



**HAL**  
open science

## Multi-classification of high-frequency oscillations in intracranial EEG signals based on CNN and data augmentation

Fatma Krikid, Ahmad Karfoul, Sahbi Chaibi, Amar Kachenoura, Anca Nica, Abdennaceur Kachouri, Régine Le Bouquin Jeannès

► **To cite this version:**

Fatma Krikid, Ahmad Karfoul, Sahbi Chaibi, Amar Kachenoura, Anca Nica, et al.. Multi-classification of high-frequency oscillations in intracranial EEG signals based on CNN and data augmentation. Signal, Image and Video Processing, 2023, Signal Image and Video Processing, 10.1007/s11760-023-02808-4 . hal-04329201

**HAL Id: hal-04329201**

**<https://hal.science/hal-04329201v1>**

Submitted on 29 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Multi-classification of High Frequency Oscillations in intracranial EEG signals based on CNN and data augmentation

Fatma Krikid<sup>1,2\*</sup>, Ahmad Karfoul<sup>2</sup>, Sahbi Chaibi<sup>1</sup>, Amar Kachenoura<sup>2</sup>, Anca Nica<sup>3</sup>, Abdennaceur Kachouri<sup>1</sup>, Régine Le Bouquin Jeannès<sup>2</sup>

<sup>1</sup> AFDE-ENIS, University of Sfax, Street of Soukra, 3038 Sfax, Tunisia

<sup>2</sup> Univ Rennes, Inserm, LTSI - UMR 1099, F-35000 Rennes, France

<sup>3</sup> CHU de Rennes, Service de Neurologie, Pôle des Neurosciences, Cliniques de Rennes, F-35000 Rennes, France

\* [fatmakrikid1@gmail.com](mailto:fatmakrikid1@gmail.com)

**Abstract**— Interictal High Frequency Oscillations (HFOs) recorded in intracranial ElectroEncephaloGraphic (iEEG) signals are reliable biomarkers for the epileptogenic zone.

Visual identification of these particular events is manually time-consuming, and is subject to clinicians' expertise. Moreover, differentiating them from other transient events such as Interictal Epileptic Spikes (IESs) presents a considerable challenge. Hence, various approaches have been developed with the aim of extracting automatically discriminant features for HFOs and IESs events. Typically, these approaches are based on machine learning (ML) algorithms, but their efficiency strongly depends on the computed features. To address this limitation, we explore deep learning (DL), as a powerful framework, for the classification of HFOs and IESs and propose a novel convolutional neural network (CNN) architecture for HFOs multi-classification. Time frequency (TF) based images, computed using the Stockwell transform of the events of interest, are used as inputs to the CNN-based approach. Furthermore, data augmentation (DA) is adopted to improve the generalization of the proposed CNN model. The numerical simulations on epileptic iEEG signals demonstrate that the proposed approach yields superior results when the DA is employed. Also, our study.

**Keywords** — Epilepsy, High Frequency Oscillations, Time-Frequency representation, Convolutional neural network, Data augmentation.

## I. INTRODUCTION

For patients suffering from drug-resistant epilepsy, surgical resection of the epileptogenic regions is a solution to limit/avoid the occurrence of epileptic seizures [1]. Various low cerebral activities [2] oscillating in the human brain such as  $\delta$  [0.1,4] Hz,  $\theta$  [4,8] Hz,  $\alpha$  [8,12] Hz and  $\beta$  [12,30] Hz. In the context of epilepsy, High Frequency Oscillations (HFOs), in the range [30,500] Hz, are potential biomarkers for estimating the epileptogenic zone (EZ) and achieving a good prognosis of such pathology [3]. In human brain, HFOs may be observed in intracranial ElectroEncephaloGraphic (iEEG) signals as non-stationary events with at least four oscillations that clearly stand out from the background activity. These particular cerebral activities are divided into four

frequency bands [4]: Gamma ( $\gamma$ [30,80] Hz), High-Gamma (H $\gamma$ [80,120] Hz), Ripples (Rs [120,250] Hz) and Fast Ripples (FRs [250,500] Hz). Typical examples of such activities are illustrated in Fig. 1.

A recent investigation [5] has demonstrated that HFOs and Interictal Epileptic Spikes (IESs) serve as complementary biomarkers for identifying the EZ. Commonly, identifying these biomarkers relies on visual inspection. But this process remains subjective and highly time-consuming. Therefore, to accurately define the EZ during presurgical diagnosis, it is crucial to explore an automated approach that can extract relevant distinctive features and efficiently classify the abovementioned biomarkers. Commonly, HFOs detection procedure is divided into two stages: detection and classification. In the present work, we only focus on the classification of HFOs and IESs events. To address this point, numerous approaches have been proposed to ensure an automatic classification of HFOs and to provide more objective analysis for clinicians. Those approaches were based essentially on supervised machine learning (ML) algorithms such as support vector machine (SVM) [4, 6],  $k$ -nearest neighbors (KNN) method, linear discriminant analysis (LDA), and logistic regression [7], or unsupervised techniques such as  $k$ -medoids [8], or Gaussian mixture model (GMM) [8]. The performance of the latter methods is subject to the discriminative power of the extracted features. Conventionally, these features are extracted either from the time domain [6], the frequency domain [8, 9] or the time-frequency (TF) domain [4, 9]. Generally, ML approaches applied to the problem of HFOs classification have provided prominent results. However, the design of such approaches is still based on manually extracted features of HFOs events which is a subjective process.

Deep learning (DL) classification has gained popularity in recent years due to the capability of neural networks to automatically extract relevant features from raw data. Additionally, DL algorithms have proved their efficacy in diverse domains such as image recognition and

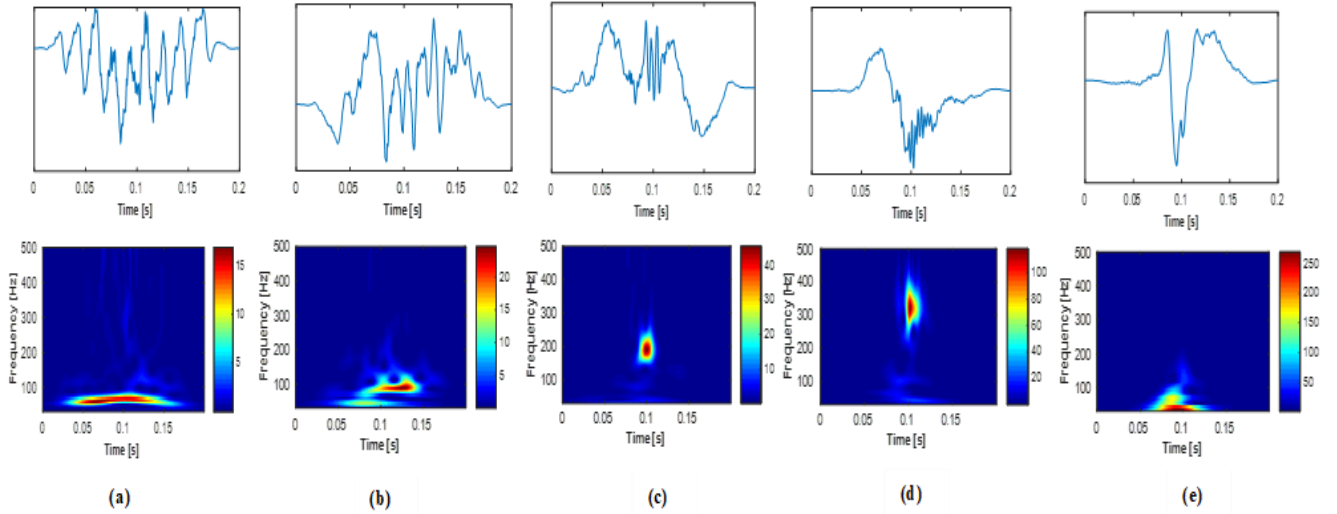


Fig. 1. HFOs events and IESs in both time domain (up) and TF domain (bottom) (a) Gamma, (b) High-Gamma, (c) Ripples, (d) Fast Ripples, (e) Interictal Epileptic Spikes.

classification [10], speech recognition [11] and neurological disorders diagnosis [12-16]. However, few studies have addressed the classification of HFOs using DL. More precisely, Zuo *et al.* [17] addressed binary classification of HFOs (Rs vs FRs) using DL approach. In fact, authors in [17] used a convolutional neural network (CNN) where a one-dimensional (1D) time signal was used at the network input. As CNNs were originally tailored for image related application [18], a 2D representation of HFOs events for CNNs was proposed in several studies [18-20], to cite a few. Lai *et al.* [18] introduced a 2D CNN-based approach for automatic classification of HFOs into Rs and FRs. Additionally, Zhao *et al.* [19] investigated the use of scalogram images with CNN for a binary classification of HFOs. More recently, Nadalin *et al.* [20] have adopted spectrogram images for HFOs binary classification, *i.e.*, IESs occurring with and without Rs. It should be noted that these DL based-approaches were limited to a binary classification (HFOs vs non HFOs). Previous methods for HFOs classification are limited in several ways. First, most of them have only considered a limited frequency range, typically [80,500] Hz. However, recent research has shown that HFOs can occur at lower frequencies beyond 30 Hz, indicating the need for a wider frequency band to be considered. Additionally, previous methods have focused on binary classification (*i.e.*, distinguishing HFOs from non-HFOs), while in practice there are different types of HFOs that can provide additional diagnostic information. To address these limitations, a DL based multi-classification instead of a binary classification of HFOs is considered in the current study. Moreover, data augmentation (DA) is introduced as a technique to enrich the training dataset and improve the robustness of the trained model. The effect of using DA

on the performance of the proposed CNN-based model is also assessed in this study.

Besides, none of the previous methods considered the whole HFOs band [30,500] Hz. To address these limitations, a DL based multi-classification instead of a binary classification of HFOs is considered in the current study. Additionally, we integrate here the use of data augmentation as a means to enrich the training dataset. The effect of using such technique on the model performance is also assessed in this study.

The remainder of this paper is structured as follows: Section II provides information about the clinical dataset. Our proposed methodology is explained in Section III. Results are discussed in Section IV. Finally, concluding remarks are presented in Section V.

## II. CLINICAL DATA

A dataset of real iEEG signals was used to evaluate the performance of the proposed approach. This dataset was recorded in the Neurology Department of the University Hospital of Rennes, France, from five patients who suffered from drug-resistant epilepsy. Our study aligns with the valuable insights contributed by earlier research in the field, as illustrated by the works of [21] and [22]. iEEG signals were recorded using 256-channel Brain Quick (Micromed, Italy) recording system with a sampling rate of 2048 Hz. Intracerebral electrodes were composed of 8 to 18 cylindrical contacts (DIXI Microdeep electrodes, length: 2 mm, diameter: 0.8 mm, 1.5 mm apart, platine-iridium). All patients gave informed consent for participation in research studies. Visual scoring of HFOs events was done by two reviewers independently. From the recorded raw iEEG signals, a

total number of 5174 marked events were labeled visually and employed in the present study. More precisely,  $\gamma$ ,  $H\gamma$ , Rs, FRs and IESs were considered in the marking procedure and annotated as  $\gamma$  (1015),  $H\gamma$  (1032), Rs (1053), FRs (1071), and IESs (1003). The distribution of events in the database is uneven among patients, as demonstrated in Table 1. To tackle this imbalance, we combined all the events across patients to assess the effectiveness of the proposed method. For more information regarding the clinical and acquisition details of the database, please refer to [4].

TABLE 1: PATIENTS CHARACTERISTICS

Patient	Event	Event number
PAQ	Rs	710
	FRs	418
	IESs	380
MED	$H\gamma$	1032
NAP	FRs	653
	IESs	623
SAN	Rs	343
LEM	$\gamma$	1015

### III. METHODOLOGY

In this study, a CNN-based approach is proposed for HFOs multi-classification. A flow chart of the proposed approach is illustrated in Fig. 2. As clearly described in this figure, two main parts are characterizing the proposed approach: data preprocessing and CNN-based multi-classification.

#### A. Data preprocessing

The first step in preprocessing the data involves taking each labeled event of interest (EOI) in our dataset and reframing it into a 200ms window. This time period (*i.e.*, 200ms) is the average of all the labeled events in our database, and it ensures that each event is adequately covered without losing any information. After that, to prevent edge effect in the time-frequency representation of each EOI, they were all multiplied by a Hanning window of 200ms length. The Stockwell transform (S-transform) was then used to generate a time-frequency representation of each windowed EOI. The goal of this representation was to capture signal variations in both time and frequency domains, simultaneously.

The Stockwell transform, commonly referred to as S-transform, is a time-frequency representation proposed by Stockwell *et al.* [23] as a combination of the continuous

wavelet transform (CWT) and short-time Fourier transform (STFT). The S-transform is comparable to the STFT except that the width of the windows varies with frequency. This makes the S-transform a good time-frequency representation leading to a high frequency resolution. In addition, in contrast to the CWT, the S-transform has absolute referenced phase information. Therefore, in this study, the S-transform is applied to both HFOs and IESs signals. Furthermore, previous research has shown that the S-transform can differentiate HFOs from similar activity [24-26].

Let  $x[n]$ ,  $n = 0, 1, 2, \dots, N - 1$ , denote discrete time series corresponding to a given continuous time signal, where  $N$  is the number of time samples. The S-transform  $S[k, m]$  of  $x[n]$  computed at a given time shift factor  $k \in \{0, \dots, M - 1\}$ , where  $M$  stands for the number of frequency bins, is defined as [24]:

$$S[k, m] = \sum_{n=0}^{N-1} x[n] \omega[k - n, m] e^{-i2\pi \frac{m}{M} k} \quad (1)$$

where  $\omega[n, m]$  refers to a specific mother wavelet function given by:

$$\omega[n, m] = \frac{|m|}{M\sqrt{2\pi}} e^{-\frac{n^2 m^2}{2M^2}} \quad (2)$$

Since CNNs are particularly well-suited for image processing, the time-frequency representation of each EOI in our dataset was transformed into an image of size  $410 \times 500$ . This process is further explained in the following subsections.

#### B. Proposed CNN-based model for HFOs multi-classification

The CNN has been found to be the best DL model to deal with image classification tasks [27-29] thanks to its ability to learn automatically the features from raw data. In this study, we propose a deep CNN architecture for HFOs and IESs multi-classification.

Furthermore, data augmentation (DA) being a powerful tool for enhancing the classification performance of CNNs, especially when the size of the dataset is relatively small [30], this technique is also investigated. Consequently, two HFOs multi-classification pipelines, HFOs multi-classification with and without DA, are proposed and illustrated in Fig. 2. Classification results of these two pipelines are then evaluated and compared. A detailed description of the proposed pipelines is given in the following subsections.

##### 1) Pipeline 1: CNN without DA

The architecture of the proposed CNN, as shown in Fig. 3(a), comprises an input layer corresponding to the 3-channel (RGB) input images which are resized into the

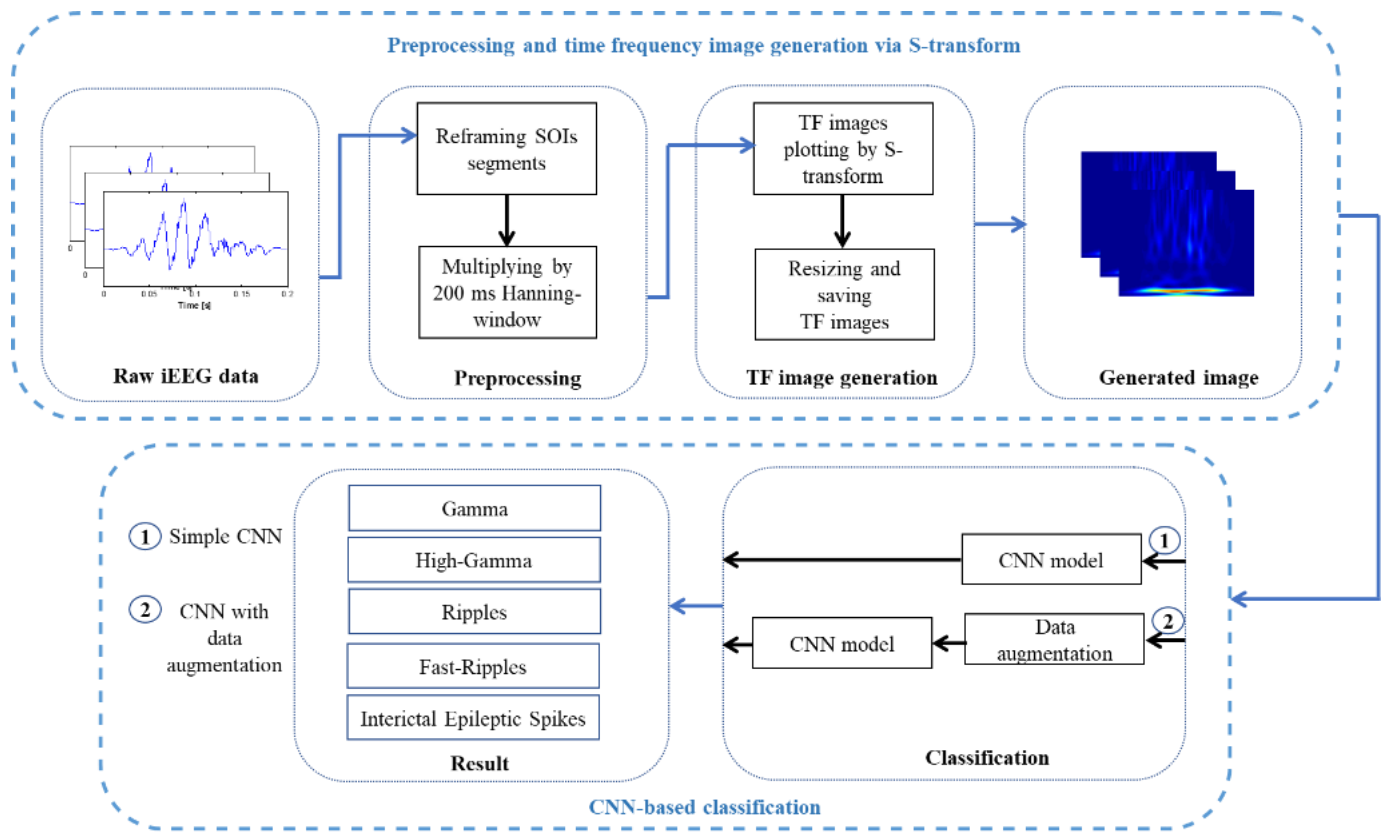
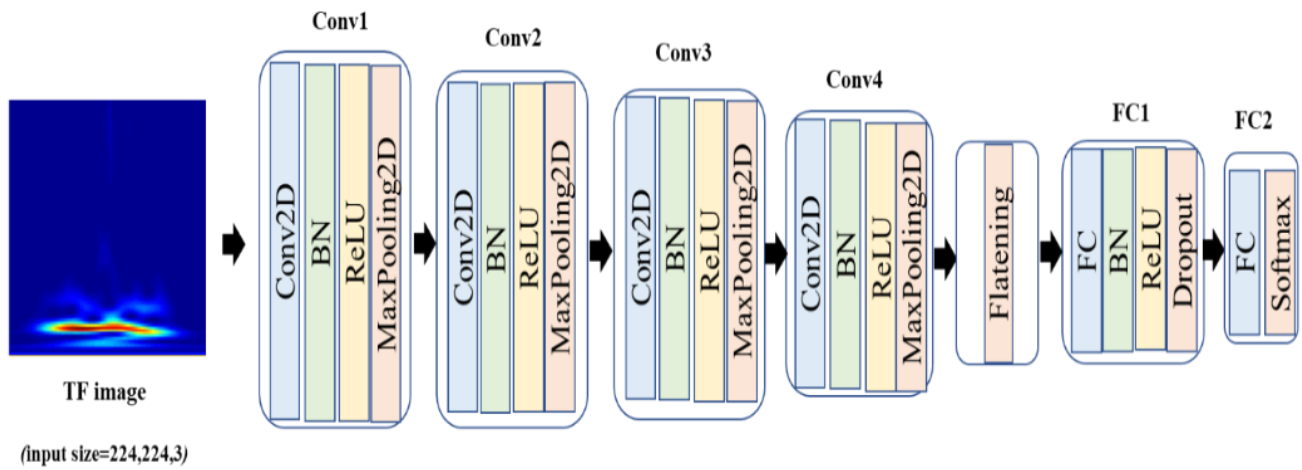


Fig. 2. Flowchart of the proposed CNN-based HFO multi-classification: (1) without data augmentation and (2) with data augmentation.



(a)

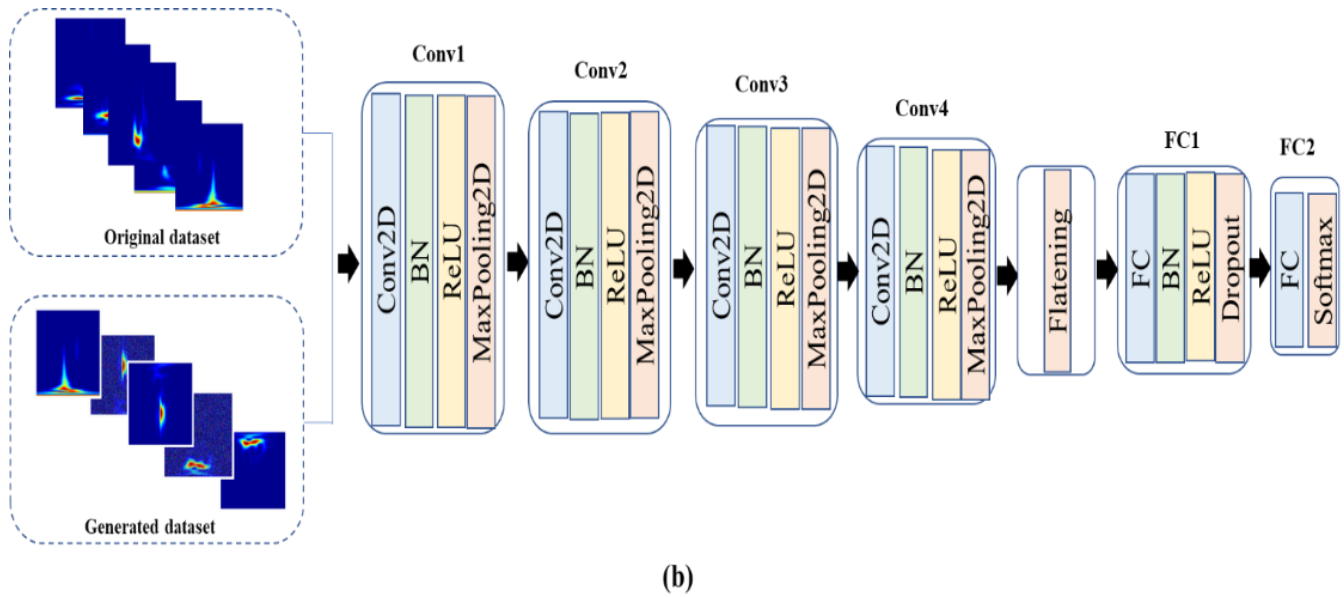


Fig. 3. CNN model architectures. (a) CNN without data augmentation (b) CNN with data augmentation.

TABLE 2  
PARAMETERS USED IN THE PROPOSED CNN MODEL

Layers No.	Layer type	Input shape	Output shape	# Trainable parameters	With BN	Activation
1	Input Layer	[224,224,8]	[224,224,8]	0	×	-
2	Conv2D	[224,224,8]	[224,224,8]	224	✓	ReLU
3	MaxPooling 2D	[224,224,8]	[112,112,8]	0	×	-
4	Conv2D	[112,112,8]	[112,112,16]	3216	✓	ReLU
5	MaxPooling 2D	[112,112,16]	[56,56,16]	0	×	-
6	Conv2D	[56,56,16]	[56,56,32]	12832	✓	ReLU
7	MaxPooling 2D	[56,56,32]	[28,28,32]	0	×	-
8	Conv2D	[28,28,32]	[26,26,64]	18496	✓	ReLU
9	MaxPooling 2D	[26,26,64]	[13,13,64]	0	✓	-
10	Flatten	[13,13,64]	[10816]	0	✓	-
11	FC1	[10816]	[25]	270425	×	ReLU
12	Dropout	[25]	[25]	0	✓	-
13	FC2 (softmax)	[25]	[5]	130	✓	-

standard size ( $224 \times 224$ ) to speed up the training process. The feature extraction process takes place automatically in a hierarchical fashion through four blocks of convolution layers each of which having a filter size  $3 \times 3$ . Each block ends with a 2D max pooling layer which represents a subsampling process. In fact, no strides are used in any of the convolution layers, and they all employ the "same padding" configuration. This means that the input and output dimensions remain the same after each convolution operation. Then, the outputs from all these layers are flattened. For the classification part, two fully connected (FC) layers contain, respectively, 25 and 5 nodes. In order to prevent the model overfitting, a dropout regularization, with a factor of 0.5, is applied. Finally, a Softmax output layer is added to predict the

final class. A ReLU activation layer is applied in each convolution layer. The Max pooling step used after the ReLU activation reduces the spatial dimensionality of the extracted feature maps and helps in extracting the most relevant features. Batch normalization (BN) layer [31] is placed between the convolution layer and ReLU layer. The role of this BN layer is (i) to normalize the feature map so that the CNN model becomes more efficient even with a small number of parameters and (ii) to promote a fast-training ability. Details on the CNN architecture parameters are given in Table 1.

## 2) Pipeline 2: CNN with DA

DA [30] is a technique to increase the size of dataset used to train the model. It consists in increasing the size



of the data by adding more synthetic patterns produced from the original ones by simply applying few transformations (flipping, rotation, translation, compression, etc.). Recently, several studies have reported the efficiency of DA in improving DL-based classification [32-34]. In the context of HFOs classification, only the study of Nadalin *et al.* [20] has employed the DA procedure to increase the training data. Fig. 3(b) illustrates the CNN-based model combined with DA.

In this paper, DA was performed, using "ImageDataGenerator" in Keras, by applying some basic image processing tool such as flipping [35]. Besides, a zero-mean Gaussian noise with standard deviation  $\sigma = 0.05$  was also employed to augment our training dataset as initially suggested in [36]. In fact, authors in [36] have recently shown that adding noise improves the performance of the model. Finally, once the data were augmented, they were shuffled with the original training data and the new training dataset was next introduced, as shown in Fig. 3(b), to the CNN model.

After applying the DA procedure, the new dataset contains approximately 2000 images per class compared to 1000 images prior to DA.

#### IV. NUMERICAL EXPERIMENTS

This section is devoted to evaluate the performance of the proposed HFOs multi-classification approach. The clinical database is firstly described. Then, the quantitative evaluation and some implementation details are given. Finally, classification performance results of the proposed framework are presented.

##### A. Quantitative evaluation

To assess the proposed method, four performance metrics are used: sensitivity (SEN), specificity (SPE), accuracy (ACC), and F1-score. They are defined as follows:

$$SEN = \frac{TP}{TP + FN} \quad (3)$$

$$SPE = \frac{TN}{TN + FP} \quad (4)$$

$$ACC = \frac{TP + TN}{TP + TN + FN + FP} \quad (5)$$

$$F1 - score = \frac{2TP}{2TP + FP + TN} \quad (6)$$

where  $TP$ ,  $TN$ ,  $FN$ ,  $FP$  stand, respectively, for the number of true positives, true negatives, false negatives and false positives.

##### B. Experimental Setup

The different experiments were conducted with a system equipped with 32 GB of memory, using MATLAB (R2017a) and Python 3 environments, and powered by an Intel® Xeon® 2.7 GHz (with 2 processors) processor. All simulations were performed and evaluated on this system. The reframing of EOIs and the generation of TF representation were accomplished using MATLAB. Tensorflow 2.6 and Keras 2.6 libraries were used to conduct the ML part. OpenCV 4.5 was employed for the image processing part.

The proposed model employed "categorical\_crossentropy" as a loss function and stochastic gradient descent with momentum (SGDM) as the optimizer. The SGDM optimizer had a momentum value of 0.9, which is the default parameter in Keras. Additionally, the model was trained using a batch size of 32 for 100 epochs, with 5 steps per epoch. The initial learning rate for the model was set to 0.0001. Furthermore, to address the issue of overfitting, a stratified K-fold (with  $K=5$ ) cross-validation technique was employed. This approach ensures a balanced distribution among classes, thus providing an advantage in maintaining data integrity. More particularly, the dataset was randomly split into five subsets ( $K_1$ ,  $K_2$ ,  $K_3$ ,  $K_4$  and  $K_5$ ). Hence, four subsets were used to train the CNN model while the remaining subset was used as test set. Moreover, the latter four subsets were divided into the ratio 80:20 for training and validation as recommended in [17]. Finally, the results from the 5-fold cross validation procedure were averaged to produce single overall classification metrics *i.e.*, SEN, SPE, ACC and F1-score for each class. It is important to mention that a separate test set was created prior to the application of DA.

##### C. Training process

Performance results of the proposed framework during the training stage for 100 epochs are illustrated in Fig. 4 and Fig. 5.

Fig. 4 shows the loss and accuracy curves during training and validation of the CNN-based model without DA. It can be seen from this figure that the model does not suffer from overfitting. After 100 epochs, as shown in Fig. 4(a), the training and validation loss values tend to 0.309 whereas the training and validation accuracies tend to 0.897 as depicted in Fig. 4(b).

Fig. 5(a) shows the loss curves during the training and validation steps of the CNN with DA, whereas Fig. 5(b) depicts the training and validation accuracy. Similarly, it can be seen that there is no overfitting problem. The validation loss reaches a value equal to 0.300 on the 58-th epoch to attain 0.245 in the last epoch. Additionally, CNN with DA shows a slight increase in the validation

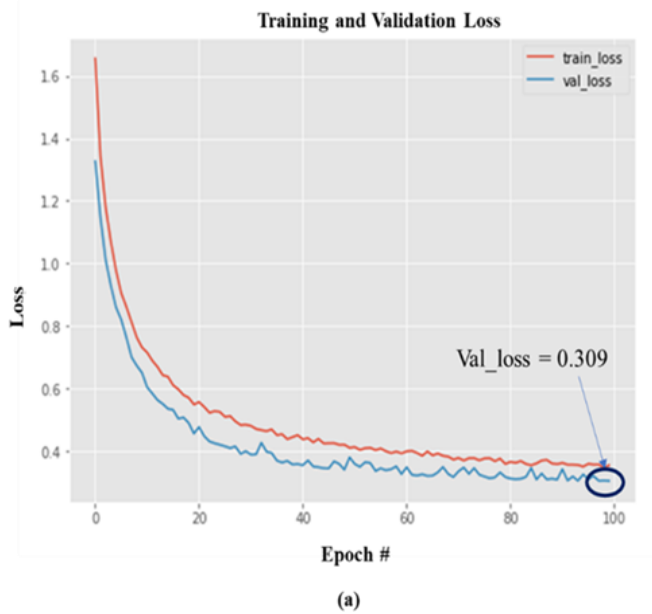


Fig. 4. Training and validation curve of CNN model without DA: (a) loss and (b) accuracy

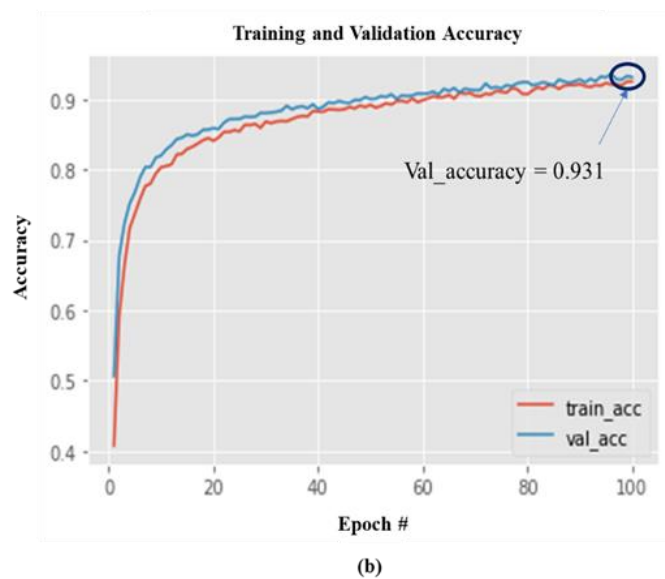
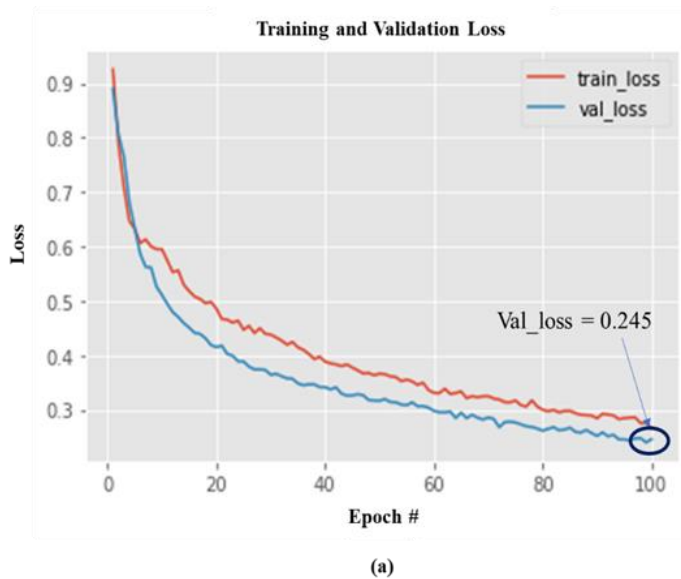


Fig. 5. Training and validation curve of CNN model with DA: (a) loss and (b) accuracy

accuracy (0.931) as illustrated in Fig. 5(b) compared to the one without DA (0.897).

Briefly, loss curves in the two cases (CNN with or without DA) decrease along with the epochs. Similarly, accuracy curves show an increase along with the epochs.

Consequently, we can conclude that, in the training phase, the proposed CNN performs well on TF images dataset.

#### D. Classification results

This subsection is devoted (i) to evaluate the performance of the learned CNN model on the test data and (ii) to compare the proposed approach with a previous ML-based model [24]. The mean and standard deviation of various classification metrics were calculated across 5-fold cross-validation and reported.



### 1) Classification performance

Table 3 entails the classification performance of the CNN model with DA compared to the CNN without DA, for HFOs and IESs classes in terms of SEN, SPE, ACC, and F1-score. It can be seen that the proposed CNN shows generally a good performance for all classes (*i.e.*  $\gamma$ ,  $H\gamma$ , Rs, FRs, IESs).

Particularly, for the CNN-based model, the  $\gamma$  class yields significantly higher performance compared to other classes. More precisely, SEN, ACC and F1-score are respectively equal to 0.950, 0.959 and 0.959. Besides, the highest value of SPE (0.973) is obtained for the IESs class. Regarding the average classification performance, the proposed CNN-based model achieves 0.814 (0.130) 0.953 (0.016), 0.925 (0.033) and 0.872 (0.085) for SEN, SPE, ACC and F1-score respectively.

Additionally, we also provide in Table 3 the classification results of the proposed CNN-based approach with DA in terms of SEN, SPE, ACC and F1-score. The highest results were obtained with the  $\gamma$  class as it the same case of CNN without DA, equal respectively to 0.950, 0.982, 0.970 in terms of SEN, SPE and ACC. Concerning the ACC, the best score is achieved by IESs class.

To investigate the global behavior of the proposed CNN model with and without DA, the average performance of all classes is illustrated in Fig. 6 in terms of SEN, SPE, ACC and F1-score. According to this figure, it can be noticed that the proposed CNN model with DA shows higher multi-classification performance in terms of abovementioned metrics compared to the case where no DA step was used. Additionally, from the results depicted in Fig. 6, a good accuracy (0.978) was obtained with DA procedure. Globally, the obtained results show the positive impact of DA on the classification performance since this procedure enables to

provide more distinguishable patterns from the original data.

### 2) Comparison with previous study

A comparative study between the proposed approach in its two versions (*i.e.*, CNN with and without DA) and a ML-based method [24], namely Support Vector Machine (SVM) was conducted. Various feature extraction techniques including TF features and image based-ones have been adopted in conjunction with the abovementioned ML algorithms. To ensure a fair comparison, the same iEEG training dataset was used for all methods assessed in this study. The overall SEN, SPE, ACC and F1-score for each algorithm were averaged and shown in Fig. 7.

According to this figure, traditional ML method (SVM) provided lower classification performance for all criteria compared to the proposed CNN-based model in its two versions. More precisely, it provided values of 0.765, 0.941, 0.906 and 0.768 for SEN, SPE, ACC and F1-score respectively, as shown in Fig. 7. This can be explained by the fact that contrary to CNN models where relevant features are automatically extracted, traditional ML methods strongly rely on a pre-selection of the relevant features in a set of manually predefined ones.

To sum up, compared to all the considered algorithms, the proposed CNN DA-based model achieved the highest ACC with a value of 0.978.

### 3) Discussion

This paper introduces a novel CNN-based approach for the multi-classification of HFOs ( $\gamma$ ,  $H\gamma$ , Rs, FRs) and IESs events. The CNN-based strategy enables the automatic extraction of discriminant features from the TF images obtained through the S-transform of these events. These features stand as the input of the CNN model for classification.

TABLE 3  
CLASSIFICATION PERFORMANCE CRITERIA FOR THE PROPOSED CNN MODEL WITH AND WITHOUT DA, THE MEAN (STD) OF SEN, SPE, ACC AND F1-SCORE ARE REPORTED

Event class	SEN		SPE		ACC		F1-score	
	CNN	CNN+DA	CNN	CNN+DA	CNN	CNN+DA	CNN	CNN+DA
$\gamma$	0.950	0.950	0.961	0.990	0.959	0.982	0.956	0.970
$H\gamma$	0.883	0.941	0.954	0.968	0.940	0.963	0.917	0.955
Rs	0.575	0.754	0.951	0.922	0.874	0.888	0.717	0.830
FRs	0.785	0.738	0.924	0.948	0.895	0.905	0.849	0.830
IESs	0.880	0.920	0.973	0.992	0.955	0.978	0.924	0.955
<b>Average</b>	0.814 (0.130)	<b>0.861 (0.092)</b>	0.953 (0.016)	<b>0.964 (0.026)</b>	0.925 (0.033)	<b>0.978 (0.039)</b>	0.872 (0.085)	<b>0.908 (0.063)</b>

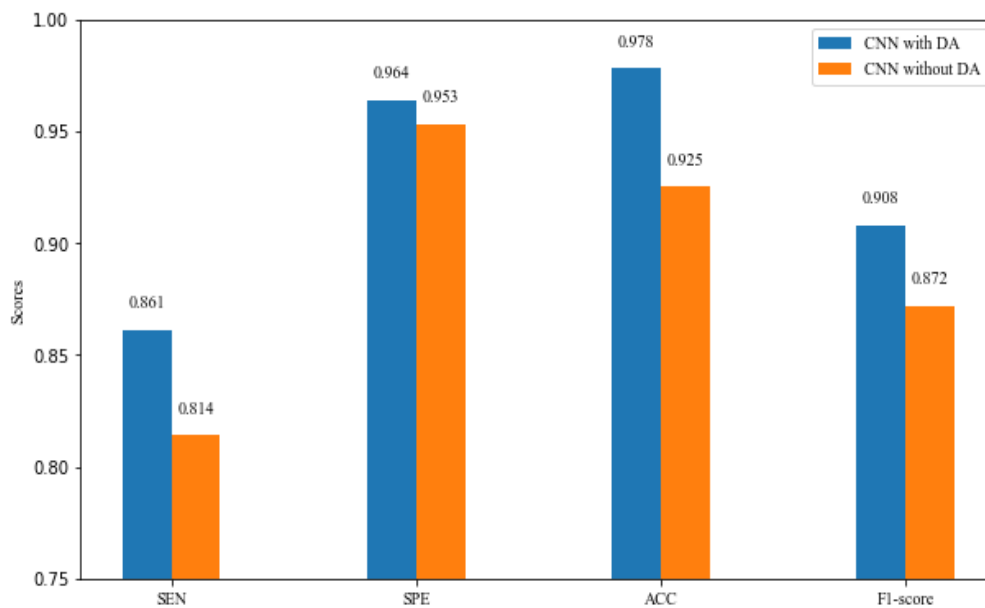


Fig. 6. Comparison between CNN with and without DA

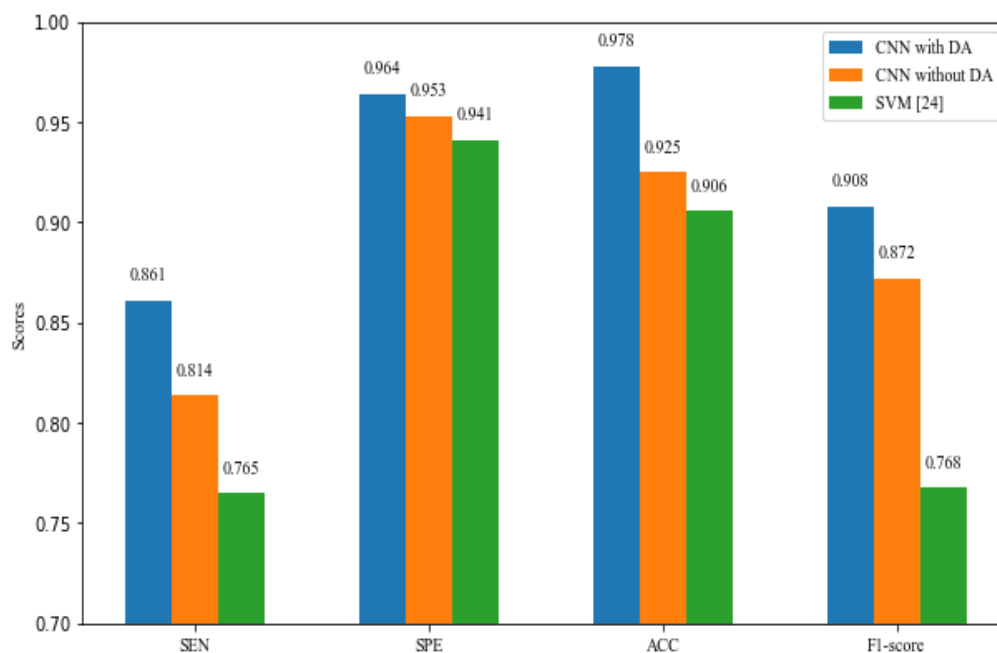


Fig. 7. HFOs and IESs multi-classification performance of the proposed CNN model in its two variants (CNN with and without DA) with traditional ML algorithm (SVM)

As shown in Fig. 6, the proposed CNN model demonstrates favorable classification performance in terms of the different evaluation metrics, such as SEN, SPE, ACC, and F1-score.

Recently, DA has become a commonly used method in many fields [33, 34]. Its main goal is to cope with the issue of limited size datasets. This is henceforth possible through the generation of new samples from the original ones using different transformations.

In this work, two factors (flipping and Gaussian noise addition) were used to augment the training TF images. Experimental results (see Fig. 6) showed, as expected, that the CNN-based model achieved better multi-classification performance in terms of SEN, SPE, ACC and F1-score when DA is employed.

To prove the effectiveness of the proposed CNN model with its two variants, we have compared its classification performance with traditional ML algorithm [24] *i.e.*, SVM. According to Fig. 7, a good behavior of the proposed CNN model in its two variants (CNN with and without DA) was observed compared to ML-based one [24]. This is probably due to the fact that standard classifier was based on manually feature extraction, which is a subjective process due to the limited knowledge of the mechanism of HFOs.

The primary contribution of this study is the utilization of a CNN-based model for the multi-classification of HFOs, which distinguishes it from previous studies that were limited to binary classification. Moreover, this research highlights the effectiveness of DA in enhancing the classification of HFOs and IESs. Typically, training a CNN model needs a significant number of samples to get good performance. Therefore, DA can be an effective approach for augmenting the diversity of the available data and, more importantly, addressing the limited size issue of clinical datasets.

When reviewing the findings of our study, it is vital to take into account certain limitations that provide valuable insights for both the present comprehension and potential enhancements. Despite the fact that DA has demonstrated, through our study, its relevance in handling limited-size database in the context of epilepsy, there is still a real need to keep collecting and incorporating more diverse and representative datasets. This would increase the ability of the employed DA to generate more diverse meaningful examples for the model training assessment and consequently to fully ensure the applicability of the proposed model to real-world scenarios and conditions.

Furthermore, enhancing the efficiency of our CNN-based model for HFOs multi-classification by refining and optimizing its architecture holds great potential. A possible way to processed is by exploring pretrained models such as VGG16 through transfer learning. This extension can notably expedite the training process, leading to faster convergence. Besides, the DL-based technique has shown, according to our results, higher HFOs multi-classification performance compared to the ML-based one. But this was obviously at the expense of the model complexity. Thus, it would be interesting to

consider instead a hybrid model that combines both DL and ML techniques. This hybrid model aims to strike the optimal balance between complexity and classifier performance, thereby offering an enhanced and optimized solution for our HFOs multi-classification task.

## V. CONCLUSION

In this paper, a new deep learning framework based on CNN for the multi-classification of HFOs and IESs was proposed. This framework was based firstly on developing a CNN-based model, and, secondly, on testing the impact of the data augmentation step on the HFOs multi-classification quality. The results showed that the best performance was obtained using CNN with data augmentation. The good behavior of the proposed framework, in terms of SEN, SPE, ACC and F1-score, was confirmed on real iEEG signals. Additionally, a comparative study was conducted to test the efficiency of our proposed framework compared to previous ML method *i.e.* SVM. Results showed the good behavior of our framework compared to the traditional ML method. To the best of our knowledge, this is the first study where a 2D deep CNN was employed for HFOs multi-classification. The DA procedure allows to compensate for the limited size of our training dataset - as collecting new clinical data is not an easy task - and improves the HFOs and IESs classification performance.

## Acknowledgment

The study presented in this paper was conducted as a component of the PHC (Partenariat Hubert Curien) Project CREDIADIC n°41711PK, CMCU Code 19G1411.

## Competing interests

The authors declare no competing interests.

## Data availability

The data that has been used is confidential.

## References

- [1] Jacobs J, Staba R, Asano E, Otsubo H, Wu J Y, Zijlmans M, Mohamed I, Kahane P, Dubeau F, Navarro V, Gotman J (2012) High-frequency oscillations (HFOs) in clinical epilepsy. *Prog Neurobio* (98) 3: 302-15. 2. [https://doi: 10.1016/j.pneurobio.2012.03.001](https://doi.org/10.1016/j.pneurobio.2012.03.001).
- [2] Peng G, Nourani M, Harvey J, Dave H (2020) Feature selection using F-statistic values for EEG signal analysis. 42th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 5963-5966.
- [3] Zijlmans M, Jiruska P, Zelmann R, Leijen F. S S, Jefferys J, G R, Gorman J (2012) High-frequency oscillations as a new biomarker in epilepsy. *Ann Neurol* 71(2): 169-178. [https:// doi: 10.1002/ana.22548](https://doi.org/10.1002/ana.22548).
- [4] Jrad N, Kachenoura A, Merlet I, Bartolomei F, Nica A, Biraben A, Wendling F (2017) Automatic Detection and Classification of High-

- Frequency Oscillations in Depth-EEG Signals. *IEEE Trans. Biomed. Eng* 64 (9): 2230-2240.
- [5] Roehri N, Bartolomei F (2019) Are high-frequency oscillations better biomarkers of the epileptogenic zone than spikes?. *Curr. Opin. Neuro.* 17(1): 213-219.
  - [6] Lachner-Piza D, Jacobs J, Bruder J C, Schulze-Bonhage A, Stieglitz T, Dümpelmann M (2020) Automatic detection of high-frequency-oscillations and their sub-groups co-occurring with interictal-epileptic-spikes. *J. Neural. Eng.* 17(1).
  - [7] Sciaraffa N, Klados M A, Borghini G, Flumeri G D, Babiloni F, Aricò P (2020) Double-step machine learning based procedure for HFOs detection and classification. *Brain*, 10 (4).
  - [8] Blanco J A, Stead M, Krieger A, Viventi J, Marsh W R, Lee K H, Worrell G A, Litt B (2010) Unsupervised classification of high-frequency oscillations in human neocortical epilepsy and control patients. *J. Neurophysiol.* 104(5): 2900-2912.
  - [9] Liu S, Sha Z, Sencer A, Aydoseli A, Bebek N, Abosch A, Henry T, Gurses C, Ince N F (2016) Exploring the time-frequency content of high frequency oscillations for automated identification of seizure onset zone in epilepsy. *J. Neural Eng.* 13(2).
  - [10] Shin H C, Roth H R, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers R M (2016) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning (2016) *IEEE tran. on med imaging.* 35(5): 1285-1298.
  - [11] Bou Nassif A, Shahin I, Attili I, Azzeh M, Shaalan K (2019) Speech Recognition Using Deep Neural Networks: A Systematic Review. *IEEE Access.* 7: 19143-19165.
  - [12] Ahad Tawhid M N, Siuly S, Wang H, Whittaker F, Wang K, Zang Y (2019) A spectrogram image based intelligent technique for automatic detection of autism spectrum disorder from EEG. *PLoS One.* 16(6). doi: 10.1371/journal.pone.0253094. eCollection 2021.
  - [13] Rashed-Al-Mahfuz M, Moni M. A, Uddin S, Alyami S A, Summers M A, Eapen V (2021) A Deep Convolutional Neural Network Method to Detect Seizures and Characteristic Frequencies Using Epileptic Electroencephalogram (EEG) Data. *IEEE Journal of Trans Eng in Health and Medicine.* 9:1-12.
  - [14] Soleimani M, Vahidi A, Vaseghi B (2021) Two-Dimensional Stockwell Transform and Deep Convolutional Neural Network for Multi-Class Diagnosis of Pathological Brain. *IEEE Transactions on Neural Systems and Rehabilitation Engineering.* 29: 163-172.
  - [15] Sun J, Cao R, Zhou M, Hussain W, Wang B, Xue J, Xiang J (2021) Hybrid deep neural network for classification of schizophrenia using EEG Data. *Sci Rep.* 11 (1):4706. [https://doi: 10.1038/s41598-021-83350-6](https://doi.org/10.1038/s41598-021-83350-6).
  - [16] Jadhav P, Rajguru G, Datta D, Mukhopadhyay S (2020) Automatic sleep stage classification using time-frequency images of CWT and transfer learning using convolution neural network. *Biocybernetics and Biomedical Engineering.* 40(1):494-504.
  - [17] Zuo R, Wei J, Li X, Li C, Zhao C, Ren Z, Liang Y, Geng X, Jiang C, Yang X, Zhang X. (2019) Automated detection of high-frequency oscillations in epilepsy based on a convolutional neural network. *Front Comput. Neurosci.* 13(6). [https://doi: 10.3389/fncom.2019.00006](https://doi.org/10.3389/fncom.2019.00006). eCollection 2019.
  - [18] Lai D, Zhang X, Ma K, Chen Z, Chen W, Zhang H, Yuan, Ding L (2019) Automated Detection of High Frequency Oscillations in Intracranial EEG Using the Combination of Short-Time Energy and Convolutional Neural Networks. *IEEE Access.* 7:82501-82511.
  - [19] Zhao B, Hu W, Zhang C, Wang X, Yao W, Liu C, Mo J, Yang X, Ma Y, Shao X, Zhang K, Zhang J (2020) Integrated Automatic Detection, Classification and Imaging of High Frequency Oscillations With Stereoelectroencephalography. *Front. Neurosci.* 14:465. [https://doi: 10.3389/fnins.2020.00546](https://doi.org/10.3389/fnins.2020.00546).
  - [20] Nadalin J K, Eden U T, Han X, Richardson R M, Chu C J, Kramer M A (2021) Application of a convolutional neural network for fully-automated detection of spike ripples in the scalp electroencephalogram. *J. of neurosciences.* 360:109239. [https://doi: 10.1016/j.jneumeth.2021.109239](https://doi.org/10.1016/j.jneumeth.2021.109239).
  - [21] Katz J, Abel TJ (2019) Stereoelectroencephalography versus subdural electrodes for localization of the epileptogenic zone: what is the evidence? *Neurotherapeutics.* 16, 59-66, Springer. DOI: <https://doi.org/10.1007/s13311-018-00703-2>.
  - [22] Peng G, Nourani M, Dave H, Harvey J (2022) Modeling and analysis of seizure network using SEEG for pre-surgery evaluation", 22nd IEEE International Conference on Bioinformatics and Bioengineering (BIBE), pp 327-332.
  - [23] Stockwell RG, Mansinha L, Lowe R P (1996) Localization of the complex spectrum: the S transform. *IEEE Trans. Signal Process.* 44(4):998-1001.
  - [24] Krikid F, Karfoul A, Chaibi S, Kachenoura A, Nica A, Kachouri A, Le Bouquin Jeannès R (2022) Classification of High Frequency Oscillations in intracranial EEG signals based on coupled time-frequency and image-related features. *Biomed. Signal Proc and. Con.* 73(3).
  - [25] Burnos S, Hilfiker P, Sürücü O, Scholkmann F, Kraysenbühl N, Grunwald T, Sarthein J (2014) Human Intracranial High Frequency Oscillations (HFOs) Detected by Automatic Time-Frequency Analysis. *PLoS One.* 9(4):e94381. doi: 10.1371/journal.pone.0094381.
  - [26] Migliorelli C, Bachiller A, Alonso J F, Romero S, Aparicio J, Jacobs-Le Van J, Mañanas M A, San Antonio-Arce V (2020) SGM: a novel time-frequency algorithm based on unsupervised learning improves high-frequency oscillation detection in epilepsy. *J Neural Eng.* 17(2):026032. doi: 10.1088/1741-2552/ab8345.
  - [27] Jmour N, Zayen S, Abdelkarim A (2018) Convolutional neural networks for image classification. *International Conference on Advanced Systems and Electric Technologies (IC\_ASET):*397-402.
  - [28] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich R (2015) Going deeper with convolutions. *IEEE conference on computer vision and pattern recognition:*1-9.
  - [29] Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. *arXiv:1409.1556*. [Online]. Available: <http://arxiv.org/abs/1409.1556>.
  - [30] Shorten C, M Khoshgoftaar T M (2019) A survey on Image Data Augmentation for Deep Learning. *J Big Data.* 60(6).
  - [31] Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32nd International Conference on Machine Learning (Lille):*448-456.
  - [32] Zhang K, Xu G, Han Z, Ma K, Zheng X, Chen L, Duan N, Zhang S (2020) Data Augmentation for Motor Imagery Signal Classification Based on a Hybrid Neural Network. *Sensors.* 20(6). doi: 10.3390/s20164485.
  - [33] Ssekidde P, Steven Eyobu O, Seog Han D, J Oyana T J (2021) Augmented CWT Features for Deep Learning-Based Indoor Localization Using WiFi RSSI Data. *Applied sciences.* 11(4).
  - [34] Garcea F, Serra A, Lamberti F, Morra L (2023) Data augmentation for medical imaging: A systematic literature review. *Computers in Biology and Medecine.* 152:106391. doi: 10.1016/j.combiomed.2022.106391.
  - [35] Khosla C, Saini B S (2020) Enhancing Performance of Deep Learning Models with different Data Augmentation Techniques: A Survey. *International Conference on Intelligent Engineering and Management (ICIEM):* 79-85.
  - [36] Moreno-Barea F J, Strazzera F, Jerez J M, Urda D, Franco L (2018) Forward Noise Adjustment Scheme for Data Augmentation. *IEEE Symposium Series on Computational Intelligence (SSCI):*728-734.