A NOVEL UNSUPERVISED AUTOENCODER-BASED HFOS DETECTOR IN INTRACRANIAL EEG SIGNALS

Weilai Li[*] *Lanfeng Zhong*[*] *Weixi Xiang* *Tongzhou Kang* *Dakun Lai*[†]

University of Electronic Science and Technology of China (UESTC), Chengdu, 611731, China

ABSTRACT

High frequency oscillations (HFOs) have demonstrated their potency acting as an effective biomarker in epilepsy. However, most of the existing HFOs detectors are based on manual feature extraction and supervised learning, which incur laborious feature selection and time-consuming labeling process. In order to tackle these issues, we propose an automatic unsupervised HFOs detector based on convolutional variational autoencoder (CVAE). First, each selected HFO candidate (via an initial detection method) is converted into a 2-D time-frequency map (TFM) using continuous wavelet transform (CWT). Then, CVAE is trained on the red channel of the TFM (R-TFM) dataset so as to achieve the goal of dimensionality reduction and reconstruction of input feature. The reconstructed R-TFM dataset is later classified by K-means algorithm. Experimental results show that the proposed method outperforms four existing detectors, and achieve 92.85% in accuracy, 93.91% in sensitivity, and 92.14% in specificity.

Index Terms— High frequency oscillation, time-frequency analysis, convolutional variational autoencoder, unsupervised detector

1. INTRODUCTION

While approximately 70% of epileptic patients have controlled their seizure with advanced antiepileptic medications, the remaining patients with drug-resistant epilepsy can get further treatment via surgical resection of the epileptogenic zone (EZ) [1]. Accurate localization of the EZ is deemed as the key to successful resection surgery. It has been discovered that the interictal High Frequency Oscillations (HFOs) in intracranial electroencephalogram are reliable biomarkers of the EZ [2]. Interictal HFOs are also used to determine the spatial extent of the seizure onset zones, which are surrogate markers of the EZ [3]. Increasing evidences have shown that comparing with epileptic spikes (below 40 Hz), the conventional biomarkers of the EZ, HFOs are proved to be more useful biomarkers in pinpointing EZ [4].

HFOs are usually defined as short-duration and low-amplitude events having two distinct constituent parts, namely ripples (Rs, 80-250 Hz) and fast ripples (FRs, 250-500 Hz) [5]. Before the invention of automatic HFOs detector, HFOs are manually detected by visual marking. Although this method is viewed as the "Gold Standard" for HFOs detection by a number of studies, the whole process of manual detection is demanding and time-consuming. It would take roughly 2 hours for a professional expert to mark all HFOs in a 2-minute recording of intracranial electroencephalogram (iEEG) within a single channel. The evident deficiencies of manual detection have spurred the development of automatic HFOs detector in the past decade [6].

The pioneering automatic HFOs detectors are based on threshold, such as the root mean square detector [7] and the line-length detector [8]. These methods are widely adopted as initial detection method for putative HFOs events (pHFOs) in subsequent studies. A majority of existing detectors are combinations of manual feature extraction and supervised learning techniques. For example, Jrad et al. (2017) [9] has proposed to use Gabor Transform to construct energy-ratio feature vectors of detected HFOs events, which are classified by multiclass support vector machine (SVM). However, these detectors have several drawbacks in HFOs detection. To be specific, the threshold detector only utilizes single feature and needs to set threshold beforehand, therefore may suffer from a low detection accuracy. On the other hand, to devise an efficacious detector based on manual feature extraction, researchers have to select among diversified features ranging from time-domain to frequency-domain and rank them by statistical test [10]. Labels of pHFOs events also need to be used if the detector uses supervised learning algorithm.

Autoencoder is a neural network architecture capable of automatic high-level feature extraction and dimensionality reduction [11]. Convolutional Variational Autoencoder (CVAE), a variant of autoencoder, can efficiently resolve the issues of data variability and noise in these data. To fully exploit the merits of this deep learning model and avoid manual annotation for training, we introduce an automatic unsupervised CVAE-based HFOs detector. This detector obviates the needs of crafting or selecting features manually and achieves dimensionality reduction in the detection of HFOs. Comparative experiment indicates that the use of CVAE can enhance the performance of our proposed detector.

The remainder of this paper is organized as follow. Section 2 describes the acquisition and preprocessing of pHFOs, Section 3 elaborates on the methodology of our proposed detector and Section 4 presents the experimental results. At last, Section 5 concludes the paper.

2. DATA ACQUISITON AND PREPROCESSING

2.1 Patients and Data Acquisition

The iEEG recordings from each adult patient were selected from the department of neurosurgery of the West China Hospital (Chengdu, China). Specifically, all raw iEEG data contained in this experiment met the same characters as follow: a continuous long-term duration

[*]These authors are equally contributed.

[†] This work was supported by the Natural Science Foundation of China (61771100). (Corresponding author: Dakun Lai)

greater than 2 hours, sample rate of 2560 Hz, data recorded with the same equipment and no hardware filter used in the data acquisition process. Filtered by the criteria, 5 patients were chosen, 14 hours' recording in total was acquired. Note that not all electrode channels can be used due to inappropriate movement of patients. All data was provided by the Institutional Review Board (IRB) of the West China Hospital, and written informed consents were obtained from all patients.

2.2 Initial HFOs Detection

The raw iEEG signals are preprocessed by multi-notch filter at 50 Hz and bandpass filtering in the frequency band of 80-500 Hz. Subsequently, using a 10ms-sliding window, the short-time energy (STE) of each frame is computed based on the filtered iEEG data [8]. The STE is defined as:

$$E^*(t) = \frac{1}{N}\sum_{k=t-N+1}^{t} x^2(k) \quad (1)$$

where $E^*(t)$ is the short time energy of the frame, which starts at the time of t. $x(k)$ is the k-th point of a pHFO, and N is the number of sample points contained in a frame. The STE estimation defines a threshold of more than 3 times of standard deviation (SD) above the average of STE signal [8]. A 150-ms raw iEEG signal segment is identified as pHFO when the STE values of three successive frames all exceed the defined threshold. The center of the pHFO is located at the middle of these three adjacent frames. The STE estimation threshold E_0 is given by the formula:

$$E_0 = E_{av} + k * SD \quad (2)$$

where k is the weight of SD and E_{av} is the mean of the STE. The parameter k need to be tuned in order to reach a balance between higher sensitivity of STE estimator and fewer number of false detected HFO events. The value of k in our experiment is equal to 5, and 4042 segments in total are acquired by applying STE estimation.

2.3 Manually Labeling by Neuroelectrophysiologists

Following the golden standard of detecting HFOs [12], all 4042 TFMs were labelled by two neuroelectrophysiologists, who are with the Department of Neurosurgery, West China Hospital of Sichuan University, China. This procedure was completed in the Matlab via a graphical user interface designed by our research group. Note that the labels are used for assessing performance of our proposed method and training other existing supervised detectors in our comparative experiments.

3. METHODLOGY

3.1. Time-frequency Analysis

Feature extraction and continuous wavelet transform (CWT) are the leading time-frequency analysis methods employed in the previous studies of HFOs [13, 14]. While feature extraction in time domain and frequency domain shrinks 1-D iEEG signals to feature scalars, CWT delineates 1-D iEEG signals on 2-D TFMs. With the purpose of preserving intact information contained in pHFOs events, we

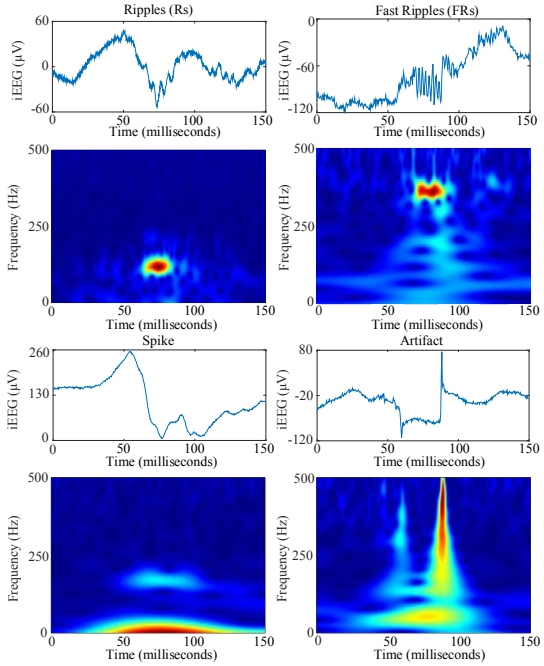

Fig. 1. Examples of TFM from each group representing R, FR, Spike, and Artifact. The first and third rows are the original iEEG signals. The second and fourth rows are the corresponding TFMs.

adopt Analytic Morse CWT function to generate TFM [14]. In contrast to other commonly used time frequency representations such as Morlet wavelet and Derivative-of-Gaussian wavelet, analytic Morse wavelet is more appropriate in depicting specific time and frequency components [15]. As shown in Fig. 1, TFM is a computed magnitude scalogram, which highlights the dominant frequency components in time-scale by red color. By thorough observation, we discover that the pHFOs can be categorized into four groups, which are R, FR, spike, and artifact, respectively. Fig. 1 presents an example of TFM from each group.

Visual identification of HFOs using TFMs not only depends on the morphology discrepancies between pHFOs, but also make use of the principal frequency component, which is tinted by red pixels. It is straightforward to distinguish Rs and FRs from false detected HFOs, as the TFMs of HFOs exhibit red isolated islands. Therefore, we extract the red channel of TFM and regard them as the replacements of TFMs. A TFM is decomposed into three grey-scale maps, which represent the red, green, and blue channel of TFM respectively in Fig. 2. This intuitively-simple but rather effectual process attains several purposes, including the reduction of dimensionality and boosting detection performance.

3.2. CVAE-based HFOs Detector

CVAE is a modified case of traditional autoencoder. It uses convolutional and pooling layers instead of the original fully-connected layers, which is more efficient for images [16, 17]. The overview of proposed CVAE detector is illustrated in Fig. 3. We train the CVAE for eliminating noise in R-TFM and extracting high-level features. Assume $X = (x_{ij})\ (i < m, j < n)$ denote the R-TFM, $Z = [z_1, z_2, \ldots, z_l]$ denote latent representation of X encoded by encoder. The image size of R-TFM is 656×875, and the value of l will be discussed in Section4.2.

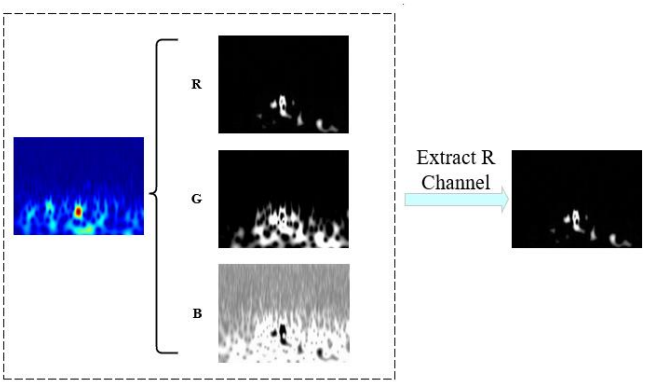

Fig. 2. Diagram of extracting red channel from TFM. The red part of the TFM shows the dominant frequency component in pHFOs and thus facilitates visual identification of HFOs.

The proposed method uses one single encoder to acquire the overall posterior inference of the R-TFM. Two convolutional layers are included in this encoder, whose kernel size are 32×32 and 16×16 respectively. We introduce $q_\Phi(Z|X)$ to represent the encoder, where Φ denotes the parameters of the encoder. N stands for the Gaussian distribution. To resolve the problem caused by sampling operation in the training of CVAE, we employ a reparameterization trick proposed by Kingma et al. [18] and Rezende et al. [17]. This procedure is expressed as:

$$\begin{aligned}(\mu, \log(\sigma)) &= Encoder(X)\\ q_\Phi(Z \mid X) &= N(Z; \mu, diag(\sigma))\end{aligned} \tag{3}$$

The decoder simply mirrors the encoder, it uses two convolutional transpose layers, which define the conditional distribution of X as $p_\theta(X|Z)$, here θ denotes the parameters of the decoder. The latent distribution prior $p(z)$ is modeled as a unit Gaussian. This network takes Z as input and outputs a set of parameters for $p_\theta(X|Z)$, then its output can be applied to reconstruct denoised X. Variational Autoencoders are usually trained by maximizing the evidence lower bound (ELBO). In order to improve the potential tightness of the ELBO, the CVAE is trained by minimizing the ELBO loss, which is defined as [18]:

$$L^{CVAE}(\theta, \phi) = E_{q(x)}\{E_{q_\phi(z|x)}[\log(\frac{p_\theta(x,z)}{q_\phi(z \mid x)}]\} \tag{4}$$

The classification of R-TFM is accomplished by K-means algorithm in view of its relatively simple mechanism, easy implementation, and fast convergence. Specifically, the reconstructed R-TFM are unfolded into arrays by vectorization of pixel matrices in a column-wise manner [19]. K-means algorithm treats vectorized version of reconstructed R-TFM (in size of 64×64) as input and categorizes reconstructed R-TFM into four group (K = 4). In order to assign these four groups of reconstructed R-TFM to specific pHFOs categories (R, FR, spike, and artifact), we adopt a HFO distinguishing feature termed spectral centroid (SC) that was capitalized in HFOs classification [20]. SC represents the frequency corresponding to the mass center of the spectrum for input data, which is pHFOs in our work. For each pHFOs, SC is defined as:

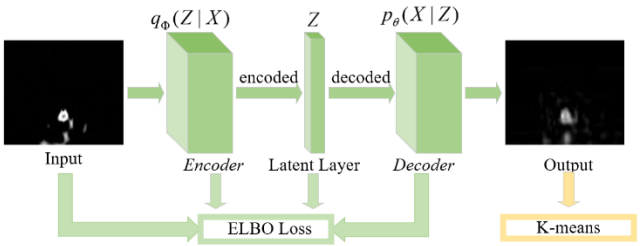

Fig. 3. Visual architecture of the proposed CVAE detector. When the training procedure of CVAE is completed, the output of CVAE is clustered by K-means algorithm.

$$SC = \frac{\sum_{k=0}^{N/2} \frac{k}{NT} |M[k]|^2}{\sum_{k=0}^{N/2} |M[k]|^2} \tag{5}$$

where T is the sampling period, and N is the number of sample points in pHFOs, respectively. $M[k]$ is a multitaper power spectral density estimate, which is expressed as:

$$M[k] = \sum_{n=0}^{N-1} w[n]x[n]e^{-j(2\pi/N)nk}, k = 0,1,2...,N \tag{6}$$

where $w[n]$ and $x[n]$ is the Hamming window and pHFOs, respectively. For each cluster of reconstructed R-TFM, the SCs of which is calculated with their corresponding pHFOs. The statistical results of SC for each group and the analysis are provided in Section4.1.

3.3 Evaluation of the Performance Metrics

As mentioned in Section2.3, true labels of the TFM dataset are used to evaluate the performance of proposed CVAE detector. True positive (TP), true negative (TN), false positive (FP), false negative (FN) are evaluated by comparing predicted labels and true labels. TP stands for the number of pHFOs that are correctly classified as HFOs (Rs and FRs), and TN is the number of pHFOs that are correctly classified as false detected HFOs (spikes and artifacts). For evaluating performance of binary classification, accuracy, sensitivity, and specificity in each repetition of the experiment are calculated by:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{8}$$

$$Specificity = \frac{TN}{TN + FP} \tag{9}$$

4. EXPERIMENTS

4.1 Analysis of K-means clustering results

Table 1 shows statistical results of SC for each group clustered by K-means algorithm. Cluster 4 has the lowest SC, while the Cluster

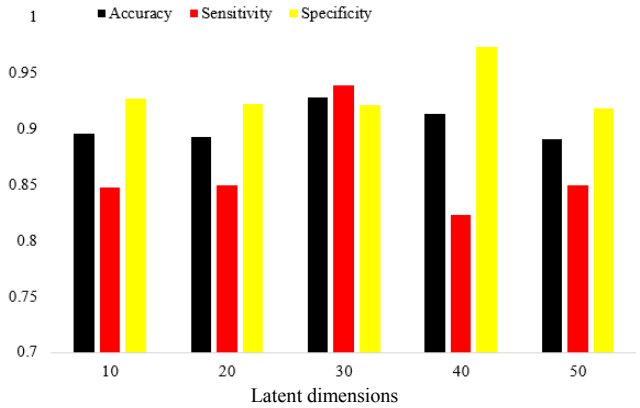

Fig. 4. Performance (accuracy, sensitivity, and specificity) of CVAE Detector with diverse latent dimensions.

Table 1. Statistical results of SC (Hz)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Median	227.7	263.8	51.6	29.1
Mean	226.3	266.4	124.8	29.7

2 has the highest SC. Cluster 1 and 3 are the intermediate of Cluster 2 and 4. According to the fact that the frequency band of R is 80-250 Hz and FR is 250-500 Hz, the clustering results indicates that Clusters 1 and 2 are comprised of Rs and FRs, respectively, Cluster 3 is consist of artifacts, and Cluster 4 is identified as spikes.

4.2 Optimization of the architecture of CVAE

We split the dataset into training, validation and test sets, with the proportion of 7:2:1. In addition, 5-fold cross validation was conducted to make sure the distribution equilibrium of our dataset.

CVAE is trained on the R-TFM dataset. To ensure the convergence of the network, we discover that after training for 200 epochs, the validation loss of CVAE stops decreasing. 200-epoch is applied as the maximal training epoch in the following experiments.

The high-level features extracted by the CVAE are relevant to the number of latent dimensions. To achieve maximum utility of CVAE, we select [10,20,30,40,50] as the candidate values for latent dimension. Fig. 4 is a histogram that shows the performance of various hidden nodes. The best result (92.96%) that averages three metrics comes from the architecture of 30 hidden nodes. With respect to specificity, the architecture of 40 hidden nodes represents the best performance (97.36%).

As shown in Table 2, the performance of detector that uses R-TFM for clustering has a great increase in all evaluation metrics compared with using TFM for clustering (three channels' pixel matrices of TFM are vectorized and cascaded into a vector). The sensitivity of clustering on the reconstructed R-TFM is nearly 11% higher than clustering on the R-TFM. By using CVAE, our detector leads to better unsupervised HFOs detection results. To further confirm the validity of extracting red channel, we also conduct experiments using reconstructed green and blue channel of TFM for clustering, and the performances of which are inferior to reconstructed R-TFM (results are not demonstrated due to space limitation).

4.3 Comparison to existing detectors

We compare our approach with other four existing supervised and unsupervised detectors. Table 2 shows the accuracy, sensitivity, and specificity performance comparisons on the same dataset. All the comparison experiments are implemented by our research group as there are no available open source codes.

Table 2. The differences of procedure and performance comparison of clustering results of TFM, R-TFM, and reconstructed R-TFM

	Procedure		
Dataset	CWT	Extract red channel	CVAE
TFM	√		
R-TFM	√	√	
Reconstructed R-TFM	√	√	√
	Clustering results of three datasets		
	ACC	SEN	SPE
TFM	80.13%	76.28%	82.68%
R-TFM	89.38%	82.99%	93.62%
Reconstructed R-TFM	92.85%	93.91%	92.14%

Table 3. Comparison of the proposed method with existing methods on the same dataset

Year	Methods	ACC	SEN	SPE
2012	RBF Neural Network [21]	79.64%	68.40%	87.08%
2017	RFB-SVM [9]	85.06%	83.80%	85.89%
2018	FCM Based EM-GMM [22]	82.93%	74.61%	88.44%
2019	2D-CNN [23]	91.26%	78.65%	99.62%
	Proposed CVAE	**92.85%**	**93.91%**	**92.14%**

Table 3 demonstrates that CVAE outperforms RBF-SVM [9] method and RBF Neural Network [21], indicating that unsupervised method is more efficient than some supervised machine learning models. Although 2D-CNN [23] achieves higher specificity, it is weaker in sensitivity compared with proposed CAVE. Unlike these methods, our approach does not require labels to train the network and is the most effective way to achieve high performance of automatic HFOs detection among the detectors mentioned in Table 3.

5. CONCLUSION

In this paper, we develop an automatic unsupervised CVAE-based HFOs detector. The deep learning model CVAE and the utilization of R-TFM are introduced in HFOs detection for the first time. Extracting the red channel of TFM removes redundant information and captures salient feature in TFM. In addition, we employ CVAE to perform automatic high-level feature extraction and reconstruction of input feature with R-TFM dataset. Extensive comparative experiments verify the effectiveness of integrating R-TFM and CVAE. A comparison with four existing HFOs detectors validates the superiority of our proposed detectors with optimal CVAE architecture.

6. REFERENCES

[1] Y. Song *et al.*, "Dysfunction of Neurovascular/Metabolic Coupling in Chronic Focal Epilepsy," *IEEE Trans Biomed Eng,* vol. 63, no. 1, pp. 97-110, Jan. 2016.

[2] M. Zijlmans, P. Jiruska, R. Zelmann, F. S. Leijten, J. G. Jefferys, and J. Gotman, "High-frequency oscillations as a new biomarker in epilepsy," *Annals of neurology,* vol. 71, no. 2, pp. 169-178, Feb. 2012.

[3] P. Modur and S. Miocinovic, "Interictal high-frequency oscillations (HFOs) as predictors of high frequency and conventional seizure onset zones," *Epileptic Disorders,* vol. 17, no. 4, pp. 413-424, Dec. 2015.

[4] J. Jacobs *et al.*, "High-frequency electroencephalographic oscillations correlate with outcome of epilepsy surgery," *Ann. Neurol,* vol. 67, no. 2, pp. 209-220, Feb. 2010.

[5] G. A. Worrell *et al.*, "High-frequency oscillations in human temporal lobe: simultaneous microwire and clinical macroelectrode recordings," *Brain,* vol. 131, no. Pt 4, pp. 928-37, Apr. 2008.

[6] M. Navarrete, J. Pyrzowski, J. Corlier, M. Valderrama, and M. Le Van Quyen, "Automated detection of high-frequency oscillations in electrophysiological signals: Methodological advances," *Journal of Physiology-Paris,* vol. 110, no. 4, pp. 316-326, Nov. 2016.

[7] 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.

[8] 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," *Clin Neurophysiol,* vol. 118, no. 5, pp. 1134-43, May. 2007.

[9] N. Jrad *et al.*, "Automatic Detection and Classification of High-Frequency Oscillations in Depth-EEG Signals," *IEEE Trans Biomed Eng,* vol. 64, no. 9, pp. 2230-2240, Sep. 2017.

[10] D. Lai *et al.*, "Channel-Wise Characterization of High Frequency Oscillations for Automated Identification of the Seizure Onset Zone," *IEEE Access,* vol. 8, pp. 45531-45543, Mar. 2020.

[11] G. E. Hinton and R. S. Zemel, "Autoencoders, minimum description length, and Helmholtz free energy," in *Proc. Conf. Adv. Neural Inform. Process. Syst,* vol. 6, pp. 3-10, Nov. 1994.

[12] M. Zijlmnas *et al.*, "High-frequency oscillations as a new biomarker in epilepsy," *Annals of Neurology,* vol. 71, no. 2, pp. 169-178, Feb. 2012.

[13] Y. Wang *et al.*, "Expert consensus on clinical applications of high-frequency oscillations in epilepsy," *Acta Epileptologica,* vol. 2, no. 1, pp. 1-10, Dec. 2020.

[14] M. Amiri, J. M. Lina, F. Pizzo, 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.

[15] G. A. Worrell, K. Jerbi, K. Kobayashi, J. M. Lina, R. Zelmann, and V. Q. Le, M., "Recording and analysis techniques for high-frequency oscillations," *Progress in Neurobiology,* vol. 98, no. 3, pp. 265-278, Sep. 2012.

[16] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," in *Proc. Int. Conf. Learn. Represent.*, 2014.

[17] D. J. Rezende, S. Mohamed, and D. Wierstra, "Stochastic Back-propagation and Variational Inference in Deep Latent Gaussian Models," *arXiv preprint arXiv:1401.4082, 2014*

[18] D. P. Kingma and M. Welling, "An introduction to variational autoencoders," *arXiv preprint arXiv:1906.02691,* 2019.

[19] V. K. Dehariya, S. K. Shrivastava, and R. Jain, "Clustering of image data set using k-means and fuzzy k-means algorithms," in *International Conference on Computational Intelligence and Communication Networks (CICN)*, Nov. 2010, pp. 386–391.

[20] J. A. Blanco *et al.*, "Unsupervised classification of high-frequency oscillations in human neocortical epilepsy and control patients," *Journal of Neurophysiology,* vol. 104, no. 5, pp. 2900-2912, Nov. 2010.

[21] M. Dümpelmann, J. Jacobs, K. Kerber, and A. Schulze-Bonhage, "Automatic 80–250 Hz "ripple" high frequency oscillation detection in invasive subdural grid and strip recordings in epilepsy by a radial basis function neural network, " *Clinical Neurophysiology,* vol. 123, no. 9, pp. 1721-1731, Sep. 2012.

[22] W. Min, T. Wan, D. Min, X. Wan, Y. Du, and J. She, "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. PP, no. 12, pp. 1-1, Oct. 2018.

[23] K. Ma *et al.*, "Automatic detection of High Frequency Oscillations (80-500Hz) based on Convolutional Neural Network in Human Intracerebral Electroencephalogram," *Annu Int Conf IEEE Eng Med Biol Soc,* vol. 2019, pp. 5133-5136, Jul. 2019.