictal activities of the epilepsy, the seizure onset is determined by video monitoring of clinical symptoms and EEG ictal onset patterns. We used ictal data to determine the seizure onset zones (SOZ). In this study, 142 SEEG probes with 1654 contacts were implemented in the patients. The average number of contacts (electrodes) implemented for each patient was 88. We analyzed 1482 contacts (electrodes) for all patients. Approximately, 67 electrodes were analyzed per patient. At least two experienced epileptologists reviewed all SEEG data and excluded noise or artifacts. The two epileptologists also reviewed all seizures recorded in the patients, and SOZ were localized by SEEG signals in 1-70 Hz during their habitual seizures. SOZ were considered as putative epileptogenic zones. The clinical true epileptogenic zones were validated on the basis of surgical outcomes. Specifically, if the resection of putative epileptogenic zones resulted in seizure freedom after surgery for at least one year (Engel surgical outcome

scale $\leq$ 1a/ILAE surgical outcome scale $\leq$ 1), the putative epileptogenic zones are considered as the true clinical epileptogenic zones.

HFOs were analyzed with band-pass filter of 80-500 Hz using EEG studio (MEG Center, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA) [27]. For the model evaluation purpose, only HFOs in the epileptogenic zones were labeled as epileptogenic HFOs in our detector. Meanwhile, the clinical epileptologists selected HFOs and BC signal segments based on the SEEG recordings and surgical outcomes. Each signal segment contains a series of 2,000 signal time points. A total of 4000 signal segments (1640 HFOs samples and 2360 BC samples) were composed of a gold standard data set for model evaluation.

B. Model Evaluation

For model evaluation, we adopt a k-fold cross validation scheme by using the gold-standard SEEG data set. Specifically, we split the whole data set into k portions. While one portion of data are reserved for model testing, the rest of k-1 portions are utilized for model training and validation. This procedure is repeated until all the portions of data are tested once.

Accuracy, sensitivity and specificity are calculated to quantify the performance of the proposed detectors. In the experiment, we evaluate true positive (TP), false positive (FP), true negative (TN), and false negative (FN) for the classification by comparing the predicted labels and ground-truth labels. Then, the accuracy, sensitivity and specificity are calculated by:

$$\begin{aligned} Accuracy &= \frac{TP + TN}{TP + TN + FP + FN} \\ Sensitivity &= \frac{TP}{TP + FN} \\ Specificity &= \frac{TN}{TN + FP} \end{aligned} \tag{21}$$

C. Compared Methods and Setting in HFOs Detection

To compare our hypergraph learning method with peer machine learning models, we also implemented k-nearest neighbors [41], neural networks (shallow layers) [42], logistic regression [43], linear and non-linear support vector machine (SVM) models [44], and our previously proposed deep learning-based SMO detector [22].

1) K-Nearest Neighbor: In the k-nearest neighbor model, we applied Euclidean distance as the distance measurement between signal segments. The number of nearest neighbors was optimized from 5 to 10.

2) Neural Networks: We designed a two-layer shallow neural networks. The input layer contains 200 nodes. We set the number of hidden nodes in the hidden layer based on empirical values [100, 50, 25, 10]. The neural network was trained with the stochastic gradient decent algorithm. Learning rate was 0.01, and training epoch was set as 500.

3) Logistic Regression: We utilized the Maximum-likelihood estimation algorithm to optimize the coefficient of logistic regression model.

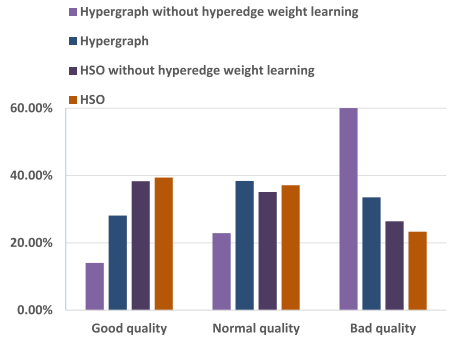

Fig. 2. The distribution of different quality of hyperedges under different hypergraph learning schemes.

4) Linear and Nonlinear Support Vector Machine (SVM): For the linear SVM model, we searched the margin penalty with empirical values $[2^{-5}, 2^{-4}, \ldots, 2^4, 2^5]$. For nonlinear SVM, we utilized the Radial Basis Function (RBF) kernel. By using a grid search method, we searched the margin penalty $[2^{-5}, 2^{-4}, \ldots, 2^4, 2^5]$, and kernel scale gamma [0.5, 1.0, 1.5, 2.0, 2.5].

5) SMO: our previous work is a 4-layer SSAE-based neural networks with an input layer, three hidden layers and an output layer. The number of nodes in the hidden layer was set to 30 based on our previous work. The building process includes pre-training, supervised learning, and fine tuning steps. A loss function with L2 regularization and sparsity regularization terms were adopted from [45]. During model optimization, we tested sparsity proportion [0.1, 0.2, 0.3, 0.4, 0.5] and L2 regularization weight [0.1, 0.2, 0.3, 0.4, 0.5]. 500 epoches were sufficient for to guarantee the convergence of the model and the learning rate was set to 0.01.

IV. Experimental Results

In this section, we conducted comprehensive experiments for evaluating our HSO detector on separating HFOs from BC signals. We first validated the SEEG data set. Then, ablation tests on HSO detector were presented. To explore the effectiveness of our proposed method under different parameter setting, the study of parameter exploration were also provided. At the end, experimental results of different machine learning methods are described in Section IV-F.

A. SEEG Data Validation

At the beginning of the experiment, we set the value of selected neighbors K to 5 and constructed 1000 hyperedge for representing the complex relationship between SEEG signal segments. Some hyperedges may bring negative effects for detecting HFOs, since the regularized term in hypergraph learning enforces the labels of signal segments connected by the hyperedge to be close, a hyperedge that contains both HFOs samples and BC samples may degrade the performance of detecting HFOs. In order to measure the effectiveness of constructed hyperedges, we divided the hyperedges into three categories: (1) Good quality hyperedges which consist of 5 HFOs samples or 5 BC samples. (2) Normal quality hyperedges which consist of 4 samples selected from HFOs or BC and 1 sample from other category. (3) Bad quality hyperedges which consist of 3 samples selected from

TABLE II
CLASSIFICATION PERFORMANCE COMPARISON BETWEEN FOUR DISTANCE MEASUREMENT APPROACHES

	Accuracy	Sensitivity	Specificity
HSO (Pearson)	87.1%	77.8%	95.5%
HSO (Spearman)	87.6%	77.4%	96.2%
HSO (Euclidean)	88.1%	75.1%	97.4%
HSO (ASCI)	**90.7%**	**80.9%**	**96.9%**

HFOs or BC and 2 samples from the other category. In our experiment, we adapted hyperedge weight learning to remove the ineffective hyperedges. During hyperedge weight learning process, the weights of some hyperedges may become negative value. To avoid the influence of these hyperedges, we set the negative values to 0 instead. Figure 2 shows the distribution of hyperedges with different quality types under different hypergraph learning schemes. In hypergraph without hyperedge weight learning, we noted that bad quality hyperedges account for a large part of hyperedges. When we adopt hyperedge weight learning scheme in traditional hypergraph, the ineffective hyperedges were removed. The total number of non-zero hyperedges droped to 502. The proportion of good quality hyperedges increased apparently. However, when the hyperedge weight learning scheme was applied in HSO, the proportion of bad quality hyperedges decreased. This is because that HSO has already benefited from the ASCI scheme. Compared to traditional hypergraph, the ASCI scheme leads to a high initial proportion of good quality hyperedges and normal quality hyperedges, improving the detection of HFOs. Figure 2 validates the effectiveness of our HSO detector.

B. Ablation Study on ASCI Scheme

During hypergraph construction, sample distance $d(x, y)$, that measures the similarity between two signal segments, plays an important role. Traditionally, Euclidean distance is a common metric to measure the distance $d(x, y)$) between two samples, such as images [26]. However, in our study, the samples are one-dimensional biological signals. Considering the unique characteristic of biological signals, Euclidean distance may not be an optimal metric to measure the similarity between two signal segments. Thus, we introduced a new ASCI mechanism to measure the similarity between SEEG signal segments. Specifically, for each signal segment, we calculated its ASCI-based distance with other signal segments, and find similar signal segments to form the hyperedge. In this way, the constructed hypergraph can detect HFOs based on a new signal-specific distance measurement approaches.

We tested the performance of HSO using four distance measurement approaches: Pearson's correlation coefficient, Spearman's rank correlation coefficient, Euclidean distance, and ASCI. For a fair comparison, we use the same experimental configuration. According to Table II, the proposed ASCI approach achieved the best performance on accuracy and sensitivity. The ASCI approach correctly detected HFOs signal with an accuracy of 90.7%, outperforming Spearman approach by 3.1%, Pearson approach by 3.6%, and Euclideanapproach

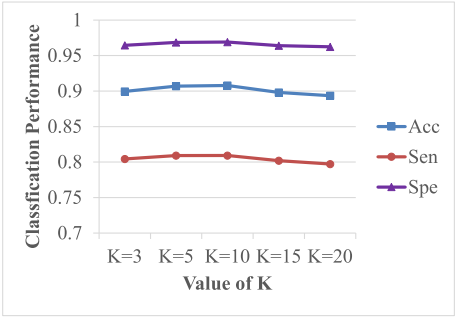

Fig. 3. Classification performance comparison with different K values.

by 2.6%. This demonstrates that our HSO benefits from the ASCI method on measuring the similarity of signal segments. Our ASCI-based HSO approach was able to achieve a sensitivity of 80.9%, which is 3.5%, 3.1 %, and 5.8% higher than the Spearman, Pearson, and Euclidean approaches, respectively. As illustrated in Table II, the Euclidean approach returned the lowest performance in terms of sensitivity. However, compared to other three approaches, the Euclidean approach achieves the best performance in terms of specificity, which relates to the detection of BC signals. The ablation experiments demonstrated that our ASCI approach provides a suitable trade-off between the detection of BC signals and HFOs.

C. Impact of Selected Neighbors K for Hyperedge Construction

In the star-expansion strategy [38], parameter K is used to select similarity-based nearest neighbors to form the hyperedge. In our experiment, we tested different value of K to optimize the classification performance of our HSO detector. In the conventional hypergraph theory, when the value of K is too small, the hyperedge only contains a few signal segments. The complex relationship between signal segments can't be fully represented by the constructed hypergraph. Conversely, when the value of K is too large, a hyperedge that connects too many signal segments has a negative effect on the discriminative ability of the constructed hypergraph.

Figure 3 shows the classification performance curve with respect to different K in our gold standard data set. Figure 3 illustrates the classification performance is not highly sensitive to parameter K. In our following experiments, we set the K to 10 to measure the classification performance of HSO detector.

D. Impact of Parameter λ and μ in the Objective Function

In the objective function, the parameter λ modulates the effect of loss term $\|\mathrm{F}-\mathrm{Y}\|^2$, while parameter μ modulates the effect of weight regularization term $\phi(w)$. Figure 4 shows the classification performance curve with respect to different λ. Figure 5 shows the classification performance variation curve with different μ values. The experiment results demonstrate that our HSO detector achieves a steady classification performance when μ and λ vary. We set parameter λ to 100 and parameter μ to 50 in the following analysis.

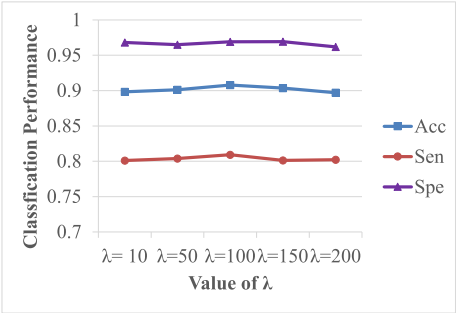

Fig. 4. Classification performance comparison with different λ values.

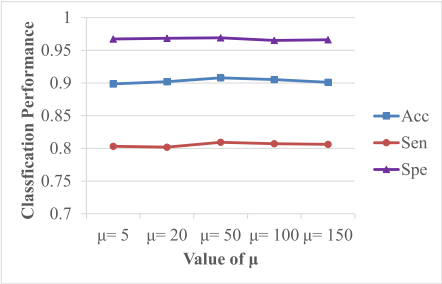

Fig. 5. Classification performance comparison with different μ values.

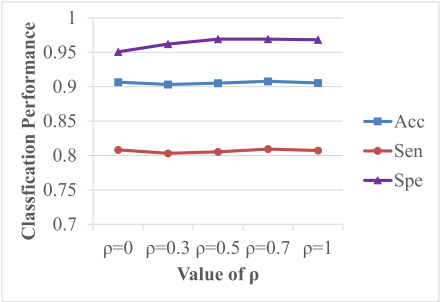

Fig. 6. Classification performance comparison with different ρ values.

E. Impact of Control Parameter ρ

In the weighted regularization term, control parameter ρ balances the sparsity on hyperedge weight matrix and classification effect. As mentioned previously, when $\rho = 0$, $\phi(w)$ controls the model complexity motivated by the success of sparse learning. When $\rho = 1$, $\phi(w)$ learns the optimal hyperedge weight. Figure 6 compares the classification performance with different ρ values. The experiment results demonstrate that our HSO detector achieves a steady classification performance, when ρ varied. Meanwhile, by varying the value of ρ, we achieve a desirable balance between the sparsity on hyperedge weight matrix and classification effect. In our experiment, we set parameter ρ to 0.7 to measure the classification performance of HSO detector in the following analysis.

F. Method Comparison

We compared the HSO detector to several machine learning classification models. Deep learning models have been demonstrated to outperform the traditional models in several applications [46]. Especially, for the application of detecting HFOs, our previous work of the SSAE-based HFOs detector [22] provides a promising way with superior performance to detect HFOs from MEG data. In this case, whether our hypergraph learning based classification model can outperform the deep learning models is yet to know. By adopting 10-fold cross validation strategy, the HSO detector was compared to multiple advanced machine learning models, including deep learning model, respectively.

TABLE III
PERFORMANCE OF DIFFERENT METHODS FOR SEVERAL MEASURES ON THE SEEG DATA SET. THE RESULTS OF OUR PROPOSED METHOD ARE DENOTED IN BOLD

	Accuracy (%)	Sensitivity (%)	Specificity (%)
K-nearest neighbors	84.7	66.1	97.4
Neural network	82.3	61.9	95.9
Logistic Regression	82.3	60.5	97.6
Linear SVM	83.1	61.5	98.1
Non-linear SVM	85.7	67.4	98.5
SMO [22]	86.1	71.2	96.5
Hypergraph [26]	88.1	75.1	97.4
HSO detector	**90.7**	**80.9**	**96.9**

Table III shows the classification performance of different methods for HFOs detection on the SEEG Data set. In our experiment, our HSO detector achieves a sensitivity of 80.9%, which improves the performance of detecting HFOs up to 19.4%. The superior performance of our HSO detector on sensitivity demonstrates utilizing the ASCI scheme to construct hypergraph can significantly enhance the classification performance. Note that although our HSO detector achieved the best performance on accuracy, the superiority of our HSO detector is not significant. Table III shows the similar performance on specificity, which reflects the proportion of true negative (the identification of BC signals). Compared to our previous work, our HSO detector outperforms the SMO method by 4.6% on accuracy, 9.7% on sensitivity, and 0.4% on specificity. Meanwhile, compared to another previous work [26], our approach HSO detector outperforms the traditional hypergraph method by 2.6% on accuracy and 5.8% on sensitivity, and achieves a low specificity. Also note that our HSO detector brings a suitable trade-off between the detection of BC signals and HFOs.

V. CONCLUSION

This work presented a hypergraph learning-based approach to automatically and precisely detect HFOs from SEEG data. While a novel ASCI-based distance metric was introduced to measure the signal morphological similarity between SEEG signal segments, a regularization term was also used to optimize the constructed hypergraph. Compared to currently available machine learning and deep learning classifiers, our detector achieved the best performance on the detection of HFOs from SEEG data. There are several future directions. One is to extend our detector to a multi-label classifier with a function to recognize additional patterns or sub-patterns (e.g., spike, ripple and fast ripple) in SEEG. Another direction is that our HFOs detector could also be applied on other iEEG technology (e.g., ECoG). Further comparisons are required

between our method and other existing approaches in classifying iEEG signals.

ACKNOWLEDGMENT

The authors are grateful to the investigators who share their SEEG data.

REFERENCES

[1] A. J. Durnford *et al.*, "Very good inter-rater reliability of engel and ILAE epilepsy surgery outcome classifications in a series of 76 patients," *Seizure*, vol. 20, no. 10, pp. 809–812, Dec. 2011.
[2] M. Cossu *et al.*, "Stereoelectroencephalography in the presurgical evaluation of focal epilepsy: A retrospective analysis of 215 procedures," *Neurosurgery*, vol. 57, no. 4, pp. 706–718, Oct. 2005.
[3] M. Cossu *et al.*, "Epilepsy surgery in children: Results and predictors of outcome on seizures," *Epilepsia*, vol. 49, no. 1, pp. 65–72, Jan. 2008.
[4] S. L. Moshé, E. Perucca, P. Ryvlin, and T. Tomson, "Epilepsy: New advances," *Lancet*, vol. 385, no. 9971, pp. 884–898, Mar. 2015.
[5] F. Bartolomei *et al.*, "What is the concordance between the seizure onset zone and the irritative zone? A SEEG quantified study," *Clin. Neurophysiol.*, vol. 127, no. 2, pp. 1157–1162, Feb. 2016.
[6] M. Cossu, F. Cardinale, L. Castana, L. Nobili, I. Sartori, and G. L. Russo, "Stereo-eeg in children," *Child's Nervous Syst.*, vol. 22, no. 8, pp. 766–778, 2006.
[7] E. Rikir *et al.*, "Electrical source imaging in cortical malformation-related epilepsy: A prospective EEG-SEEG concordance study," *Epilepsia*, vol. 55, no. 6, pp. 918–932, Jun. 2014.
[8] A. Miao *et al.*, "Using ictal high-frequency oscillations (80–500Hz) to localize seizure onset zones in childhood absence epilepsy: A MEG study," *Neurosci. Lett.*, vol. 566, pp. 21–26, Apr. 2014.
[9] J. Xiang *et al.*, "Frequency and spatial characteristics of high-frequency neuromagnetic signals in childhood epilepsy," *Epileptic Disorders*, vol. 11, no. 2, pp. 113–125, Jun. 2009.
[10] M. A. V. Klooster *et al.*, "High frequency oscillations in the intra-operative ECoG to guide epilepsy surgery ('The HFO Trial'): Study protocol for a randomized controlled trial," *Trials*, vol. 16, no. 1, p. 422, Sep. 2015.
[11] N. E. C. van Klink *et al.*, "High frequency oscillations in intra-operative electrocorticography before and after epilepsy surgery," *Clin. Neurophysiol.*, vol. 125, no. 11, pp. 2212–2219, Nov. 2014.
[12] P. Modur, "High frequency oscillations and infraslow activity in epilepsy," *Ann. Indian Acad. Neurol.*, vol. 17, no. 5, p. 99, 2014.
[13] P. Modur and S. Miocinovic, "Interictal high-frequency oscillations (HFOs) as predictors of high frequency and conventional seizure onset zones," *Epileptic Disorders*, vol. 17, no. 4, pp. 413–424, Dec. 2015.
[14] L. Meng, "A magnetoencephalography study of pediatric interictal neuromagnetic activity changes and brain network alterations caused by epilepsy in the high frequency (80–1000 Hz)," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 3, pp. 389–399, Mar. 2019.
[15] J. Guo *et al.*, "Automatic and accurate epilepsy ripple and fast ripple detection via virtual sample generation and attention neural networks," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 8, pp. 1710–1719, Aug. 2020.
[16] J. Xiang *et al.*, "Volumetric imaging of brain activity with spatial-frequency decoding of neuromagnetic signals," *J. Neurosci. Methods*, vol. 239, pp. 114–128, Jan. 2015.
[17] J. Xiang *et al.*, "Accumulated source imaging of brain activity with both low and high-frequency neuromagnetic signals," *Frontiers Neuroinform.*, vol. 8, p. 57, May 2014.
[18] N. van Klink *et al.*, "Automatic detection and visualisation of MEG ripple oscillations in epilepsy," *NeuroImage, Clin.*, vol. 15, pp. 689–701, Jan. 2017.
[19] E. Tamilia, J. R. Madsen, P. E. Grant, P. L. Pearl, and C. Papadelis, "Current and emerging potential of magnetoencephalography in the detection and localization of high-frequency oscillations in epilepsy," *Frontiers Neurol.*, vol. 8, p. 14, Jan. 2017.
[20] L. R. Quitadamo, R. Mai, F. Gozzo, V. Pelliccia, F. Cardinale, and S. Seri, "Kurtosis-based detection of intracranial high-frequency oscillations for the identification of the seizure onset zone," *Int. J. Neural Syst.*, vol. 28, no. 07, Sep. 2018, Art. no. 1850001.
[21] S. Burnos *et al.*, "Human intracranial high frequency oscillations (HFOs) detected by automatic time-frequency analysis," *PLoS ONE*, vol. 9, no. 4, Apr. 2014, Art. no. e94381.
[22] J. Guo *et al.*, "A stacked sparse autoencoder-based detector for automatic identification of neuromagnetic high frequency oscillations in epilepsy," *IEEE Trans. Med. Imag.*, vol. 37, no. 11, pp. 2474–2482, Nov. 2018.
[23] D. Zhou, J. Huang, and B. Schölkopf, "Learning with hypergraphs: Clustering, classification, and embedding," in *Proc. 19th Int. Conf. Neural Inf. Process. Syst.* Cambridge, MA, USA: MIT Press, 2006, pp. 1601–1608.
[24] Y. Gao, M. Wang, D. Tao, R. Ji, and Q. Dai, "3-D object retrieval and recognition with hypergraph analysis," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 4290–4303, Sep. 2012.
[25] J. Yu, D. Tao, and M. Wang, "Adaptive hypergraph learning and its application in image classification," *IEEE Trans. Image Process.*, vol. 21, no. 7, pp. 3262–3272, Jul. 2012.
[26] Y. Gao, M. Wang, Z.-J. Zha, J. Shen, X. Li, and X. Wu, "Visual-textual joint relevance learning for tag-based social image search," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 363–376, Jan. 2013.
[27] P.-P. Song, J. Xiang, L. Jiang, H.-S. Chen, B.-K. Liu, and Y. Hu, "Dynamic changes in spectral and spatial signatures of high frequency oscillations in rat hippocampi during epileptogenesis in acute and chronic stages," *Frontiers Neurol.*, vol. 7, p. 204, Nov. 2016.
[28] Z. Zhang, J. Wang, and H. Zha, "Adaptive manifold learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 2, pp. 253–265, Feb. 2012.
[29] X. Peng, G. Long, T. Shen, S. Wang, J. Jiang, and M. Blumenstein, "Temporal self-attention network for medical concept embedding," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2019, pp. 498–507.
[30] X. Peng, G. Long, T. Shen, S. Wang, and J. Jiang, "Self-attention enhanced patient journey understanding in healthcare system," 2020, *arXiv:2006.10516*. [Online]. Available: http://arxiv.org/abs/2006.10516
[31] X. Zheng *et al.*, "Rethinking performance estimation in neural architecture search," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11356–11365.
[32] M. Lin *et al.*, "HRank: Filter pruning using high-rank feature map," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1529–1538.
[33] M. Ahsan, M. I. Ibrahimy, and O. O. Khalifa, "EMG signal classification for human computer interaction: A review," *Eur. J. Sci. Res.*, vol. 33, no. 3, pp. 480–501, 2009.
[34] N. van Klink, A. Hillebrand, and M. Zijlmans, "Identification of epileptic high frequency oscillations in the time domain by using MEG beamformer-based virtual sensors," *Clin. Neurophysiol.*, vol. 127, no. 1, pp. 197–208, Jan. 2016.
[35] L. P. Andrade-Valenca, F. Dubeau, F. Mari, R. Zelmann, and J. Gotman, "Interictal scalp fast oscillations as a marker of the seizure onset zone," *Neurology*, vol. 77, no. 6, pp. 524–531, Aug. 2011.
[36] C. Papadelis *et al.*, "Interictal high frequency oscillations detected with simultaneous magnetoencephalography and electroencephalography as biomarker of pediatric epilepsy," *J. Visualized Exp.*, no. 118, Dec. 2016, Art. no. e54883.
[37] J. Lian, G. Garner, D. Muessig, and V. Lang, "A simple method to quantify the morphological similarity between signals," *Signal Process.*, vol. 90, no. 2, pp. 684–688, Feb. 2010.
[38] J. Y. Zien, M. D. F. Schlag, and P. K. Chan, "Multilevel spectral hypergraph partitioning with arbitrary vertex sizes," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 18, no. 9, pp. 1389–1399, Sep. 1999.
[39] D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Schölkopf, "Learning with local and global consistency," in *Proc. Adv. Neural Inf. Process. Syst.*, 2004, pp. 321–328.
[40] C. Zu *et al.*, "Identifying disease-related subnetwork connectome biomarkers by sparse hypergraph learning," *Brain Imag. Behav.*, vol. 13, no. 4, pp. 879–892, Aug. 2019.
[41] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *Amer. Statistician*, vol. 46, no. 3, pp. 175–185, Aug. 1992.
[42] S. Haykin, "Neural networks, a comprehensive foundation," Macmilan, New York, NY, USA, Tech. Rep., 1994.
[43] A. J. Dobson, "An introduction to generalized linear models," Tech. Rep., 1990.
[44] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
[45] B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: A strategy employed by v1?" *Vis. Res.*, vol. 37, no. 23, pp. 3311–3325, Dec. 1997.
[46] Y. Bengio and Y. LeCun, "Scaling learning algorithms towards AI," *Large-Scale Kernel Mach.*, vol. 34, no. 5, pp. 1–41, 2007.