ABSTRACT  High frequency oscillations (HFOs) in intracranial electroencephalography (iEEG) recordings are a promising clinical biomarker that can help define the epileptogenic regions in the brain. The aim of this study is to characterize the spatial and temporal distribution of HFOs in channel-wise instead of event-level as usual and to develop an automated the seizure onset zone (SOZ) identification by using a support vector machine (SVM) approach on the channel-wise features in a short-term recording. In this work, five consecutive patients with medically intractable epilepsy were enrolled. For each patient, ten-minute segments were defined from two hours of iEEG recordings during sleep state. A total of 17 channel-wise features including 6 rate-based, 6 duration-based, 3 amplitude-based, and 2 power-based features of HFOs were extracted from each 10-min segment, which including ripples (Rs, 80-250 Hz) and fast ripples (FRs, 250-500Hz) were detected automatically using validated detectors. Each channel-wise feature was ranked by using the Student’s t-test method and the most distinctive features were selected to explore the characteristics of HFOs in each channel. A supervised-learning based SVM classifier with the selected channel-wise features or their combinations was developed to identify each channel within the independently clinician-defined SOZ or not. Over 3,816 channel-10-min segments of iEEG recordings, the evaluated accuracy, sensitivity, and specificity of the proposed approach with the optimal combination of top five ranked features for SOZ identification are 86.6%, 73.0%, and 94.1%, respectively, for ten-fold cross-validation, and 86.0%, 79.2 %, and 91.8%, respectively, for the leave-1-out cross-validation. Compared with the recently reported SOZ detectors based on event-wise feature of HFOs, the channel-wise features and the combination with machine learning approach demonstrate its feasibility in SOZ identification with a relative higher performance and potentially reduce the time needed currently for long-term recording and manual inspection.

INDEX TERMS  Epilepsy, high frequency oscillations (HFOs), intracranial electroencephalograms (iEEG), machine learning, seizure onset zone.

I. INTRODUCTION  
Epilepsy is one of the most frequent chronic neurological diseases affecting an estimated number of 65 million people of worldwide and occurs in all age ranges [1], [2]. The current gold standard for identification of the epileptogenic zone is the seizure onset zone (SOZ) [3]–[8]. However, identifying this brain region is challenging. The current surgical outcomes are unfavorable 40-50% of well-selected patients, which has led to a search for new biomarkers that could accurately and effectually identify the SOZ and therefore result in better postsurgical outcomes [9]. Recently, a number of studies reported that interictal high frequency oscillations (HFOs, 80 - 500 Hz) in the recorded intracranial electroencephalography (iEEG) have been shown to be a promising biomarkers to identify the epileptogenic zone by...
using brief segments of data between seizures, and has altered the traditional EEG frequency bands which were thought clinically relevant (up to 40 or 70 Hz) [10]–[16]. However, multiple recent studies also reported that HFOs are present not only in epileptic cerebral area but also in non-epileptic brain regions often including visual cortex, motor cortex and language areas [17]–[20]. The presence of these physiologic events complicates the clinical use of HFOs as biomarkers for the delineation of SOZ and increase the risk of injury to functional areas.

The characterization of HFOs is an increasingly important measure of SOZ detection at present [21], [22]. Several earlier studies have attempted to characterize and distinguish HFOs discovered inside and outside the SOZ based on manual extraction of event-wise features in time domain [21]–[23], spectrum domain [23], time-frequency domain [25], and/or nonlinear domain [26]–[28]. Most of these studies focused on exactly quantifying waveform morphology using these features from a statistical point of view or employing an unsupervised method, after clustering them according to their known origin. Although the main finding of stereotyped HFOs showing temporal [23], [29] and/or spatial variability [30] may help elucidate mechanisms of seizure generation, the analysis of HFO events form one by one has limited use in epilepsy diagnosis and planning neurosurgery. Meanwhile, a large portion of the current clinical HFO literature is based upon a channel-wise strategy of analyzing the rate of HFO occurrence (detected either manually or using software) to classify the channels that are in the SOZ [29]–[33], which allows for experts to process and count HFOs in fixed durations manually. Compared with normal physiologic HFOs, pathological HFOs may not only have distinctive shapes but occur more often during slow-wave sleep than in wakefulness [23], [31], and epileptiform HFOs rates were proven significantly higher within the SOZ than in non-SOZ regions, especially during interictal periods in a deep sleep state [13], [14]. Some also attempted to differentiate between epileptic and non-epileptic HFOs based on others features besides their rate [34], [35], [41]. The investigated features are related to the duration, amplitude, frequency of the HFOs. However, these approaches have predominantly utilized a single rate or an arbitrary partial combination of the abovementioned features to identify SOZs and have not considered the inter-patient variability nor the high-dimensional dynamics of the epileptic activity.

In our previous study [36], we proposed an automated HFOs detector based on deep convolutional neural networks (CNN), which is capable of automatically extracting the shared features of HFOs events of different patients and shown a higher accuracy for distinguishing real HFOs as well as Rs and FRs from false HFOs such as noise and spike. Based on our preliminary inspection of count of identified HFOs in each channel and the occurrence time of each HFOs event, both of them were found a greater variability between within the SOZ and other regions. As such, we hypothesize that it may be possible to reduce the variability across patients by combining the complementary values contained with different electrophysiological characteristics of HFOs and thereby improve the generalization capability of SOZ localization. Although each feature constitutes specific electrophysiological information about the epileptogenicity of brain tissue and might add predictive value when used in unison with other features, it is a new challenge to design more sophisticated models for any conclusion regarding the practical significance of the feature differences and their reflection in the usefulness of SOZ identification, by effectively utilizing multiple electrophysiological features.

Following up on our previous work, this study assessed the feasibility of using machine learning approach on seventeen, channel-wise characteristics of HFOs instead of features of each event to automatically identify channels recorded from the seizure onset zone in patients with drug-resistant epilepsy. To our best knowledge, the present study provided the first attempt to systematic investigation of HFOs regarding automated identification of SOZ with their traditional and customized characteristics representing electrophysiological knowledge, including the rate, duration, amplitude, and power of channel-wise HFOs instead of each event. By automatically localizing the zone of seizure onset or at least estimating a region of interest containing or closed to the epileptogenic zone before surgery procedure, the investigated machine learning based SOZ technique would provide a useful and applicable pre-surgery guideline for operator, and potentially reduce the time needed for long-term recording and manual inspection.

II. MATERIAL AND METHODS

A. DATA ACQUISITION

1) PARTICIPANTS

Five consecutive patients with medically intractable epilepsy were included in the study (three male, two female patients) from 2016 to 2017 at the department of neurosurgery of the West China Hospital (affiliated to Sichuan University, Chengdu, China). All patients underwent a comprehensive presurgical evaluation by the doctors including a detailed history, routine scalp EEG, and imaging (MRI and CT of brain), and subsequently been implanted with intracranial electrodes for a long-term iEEG recording. This study was approved by the Institutional Review Board (IRB) of the West China Hospital, where data collection was performed, and written informed consent was obtained from all patients.

2) INTRACRANIAL EEG RECORDING

The intracranial EEG of each patient was recorded from subdural electrodes formatted in subdural silastic grids (4 mm diameter electrode contacts, 10 mm inter-electrode spacing) which all referenced to a bipolar montage along the length of the grid. The duration of clinical monitoring (24–208 hours) and the location and number of implanted electrodes (46 - 110) were determined in accordance with clinical considerations. For each patient, a continuous
long-term intractable EEG recording was performed in a sampling rate of 4,096 Hz by using a multichannel bio-amplifier with XLTEK EMU128FS system (Natus neurology, USA). Note that no hardware filter was used in the acquisition procedure, except the high-pass filter with 0.1 Hz as a cutoff frequency classically used to remove the offset of the baseline.

3) CLINICAL SOZ LOCALIZATION
Clinical reports abovementioned were used for clinical SOZ localization. The seizure onset channels of each patient were determined visually by two neurologists based on the well known knowledge of the seizure onset channels with the earliest iEEG seizure discharges [17], [30], [35]. In case of multiple SOZs, the seizure onset channels from all seizures were marked as SOZ and treated equally. Meanwhile, the remaining non-spiking channels with normal activity were referred to as recording from the normal zone (NoZ). The same approach has been used in previous studies [33]–[35]. In this way, both two corresponding electrodes of one marked SOZ channel were localized and all SOZ electrodes were then annotated as the gold standard to validate our proposed method.

B. DATA ANALYSIS
1) ALGORITHM FRAMEWORK
A full set of algorithms for the automated detection of HFOs in iEEG data, the evaluation of their channel-wise features and the identification of the candidate SOZ were developed and programmed with the help of the MATLAB software package (MATLAB 2010, the Mathworks Inc., Natick, MA). It consists of five main blocks of functions as shown in Fig. 1:
(i) data pre-processing and segmentation;
(ii) CNN-based HFOs detection that was presented in our previous work [36];
(iii) candidate channel-wise features extraction of HFOs, such as the rate, duration, amplitude, and power of HFOs;
(iv) informative features selection by conducting on a student t-test on all candidate channel-wise features;
(v) machine-learning based classifications and automated identification of SOZ.

2) DATA PRE-PROCESSING
In this work, the raw data recorded was exported in European Data Format (EDF) without being filtered and subsequent data segment selection for each recording was carried out using bipolar montages with neighboring electrodes along the length of the grid with the help of the software NeuroWorks 8 (Natus Medical Incorporated, Canada). Visual review of individual channels was performed and channels with excess line noise (50 Hz), motion artifacts, and containing no visible EEG signal were removed before analysis. The total number of the selected channels was 318 (a total of 395 electrodes referring to a bipolar montage) in five patients. Then, we randomly selected two hours of continuous raw data during sleep, which were at least 60 min prior to the onset of a seizure. Each of the selected iEEG signals was cut continuously into 10 minute segments (a total of 60 10-min segments for five patients). Moreover, each data segment was filtered with a 4-order Buterworth bandpass filter to reconstruct HFOs both of the Rs (80 - 250Hz) and FRs (250 - 500Hz) events. After filtering, all filtered segments of these two frequency ranges were normalized for subsequent data analysis such as automatic HFO detection and SOZ localization.

3) AUTOMATED HFOs DETECTION
Automated HFOs detection is a crucial step to get a more complete overview of the HFOs characteristics [27], [41], [42] and to further investigate their relationship with epilepsy, especially with the assessment of the SOZ in the context of continuous monitoring. In this work, the automated HFOs
detection was performed by using our recently reported HFOs detector with high diagnostic sensitivity and specificity [36]. The algorithm employs the combination of short-time energy (STE) estimation and the CNN classifier to detect HFOs in all recorded channels of each 10-min data segment [36]. Meanwhile, the time occurrence of each transient of HFOs is accurately identified so as to further assist the extraction of duration-based features in this work. Unlike what we’ve reported before, when detecting HFOs event, each 10-min segment of the recorded iEEG was further divided into numerous of successive 5-ms frames instead of 10-m for HFOs event detection for better temporal resolution. Once energy values of five adjacent frames all exceed a certain threshold level, one candidate HFOs event with different duration can be identified individually, which may span numerous of 5-ms frames (at least 5 frames). The duration of each HFOs event in time window is correspondingly determined between the beginning time of the first frame and the ending time of the last frame. Note that the temporal resolution of the duration of each HFOs event was 5 ms in this work.

4) CHANNEL-WISE FEATURES EXTRACTION

After collected a pool of HFOs, a total of 17 novel channel-wise characteristics of HFOs were proposed to be derived from all HFOs, Rs, and FRs in each 10-min segment data. There were 6 rate-based, 6 duration-based, 4 amplitude-based and 2 power-based features, which are summarized in Table 1 and described as followings:

- **Rate-Based Features** ($f_{H}(x_{ij})$, $f_{R}(x_{ij})$, $f_{FR}(x_{ij})$, $Rif_{H}(x_{ij})$, $Rif_{R}(x_{ij})$, and $Rif_{FR}(x_{ij})$): The channel-wise feature of absolute HFOs rate $f_{H}(x_{ij})$ characterizes the absolute count $x$ of HFOs per min of the $i^{th}$ 10-min segment in the $j^{th}$ channel, and the corresponding relative ratio $Rif_{H}(x_{ij})$ is computed between the absolute and the maximum rate of HFOs in each channel. For the Rs and FRs events,
TABLE 2. Optimal channel-wise features for SOZ indentification by using the student t – test ranking method.

| Category       | No. of input features | No. of optimal features | Description of optimal feature | ranking |
|----------------|-----------------------|-------------------------|-------------------------------|---------|
| Rate           | 6                     | 3                       | Rif_{ij}(x_{ij}): Relative ratio of HFOs rate | 1       |
|                |                       |                         | Rif_{ij}(x_{ij}): Relative ratio of Rs rate | 2       |
|                |                       |                         | f_{ij}(x_{ij}): Absolute Rs rate | 5       |
| Duration       | 6                     | 2                       | Rif_{ij}(t_{ij}): Ratio of duration of Rs | 3       |
|                |                       |                         | Rif_{ij}(t_{ij}): Ratio of duration of HFOs | 4       |
| Amplitude      | 3                     | 1                       | -                             | -       |
| Power          | 2                     | 0                       | 0                             | -       |
| Total          | 17                    | 5                       | -                             | -       |

one absolute rate and one corresponding relative ratio are also derived for each data segment in each channel, respectively.

- **Duration-Based Features** ($\text{Rif}_{ij}(t_{ij})$, $\text{Sdf}_{ij}(x_{ij})$, $\text{Rif}_{ij}(t_{ij})$, $\text{Rif}_{ij}(t_{ij})$, $\text{Avf}_{ij}(x_{ij})$, and $\text{Avf}_{ij}(x_{ij})$): Based on the detected duration $t$ of each HFOs event of the $i$th-10 min segment in the $j$th channel, the channel-wise feature of ratio of HFOs duration is defined as the ratio between the summary of all HFOs' duration of the $j$th channel and the total duration of all channels for the $i$th-10 min segment in each patient. And, the feature of average duration of Rs $\text{Avf}_{ij}(t_{ij})$ is computed as the average in the ratio between the summary of all Rs’ duration of each channel and the number of Rs. The feature of average duration of FRs ($\text{Avf}_{ij}(t_{ij})$ ) is similar to that of Rs. Additionally, the corresponding standard deviation of HFOs duration is computed as following equation:

$$Sdf_{ij}(t_{ij}) = \sqrt{\frac{\sum_k (t_k - t_{k\text{mean}})^2}{K}} \quad (1)$$

where $t_k$ is the duration of the $k$th HFOs of the $i$th segment in the $j$th channel, $t_{k\text{mean}}$ is the average duration of the total HFOs events $K$.

- **Amplitude-Based Features** ($\text{AvSdf}_{ij}(a_{ij})$, $\text{AvSdf}_{ij}(a_{ij})$, and $\text{AvSdf}_{ij}(a_{ij})$): Considering each HFOs event $k$ containing numerous oscillations $n$, the standard deviation of amplitude of each HFOs event can be defined to characterize partially its morphology. As such, the channel-wise average of standard deviation of oscillations amplitude is customized as one of channel-wise feature $\text{AvSdf}_{ij}(a_{ij})$ as following:

$$A_v Sdf_{ij}(a_{ij}) = \sqrt{\frac{\sum_n (a_n - a_{k\text{mean}})^2}{N}} / K \quad (2)$$

where $a_n$ is the amplitude of one oscillation of the $k$th HFOs of the $i$th segment in the $j$th channel, $N$ is the total number of the oscillation of the average duration of the total $K$ HFOs events, $K$ is the total number of the HFOs events of the $i$th segment in the $j$th channel.

- **Power-Based Features ($\text{Avf}_{ij}(p_{ij})$ and $\text{Avf}_{ij}(p_{ij})$):** For the frequency domain, the channel-wise average energy $p$ of a total of $K$ Rs (80-250 Hz) in the $i$th segment of the $j$th channel was computed as following, and the feature $\text{Avf}_{ij}(p_{ij})$ of the FRs (250-500 Hz) was also obtained with a similar equation:

$$A_v f_{ij}(p_{ij}) = \sum_k p^2 / K \quad (3)$$

5) FEATURES SELECTION

Considering the complex electrophysiological mechanism of epilepsy, a common limitation of individually using the above proposed channel-wise features is that they may not provide a high enough sensitivity, specificity, or either both of them. Thus, we further hypothesis is that, if such features reflect different phenomena of an underlying neuro activity, they might add complementary information to each other, consequently, a combined informative features might improve the capability of the proposed method for identification of SOZ. Therefore, the most distinctive features were selected and combined for subsequent classification by using a feature ranking/selection algorithm. In this work, the Student’s $t$-test was used for this purpose [37]. The p-value obtained is used to rank all of the extracted channel-wise features, higher p-value indicate better ranking. The selected channel-wise characteristics including the traditional and customized features associated with HFOs were computed for each 10-min data segment of each channel.

6) MACHINE LEARNING BASED CLASSIFICATION

The supervised-learning based method of support vector machine (SVM) with radial basis function (RBF) kernel function [38] was used in this work to differentiate between electrodes of located in the clinical SOZ and those out of this area, which was suggested by recent works [17], [35] and shown a superior performance in our pilot study among four standard machine-learning techniques such as k-nearest neighbors (k-NN), decision tree (DT), and SVMs with linear and polynomial kernel function. And, two different experimental approaches of cross-validation were carried out, as shown in Fig. 2(A).

All of 17 extracted features used for the classification were divided into four categories such as rate-based, duration-based, amplitude-based, and power-based features. These channel-wise features were extracted from each 10-min segment (epoch) of a channel, and 12 feature matrices with number of channels by 17 were obtained from the 2 hours recording of each channel in each patient. Note that one cell of the feature matrix was set to be zero if there is no any HFOs events detected in this segment. A total of 60 10-min segment data was analyzed to distinguish a total of 395 electrodes between the SOZ and NoZ electrode.
The overall analytic scheme of these two cross-validation protocols is illustrated in Fig. 2(B). One is a ten-fold cross validation, where the dataset was divided into an 80% training set and a 20% testing set. Then, a leave-1-out cross-validation was carried out. For every patient in the dataset, the data from the rest of the patients was employed as the training data and respective patient’s data used as the testing data. For each of the cross-validation iteration, training and testing datasets were generated by employing one of the cross-validation methods. These two validation approaches allow the calculation of generalizable performance metrics for the analyzed classifier from the view of balance degree of a dataset in case of the ten-fold cross and the view of feasibility in a new patient in case of the leave-1-out cross [35]. With combinations of same category features the trained SVM classifier with optimized model parameters was used to identify automatically the SOZ or NoZ channels in the testing set and their corresponding performance was therefore evaluated. Moreover, to evaluate the significance of the improvement in the classification results when including various combinations of selected features, we performed a series of tests on the SVM classifier with various selected features combination, including combinations of the top one, the top two, the two three, ……, the top 16 informative features and all of the 17 extracted channel-wise features.

C. STATISTICAL ANALYSIS
All detected HFOs were used to identify the seizure onset areas by using the SVM based classifier as mentioned above.

The performance of the proposed classification when using different feature subsets was evaluated by identifying the classes of the test dataset and comparing the identifications against the clinical gold standard annotations of the test dataset, as shown in Fig. 2(C). A channel was considered true positive \( C_{\text{SOZ}} \) if it was overlapped with the seizure onset site marked by clinicians and was considered false positive \( C_{\text{NoZ}} \) if it lied outside of the SOZ region. This performance was conducted using standard performance metrics such as the area under curve (AUC), sensitivity, specificity, and accuracy, which were defined as:

\[
\text{Sensitivity} = \frac{C_{\text{SOZ}} \text{ in SOZ}}{C_{\text{SOZ}} \text{ in SOZ} + C_{\text{NoZ}} \text{ in SOZ}}
\]

\[
\text{Specificity} = \frac{C_{\text{NoZ}} \text{ out SOZ}}{C_{\text{NoZ}} \text{ out SOZ} + C_{\text{SOZ}} \text{ out SOZ}}
\]

\[
\text{Accuracy} = \frac{C_{\text{SOZ}} \text{ out SOZ} + C_{\text{NoZ}} \text{ out SOZ}}{\text{number of tested electrodes}}
\]

III. RESULTS
A. HFOs DETECTION AND CHANNEL-WISE FEATURES SELECTION
As shown in Fig. 3(A), examples of automatic detection of HFOs (Rs and FRs), spike, and noise from two patients are given in raw signal, filtered signals after 80-250 Hz filter and 250-500 Hz filter, and two dimensional time-frequency maps. HFOs events consisted of the Rs and FRs events, and false HFOs including noise and spike were excluded in this study. Over 3,816 channel-10-min segments of iEEG recordings...
(5 patients with on average 54 channels per patient, 2 h per channels) were analyzed in this study. A total of 4,014 HFOs was detected automatically, including 2,530 events within the SOZ area and 1,484 outside it. And, 2,375 Rs and 155 FRs were found in the SOZ region, while there were 1,460 Rs and 24 FRs from norm brain region.

Fig. 4 shows the evaluated performances of combination of various ranked features, where a total of seventeen features combinations. Specifically, the proposed 17 channel-wise features of HFOs were computed for each channel-10-min segment, and ranked by the obtained $p$ value of the Student $t$-test method. To find most informative feature or features combination, an optimization of combination of inter-category ranked features for SOZ identification was carried out by using the ten-fold and leave-1-out cross-validations. As shown in Fig. 4, we found that the point at the top five ranked features has a tradeoff and optimal performances including the AUC, the accuracy, the sensitivity, and the specificity. For combinations from that to the subsequent five, obtained performance values of four metrics almost keep the same with that of the method combined with the top five ranked features, whereas the corresponding computational complex would be increase as the increasing number of features. Moreover, after the point of the ten ranked features, performances of SOZ localization show a drastically reduced and unstable. In this case, the optimal features combination is determined for the proposed method in this work, including three rate-based features and two duration-based features. Referring the Tables 1 and 2, The ID and name of these features are No. 4, $R_i f_{iH}(xij)$: relative ratio of HFOs rate, No. 5, $R_i f_{iR}(xij)$: relative ratio of Rs rate, No. 8, $f_{R}(xij)$: absolute Rs rate, No. 7, $R_i f_{iR}(tij)$: Ratio of duration of Rs, and No. 2, $R_i f_{iH}(tij)$: Ratio of duration of HFOs. As shown in Fig. 3(B), examples of these top five ranked features are given in each color coded channel of patients 2 and 3, respectively. One color bar is given for each feature individually, where a channel color-coded with a hotter color means a higher possibility within the SOZ.

**FIGURE 3.** (A) Examples of automatic detection of HFOs (Rs and FRs), spike, and noise are given in raw signal, filtered signals after 80-250 Hz filter and 250-500 Hz filter, and two dimensional time-frequency maps in two patients. (B) Examples of top five ranked features are given in each color coded channel of patients 2 and 3 with the leave-1-out cross-validations, where a channel color-coded with a hotter color means a higher possibility within the SOZ.

**B. SOZ Identification and Channel-Wise Features Combination**

As shown in Table 3, various combinations of multiple channel-wise features were input into the SVM classifier.
to determine whether the performance of SOZ identification can be improved and the inter-patient variability can be reduced. Four combinations of same category features are the rate-based SVM with 6 features, duration-based SVM with 6 features, amplitude-based SVM with 3 features, and power-based SVM with 2 features. Importantly, one combination with the selected top five ranked features and one combination with all 17 features are both evaluated. For each group, obtained performance metrics of the AUC, the accuracy, the sensitivity, and the specificity by using the ten-fold and leave-1-out cross-validations described previously are given in Table 3 separately. Our results also indicate that both the rate-based and duration features combination show an obvious superiority for SOZ identification in comparison with the amplitude-based and power-based combination, and it shows a good agreement with previous studies [34], [35]. Interestingly, the proposed algorithm with the selected top five ranked features (or called an optimal combination of inter-category features) resulted in higher ten-fold cross-validation performances of 86.6%, 73.0%, and 94.1% for the accuracy, the sensitivity, and the specificity, respectively. For the corresponding leave-1-out protocol, the obtained performances are of 86.0%, 79.2%, and 91.8%, respectively. Lastly, based on these results further SOZ identification with the optimal feature combination was performed in each patient. Over 227 channels from five patients, 15 channels were successfully identified within the SOZ region. Two examples are shown in Fig. 5, where the best and worst good-fit between our estimated SOZ and the clinical SOZ localization were found in patients 4 and 1, respectively. As shown in Fig. 5, examples of SOZ identification with our proposed method on the combination of top five ranked channel-wise features

### TABLE 3. Cross-validated classification performance for SOZ identification by using SVM-based classifiers with combinations of same category features and inter-category optimal features, by using both of the ten-fold and leave-1-out cross-validation experiments.

| Features combination (number) | Cross-validation Method | AUC   | Sensitivity | Specificity | Accuracy |
|-------------------------------|------------------------|-------|-------------|-------------|----------|
| Rate (6)                      | Average ten-fold       | 0.637 | 68.3%       | 92.8%       | 84.1%    |
|                               | Leave-1-out            | 0.675 | 77.2%       | 89.1%       | 82.7%    |
| Duration (6)                  | Average ten-fold       | 0.505 | 52.0%       | 96.7%       | 82.0%    |
|                               | Leave-1-out            | 0.554 | 62.3%       | 93.0%       | 83.8%    |
| Amplitude (3)                 | Average ten-fold       | 0.153 | 15.7%       | 97.3%       | 69.3%    |
|                               | Leave-1-out            | 0.144 | 14.8%       | 97.8%       | 65.9%    |
| Power (2)                     | Average ten-fold       | 0.174 | 27.3%       | 80.8%       | 60.0%    |
|                               | Leave-1-out            | 0.064 | 32.3%       | 71.9%       | 61.7%    |
| All features (17)             | Average ten-fold       | 0.248 | 31.9%       | 77.6%       | 62.1%    |
|                               | Leave-1-out            | 0.141 | 18.8%       | 93.9%       | 63.1%    |
| Optimal features (5)          | Average ten-fold       | 0.689 | 73.0%       | 94.1%       | 86.6%    |
|                               | Leave-1-out            | 0.716 | 79.2%       | 91.8%       | 86.0%    |
from one 10-min-segment iEEG are presented with the leave-1-out cross-validations. The result in left panel shows a good agreement between our estimated SOZ electrodes (red circle) and clinical marked SOZ region (blue rectangle). Whereas a slight mismatch between of them also was found in this study as shown in right panel of the Fig. 5, two estimated SOZ electrodes were not correctly identified in the SOZ area and two SOZ electrodes were missed in our estimated SOZ, which would lead to a corresponding decrease of the performance.

### C. COMPARISON TO EXISTING WORKS

Furthermore, compared with recent studies that are most close to our proposed method [17], [26], [32]–[35], our proposed method shows an comparable performance in accuracy and specificity, such as 86.6% and 94.1% for ten-fold test, 86.0% and 91.8 for leave-1-out test, as shown in Table 4. Meanwhile, the evaluated sensitivity around 70-80% is similar to those reported works [32]–[35] and a little bit lower than that reported work [33], while our method demonstrated an improvement in both accuracy and specificity. However, one should be cautious that this comparison of performance to existing works is only a relatively qualitative analysis to assess the feasibility of this proposed method with a similar level.

### IV. DISCUSSIONS

#### A. MAIN CONTRIBUTIONS OF THIS STUDY

The present study reports a SVM based method for automated identification of SOZ and NoZ electrodes using combination of multiple channel-wise features extracted from interictal
iEEG data, which were collected in a clinical setting. Importantly, our study, to our best knowledge, is the first attempt to systematically extract and optimize channel-wise features of HFOs rather than that of each event, which represent higher dimensional information and distributions of HFOs including rate, duration, amplitude, and power. The optimal combination of engineering and electrophysiological features of HFOs provides a way of reducing variability across patients and redundant information among features to improve the performance of SOZ localization. With experiments in thousands HFOs detected from hundreds of electrodes in five patient with medically intractable epilepsy, the automated identification of SOZ presented in this work show a solidly consistent in all patients, and the spatial distribution of identified SOZ channels showed agreement with clinician-define SOZ. Compared to other recent reported methods for SOZ identification, our proposed method with the combination of optimized channel-wise features has an obvious higher performance in terms of the accuracy, the sensitivity, and specificity. In addition to the above specific contributions, we evaluate the usefulness of these channel-wise features as well as the proposed method through the performance of a machine learning based classification algorithm, instead of simply assessing statistical significance of the difference, which can be meaningless from a practical point of view. Besides that, the ability to perform SOZ identification automatically within a maximum monitoring duration of two hours offers a considerable benefit when considering its clinical applicability.

B. THE VALUE OF FEATURE SELECTION AND COMBINATION

These obtained results also highlight the need for further research to clarify the differences in significance of SOZ identification among all extracted channel-wise features as well as various combinations of them. Previous works have focused on SOZ detection algorithm with event-wise features of HFOs or an arbitrary partial combination of them. Although, the most widely reported feature for SOZ identification is the rate of HFOs as well as the Rs and the FRS [29]–[33], the generalization capability of such single feature has been insufficient, primarily because of the variability across patients [21], [22], spatial variation [30], and temporal changes [23], [29] of HFOs events in clinical practice. In the analyses of two cross-validations on seventeen proposed individual channel-wise features and four inter-category feature combinations in this work, both experiments demonstrated significant differences in performance of SOZ identification, as shown in Table 3. The feature rank of significance for SOZ identification is given as p-value calculated with the Student t-test method, as shown in Table 1. And, the optimization of combination of various ranked features is presented in Fig. 4. Moreover, this study demonstrated that supervised machine learning method on combining optimal channel-wise features can be more accurate in performing SOZ identification than those either of supervised method with arbitrary combination of features [17], [34], [35] or of unsupervised method with single feature [26], [32], [33], essentially by reducing the inter-patient variability. The analyses highlights the importance of considering the feature selection and suggests an optimal combination with the top ranked features in order to accurately and efficiently identify the SOZ contacts from normal brain tissue.

C. IDENTIFY SOZs WITH SHORT RECORDINGS DURING SLEEP

Epilepsy is characterized by recurrent unprovoked seizures that are occasionally excessive and abnormal electrical discharges of cerebral neurons. Since the apparent random nature of ictal events, epileptic patients is usually underwent a long-time iEEG monitoring for pre-surgical evaluation. The surgical intervention of epilepsy aims at removing the entire epileptogenic zone, defined as the region which is indispensable for generating seizures. However, identifying this brain area by visually inspecting the seizure onset in a mass iEEG recordings is challenging and time consuming, as seizures are typically short events with changes in awareness, changes in feelings or sensations, and sometimes there is no any seizure even during more than ten days monitoring [5]–[7]. Frequency based analysis of electrical signals recorded from the brain are popular among the neuroscience community due to the inherent spectral encoding of brain activities. Recently, interictal SOZ localization have attracted many research groups’ interest in the context of the novel biomarker of HFOs, with the main objective of finding and validating of a single feature or an arbitrary combination of them that can be used in all patients [25], [34], [35]. As such, the possible value of HFOs recorded interictally is of special interest, since it focuses on collecting high frequency oscillations during interictal time rather than requiring to record seizures. With the optimization of combination of channels of features, interictal HFOs analyses in two hours iEEG data during a sleep-wake cycle have been demonstrated useful for SOZ identification in this work. This result suggests that it would be technically possible for our proposed method for SOZ localization to be useful in a short-term recording in clinical.

D. LIMITATIONS AND FUTURE WORKS

At present there is no available biomarker of the epileptogenic zone to reliably identify relevant pathologic tissue. Although this study demonstrated an obvious improvement in the performance of the proposed method for identifying SOZ by using a set of optimal channel-wise features of HFOs, in comparison with other recent works, there is still a long way to go before it to be used clinically because of several important limitations as followings. First, the relatively small patient data was collected in the present work. Collecting a great number of clinical data from patients with various pathological epilepsies, various age groups, during different sleep states, and simultaneously functional imaging would be interested in our future work. Second, one should be cautious
that the placement of subdural electrodes was determined entirely by the patient’s clinical indication, inevitably leading to an undersampling of nonpathologic brain tissue, as well as uneven sampling of anatomic regions. Future studies on retrospective evaluation between estimated SOZ with the proposed method and the corresponding outcome of epilepsy surgery and further prospective investigations are of much importance for HFOs as well as these channels-wise features to be established as useful clinical tools for diagnosis and treatment of epilepsy. Meanwhile, further works on other machine learning and/or deep learning approaches are of potential for automatically extracting and selecting features for better performance, which can automatically learns and discovers the complex features from the raw data and performs the classification in an end-to-end manner [36], [40].

V. CONCLUSION

The current gold standard for identification of the epileptogenic zone is the seizure onset zone (SOZ). However, HFOs are complex dynamic phenomena and identifying channels with HFOs within the SOZ or not is challenging. This study suggested that characterization of the spatial and temporal distribution of HFOs in channel-wise instead of event-level as usual and their integration with machine learning approach could serve as an efficient and useful tool for automatic identification of SOZ. By comparing with various combinations of the channel-wise features in same category and inter-category, to our best knowledge, we investigated firstly and systematically the temporal and spatial distribution of HFOs in channel-wise regarding automated identification of SOZ with their traditional and customized characteristics representing electrophysiological knowledge, including the rate, duration, amplitude, and power of the HFOs. Our results indicate that the proposed channel-wise features and the combination with machine learning approach can automatically localize the zone of seizure onset or at least estimating a region of interest containing or closed to the epileptogenic zone before surgery procedure, and would provide a useful and applicable pre-surgery guideline for operator and potentially reduce the time needed for long-term recording and manual inspection.

ACKNOWLEDGMENT

The authors would like to acknowledge Biomedical Imaging and Electrophysiology Laboratory (BMI-EP), University of Electronic Science and Technology of China, for providing computational and biomedical equipments.

REFERENCES

[1] World Health Organization. Epilepsy. Accessed: Apr. 2, 2018, [Online]. Available: https://www.who.int/mediacentre/factsheets/fs999/en/
[2] W. C. Stacey, “Seizure prediction is possible—now let’s make it practical,” EmBioMedicine, vol. 27, pp. 3–4, Jan. 2018.
[3] P. Kwan and M. J. Brodie, “Early identification of refractory epilepsy,” New England J. Med., vol. 342, no. 5, pp. 314–319, Feb. 2000.
[4] F. Rosenow, “Presurgical evaluation of epilepsy,” Brain, vol. 124, no. 9, pp. 1683–1700, Sep. 2001.
[5] D. Wang, D. Ren, K. Li, Y. Feng, D. Ma, X. Yan, and G. Wang, “Epileptic seizure detection in long-term EEG recordings by using wavelet-based directed transfer function,” IEEE Trans. Biomed. Eng., vol. 65, no. 11, pp. 2591–2599, Nov. 2018.
[6] G. Wang, D. Ren, K. Li, D. Wang, M. Wang, and X. Yan, “EEG-based detection of epileptic seizures through the use of a directed transfer function method,” IEEE Access, vol. 6, pp. 47189–47198, 2018.
[7] G. Wang, Z. Sun, R. Tao, K. Li, G. Bao, and X. Yan, “Epileptic seizure detection based on partial directed coherence analysis,” IEEE J. Biomed. Health Inform., vol. 20, no. 3, pp. 873–879, May 2016.
[8] G. Wang and D. Ren, “Effect of Brain-to-Skull conductivity ratio on EEG source localization accuracy,” BioMed Res. Int., vol. 2013, pp. 1–10, 2013, doi: 10.59346.
[9] K. Malgren and A. Edelvik, “Long-term outcomes of surgical treatment for epilepsy in adults with regard to seizures, antiepileptic drug treatment and employment,” Seizure, vol. 44, pp. 217–224, Jan. 2017.
[10] A. Bragin, J. Engel, C. L. Wilson, I. Fried, and G. W. Mathern, “Hippocampal and entorhinal cortex high-frequency oscillations (100–500 Hz) in human epileptic brain and in kainic acid-treated rats with chronic seizures,” Epilepsia, vol. 40, no. 2, pp. 127–137, Feb. 1999.
[11] M. Zijlmans, P. Jirsa, K. Leijten, J. G. R. Jefferys, and J. Gotman, “High-frequency oscillations as a new biomarker in epilepsy,” Ann. Neurol., vol. 71, no. 2, pp. 169–178, Feb. 2012.
[12] J. Jacobs, P. LeVan, R. Chander, J. Hall, F. Dubeau, and J. Gotman, “Interictal high-frequency oscillations (80–500 Hz) are an indicator of seizure onset areas independent of spikes in the human epileptic brain,” Epilepsy, vol. 49, no. 11, pp. 1893–1907, Nov. 2008.
[13] J. Jacobs, R. Staba, E. Asano, H. Otsubo, J. Y. Wu, M. Zijlmans, I. Mohamed, P. Kahane, F. Dubeau, V. Navarro, and J. Gotman, “High-frequency oscillations (HFOs) in clinical epilepsy,” Prog. Neurobiol., vol. 98, no. 3, pp. 302–315, 2012.
[14] J. D. Jirsch, “High-frequency oscillations during human focal seizures,” Brain, vol. 129, no. 5, pp. 1593–1608, Apr. 2006.
[15] P. A. House, “High-frequency electroencephalographic oscillations correlate with outcome of epilepsy surgery,” Yearbook Neurol. Neurosurg., vol. 2010, pp. 227–229, 2010.
[16] T. Fedele, S. Burnos, E. Boran, N. Krayenbühl, P. Hülfter, T. Grunwald, and J. Sarnthein, “Rejection of high frequency oscillations predicts seizure outcome in the individual patient,” Sci. Rep., vol. 7, no. 1, Oct. 2017, Art. no. 13836.
[17] J. Cimbalkin, B. Brinkmann, V. Kremen, P. Jurak, B. Berry, J. V. Gompel, M. Stead, and G. Worrell, “Physiological and pathological high frequency oscillations in focal epilepsy,” Ann. Clin. Transl. Neurol., vol. 5, no. 9, pp. 1062–1076, Aug. 2018.
[18] A. Matsumoto, B. H. Brinkmann, S. Matthew Stead, M. Tucewicz, W. R. Marsh, F. Meyer, and G. Worrell, “Pathological and physiological high-frequency oscillations in focal human epilepsy,” J. Neurophysiol., vol. 110, no. 8, pp. 1958–1964, Oct. 2013.
[19] R. Sakuraba, M. Iwasaki, E. Okumura, K. Jin, Y. Kakisisa, K. Kato, T. Tominaga, and N. Nakasato, “High-frequency oscillations are less frequent but more specific to epileptogenicity during rapid eye movement sleep,” Clin. Neurophysiol., vol. 127, no. 1, pp. 179–186, Jan. 2016.
[20] T. Nagasawa, C. Juhász, R. Rothermel, K. Hoechtstetter, S. Sood, and E. Asano, “Spontaneous and visually driven high-frequency oscillations in the occipital cortex: Intracranial recording in epileptic patients,” Hum. Brain Mapping, vol. 33, no. 3, pp. 569–583, Mar. 2011.
[21] H. Akawadi, N. Guanwardi, L. Goncharova, D. Spencer, J. L. Gerrard, H. Zaveri, R. B. Duckrow, H. Blumenfeld, and L. J. Hirsch, “The spatial and signal characteristics of physiologic high-frequency oscillations,” Epilepsia, vol. 55, no. 12, pp. 1986–1995, Dec. 2014.
[22] K. Kerber, M. Dümpelemann, B. Schelter, P. Le Van, R. Korinthenberg, A. Schulze-Bonhage, and J. Jacobs, “Differentiation of specific ripple patterns helps to identify epileptogenic areas for surgical procedures,” Clin. Neurophysiol., vol. 125, no. 7, pp. 1339–1345, Jul. 2014.
[23] A. Pearce, D. Wulsin, J. A. Blanco, A. Krieger, B. Litt, and W. C. Stacey, “Temporal changes of neocortical high-frequency oscillations in epilepsy,” J. Neurophysiol., vol. 110, no. 5, pp. 1167–1179, Sep. 2013.
[24] U. Malinowska, G. K. Bergey, J. Harezlak, and C. C. Jou, “Identification of seizure onset zone and preictal state based on characteristics of high frequency oscillations,” Clin. Neurophysiol., vol. 126, no. 8, pp. 1505–1513, Aug. 2015.

D. Lai et al.: Channel-Wise Characterization of HFOs for Automated Identification of the SOZ
[25] K. Kanazawa, R. Matsumoto, H. Imamura, M. Matsushita, T. Kikuchi, T. Kunieda, N. Mikuni, S. Miyamoto, R. Takahashi, and A. Ikeda, “Intracranially recorded ictal direct current shifts may precede high frequency oscillations in human epilepsy,” Clin. Neurophysiol., vol. 126, no. 1, pp. 47–59, Jan. 2015.

[26] E. E. Geerdesema, G. H. Visser, D. N. Velts, S. P. Claus, M. Zijlmans, and S. N. Kalitzin, “Automated seizure onset zone approximation based on nonharmonic high-frequency oscillations in human interictal intracranial EEGs,” Int. J. Neural Syst., vol. 25, no. 05, Jun. 2015, Art. no. 1550015.

[27] J. A. Blanco, M. Stead, A. Krieger, J. Viventi, W. R. Marsh, K. H. Lee, G. A. Worrell, and B. Litt, “Unsupervised classification of high-frequency oscillations in human neocortical epilepsy and control patients,” J. Neurophysiol., vol. 104, no. 5, pp. 2900–2912, Nov. 2010.

[28] 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, Jul. 2018, Art. no. 1850001.

[29] S. V. Gliske, Z. T. Irwin, C. Chestek, G. L. Hegeman, B. Brinkmann, O. Sahger, H. J. L. Garton, G. A. Worrell, and W. C. Stacey, “Variability in the location of high frequency oscillations during prolonged intracranial EEG recordings,” Nature Commun., vol. 9, no. 1, Jun. 2018, Art. no. 2155.

[30] H. Guragain, J. Cimbalkin, M. Stead, D. M. Gropp, B. M. Berry, V. Krennin, D. Kenney-Jung, J. Britton, G. A. Worrell, and B. H. Brinkmann, “Spatial variation in high-frequency oscillation rates and amplitudes in intracranial EEG,” Neurology, vol. 90, no. 8, pp. e639–e646, Jan. 2018.

[31] R. J. Staba, C. L. Wilson, A. Bragin, I. Fried, and J. Engel, “Quantitative analysis of high-frequency oscillations (80–500 Hz) recorded in human epileptic hippocampus and entorhinal cortex,” J. Neurophysiol., vol. 88, no. 4, pp. 1743–1752, Oct. 2002.

[32] P. M. Murphy, A. J. von Paternos, and S. Santaniello, “A novel HFO-based method for unsupervised localization of the seizure onset zone in drug-resistant epilepsy,” in Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Sepgwoip, South Korea, Jul. 2017, pp. 1054–1057.

[33] P. M. Murphy, A. J. von Paternos, and S. Santaniello, “A novel HFO-based method for unsupervised localization of the seizure onset zone in drug-resistant epilepsy,” in Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Sepgwoip, South Korea, Jul. 2017, pp. 1055–1057.

[34] S. Liu, C. Gurses, Z. Sha, M. M. Quach, A. Sencer, N. Bebek, D. J. Curry, S. Prabhu, S. Tummala, T. R. Henry, and N. F. Ince, “Stereotyped high-frequency oscillations discriminate seizure onset zones and critical functional cortex in focal epilepsy,” Brain, vol. 141, no. 3, pp. 713–730, Jan. 2018.

[35] S. Liu, C. Gurses, Z. Sha, M. M. Quach, A. Sencer, N. Bebek, D. J. Curry, S. Prabhu, S. Tummala, T. R. Henry, and N. F. Ince, “Stereotyped high-frequency oscillations discriminate seizure onset zones and critical functional cortex in focal epilepsy,” Brain, vol. 141, no. 3, pp. 713–730, Jan. 2018.

[36] N. von Ellenrieder, B. Frauscher, F. Dubeau, and J. Gotman, “Interaction with slow waves during sleep improves discrimination of physiologic and pathologic high-frequency oscillations (80-500 Hz),” Epilepsia, vol. 57, no. 6, pp. 869–878, May 2016.

[37] Y. Varatharajah, B. Berry, J. Cimbalkin, V. Krennin, J. Van Gompel, M. Stead, B. Brinkmann, R. Iyer, and G. Worrell, “Artificial intelligence with real-time intracranial EEG monitoring to automate interictal identification of seizure onset zones in focal epilepsy,” J. Neurol. Eng., vol. 15, no. 4, Jun. 2018, Art. no. 046035.

[38] D. Lai, X. Zhang, K. Ma, Z. Chen, W. Chen, H. Zhang, H. Yuan, and L. Ding, “Automated detection of high frequency oscillations in intracranial EEG using the combination of short-time energy and convolutional neural networks,” IEEE Access, vol. 7, pp. 82501–82511, 2019.

[39] J. F. Box, “Guinness, gosset, Fisher, and small samples,” Stat. Sci., vol. 2, no. 1, pp. 45–52, Feb. 1987.

[40] Y. Li, C. Lin, and W. Zhang, “Improved sparse least-squares support vector machine classifiers,” Neurocomputing, vol. 69, nos. 13–15, pp. 1655–1658, Aug. 2006.

[41] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, pp. 436–444, May 2015.

[42] G. P. Ren, J. Q Yan, X. X Yu, “Automated detector of high frequency oscillations in epilepsy based on maximum distributed peak points,” Int. J. Neural Syst., vol. 28, no. 1, 2018, Art. no. 1750029.

[43] M. Amir, J.-M. Lina, F. Puzzo, and J. Gotman, “High frequency oscillations and spikes: Separating real HFOs from false oscillations,” Clin. Neurophysiol., vol. 127, no. 1, pp. 187–196, Jan. 2016.

[44] R. Zuo, J. Wei, X. Li, C. Li, C. Zhao, Z. Ren, Y. Liang, X. Geng, C. Jiang, X. Yang, and X. Zhang, “Automated detection of high-frequency oscillations in epilepsy based on a convolutional neural network,” Frontiers Comput. Neurosci., vol. 13, Feb. 2019, Art. no. 6.
TONGZHOU KANG received the bachelor’s degree in microelectronics science and technology from Hubei University, Hubei, China, in 2019. He is currently pursuing the master’s degree with the School of Electronic Science and Engineering, University of Electronic Science and Technology of China (UESTC). His current research interest includes signal processing.

HAN YUAN received the B.S. degree in biomedical engineering from Tsinghua University, Beijing, China, in 2005, and the Ph.D. degree in biomedical engineering from the University of Minnesota, Minneapolis, MN, USA, in 2010. She is currently an Assistant Professor with the Stephenson School of Biomedical Engineering, The University of Oklahoma, Norman, OK, USA. Her research interests include multimodal functional neuroimaging and neuromodulation.

LEI DING (Member, IEEE) received the B.E. degree (Hons.) in biomedical engineering from Zhejiang University, China, in 2000, and the Ph.D. degree in biomedical engineering from the University of Minnesota, in 2007. He is currently a Presidential Professor of biomedical engineering and neuroscience, the Director of the Institute for Biomedical Engineering, Science, and Technology, The University of Oklahoma, and an affiliated Faculty Member with the Laureate Institute for Brain Research. He has made significant original contributions to research in functional neuroimaging, noninvasive neuromodulation, brain–computer interface, and imaging biomarkers for neuropsychiatric disorders, and has published more than 120 peer-reviewed articles in these areas. He received a Dissertation Fellow from the University of Minnesota for his Ph.D. degree. He has been a member of the IEEE Engineering in Medicine and Biology Society (EMBS), since 2001. He is also active in professional activities, such as ISBEM, OHBM, and SfN. He was a recipient of the National Science Foundation CAREER Award and the Early Career Achievement Award of 2016 from the IEEE-EMBS. His work was selected as one of only 13 Technologies of the Year by Smithsonian Innovation Festival showcased at the National Museum of American History. He serves as an Associate Editor for the IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING and IEEE ACCESS and as a Reviewer for more than 40 prestigious journals. He serves on editorial boards for several other biomedical journals. He has been a member of the IEEE-EMBS Technical Committees on BSP and BIIP and an Associate Editor at the EMBS Conference Editorial Board, since 2008. He has been serving as the Conference Program Chair/Co-Chair, the Theme Chair/Co-Chair, the Track Chair/Co-Chair, or the Session Chair/Co-Chair on EMBS annual conferences and/or other EMBS special topic conferences, since 2007.

***