



ISSN: 2454-9940



**INTERNATIONAL JOURNAL OF APPLIED
SCIENCE ENGINEERING AND MANAGEMENT**

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.in

www.ijasem.in

Evaluation of High Frequency Oscillation with Cross Validation and End-to-End Bi-Branch Neural Network

APPARAO KAMARSU Dr. KANAKA DURGA RETURI Dr. VAKA MURALI MOHAN

Abstract

In order to accurately localize epileptic foci and predict the prognosis of drug-refractory epilepsy, high-frequency oscillation (HFO) detection is crucial. Clinicians may save time and effort by investigating a high-performance automated detection approach for HFOs. Existing approaches struggle to fulfill the needs of clinical application because of their narrow analytical viewpoint and straightforward model creation. To achieve full-stack automated detection of HFOs, the results from two separate branches are combined. We have validated our technique on five people with intractable epilepsy. High sensitivity (94.62%), specificity (92.7%), and F1-score (93.33%) were obtained using the suggested technique during intravalidation; high sensitivity (92.00%), specificity (88.26%), and F1-score (89.11%) were obtained during cross-validation. Based on the obtained data, it can be concluded that the suggested approach is superior to the two conventional detection paradigms, namely the single signal strategy and the single time frequency diagram strategy. In addition, when it comes to visual analysis vs. automated detection, the kappa value is a solid 0.795 on average. Meanwhile, /e approach has impressive generalization ability and excellent consistency with the benchmark.

Introduction

Preoperative evaluation of patients with drug-refractory epilepsy relies on precise localization of the epileptogenic zone (EZ) [1-3]. A high prognosis is strongly connected with surgical excision of the channel with a high incidence of high frequency oscillations (HFOs) [4, 8], making the detection of an

HFO signal with a frequency of 80-500 Hz of major value for correct localization of EZ. The synchronized transient of neurons is reflected in HFOs, which are a kind of spontaneous electroencephalogram gram pattern [9]. Rs (Ripples,

80-250 Hz), FRs (Fast Ripples, 250-500 Hz), and VHFOs (very high frequency oscillations, 1000-2500 Hz) are the three categories into which /eye fall according to frequency [10, 11]. Currently, the gold standard in clinical diagnosis is the visual interpretation of HFOs by physicians using long-term stereo electroencephalography (SEEG) and video recordings [12-15]. Nonetheless, one patient since continuous monitoring is often required for many days to a week, manual processing of such a massive quantity of data is beyond the reach of clinical staff

Software Developer, Mann-India Technologies Pvt.Ltd., Noida, Uttar Pradesh, India
Dean Academics & HoD, CSE, Malla Reddy College of Engineering for Women,
Maisammaguda, Medchal, Hyderabad, Telangana, India
durga1210@gmail.com
Principal & Professor of CSE, Malla Reddy College of Engineering for Women,
Maisammaguda, Medchal, Hyderabad, Telangana, India

There have been several reports of automated HFO detectors from a variety of investigations. Both the band-pass signal and the time-frequency diagram have been used traditionally for HFO detection. Over the last decade [17–24], researchers have examined single-step detection techniques with varied signal characteristics, such as the Teaser energy operator, wavelet entropy, fuzzy entropy, short-time energy, and so on. While these techniques can identify HFOs, they have trouble distinguishing them from other abnormalities such as spikes, pulse-like artifacts, and harmonically rich signals [20]. Two-stage approaches [25-29] are presented as a solution to this issue, with the first stage consisting of a detector, and the second stage consisting of either a supervised classifier or an unsupervised clustering to further investigate the signal's features. To be more specific, the two-stage automatic de section paradigm has always been used in terms of the time-frequency diagram, and a stacked convolution neural network (CNN) [28] or a stacked demising auto encoder [29] can identify genuine HFOs from candidate events isolated from background activities. In instance, following an initial detector [25, 30-32], the basic 2d-CNN structure is typically employed to extract time-frequency picture features. Yet, it remains difficult to properly understand information that is worth paying attention to in signals or temporal frequency pictures.

Substances and Techniques

Materials

Acquisition of Data. Information on the five patients with drug-resistant epilepsy who participated in the trial is provided in Table 1 below. Prior to surgery, all patients have had regular scalp EEGs and other imaging studies (including MRI and PET). Simultaneously, physicians from various backgrounds discussed each case. When considered with preoperative diagnostic assessment,

Table 1: Clinical characteristics of five patients

Patient no., gender	Age	Epilepsy duration	Surgical pathology	Implantation sites	No. of selected/total channels (n)
P1, male	21	7	HS	LH, LA, LBF, LIF	14/78
P2, female	33	5	HS	LH, LA, LBF, LIF, RH	24/77
P3, female	36	20	HS	LH, LA, LFO, LAT, RH	17/84
P4, male	33	7	HS	RH, RA, RBF	19/77
P5, female	38	23	HS	LH, LA, LOF, LFO, RH	15/84
Mean ± SD	32 ± 6	12 ± 8	—	—	18 ± 4/80 ± 3

Hippocampus sclerosis, anatomically speaking, is a term used in the field of surgical pathology. Left hippocampus, left amygdale, left basis frontal, left inferior frontal, left front parietal operculum, left anterior temporal, left orbit frontal, right

hippocampus, right amygdale, right basis frontal, right orbit frontal. Number of channels with epileptic activity as a percentage of all recorded channels (n).

Stereotactic EEG and electrode planning have demonstrated promise in locating epileptic hotspots [37]. A robotic stereotactic surgery assistant workstation (robotied stereotactic assistant, ROSA, from a French company named ManTech) was used to import enhanced thin-scan MRI and thin-scan CT (layer thickness 1 mm, interval 1 mm) for data fusion, and subsequently, an electrode implantation plan was developed. We utilized 0.8 mm SEEG electrodes with 5-18 connections (2 mm long, 1.5 mm apart) as illustrated in Figure 1. (A). Figure 1 displays a typical electrode reconstruction of the medial temporal lobe (b). Distance, or range, All of the patients' SEEG data is recorded at 2048 Hz and saved as.eeg files; the Nat us Neuroworks program is used to examine the patients' videos. We randomly choose the SEEG data of each patient's waking and sleeping periods for 2 hours from their corresponding video records. There is at least one hour between the time of the chosen record and the patient's onset. /e EDF format and the channel distribution for each patient are also stored. Please take note that no preprocessing steps were taken prior to exporting this data.

Preparing the Data.

incipient pauses Segmenting the data, reversing the polarity, removing defective channels, applying filters (such as band-pass and notch), and so on are all necessary steps in preparing the SEEG data of each subject. Figure 2 depicts the whole procedure. For the most part, the acquired SEEG were large-scale long-term recordings; 2 hours of data takes up around 8 G of memory, but the computer and software can work with only so much space. Since the segmenting signals may provide the same data as the lengthier recordings, a portion of them had to be intercepted for the experiment; typically, 30-minute signal segments were acquired for processing. Since the original SEEG requires a polarity conversion procedure, the bipolar data, whose waveforms and amplitudes are often less distorted, may now be used. We eliminated obviously disrupted channels and empty electrodes before to utilizing them since interference is a common occurrence during EEG recording and because empty electrodes are often used in clinical practice. Additionally, signal acquisition gear operates on a 50 Hz mains power supply, necessitating the adoption of a 50 Hz frequency-doubling notch filter to eliminate power-frequency interference and frequency-doubling noise.

Finally, the band-pass filter keeps the signals between 80 and 500 Hz because, relative to the lower-frequency background noise, the HFOs are more energetic in this range. The preprocessing uses a number of programs, including Any Wave, EDFBrowser, and others. In particular, we utilize EDFBrowser to segment the original, discontinuous EEG data into continuous signal fragments; we then use AnyWave to conduct polarity conversion, poor channels elimination, band-pass filtering, and notch filtering on these signals.

Robotic Visual Labeling

The first detector in the two-stage automated detection process performs a preliminary screening of the data acquired and preprocessed in the aforementioned procedure to get the HFO candidate event set. A certain percentage of the samples in this pool were likely to be erroneous, meaning they were not HFOS. In light of this, we enlisted the aid of clinical specialists in determining which HFOs were legitimate and which were not, using their judgment as the benchmark in this research. Two licensed medical experts perform the /e marking process. Our database contains actual HFOs with frequencies between 80 and 500 Hz; their ripple frequency varies from 80 to 250 Hz, and their fast ripple frequency spans from 250 to 500 Hz. There is also a variety of non-HFO artifacts in the HFO database. The original EEG data, a band-pass signal spanning 80–500 Hz, and a time–frequency diagram generated using the wavelet transform are all used as references in the /e marking procedure. After filtering, if the signal's amplitude is much greater than the baseline and there is an islanding effect in the time-frequency diagram, it is determined to have met the HFO signal standard and is filed away accordingly. However, the negative samples that do not fit the standard will also be labeled and stored. In /us, we have established both the positive and negative sample data sets. Waveforms of various common signals are seen in Figure 3. We classify items (a) and (b) as genuine HFOs, whereas items (c), (d), and (e) are examples of three distinct kinds of artifacts that make up the non HFOs collection. As can be seen in the center row, filtering transient signals (such as epileptic spikes, sharp waves, and sharp artifacts) or signals with harmonics may result in "fake" high-frequency oscillation events, with waveforms in the filtered signal that seem similar to those of the true HFOs.

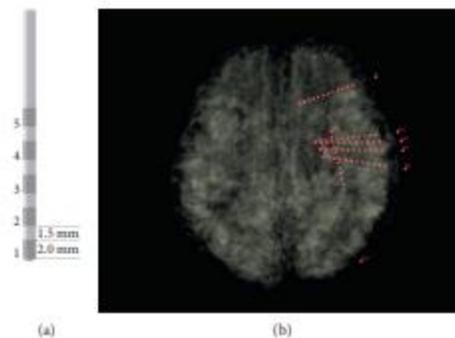
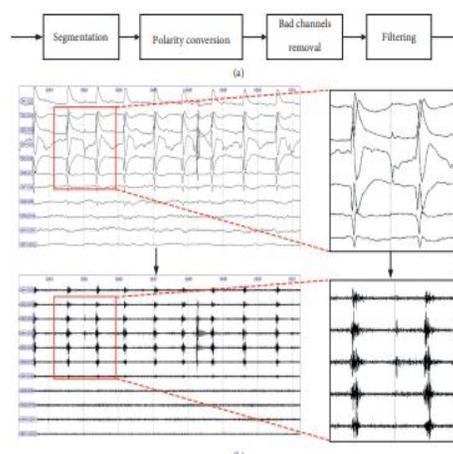


Figure 1: SEEG deep brain electrode. (a) SEEG electrode. (b) Schematic diagram of typical medial temporal lobe electrode reconstruction.



Preparing the Data, Fig. 2. Preprocessing, in this context, refers to (a) the procedure through which data is prepared. (b) The preprocessed data results; the original waveform is shown in the top row, and the processed version is shown in the bottom row. The right column displays information from the red box.

Initial detector mistake, etc. However, as shown in the bottom row, the temporal frequency diagram for these fake events is considerably different from that of genuine HFO events. The time-frequency plot of spike-type artifacts, often created by a band-pass filter, typically displays a "island-" like increasing trend, whereas the time-frequency plot of harmonic-type abnormalities displays dispersed high energy over the whole frequency range [38]. In order to guarantee impartiality and efficiency, two physicians independently labeled a sufficient number of samples, compared their findings, and settled on a single labeling principle. Everyone who applied for a position was given a label at last. There were 16167 records created, of which 7754 were considered

positive samples and 8413 were considered negative samples. The data is shown in Table 2.

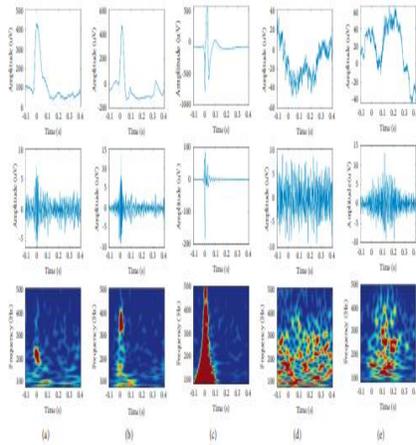


Fig. 3 is an example of visual marking. The raw data may be found in the top row, the filtered signal (80-500 Hz) in the center, and the time-frequency representation in the bottom row. Ripple (a), rapid ripple (b), artifact 1 (band-pass filter-induced high-frequency transients), artifact 2 (amplitude greater than global background but not statistically different from local activity), and artifact 3 (harmonics of low frequency no sinusoidal signals) are all examples of artifacts.

TABLE 2: Clinical database of HFOs.

Patient no.	Data batch ¹	Positive ²	Negative ³	Total
Pt1	1	2905	2916	5821
Pt2	1	2411	2343	4754
Pt3	1	401	517	918
	2	412	604	1016
Pt4	1	412	555	967
	2	488	508	996
Pt5	1	225	533	758
	2	500	437	937
Total	—	7754	8413	16167

¹Data batch represents different batches/times of data collection. ²Number of positive samples. ³Number of negative samples.

Synopsis of the Whole System.

Figure 4 depicts the whole procedure of our study's recommended technique. At first, the raw SEEG signals of patients were gathered and preprocessed in various ways (segmentation, filtering, etc.), and a threshold detector was used to establish a preliminary candidate event set of clinical high-frequency oscillation. To create 2d time-frequency pictures, the data were processed using continuous wavelet transforms (CWT). We built our own private data set by combining the time-frequency diagram, the raw signal, and the filtered signal with the neuroscientists'

visual markings distinguishing genuine from spurious HFOs. Finally, a bi-branch fusion model for automated HFO identification was developed. The model's input data was split between the two modalities, and from there, a 1d-ResNet-LSTM hybrid network was built for the signal branch and a 2d-ResNet-CBAM hybrid network was built for the Topic branch. First, we fused the output of the two portions by building a fusion module, in which the multilayer perceptron (MLP) classifier is used to identify the results from the feature learning of the two modalities' data. In order to make the training as efficient as possible, we decided to use a synchronous approach. After the model was trained, it was put through its paces with some test data to see how well it did at making classifications.

Nine-Threshold Detector-Based Primary Detector. As a first step in our automated detection technique, we developed a detector that is as good as possible at picking up HFOs in a distant recording, which means that our underlying algorithm must exhibit both high sensitivity and low specificity. When the band-pass filtered signal exhibits at least three continuous peaks greater than 3-5 standard deviations, the signal is regarded to be an authentic HFO [39]. Clinicians should just visually designate the signal inside the suspected event set based on the first detection. The effectiveness of HFO labeling may be enhanced by using /is technique.

This is the basic threshold-based detection algorithm:

- (1) Using the filtered signals, we determined the subject's standard deviation (SD) for each channel, set the threshold at 2.5 times the peak value, and tallied the number of times each signal's peak value was over the threshold.

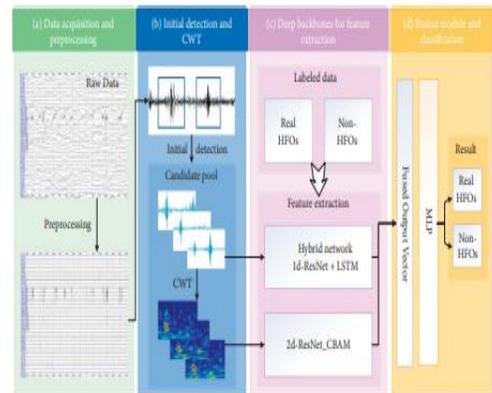


Figure 4: Flowchart of the Suggested Procedure To begin with, we must collect

and prepare the data. b) The preliminary detection and CWT. Extraction of features from deep neural networks (c). Classification and fusion module every one of the 128 sampled locations had the number of times it took to cross over from the threshold to the peak value recorded, along with the coordinates of the sites where this number was larger than three did. A Hilbert transform was used to extract the envelope of the ω signal, and a threshold value of three times the median of the background signal was used to record any positions that exceeded the threshold. For the location that satisfies all three criteria at once, a total of 0.5 seconds of signal is retrieved as a possible HFO: 0.1 seconds before the location and 0.4 seconds after it. Case in point of applying this preliminary detection is shown in Figure 5.

It is a fusion model with two branches.

A pool of HFO suspicious events was established after the first detector was run. Overall, the model is fed data from the 80-500 Hz band-pass signal and the 2d time-frequency plot. In order to extract the high-order characteristics of separate modal data, two deep backbone networks were constructed (1d Resnet + LSTM and 2d-ResNet CBAM) for the two inputs, respectively. To be more precise, the Topic branch is implemented using a 2d-ResNet that has a CBAM module implanted behind each convolution block, while the signal branch is constructed as a hybrid network with a 1d-ResNet and an LSTM linked in parallel. Both branches' output vectors were combined into one. In the end, a multilayer perception was employed to distinguish between genuine and fake HFOs. Figure 6: 2.4.1. Signal Branch, showing the general design of the bi-branch model. We employ band-pass filtered sounds between 80 and 500 Hz as the model's input data to eliminate low frequency in dereference [25] and the influence of a small quantity of irrelevant frequency bands greater than 500 Hz on HFO identification [40, 41]. See Figure 7 for a visual representation of the signal trunk's overall architecture. To achieve deep feature learning from each signal branch, we used a hybrid network

based on a 1d-ResNet34 and LSTM, with the two networks being linked in simultaneously.

First, the morphological features of the signal are modeled with the help of the 1d-ResNet34. To accomplish this goal, we use a specific sort of convolution neural network (Resnet) that relies on a residual link. Assuming that the mapping relation to be solved is $H(x)$, the residual network decomposes it into two components, as follows:

$$H(x) = F(x) + x,$$

In addition, F denotes the residual function (\bullet) [42]. Learning an identity mapping $H(x) = x$ at the top level of the network is the same as reducing the residual component to zero, or $F(x) = 0$. In this configuration, the element's input is simply added to its output, bypassing any intermediate stages. To be more specific, we adapted the 2d-ResNet model presented by The et al. [42] and recast it as a 1d model. ω the 1d-ResNet model has five steps. After the first stage of a 7×7 convolution with a 2×2 stride, the feature map is just a quarter the size of the input. The last four steps include stacking the remaining four bricks. For ResNet34, the fundamental residual block specified by The et al. [42] is used. Two 33 convolutions are placed on top of one another to form each block. Block stacks number [3, 4, 6, and 3].

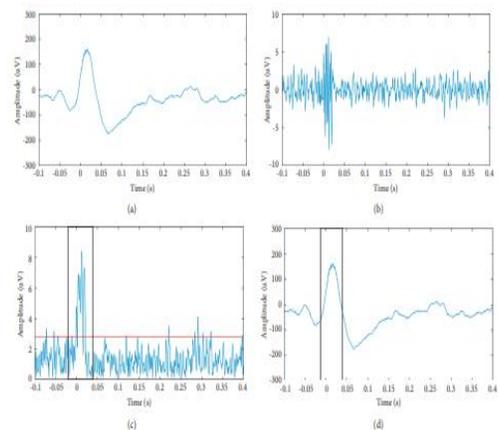
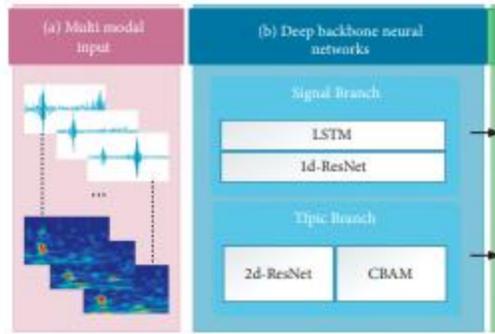


Fig. 5 is an unprocessed example of a first detection. B. signal using a band-pass filter (80–500 Hz). (c) A Hilbert transform envelope plot, where the red line indicates the median value of background activity and the black box indicates an excess. Black

box at (d)/ displays first detection findings (e).



An overview of the bi-branch feature fusion model is provided in Figure 6. A flood of data from several places. This is because (b) deep backbone networks. Module for fusing features and categorizing them.

To understand temporal correlation, however, (b) the network design incorporates a variation of RNN (LSTM). For feature extraction, we employ a multi-layer long short-term memory (LSTM) block with 100 hidden units.

(c) On the basis of the spatial modeling of signals by CNN (1d-ResNet), complemented by the temporal modeling by RNN (LSTM), a hybrid network is developed to represent signal attributes from multiple angles. The signal's higher-order features are captured faithfully. Two-dimensional output vectors are fused and spliced to produce the multiperspective fusion features of the filtered signals.

In the Topic Department. The model can distinguish between noise and true HFOs in the time-frequency image if a Topic branch is created. Time-frequency transformation must be performed first on the filtered SEEG signal from the patient. Epilepsy telltale EEG signs.

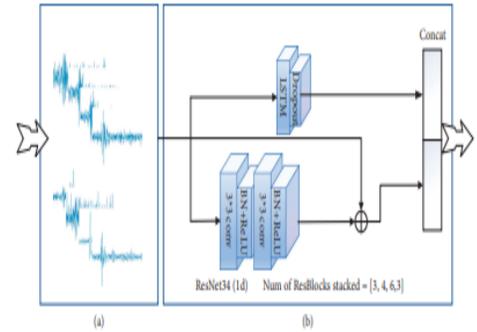


Illustration of the model's signal-branch structure (Figure 7). Band-pass filtered signal (a). Specifically, (b) 1d-ResNet + LSTM's /e-centric architecture.

A stochastic signal that fluctuates both temporally and spatially. Multiresolution analysis of random no stationary EEG data is ideally performed using the /e wavelet transform. The signal is transformed using a Morley wavelet, and its wavelet basis function looks like this:

$$\Psi_{a,b}(t) = \exp\left(i\omega_0 \frac{t-b}{a}\right) \exp\left(-\frac{(t-b)^2}{2a^2}\right).$$

Where zero is the fundamental frequency, and is the wavelet basis stretching amount and b is the wavelet basis translation amount. The formula of the wavelet transform may be obtained from the characteristics of the Fourier transform as

$$\begin{aligned} \Psi f(a,b) &= \int_{-\infty}^{+\infty} f(t) \bar{\Psi}_{a,b}(t) dt \\ &= \int_{-\infty}^{+\infty} f(t) \exp\left(-i\omega_0 \frac{(t-b)}{a}\right) \exp\left(-\frac{(t-b)^2}{2a^2}\right) dt. \end{aligned}$$

As can be seen in Figure 8, we also constructed a deep backbone network for the Topic branch. Overall, this fork takes 2d-ResNet as its foundational framework and integrates a CBAM component into it. In the field of image classification, Resnet is a popular foundational network. According to the previous description, it supplements the traditional CNN with a residual block. Based on these findings, the research presents the CBAM module to enhance the model's performance even more. Features are extracted from the whole branch using a 2d-ResNet50, and the model as a whole consists of five steps. Resnet50, in contrast to signal branch, employs the unique residual blocks described by The et al. [42]. Each unit employs dimension reduction, convolution processing, and dimension restoration across three

convolution layers (1 1, 3 3, 1 1). A proper stacking of blocks would l

ook like this: [three, 4, 6, 3]. as an added bonus, the 2d-ResNet incorporates CBAM after every single block. The input feature graphs are placed via MLP after being processed by global maximum pooling, global average pooling based on width, and global maximum pooling based on height. The final CA feature map was created by combining the element wise added MLP output features with sigmoid activation. Woo et al. [43] provide the following formula for Computing it.

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \\ = \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))),$$

When the sigmoid function is, The Rely activation function, W0, determines the common MLP weights, W0 RC/arc and W1 RC/arc/r. The spatial attention (SA) module may perform either global maximum pooling or global average pooling of input before summing the values along the channel dimension. We utilize a convolution procedure to reduce the dimension to 1. When everything is said and done, the sigmoid is what is employed to generate the SA characteristic graph. Woo et al. [43] provide the following formula for determining it.

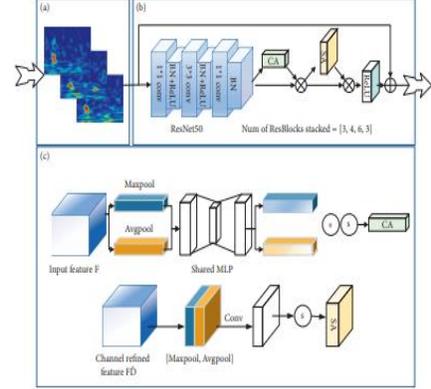
$$M_s(F) = \sigma(f^{2 \times 2}([\text{AvgPool}(F); \text{MaxPool}(F)])) \\ = \sigma(f^{2 \times 2}([F_{avg}^c; F_{max}^c])).$$

The CA module and SA module are then embedded in the back of each block of 2d ResNet50, creating a series-connected embedded network (2d-ResNet CBAM) that efficiently learns the information that is helpful to classification from the input image while simultaneously suppressing some unnecessary information.

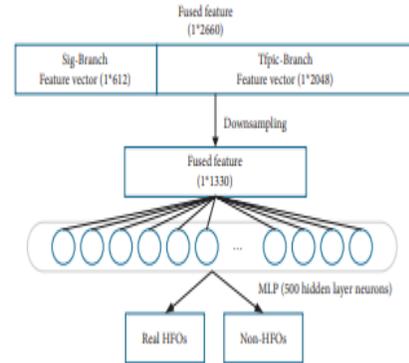
Fusion-related components.

The primary contribution of this work is a suggestion for integrating filtered SEEG signals with features of temporal frequency maps. The feature fusion module's overall structure is shown in Figure 9. A candidate

event's deep features in the 80-500 Hz band-pass signal may be obtained through the signal branch, and the event's deep features in the time-frequency image can be obtained via the Topic branch. We combine features from two routes into one cohesive whole.



Internal structure of the Topic tree model (Figure 8). In particular, (b) /e 2d-structure time-frequency plots. For Resent's CBAMs The sigmoid function (S) is indicated in (c) CBAM, where the channel attention (CA) module and the spatial attention (SA) module are shown in the top and bottom rows, respectively.



Vector for the proposed occurrence. The output of the signal branch has a size of 1612, the Topic branch's output has a dimension of 12048, and the combined output vector's dimension is 12660. After that, a half-sampling was used to get a single candidate event's fusion output vector in 1 1330 dimensions. Finally, all potential HFO events are used to train a multi-layer perception, which then determines whether each event is an HFO. The following formula may be used to calculate the combined output of the H and O layers:

$$\begin{aligned}
H &= XW_h + b_h, \\
O &= HW_o + b_o \\
&= XW_hW_o + b_hW_o + b_o.
\end{aligned}$$

Where b_h and b_o represent the bias of the hidden layer and output layer weight matrices, respectively, and X represents the input vector. Additionally, the following specification of the binary cross entropy loss function is utilized during training:

$$L_{\text{logistic}}(\hat{y}, y) = -y \log \hat{y} - (1 - y) \log (1 - \hat{y}),$$

In this equation, y represents the target value, \hat{y} the margin of error, and y the actual data point.

Results

Measurements against which outcomes may be evaluated.

Accuracy (ACC), sensitivity (SEN), specificity (SPE), precision (PRE), false discovery rate (FDR), and F1-score were used as measures to evaluate the suggested model's performance in this research. We also developed the SEN SPE-score to show that the proposed method achieves a better balance of sensitivity and specificity. Most indicators are computed using the confusion matrix. In the confusion matrix, a TP indicates that the predicted category and the true category are both P, whereas a TN shows that the predicted and true categories are conflicting with one another and are both N. A false positive (FP) happens when the predicted category is P when the true category is N, as opposed to a false negative (FN), which occurs when the anticipated category is N when it should have been P. To exactly calculate these metrics, use this algorithm:

$$\text{Accuracy (ACC)} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{Sensitivity (SEN)} = \frac{TP}{TP + FN}$$

$$\text{Specificity (SPE)} = \frac{TN}{TN + FP}$$

$$\text{Precision (PRE)} = \frac{TP}{TP + FP}$$

$$F1 \text{ score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$FDR = \frac{FP}{TP + FP} = 1 - \text{Precision}.$$

To assess how well a model strikes a compromise between these two indications, we also compute their harmonic mean (SEN SPE-score) using the formula:

$$\text{SEN_SPE - score} = \frac{2 * \text{SEN} * \text{SPE}}{\text{SEN} + \text{SPE}}$$

Table 3 displays the optimum values for all factors utilized in the experiment, including the parameter settings, training strategy, and experimental environment. The LSTM hidden size was set experimentally to 100, the LSTM layer count was set to 2, and the number of hidden layer neurons in the fusion module was set to 500 in order to provide enough model capacity and avoid overfitting simultaneously. Additionally, in both cases, the number of stacked Resnet blocks was determined to be [3, 4, 6, 3].

TABLE 3: Training parameters.

Step	Parameters	Setting
Signal branch	LSTM hidden size	100
	LSTM num layers	2
TFpic branch	Numbers of blocks stacked	[3, 4, 6, 3]
	MLP number of hidden layer neurons	500
Overall	Initial learning rate	0.01
	Batch size	32
	Training epochs	60

Signal branch and Topic branch that adhere to resnet is traditional design tenets. We set the network training hyper parameters to an initial learning rate of 0.01, a batch size of 32, and a total number of training epochs of 60. In particular, the values of each parameter were recorded at the time the model produced the best result in the validation set; this served as the basis for selecting the hyperactive parameters. As will be shown in the following section, we split up the experimental data into a training set, validation set, and test set. The model

was trained on the training set, validation set results were used to inform parameter selection, and the test set was used to verify our hypotheses. We used a bi-branch synchronous training strategy, which requires a one-to-one correspondence between the input data of the bi-branch and the output data of the model, to feed the band-pass filtered signal and time-frequency diagram into the model. When errors are produced at the output, the weight matrix is updated by cycling back through the input data of the two modes until the process is complete. The following details the conditions of our experiments. All of our training and testing was done on an HP Z8G4 graphics workstation outfitted with a single Nvidia GeForce RTX 2080Ti and 12G of dedicated video memory.

Pairs of Exercises for Practice and Evaluation.

To make the most of our database, we employ two validation strategies, intrasubject validation and cross-subject validation, and we divide the data set into five folds to generate five groups of experimental data, one for each patient record.

Validation within a single subject.

Figure 10 depicts the data partitioning and intrasubject validation implementation. To be more specific, we took 80% of the data, randomly shuffled it, and then used that as the training set, 10% of the data as the validation set, and the remaining 10% as the test set.

Validation across Subjects.

The study presented here marks a significant advance in the field by successfully applying the model to cross-patient testing. In what follows, we will demonstrate the findings with particular emphasis on /e.

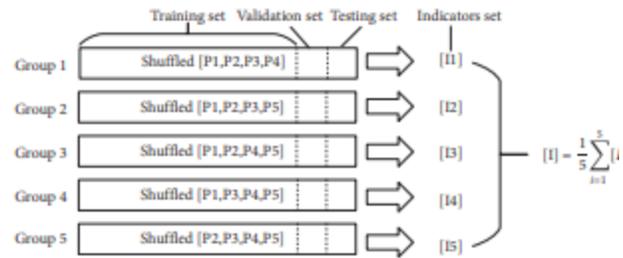


Figure 10: Intrasubject validation

Figure 11 displayed the specific data distribution and application. In particular, we employed the leave-one-out cross-validation technique, wherein 80% of the subsets served as training set and 20% as validation set and each of the 5 subsets served as test set in turn. After obtaining five retrained models using five distinct training sets, the models' respective performance was evaluated using test sets. At last, we took an average across all five data sets to see how well they performed.

The Results of Our Approach

Validation performed within the subject. Based on the data from 5 clinical patients' SEEG recordings, we used intrasubject validation to determine an average sensitivity of 94.62%, specificity of 92.70%, precision of 92.12%, accuracy of 93.62%, FDR of 7.88%, F1-score of 93.33%, and SEN SPE-score of 93.63% (Table 4).

Carrying out a Validation across Subjects.

Five sets of experiments were conducted to validate the results from one set of subjects on another. /e results showed that the average sensitivity was 92.00%, the average specificity was 88.26%, the average precision was 86.86%, the average accuracy was 89.76%, the average FDR was 13.14%, the average F1-score was 89.11%, and the avg enrage of SEN SPE-score was 89.87%, as shown in Table

Talking About

Important Findings from Our Research.

Clinical analysis of HFOs still relies mostly on visual assessment. Patients' EEGs may exhibit a high-frequency oscillating rhythm, which can be visually identified by experts. However, a new automatic detection method is required because of the difficulty of visual labeling and the high subjectivity and inconsistency among experts. Since 2002, different HFO automatic detection methods have been reported in different studies. Our work has made further improvement and innovation on the basis of predecessors. Its main contribution lies in the following three aspects.

Bi-Branch Fusion Model Realizes Complementary Advantages.

Most early proposed detectors extract features from waveforms of EEG signals to enable automatic classification. This leaves the problem of poor specificity and is prone to misdiagnosis in clinical practice because there are many artifacts and other signals in the band-pass filtered signals that cannot be distinguished from the real HFOS. Time-frequency diagrams can help with this, but they are not good enough for use in the clinic because signal classification based solely on visuals has low sensitivity and is prone to missed diagnoses. To address this problem, we proposed a novel approach wherein the filtered band-pass signal and time-frequency image are used as input data to a model, the model is established as a bi-branch deep learning model, the model's output is fused from the two branches, and HFOs and non-HFOs are automatically classified. The advantages of CNN and RNN are combined in the hybrid network (1d Resent + LSTM) used in the signal subfield. While convolution neural networks (CNNs) are responsible for feature extraction in the morphological space of signals, recurrent neural networks (RNNs) are responsible for feature extraction in the temporal dimension of signals. The 2d-ResNet CBAM model, trained on the Topic dataset, paid more attention to the informative features of the time-frequency image, allowing it to distinguish between HFOs and other

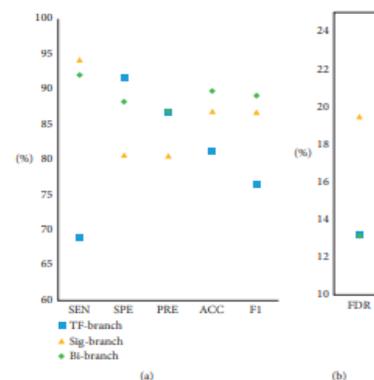
sources of noise in the Topic frequency spectrum.

Strong generalization results can be attributed to thorough cross-validation.

Because of this early emphasis on signal detection performance, most experiments split the candidate pool into a training set and a test set through a random process [21, 27, 43, 45, and 46]. To make the best decision for a new patient, it is helpful to apply the a priori knowledge gained from previous cases. Therefore, it is essential that the model account for

TABLE 8: Classification performance with different structures.

Group	Model	Confusion matrix					Evaluation metrics (%)				
		TN	FP	FN	TP	SEN	SPE	PRE	ACC	FDR	F1
1	Bi-	838	132	50	675	93.10	86.39	83.64	89.26	16.36	88.12
	TF-	866	104	254	471	64.97	89.28	81.91	78.88	18.09	72.46
	Sig	768	202	69	656	90.48	79.18	76.46	84.01	23.54	82.88
2	Bi-	954	109	120	780	86.67	89.75	87.74	88.33	12.26	87.20
	TF-	980	83	429	471	52.33	92.19	85.02	73.92	14.98	64.79
	Sig	886	177	92	808	89.78	83.35	82.03	86.30	17.97	85.73
3	Bi-	865	256	30	783	96.31	77.16	75.36	85.21	24.64	84.56
	TF-	974	147	243	570	70.11	86.89	79.50	79.83	20.50	74.51
	Sig	815	306	22	791	97.29	72.70	72.11	83.04	27.89	82.83
4	Bi-	2282	61	41	2370	98.30	97.40	97.49	97.85	2.51	97.89
	TF-	2308	35	363	2048	84.94	98.51	98.32	91.63	1.68	91.14
	Sig	2098	245	18	2393	99.25	89.54	90.71	94.47	9.29	94.79
5	Bi-	2642	274	418	2487	85.61	90.60	90.08	88.11	9.92	87.79
	TF-	2656	260	793	2112	72.70	91.08	89.04	81.91	10.96	80.05
	Sig	2288	628	174	2731	94.01	78.46	81.30	86.22	18.70	87.20



Average results from classifying three different building types are shown in Figure 12. SEN, SPE, PRE, ACC, and F1; b) FDR.

Capacity for extrapolation to other patients. It is impossible to prevent data leakage in intrasubject validation since the same patient's data will appear in both the training set and the test set. Since the data is not fully "unseen," the model's performance is unlikely to improve enough when applied to a new patient to fulfill clinical demands, even if /is technique is used to verify the model's accuracy. Taking into account the

reality of the case, we used the leave-one-out (LOI) technique of cross-validation, which involves partitioning data sets into two groups—a training set and a testing set—that are kept secret from one another. Whether or whether we use cross-validation, our detector outperforms similar methods in every metric [28]. With such high generalization performance, our technique is better suited as a diagnostic aid in clinical settings.

Table 9: /e range of gains and losses in various indicators.

Methods	SEN	SPE	PRE	ACC	FDR	F1
Loss ¹	2.17%	3.33%	—	—	—	—
Improvement ²	—	—	3.22%	5.73%	3.22%	7.47%

¹Loss degree of SEN and SPE using bi-branch model compared with 2 single-branch models. ²Improvement degree of SEN and SPE using 2 single-branch models compared with the bi-branch model.

Validation of Visual and Robotic Desections for Agreement.

he study's gold standard for clinical usage is physicians' visual marking of HFOs. For this reason, it is essential that the suggested detection technique shows excellent consistency with the reference approach. It is possible to compare the reliability of various approaches using Cohen's kappa coefficient. The confusion matrix is used in its calculation, with the precise formula being as follows:

$$\text{kappa} = \frac{p_0 - p_e}{1 - p_e}$$

Where p_0 is the overall accuracy predicted by the model, given $(a_0 b_0 + a_1 b_1)/n$, where a_0 and a_1 are the quantities predicted by the model to be 0/1, and b_0 and b_1 are the actual quantities of the two classes 0/1. Its values might be anything from -1 to 1, however they are often between 0 and 1. Consistency levels range from 0.0 to 0.20, with 0.21 to 0.40 indicating a low level of consistency, 0.41 to 0.60 indicating a medium level, 0.61 to 0.80 indicating a high level of consistency, and 0.81 to 1 indicating an almost perfect level of consistency. We compared the novel

detection approach to the gold standard and computed Cohen's kappa coefficient. The coefficients ranged from 0.784 to 0.765 to 0.708 to 0.957 to 0.762 to 0.762, with an overall average of 0.795. The results of the tests conducted on the approach provided in this study show remarkable congruence with the opinions of medical professionals.

Conclusions

Our results demonstrate that the two-modal data-based bi-branch model may be employed as a reliable automated HFO identification method, with strong generalization performance across a variety of patients. Hence, the suggested approach is excellent for use in the clinic. The following benefits are most common: (1) The proposed bi-branch model, in terms of model architecture, combines the benefits of SEEG signal and time-frequency image for HFOs detection, and it employs two separate backbone networks (1d-ResNet + LSTM and 2dResNet CBAM) to automatically and concurrently extract the features of the two modes. High classification accuracy, sensitivity, and specificity are achieved by combining the results of two branches; (2) the approach has been clinically validated in terms of both intrasubject and cross-subject validation. Our suggested technique achieved an excellent performance with high accuracy, sensitivity, specificity, etc., particularly for study employing the cross-subject validation, which is adequate for practical clinical applications.

References

[1] J. Cimbalnik, B. Brinkmann, V. Kremen et al., "Physiological and pathological high frequency oscillations in focal epilepsy," *Annals of clinical and translational neurology*, vol. 5/9, pp. 1062–1076, 2018.

[2] G. Wang, D. Wang, C. Du et al., "Seizure prediction using directed transfer function and convolution neural network on intracranial eeg," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 12, pp. 2711–2720, 2020.

[3] M. Cotic, Y. Chinvarun, M. del Campo, P. L. Carlen, and B. L. Bardakjian, "Spatial coherence profiles of ictal highfrequency oscillations correspond to those of interictal

- low-frequency oscillations in the ecog of epileptic patients,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 1, pp. 76–85, 2014.
- [4] T. Fedele, S. Burnos, E. Boran et al., “Resection of high frequency oscillations predicts seizure outcome in the individual patient,” *Scientific Reports*, vol. 7, no. 1, pp. 13836–13910, 2017.
- [5] M. Demuru, S. Kalitzin, W. Zweiphenning et al., “The value of intra-operative electrographic biomarkers for tailoring during epilepsy surgery: from group-level to patient-level analysis,” *Scientific Reports*, vol. 10, no. 1, pp. 14654–14718, 2020.
- [6] K. Remakanthakurup Sindhu, R. Staba, and B. A. Lopour, “Trends in the use of automated algorithms for the detection of high-frequency oscillations associated with human epilepsy,” *Epilepsia*, vol. 61, no. 8, pp. 1553–1569, 2020.
- [7] Z. Chen, M. I. Maturana, A. N. Burkitt, M. J. Cook, and D. B. Grayden, “High-frequency oscillations in epilepsy: what have we learned and what needs to be addressed,” *Neurology*, vol. 96, no. 9, pp. 439–448, 2021.
- [8] C. Jiang, X. Li, J. Yan et al., “Determining the quantitative threshold of high-frequency oscillation distribution to delineate the epileptogenic zone by automated detection,” *Frontiers in Neurology*, vol. 9, p. 889, 2018.
- [9] K. Zhang, W. Shi, C. Wang et al., “Reliability of eeg microstate analysis at different electrode densities during propofol-induced transitions of brain states,” *NeuroImage*, vol. 231, Article ID 117861, 2021.
- [10] A. J. Hincapie, A.-S. Hincapie, and B. Frauscher, “Localization of the epileptogenic zone using high frequency oscillations,” *Frontiers in Neurology*, vol. 10, p. 94, 2019.
- [11] L. Li, M. Patel, J. Almajano, J. Engel Jr., and A. Bragin, “Extrahippocampal high-frequency oscillations during epileptogenesis,” *Epilepsia*, vol. 59, no. 4, pp. e51–e55, 2018.
- [12] J. Parvizi and S. Kastner, “Promises and limitations of human intracranial electroencephalography,” *Nature Neuroscience*, vol. 21, no. 4, pp. 474–483, 2018.
- [13] P. Jayakar, J. Gotman, A. S. Harvey et al., “Diagnostic utility of invasive eeg for epilepsy surgery: indications, modalities, and techniques,” *Epilepsia*, vol. 57, no. 11, pp. 1735–1747, 2016.
- [14] M. Fahimi Hnazaee, B. Wittevrongel, E. Khachatryan et al., “Localization of deep brain activity with scalp and subdural eeg,” *NeuroImage*, vol. 223, Article ID 117344, 2020.
- [15] P. Cuisenier, B. Testud, L. Minotti et al., “Relationship between direct cortical stimulation and induced high-frequency activity for language mapping during seeg recording,” *Journal of Neurosurgery*, vol. 134, no. 3, pp. 1251–1261, 2020.
- [16] J. Cimbalnik, A. Hewitt, G. Worrell, and M. Stead, “The cs algorithm: a novel method for high frequency oscillation detection in eeg,” *Journal of Neuroscience Methods*, vol. 293, pp. 6–16, 2018.
- [17] Y. Du, B. Sun, R. Lu, C. Zhang, “A method for detecting high-frequency oscillations using semi-supervised k-means and mean shift clustering,” *Neurocomputing*, vol. 350, pp. 102–107, 2019.
- [18] T. Wan, M. Wu, X. Wan, and Y. Du, “Automatic detection of high frequency oscillations based on fuzzy entropy and fuzzy neural network,” in *Proceedings of the 2016 35th Chinese Control Conference (CCC)*, pp. 5027–5032, IEEE, Chengdu, China, July 2016.
- [19] M. Wu, T. Wan, M. Ding, X. “A new unsupervised detector of high-frequency oscillations in accurate localization of epileptic seizure onset zones,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 12, pp. 2280–2289, 2018.
- [20] A. B. Gardner, G. A. Worrell, E. Marsh, D. Dlugos, and B. Litt, “Human and automated detection of high-frequency oscillations in clinical intracranial eeg recordings,” *Clinical Neurophysiology*, vol. 118, no. 5, pp. 1134–1143, 2007.