Clustering of High Frequency Oscillations HFO in Epilepsy Using Pretrained Neural Networks

Zayneb Sadek, Abir Hadriche and Nawel Jmail

October 26, 2022

Clustering of high frequency oscillations HFO in epilepsy using pretrained neural networks

Zayneb Sadek[(1)], Abir Hadriche[(1,2)] and Nawel Jmail[(1,3)]
Digital Research Center of Sfax[(1)], Regim Lab, ENIS, Sfax University, Tunisia [(2)], Miracl Lab[(3)]
Sfax University, Tunisia
zayneb.sadek123@gmail.com [(1)], Abir.hadriche.tn@ieee.org[(2)], ,naweljmail@yahoo.fr[(3)]

Abstract. High frequency oscillations (HFOs) have been presented as a promising clinical biomarker of regions responsible of epileptic seizure onset zone (soz) and thus a potential aid to guide epilepsy surgery. Visual identification of HFOs in long-term continuous intracranial EEG (iEEG) is cumbersome, due to their low amplitude and short duration. The objective of our study is to improve and automate HFO detection by developing analysis tools based on an unsupervised clustering method. First, we used a temporal basis set from Jmail et al 2017 while exploiting the time-frequency content of iEEG data. Subsequently, we used a CNN (resnet 18) feature extractor. Then, we applied the clustering method based on reducing the events dimension per frame while preserving the distance between points when displaying from high-dimensional space to a low-dimensional one. The clustering method (Deep Cluster) is based on a standard k-means clustering algorithm. This algorithm successfully isolated HFOs from artifacts, peaks and peaks with ripples. Using this algorithm, we were able to locate the seizure onset area.

Keywords: epileptic seizure, intracranial EEG (iEEG), high frequency oscillations (HFOs), CNN, K-means.

1 Introduction

Epilepsy is characterized by increased electrical activity in the brain, resulting in temporary disruption of communication between neurons. In addition, Intracranial EEG (iEEG) is recorded as a traumatic technique where electrodes are placed directly on the brain to visualize cortical sub-zones interactions. iEEG has been widely used by neurologists to detect seizure onset area and accurately identify seizure onset [1, 2, 3,4]. In epilepsy, IEEG signals depict spikes, oscillations at different frequencies, and superimposed spikes and oscillations [5].
Pathological high-frequency oscillations (HFOs) between 80 and 500 Hz have recently been proposed as a potential biomarker of the onset zone of epileptic seizures and have shown superior accuracy to interictal epileptiform discharges. HFOs are

field potentials that reflect short-term synchronization of neuronal activity [6]. Further studies proposed sub-HFOs as ripples (80-250Hz) and fast ripples (FR, 250-500Hz) [7,8].
In fact, HFOs appear as transients with a low amplitude in intracranial EEG. Therefore, visual identification of these events remains such a difficult task. This challenge required a robust and high precision automatic detector.
However, most studies of HFO detections have been performed by simple thresholding, defining context energy by calculating root mean square (RMS) [9], or signal line length [10],[11].
In semi-automatic detectors, visual inspection is performed after initial detection [12],[13], while in fully automatic detectors, supervised classifiers or advanced signal processing steps are required [14],[15].
Therefore, we propose to study an automatic unsupervised clustering detection of HFOs to map the detected events and investigate their spatial diffusion. First, we proposed a database of simulated IEEG data (in the HFO frequency range) where we evaluated them for different constraints (SNR, overlap rate, relative amplitude and frequency range). Second, explored the spectrum of iEEG signal in high frequency band. These results would help us to identify the channel with a maximum of HFOs.
Thus, in the first section, we described our simulated data and the method used for clustering. In the second section, we exposed our obtained results and finally we concluded and discussed our results.

2 Materials and methods

2.1 Materials

All signal processing steps of our study is processed in Matlab software (Mathworks, Natick, MA).

Simulated data: is obtained by a combination of a peak and HFO shapes as real IEEG signal, sampled at 512Hz and of duration 2s, with 1024 samples. Thanks to different tests, we have prepared five classes of signals composed of spikes, HFO (Ripple, Fast Ripple) and superimposed ripple and spike. By varying different parameters: relative amplitudes, frequency of oscillations, signal to noise ratio (SNR) and overlap rate: we obtained 120 sets of simulated data composed of spikes and HFO events in this range [85,105,200,350,450] Hz (ripples and fast ripples).

2.2 Methods

In this paper we propose a solution for automatic detection using unsupervised clustering. The schematic diagram shows our clustering.

Fig. 1 depicted three steps in the clustering pipeline. In the first step, we recovered a simulated IEEG time base where it is presented of mixed of pure HFO combinations, artifact, spike, and spiky event with ripples. These events produce a complex shape in which it is difficult to distinguish the basic elements. In the second step, we presented the frequency plan of the studied events. In the last step, we used clustering algorithm based a CNN for feature extraction as well as a PCA method and t-SNE for dimension reduction. then, we applied a standard K-means clustering algorithm.

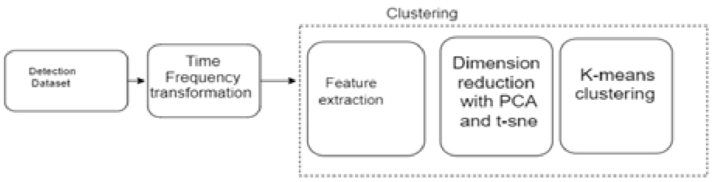

Fig. 1. clustering steps of HFO.

The approaches described above solve the clustering problem of pathological HFOs.
In fact, clustering is a quiet difficult task since original signal are not suitable for algorithms and require a step of feature extraction, i.e. a feature description vector.

CNN Feature extraction

A convolutional neural network [16, 17] is a neural network that uses a mathematical operation called convolution or convolution product (a linear operation). Each convolutional neural network contains at least one convolution layer. Let f and g be two functions defined on R, the convolution product between f and g is denoted f $*$g and it is defined by the following equation:

$$s(x) = (f * g)(x) = \int_{-\infty}^{+\infty} f(t)g(x-t)dt \qquad (1)$$

In machine learning, the input is always a multidimensional array of data, and the kernel is always a multidimensional array of parameters that will be adapted by the learning algorithm. Convolution is always used with a dimension greater than 1. The most used convolution is a 2D convolution. In this case, for an image input I and for a kernel K, the discrete convolution is written:

$$s(i,j) = (k * I)(i,j) = \sum_m \sum_n I(i-m, j-n)k(m,n) \qquad (2)$$

The convolutional neural network (CNN) is widely used in automatic image extraction feature. The feature dimension extracted by this method is very high, and several features have strong correlations. CNN is respectively composed by convolution layer, clustering layer, fully connected layer and classifier.
Features extracted by the convolution layer in CNN structure are called feature map. The design of convolutional neural network architectures vary, but in general terms they can be divided into several functional blocks (a block of convolution layers, a block of pooling layers, and a block of fully linked layers).
The most common output of a neural network in classification tasks is the membership probability of the input image, obtained using a function (softmax), which is a generalization of the logistic function of multidimensional case.

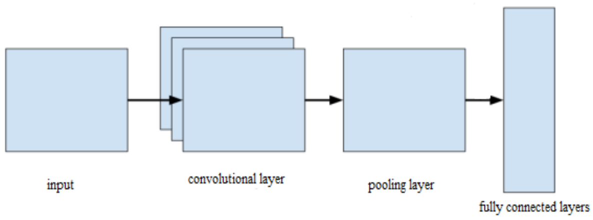

Fig. 2. Schematic of image feature extraction from a CNN.

Dimensionality reduction methods

Dimensionality reduction methods preserve the distance between points when going from a high-dimensional space to a low-dimensional space. Typical representatives are depicted using principal components analysis (PCA) and multidimensional scaling (MDS). The feature space dimensionality reduction problem can be solved using linear methods, such as the principal component analysis method (PCA).
Assessment by t-SNE
The methods used are stochastic neighbor method with t-distribution (t-SNE) and uniform spatial approximations and projections (UMAP) [18]. In general, the application of these methods produces similar results, but UMAP is more computationally efficient than t-SNE.
Methods for finding nearest vectors in high-dimensional spaces by K-means.
The most popular algorithms for solving this kind of problems are the approximate nearest neighbor method. in [19] the set of algorithms based on a modified K-means method and allows calculations on GPU.
Silhouette Scoring

is a metric for evaluating a clustering algorithm, calculated from two scores, a and b.
a is the average distance between a sample and all other points in the same cluster and b is the average distance between a sample and all other points in the nearest cluster. The silhouette score of a sample is calculated by the following formula:

$$s = \frac{b-a}{\max(a,b)} \qquad (3)$$

3 Results

Our result represents the clustering approach using spectral representation of iEEG to isolate ripples and other events such as spikes and contaminated noise.
In Fig. 3 we depicted the simulated data set in the frequency plan. At each line, we find events in the frequency plan. As well as pure HFOs in the form; Ripples in the frequency range [80-250Hz], fast Ripples in [250-450Hz], spike, artifacts (high noise) and a spike event with a ripple.

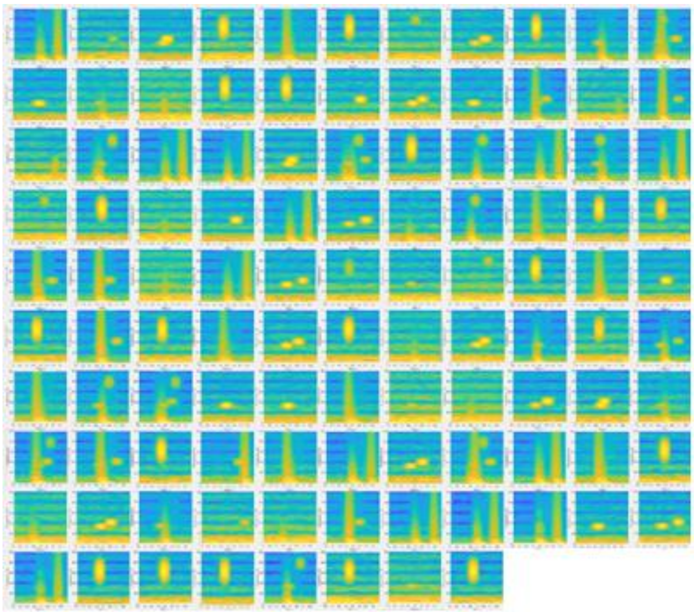

Fig. 3. Time-frequency distribution of simulated IEEG data.

Fig. 3. shows the spatial pattern of HFO distribution. These Color Maps represents a state dispute (Spike, Ripple,Fast ripple and Artifact) from 2D t-SNE visualization after different feature extractors.

According to [20] our results show that extracted features with ResNet18 have a better contribution to the test between VGG16, VGG 19, ResNet50, Resnet101 and Resnet34.
 Therefore, it is expected that the feature representation will be more compact and well separated as we progress through the used pipeline.

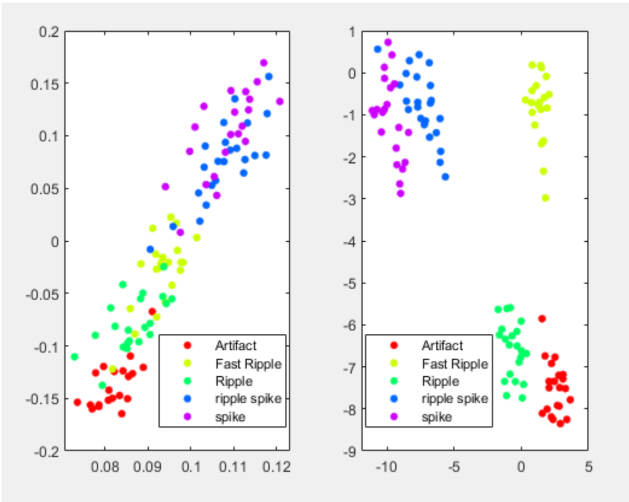

Fig. 4. 2D t-SNE visualization of extracted features using resnet18.

Fig.5. represents Spatial distribution of simulated IEEG data after initial detection and final clustering using k-means of extracted CNN feature (ResNet18).

Indeed, the clustering result presents 5 classes, where the upper line is a frequency plan of artifact. Thus, in the second line we found a class of peaks and a peak event with a ripple. In the third line, there is a well-ranked fast ripple class. Then in the following line, we got a class of spike overlaid, a little bit with spike ripple. The bottom line represents the perfectly ranked ripple frequency plan.

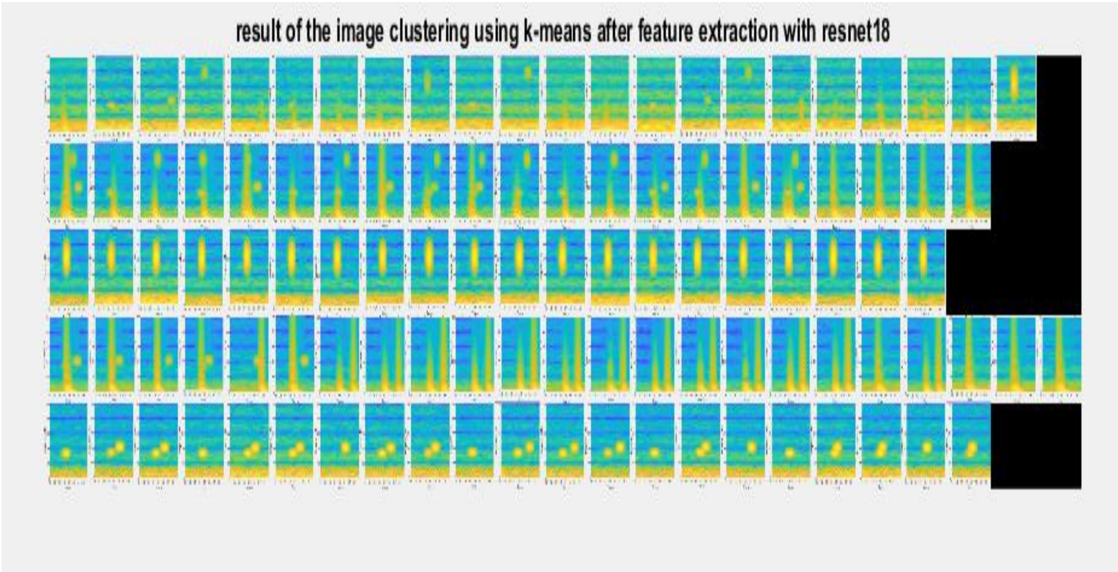

Fig. 5. Spatial distribution of HFOs after final grouping.

in Fig. 6. we depicted the HFOs clustering algorithm silhouette score. Note that the silhouette score can range from -1 to +1. A silhouette score of -1 means incorrect clustering and +1 means correct and very dense clustering. A silhouette score of 0 means the clusters overlap.

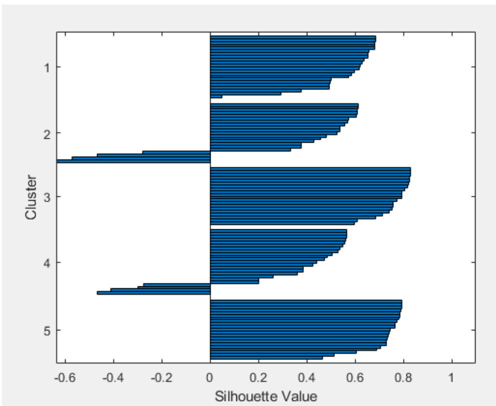

Fig. 6. silhouette score of HFOs clustering algorithm.

4 Conclusions

In this study, we applied the HFO clustering approach to detect IEEG events such as epileptic spikes, ripples, and fast ripples in intracranial EEG datasets (simulated IEEG). The HFO clustering network described follows time frequency representation of simulated iEEG. This network is created through a novel automated approach based on sequential feature extraction through pre-trained neural networks and t-SNE methods for dimension reduction.
First, we applied a CNN feature extraction (resnet18). then we used the t-SNE method for downsizing. The algorithm succeeded in isolating a selection of events from five classes, namely spikes, ripples and fast ripples, contaminated noises as well as spikes with superimposed ripples.
The clustering result should be recommended as a guide to identify the channel with a maximum number of HFOs which subsequently can be identified as a strong indicator at seizure onset. As perspective, we propose to use our results in order to localize HFO and define their networks connectivity for an accurate definition of epileptogenic zones, which would imply a significant impact on surgical intervention for drug-resistant patients in order to delineate epileptogenic tissue.

References

[1] Zentner, Surgical treatment of epilepsies, no. Springer, 2002.
[2] Hadriche, A., ElBehy, I., Hajjej, A., Jmail, N. (2022). Evaluation of Techniques for Predicting a Build Up of a Seizure. In: Abraham, A., Gandhi, N., Hanne, T., Hong, TP., Nogueira Rios, T., Ding, W. (eds) Intelligent Systems Design and Applications. ISDA 2021. Lecture Notes in Networks and Systems, vol 418. Springer, Cham. https://doi.org/10.1007/978-3-030-96308-8_76
[3] Hadriche.A, Behy.I, Necibi.A, Kachouri.A,Ben Amar.C Jmail.N, :Assessment of Effective Network Connectivity among MEG None Contaminated Epileptic Transitory Events.Computational and Mathematical Methods in Medicine(2021). https://doi.org/10.1155/2021/6406362

[4] J. Nawel, H. Abir, B. Ichrak, N. Amal and B. A. Chokri, "A Comparison of Inverse Problem Methods for Source Localization of Epileptic Meg Spikes," *2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)*, 2019, pp. 867-870, doi: 10.1109/BIBE.2019.00161.

[5] Jmail N., Gavaret M., Wendling F, Kachouri A, Hamadi G, Badier JM.:A comparison of methods for separation of transient and oscillatory signals in EEG, J. Neurosci. Methods, pp. 273-289..193, (2011).

[6]Zelmann, F. Mari, J. Jacobs, M. Zijlmans, F. Dubeau and J. Gotman, "A comparison between detectors of high frequency oscillations", Clinical Neurophysiology, 123, no. 1, pp. 106-116, 2012.

[7] J. Engel, A. Bragin, R. Staba and I. Mody, "High-frequency oscillations: what is normal and what is not?", Epilepsia, 50, no. 4,pp. 598-604, 2009.
[8] C. A. Schevon, A. J. Trevelyan, C. E. Schroeder, R. R. Goodman, G. McKhann and R. G. Emerson, "Spatial characterization of interictal high frequency oscillations in epileptic neocortex", Brain, 132, no. Pt 11, pp. 3047-59, 2009.
[9] G. A. Worrell, A. B. Gardner, S. M. Stead, S. Hu, S. Goerss, G. J. Cascino, F. B. Meyer, R. Marsh and B. Litt, "High-frequency oscillations in human temporal lobe: simultaneous microwire
[10] 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, 104, no. 5, pp. 2900-12, 2010.

[11] B. G. Andrew, W. Greg A, M. Eric, D. Dennis and L. Brian, "Human and Automated Detection of High-Frequency Oscillations in Clinical Intracranial EEG Regcordings", Clin Neurophysiol, 118, no. 5, pp. 1134-1143, 2007.
[12] C. A. Schevon, A. J. Trevelyan, C. E. Schroeder, R. R. Goodman, G. McKhann and R. G. Emerson, "Spatial characterization of interictal high frequency oscillations in epileptic neocortex", Brain, 132, no. Pt 11, pp. 3047-59, 2009.
[13] T. Akiyama, H. Otsubo, A. Ochi, T. Ishiguro, G. Kadokura, R. Ramachandrannair, S. K. Weiss, J. T. Rutka and O. Carter Snead, "Focal cortical high-frequency oscillations trigger epileptic spasms: confirmation by digital video subdural EEG", Clin Neurophysiol,
116, no. 12, pp. 2819-25, 2005.
[14] B. Crépon, V. Navarro, D. Hasboun, S. Clemenceau, J. Martinerie, M. Baulac, C. Adam and M. Le Van Quyen, "Mapping interictal oscillations greater than 200 Hz recorded with intracranial macroelectrodes in human epilepsy", Brain, 133, no. Pt 1, pp. 33-45, 2010.
[15] S. Burnos, P. Hilfiker, O. Sürücü, F. Scholkmann, N. Krayenbühl, T. Grunwald and J. Sarnthein, "Human intracranial high frequency oscillations (HFOs) detected by automatic time-frequency analysis",PLoS One, 9, no. 4, pp. e94381, 2014.
[16] I. Goodfellow, Y. Bengio, and A. Courville. Deep Learning. MIT Press, 2016. 43
[17] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11) :2278–2324, Nov. 1998. 35, 43, 46, 49.
[18] McInnes, Leland, John Healy, and James Melville. "Umap: uniform manifold approximation and projection for dimension reduction". arXiv preprint arXiv:1802.03426 (2018).
[19] Chan, David M., et al. "t-SNE-CUDA: GPU-accelerated t-SNE and its applications to modern data". Preprint arXiv arXiv:1807.11824(2018).
[20] Benchmarks for popular CNN models. URL: https://github.com/jcjohnson/cnn-benchmarks.