A channel-wise attention-based representation learning method for epileptic seizure detection and type classification

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC0

SUBMISSION DATE / POSTED DATE

27-11-2021 / 06-12-2021

CITATION

Baghdadi, Asma; fourati, Rahma; Aribi, yassine; Daoud, Sawsan; Dammak, Mariem; Mhiri, Chokri; et al. (2021): A channel-wise attention-based representation learning method for epileptic seizure detection and type classification. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.17087147.v1>

DOI

[10.36227/techrxiv.17087147.v1](https://doi.org/10.36227/techrxiv.17087147.v1)

A channel-wise attention-based representation learning method for epileptic seizure detection and type classification

Asma Baghdadi, *Member, IEEE*, Rahma Fourati, *Member, IEEE*, Yassine Aribi, *Senior Member, IEEE*,
Sawsan Daoud, Mariem Dammak, Chokri Mhiri, Patrick Siarry, *Senior Member, IEEE*
and Adel M. Alimi, *Senior Member, IEEE*

Abstract—Epilepsy affect almost 1% of the worldwide population. An early diagnosis of seizure types is a patient-dependent process which is crucial for the treatment selection process. The selection of the proper treatment relies on the correct identification of seizures type. As such, identifying the seizure type has the biggest immediate influence on therapy than the seizure detection, reducing the neurologist's efforts when reading and detecting seizures in EEG recordings. Most of the existing seizure detection and classification methods are conceptualized following the patient-dependent schema thus fail to perform well with unknown cases. Our work focuses on patient-independent schema for seizure type classification and pays more attention to the explainability of the underlying attention mechanism of our method. Using a channel-wise attention mechanism, a quantification of the EEG channels contribution is enabled. Therefore, results become more interpretable and a visualization of brain lobes contribution by seizure types is allowed. We evaluate our model for seizure detection and type classification on CHB-MIT and the recently released TUH EEG Seizure, respectively. Our model is able to classify 8 seizure types with an accuracy of 98.41%, directly from raw EEG data without any preprocessing. A case study showed a high correlation between the neurological baselines and the interpretable results of our model.

Index Terms—Seizure type classification, EEG feature learning, patient-independent, attention mechanism, Deep LSTM, explainable model.

I. INTRODUCTION

EPILEPSY is a neurological disease that manifests with irregular and sudden discharges of neurons in the brain. Affecting almost 1% of the worldwide population, it impacts the quality of life of these persons negatively. Neurologists use medications to control seizures. While it might work for

some, this method might not have the same effect for a patient with uncontrolled seizures.

Seizures manifest in different forms such that each type requires a specific treatment. The initialisation of the treatment procedure relies on the correct identification of seizures type. Seizure types are classified by the International League Against Epilepsy based on the manifestation symptoms. Neurologists perform the identification process using electroencephalography recording combined with videos. Thanks to their expertise, a correct identification of the seizure attack is usually made. Diagnosing correctly the seizure type provide accurate prognosis information, thus help the neurologists to select the adequate drug therapy. However, it remains challenging, labor-intensive, and time-consuming. It usually involves the monitoring of several real-time seizures of a patient, needing a continuous EEG recording [1].

Similar clinical features are the main contributing elements in inaccurate distinguishing characteristics, as clinical and EEG symptoms may be similar (comparable) for both focal and generalized seizures [2]. Recently, several researches have demonstrated that for some cases an experienced neurologist can have troubles in recognizing the correct type of seizure [3]. The variability of a seizure manifestation between different patients as well as for the same patient overtime, further complicates the clinical diagnosis.

EEG artifacts must be detected over the whole record which complicates the task of seizure bio-markers identification. An automated seizure classification and detection system assisting professionals can greatly improve the time-consuming clinical EEG diagnosis and its volatility.

Patient-specific approaches are widely used for seizure detection and have showed good performances since a model is trained for each patient. This model are conceptualized to be tested using samples from the same patient and researchers don't notice that the obtained scores are specifically related to the patient in concern. The advantage of a patient-independent method is that the model can generalize better due to the training using heterogeneous EEG data. The model learn different characteristics thus it performs well in detecting ictal and non-ictal segment for a new patient

In this context, our work proposes a patient-independent method for epileptic EEG representation learning and seizure type classification. An attention layer is integrated which learns channel-wise weights from multi-channel raw EEG

A. Baghdadi, R. Fourati and A. M. Alimi are with the REsearch Groups in Intelligent Machines (REGIM Lab), University of Sfax, National Engineering School of Sfax (ENIS), BP 1173, Sfax, 3038, Tunisia.

A. M. Alimi is with the department of Electrical and Electronic Engineering Sciences, Faculty of Engineering and the Built Environment, University of Johannesburg, South Africa.

E-mail: {asma.baghdadi, rahma.fourati, adel.alimi}@ieee.org

Yassine Aribi is with the University College of Ranyah Taif University, Sciences and Technology department, Taif, Saudi Arabia.

E-mail: yaliraibi@tu.edu.sa

Sawsan Daoud, Mariem Dammak and Chokri Mhiri are with the department of neurology, Habib Bourguiba, university hospital, Sfax, Tunisia. Neuroscience laboratory, Sfax university, LR-12-SP-19, Sfax, Tunisia

Patrick Siarry, is with Laboratoire Images, Signaux et Systèmes Intelligents (LISSI), Université de Paris 12, 94010 Créteil, France.

E-mail: siarry@univ-paris12.fr

signal. To the best of our knowledge this is the first application of a raw-based channel-wise attention mechanism applied on seizure type classification.

This work presents four main contributions: (1) EEG feature learning with deep LSTM model which does not rely on the knowledge expert for the task of handcrafted feature extraction. (2) A channel-wise attention mechanism for the analysis of the most relevant channels in EEG-based epilepsy task. (3) An extensive validation of the proposed method on seizure detection and seizure type classification. (4) Explainable brain region analysis through a case study where a correlation between the seizure type and the localization of the highest active channels is established.

II. RELATED WORK ON EEG-BASED SEIZURE DETECTION AND TYPE CLASSIFICATION

Generally, seizure type classification was realized by extracting relevant features and using ML techniques for the classification. However, raw EEG signal are rich with spatial and temporal features that can be learnt and classified for the aim of seizure detection and classification. In the literature few public and free datasets exist. For instance, CHB-MIT [4] and TUSZ [5] are widely used free benchmarks. We first describe these datasets and next overview the related work on them.

A. Existing epileptic benchmarks

The scalp long-term EEG dataset CHB-MIT [4] contains data gathered from 23 pediatric patients admitted in the Boston Children's Hospital suffering from intractable seizures. EEG Recordings consist of 983 hours expertly labeled to mark the onsets' start and end in epochs with ictal activities. The concern with this data is the duration of these seizures, which are minor in comparison with the total EEG duration of each case, causing a very unbalanced data distribution thereby generating a very challenging classification. 256 Hz was the sampling frequency for all recorded signals, with a 16-bit resolution. The international 10-20 electrode positioning system was chosen.

The Temple University Hospital SeiZure corpus (TUSZ) dataset v1.5.1 [5] comprises sEEG data from 1185 sessions of which 1029 sessions containing seizures, for a total of 2370 seizure event recordings as specified in the summary sheet annexed to the dataset. TUSZ covers a total of 8 seizures types as depicted in table II indicating the total number of events per type. The used version consists of EEG recordings resampled to 250 Hz.

TABLE I
DESCRIPTION OF TUSZ AND CHB-MIT DATASETS

Features	TUSZ	CHB-MIT
No. of patients (sEEG)	592	23
No. of patients with seizures	202	23
No. of seizure recordings	2370	185
No. of seizure classes	8	2

TABLE II
DETAILS OF ALL 8 TYPES OF SEIZURE INCLUDED IN THE TUSZ DATASET

Seizure code	Seizure type	Total recording (variable durat ^{o*})	Nb. of samples
FNSZ	Focal Non-Specific	1536	17021
GNSZ	Generalized Non-specific	408	9052
CPSZ	Complex Partial	277	5244
SPSZ	Simple Partial	49	365
ABSZ	Absence	50	74
TNSZ	Tonic	18	83
TCSZ	Tonic Clonic	30	381
MySZ	Myoclonic	2	258

*: The duration of seizure segments in each session are variable. The segmentation en 5 seconds segments can change the data distribution over classes

B. Related work on CHB-MIT and TUSZ datasets

Attention mechanisms has attracted extensive interest in clinical diagnosis, thanks to their capability of extracting feature as in [6] and providing a model interpretability [14]. Most of the recent seizure detection and type classification methodes are designed with patient-specific scheme since they depend on one patient data. However, patient-independent scheme is more interesting and suitable for real context yet more challenging.

Zhang et al. presented in their work [12] a patient-independent method for seizure detection. Raw EEG signals from TUSZ dataset are decomposed into seizure and patient representation in a latent space. The model is composed of two branches. The first branch is an attention-based Convolutional Neural Network (CNN) for seizure detection. The second branch is a CNN for the patient detection. An adversarial learning between the two branches allows the diagnosis of the patient's seizure state and the detection of the patient's identity. By analysing the attention weights, the authors found that T5 channel is important across patients, which is consistent with neurological findings.

Yuan et al [6] [15] favor the critical EEG channels by the mean of their energy passed to a Stacked AutoEncoders (SAEs) as an attention module. They proposed for the first time an method based-on an attention mechanism for biosignal channel selection in healthcare. The proposed method learns a global representation issued from EEG spectrograms of all channels and a local one corresponding to the local view representation treating the signals channel by channel. The channel-aware attention mechanism consists of calculating a global context vector based on the energy of each channel. The method was evaluated using combined samples from 9 cases included in the CHB-MIT dataset. It achieves 97.85% as an F1-score value which outperforms other tested baselines. The authors in [15] added a case study to evaluate their finding and showed that the highest attention scores calculated for one case are scores of 2 channels located at the seizure area.

Table III illustrates recent epileptic methodes. CHB-MIT is widely used for seizure detection. While TUSZ is also used for seizure detection as well as seizure type classification since it is decomposed into 3, 7 or even 8 seizure types.

In certain instances, comparing different methods can be quite challenging since they are validated using different

TABLE III
EXISTING EEG-BASED SEIZURE DIAGNOSIS WORKS

Ref	Year	Method	Dataset	Data	(#) Seizure types	Performance
[6]	2018	<i>Att_{loc}</i> and <i>Att_{glob}</i> attention mechanisms	CHB-MIT	EEG Spectrograms	Seizure detection: Ictal Vs Interictal	Acc:96.61
[7]	2019	Support Vector Machine	TUSZ	Mel Frequency Cepstral Coefficients, Hjorth Descriptor, Independent Component Analysis	(3) Focal non-specific, Generalized non-specific, Tonic-clonic	Acc: 91.4%
[8]	2019	K-Nearest Neighbors	TUSZ	2 pre-processing methods based on FFT	(7) Focal non-specific, Generalized non-specific, Simple-partial, complex-partial, Absence, tonic, tonic-clonic	F1: 90.10%
[9]	2019	CNN	TUSZ	multi-spectral feature from EEG spectrograms	(8) focal non-specific, generalized non-specific, simple-partial, complex-partial, absence, tonic, tonic-clonic, myoclonic	Acc: 84.06%
[10]	2019	Neural Memory Network	TUSZ	Intensity values of STFT	As above	F1:94.50%
		Hybrid Bilinear	TUSZ	As above	As above	F1:97.40%
[11]	2020	Extreme Gradient Boosting	TUSZ	multi-spectral feature from EEG spectrograms	As above	F1: 70.70%
		Ensemble CNN's	TUSZ		As above	F1: 98.40%
[12]	2020	Adversarial Learning using CNNs	TUSZ	Raw EEG	Seizure detection: Normal Vs Seizure	Acc: 80.50
[13]	2021	Unified multi-level spectral-temporal feature extraction	CHB-MIT	PCA-CSP MMSE and	Onset detection	Acc:97.80

datasets with a variable amount of data across classes. In the domain of deep learning based methods To meet the aims of seizure type classification based on deep learning approaches, the authors in [11] designed an ensemble of three DenseNet-based CNN's trained on the TUSZ dataset to reach an average F1-score of 98.40%. The proposed architecture in [11] consists of 45.94M trainable parameters, while the hybrid bilinear structure proposed in the study of [10] contains only 1.2M trainable parameters. In the latter, myoclonic seizures was discarded from the classification task due to the insufficient number of samples on this class. Both [9] and [16] proposed solutions for a multi-class classification problem evaluated on the TUSZ dataset, achieving 84.06% and 94.05%, respectively. As observable from Table III, the bilinear model proposed by [10] achieved better performance on the 8 seizure type classification problem. Machine learning methods, including the k-Nearest Neighbor (KNN) proposed in [8] and the Support Vector Machine (SVM) proposed in [7], exhibit acceptable performance (90.70% and 91.40%), nevertheless, achieved through undesirable extensive feature engineering.

Often in seizure detection, models trained on a single patient perform better than general models trained on multiple patients data. This is partly because there is a large variation between human brains, and partly because there is not necessarily any correspondence between device channels across patients.

In work [17], the author has discussed the difference be-

tween machine learning techniques and deep learning method in distinguishing patients under antiepileptic drugs and those taking no medications, as well as between the two anticonvulsants. The method was validated on TUSZ dataset since it is the largest available dataset [5] The comparison invoked in the work [18] shows that a small difference exists between the used ML techniques and deep model in achieving a moderate accuracy rate for medication use detection. In addition, deep models are less time consuming then ML techniques.

It is worthy that seizure type classification is more important than seizure detection, since it guides the neurologist in epilepsy treatment. Despite its importance, raw EEG data-based works have rarely been attempted for seizure type classification. Inspired by the advantages of deep learning approach, in this work a novel LSTM based on attention mechanism for both seizure detection and seizure type classification is proposed. Our novelty resides in the explainability of our method in analyzing the contribution of each channel in the final decision of the diagnosis through the visualization of the raw EEG data and the corresponding learnt temporal attention-based representation. Our method is based on raw EEG data, i. e. no preprocessing is performed and this is mainly due to the importance of the artifact in EEG-based epilepsy task.

III. ATTENTION-BASED DEEP LSTM MODEL

This section provides details about the proposed method. First, the LSTM model is presented. Then, the attention-based deep LSTM model is explained.

A. LSTM model

Over the last years, deep learning networks were used for EEG classification tasks, including seizure detection [19], emotion recognition [20] [21] [22], and classification of motor (imagery) tasks. Various studies showed that LSTMs outperform other models like decision trees, support vector machines used in our previous work [23], logistic regressions, random forest classifiers, naïve Bayes, feedforward neural networks, deep belief networks, and even CNNs for some tasks [24]. The superior performance of LSTMs' EEG classification over other models is likely due to their ability to account for time dependencies. As EEG is a time-series data, preserving temporal characteristics might significantly improve the model's accuracy.

The LSTM architecture like an RNN network contains 3 main layers. The strength of LSTM came from its hidden layer. This latter contains special blocks called memory blocks. The input and output gates of these blocks perform the control by the activation functions. The revised version of LSTM added a forget gate to the memory blocks. An LSTM network finds the mapping from input sequence $x = (x_1, x_2 \dots x_T)$ to the output sequence $y = (y_1, y_2 \dots y_T)$ by figuring out the network unit activations using the following equations:

$$i_t = \sigma(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_{fx}x_t + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f) \quad (2)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g(W_{cx}x_t + W_{cm}m_{t-1} + b_c) \quad (3)$$

$$o_t = \sigma(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_{t-1} + b_o) \quad (4)$$

$$m_t = o_t \odot hc_t \quad (5)$$

$$y_t = \phi(W_{ym}m_t + b_y) \quad (6)$$

In the above equations, W represents the weight and W_{ix} is the maximum weight of the input gate to the input. W_{ic} , W_{fc} and W_{oc} are the diagonal weights of peepholes connections. The majority of the architectures consisted of one or two LSTM layers, followed by one or two fully connected layers. Input to the LSTM comprised mostly features extracted from EEG signals. However, the signal itself and EEG images (spectrograms) were also used [25].

Several EEG-based studies compared the use of handcrafted features to the raw EEG signal as input for the LSTM model. Usage of the signal itself is consistently and massively under-performed in these comparisons [26], [27], and [28]. These studies reported an accuracy rate of $50.00\% \pm 1.50$ when applying their methods on raw EEG data, while using artifact removal techniques to reduce noise slightly improved the performances. Our work shows that even with raw EEG data and without any preprocessing, our LSTM-att is able to achieve our objective in classifying epileptic seizures.

B. Attention mechanism for multi-channel epileptic signals

Attention mechanism was first introduced by [29] in Encoder/Decoder based on LSTM units for textual sequence translation. They suggest that relative importance should be given to each input words, as well as taking into account the context vector. In a follow up work [30], channel-wise attention-based CNN demonstrates superior performance in image captioning because it has the ability to change different channels' weight in order to explore feature map information. More specifically, it has the ability to gather additional important information about channels.

on the other side, the contribution of different EEG channels varies in seizure diagnosis [31]. Thus, an attention mechanism is introduced for channel importance learning and to pay different attention to various brain lobes. As abovementioned, the attention mechanism allows modeling of dependencies among EEG channels [12] and has shown success in some research topics [17] [32] [33].

Meanwhile, the excellent temporal feature learning ability of Recurrent Neural Networks (RNNs) has been extensively used in research areas such as speech recognition [34], language modeling [35], diseases prediction [27] and many others. Thus, we propose an attention-based LSTM model to automatically extract discriminative information from the received temporal multi-channel EEG data. The Fig.1 depicted all block of the proposed method. First, to explore the importance among the different channels of EEG signal, a channel-wise attention-based mechanism is employed as shown in the left block of the structure diagram of Fig. 1. In the case of seizure detection or seizure type classification, some channels may not contribute to the final decision and thus add redundant information and demean the method capabilities.

The adapted channel-wise mechanism takes into consideration the information of all channels and assigns weights to different channels based on their importance. This mechanism will allow to explain the contribution of each channel to the final decision. Consider that, $X = [X_1, X_2, \dots, X_n]$ represents EEG samples, and $X_i = [x_1, x_2, \dots, x_k]$ where x_k represents the k^{th} channel of EEG sample X_i , and k is the total number of channels of each sample. In this model, the attention scores are directly learnt from the EEG sample. The attention layer, shown in Fig. 1, is used to generate attention weights for each channel and then executes an element-wise multiplication with the output of the dense layer of the LSTM block. In the attention block, the original data are inputted into a fully connected layer where the parameters W and b are initialised for all channels. The attention matrix is element-wisely multiplied by the original inputs. The outputs of the attention block are multiplied by the output of the dense layer of the LSTM block. Then, the attention-based temporal features are passed to a dense layer with the suitable activation respecting the classification task, i. e. sigmoid for seizure detection or softmax for seizure type classification to get the label of the EEG sample.

The attention layer is computed using the following equations:

$$Y_1 = f_{nor}(X_0) \quad (7)$$

Fig. 1. The structure diagram of our attention-based LSTM model

$$Y_2 = w_{al} * Y_1 + b_{al} \quad (8)$$

$$Y_3 = f_{tens}(X_2) \quad (9)$$

$$Y_4 = \sigma(Y_3) \quad (10)$$

Here, X_0 denotes an input of size $(N_{samples}, N_{timesteps}, N_{channels})$. Symbols $(N_{samples}, N_{timesteps}$ and $N_{channels})$ represent the number of samples, the number of time steps, and the number of EEG channels, respectively. Y_1 is a normalized matrix of X_0 size. W_{al} a weight of size $(N_{channels}, N_{channels})$, a bias b_{al} of $(N_{samples}, N_{timesteps}$ and $N_{channels})$, and Y_2 have the same size as Y_1 . A symbol $\sigma(\cdot)$ represents a nonlinear activation function, which transforms the importance of channels to probability distribution, like softmax(.) and sigmoid(.). Y_2 is a tensorized matrix of Y_1 .

Functions $f_{nor}(\cdot)$ and $f_{tens}(\cdot)$ are to normalize and tensorize a matrix.

$$att_{scores} = Y_3 = [att_1; att_2; \dots; att_k] \quad (11)$$

The middle block of the Fig. 1 shows the temporal feature learning module, which comprises a two-layer LSTM. The LSTM network can learn the context information of the sequence thanks to its recurrent structure [36]. The predicted seizure class Y_5 is related to the last dense layer (D2) and the learnt attention scores:

$$Y_5 = Dense_{out} \odot att_{scores} \quad (12)$$

The symbol $\odot$ means an element-wise multiplication between tensors.

IV. EXPERIMENT DESIGN

A. Data preparation

Since the TUSZ data contains numerous sampling rates and to ensure constant input dimension to the neural network, we

used the 250 Hz re-sampled version of the dataset. Twenty-two common channels were selected and readjusted based on the 10–20 international system of scalp EEG placements. In this study, we keep the re-sampled EEG data unchanged. The data is structured into an adequate shape to be fed into the proposed model. The sample duration is fixed to 5-seconds. More specifically, the dimension of each sample becomes $[22 \times 1250]$ where 22 is the number of EEG channels and 1250 is the number of time steps.

In the case of CHB-MIT dataset, recordings are 23-channels signals with almost 18 common channels to be used in this work: FP1, T7, P7, O1, F3, C3, P3, FP2, F4, C4, P4, O2, F8, T8, P8, FZ, CZ and PZ. To the aim of seizure detection task, we extracted ictal and non-ictal segments from the 24 cases included in the dataset. Since, every case can have one or multiple seizures, we extracted all existent seizures for each case. Then, we selected a balanced amount of data to be used for the model training and evaluation. A total of 18,320 segments is used with a dimension of $[18 \times 1280]$.

B. Training and evaluation

In order to extensively evaluate the proposed model performance, a stratified five-fold cross validation is used. We randomly split the dataset into five-folds, where each fold maintains the proportional distribution of classes. Adam optimizer is chosen for the LSTM training with batch size value 20 and for 100 epochs. The TUSZ dataset have an unbalanced class distribution as shown in Table II). The minor class, absence seizures, has 6 minutes of recording. In this case, accuracy by itself is unlikely to evaluate the model. The Precision, Recall and F1-score are thus added as metrics to evaluate the performance of our proposed model. As a regularization technique, we employed the Early stopping mechanism. Hence, we avoided an over-fitting during the training process. This technique consists on the monitoring of the validation loss, if the latter does not improve within 10 epochs, the training process is stopped.

C. Hyper-parameters fitting

In our work, a grid search for all parameters was adopted as depicted in Table IV.

TABLE IV
PROPOSED LSTM ARCHITECTURE PARAMETERS

Parameters	Range	Best Value
LSTM Layers	[1-2]	2
Dropout Layers	[0-2]	0
Learning rate	[0.001,0.002,0.005,0.01,0.1,0.2,0.3]	0.001
Memory units	[50,100,150,200,250]	250
Dropout probability	[0.2,0.3]	None
Dense activation	['softmax', 'relu', 'tanh', 'sigmoid', 'linear']	Softmax
Optimizer	None	Adam
Epochs	[20, 50, 100]	100
Batch size	[10,20,50]	20

This technique prompted the best accuracy for all possible combination of parameters. It is a time consuming step, but it insures that better fitting is used for the final model. Next, we report the hyper-parameter settings in detail. The input EEG sample has a shape of $[N_{timesteps}; N_{channels}]$. The number of units on the first and second LSTM layers was tuned to a range of [50,100,150,200,250]. Dropout probability of each layer tested to be between [0.0,0.2,0.5]. The two FC layers have $N_{features}$ and $N_{classes}$ hidden neurons respectively, and several activation functions were tested ['softmax', 'relu', 'tanh', 'sigmoid', 'linear']. The attention layer has $N_{channels}$ hidden neurons corresponding to the input channels.

A categorical cross-entropy loss function is used in our model for multi-class classification, while binary cross-entropy is used for seizure detection. Considering the limited computational resources available in this study, we chose to use the Adam optimizer and omit other optimizers in the gridsearch parameters.

We tested our method for two classification problems on two datasets: TUSZ and CHB-MIT. Our model was implemented using Python 3.7.9 and Keras 2.3 with the Tensorflow-gpu 2.1.0. The model was run on a NVIDIA GEFORCE GTX 960M. The average Accuracy, Precision, Recall and F1-scores were reported. AUC scores are reported only for seizure detection on CHB-MIT dataset.

V. EXPERIMENTAL RESULTS

In this section, we will discuss the performance results of the proposed model based-on raw EEG data: attention-based deep LSTM, as described in section III-B and the basic deep LSTM model section III-A. Two different classification problems based-on EEG signals are addressed to evaluate whether the integration of the proposed channel-wise attention mechanism is attractive for either method performance improvement and results explainability. Since the considered datasets are imbalanced in their nature, Accuracy, Precision, Recall and F1-score, the average and the standard deviation are used as evaluation metrics.

Fig. 2. LSTM-att model performance in term of accuracy and loss for seizure detection with TUSZ dataset

A. Evaluation on TUSZ for seizure detection and type classification

In seizure detection problem, the LSTM-att achieves an accuracy of $96.78 \pm 0.21\%$ which outperforms the basic LSTM model of approximately 9.46% for the TUSZ dataset as illustrated in Table V. To add, our LSTM-att reached a value 0.976 for AUC metric. Note that, the higher the AUC, the better the model is at distinguishing between patients with the seizure and no seizure. According to the training and validation loss curves depicted in Fig.2, the LSTM-att model does not suffer from overfitting problem.

For seizure type classification, the imbalanced issue is more highlighted than the seizure detection. In this case, comparison is made using F1-score which is the harmonic mean of Precision and Recall and gives a better measure of the incorrectly classified cases than the accuracy metric. For instance, the LSTM model achieves 77.55% as F-score value while the LSTM-att model improved it by 19.32% as shown in Table VI.

The confusion matrix in Fig 3 highlights the classification performance of the LSTM-att model on 8 seizure types from TUSZ dataset. For example, Myoclonic Seizure (MYSZ) is confused with Focal Non-specific Seizure (FNSZ) and Complex Partial Seizure (CPSZ) with 1.39% and 2.78%, respectively. The highest accuracy achieved is with the Tonic seizure (TNSZ) class. However, the lowest accuracy is obtained for ABSZ class. The low count of absence seizures in the TUH dataset can account for this comparability, with just six recording minutes for the model to learn from..

Fig. 3. Confusion matrix of our LSTM-att model validated on TUSZ

According to the aforementioned results, it was shown that the model trained using learnt feature in combination with

TABLE V
RESULTS ON TUSZ DATASET FOR SEIZURE DETECTION

Class	LSTM				LSTM-att			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
Ictal	88.40	88.40	88.45	88.42	97.01	97.01	96.39	96.69
Non-ictal	86.25	86.25	86.72	86.48	96.56	96.56	97.24	96.89
Mean	87.32	87.08	87.39	87.59	96.78	96.86	96.50	96.60
Std	0.14	0.63	0.78	0.44	0.21	0.64	0.60	0.54
	AUC=0.913				AUC=0.976			

TABLE VI
RESULTS ON TUSZ DATASET FOR SEIZURE CLASSIFICATION

Class	LSTM				LSTM-att			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
GNSZ	96.00	96.00	84.56	89.91	98.48	98.48	98.57	98.52
FNSZ	83.01	83.01	93.14	87.78	98.11	98.11	98.86	98.48
CPSZ	83.78	83.78	94.57	88.85	98.95	98.95	98.38	98.66
ABSZ	81.25	81.25	87.00	84.14	93.75	93.75	100.00	96.77
SPSZ	84.85	84.85	38.36	52.83	97.30	97.30	98.63	97.96
TCSZ	100.00	100.00	53.85	70.00	98.11	98.11	100.00	99.05
TNSZ	80.21	80.21	87.64	83.76	100.00	100.00	87.50	93.33
MYSZ	100.00	100.00	43.42	60.55	95.83	95.83	90.79	93.24
Mean	88.32	88.32	72.14	77.55	98.41	97.86	96.02	96.87
Std	0.26	0.26	1.30	0.96	0.18	0.90	1.06	0.78

calculated weights generated by the attention layer produced higher accuracy compared to the basic LSTM model. For instance, these improvements show that the channel weights representing their contribution scores compliment the learnt features in better discriminating seizure classes.

In comparison with state-of-the-art methods as illustrated in Table VII, our LSTM-att is the first work to consider seizure detection and type classification on TUSZ dataset using the channel-wise attention mechanism and LSTM model directly fed with EEG raw data. For seizure detection, our proposed model improved the AUC scores and accuracy with 2.68% and 16.28% compared to CNN fed raw data [12]. In the literature, there is no work for seizure type classification on TUSZ dataset using raw EEG data. Consequently, a comparison with feature-based methods [9] [11] is handled, where our model greatly outperforms them.

B. Evaluation on CHB-MIT for Ictal vs Non-ictal classification

Seizure detection problem consists in ictal vs non-ictal classification task. This part is validated on data from the CHB-MIT dataset. The latter does not allow us to validate the model for seizure type classification due to the lack of seizure type labeled data. Only the start and the end of each onset is indicated for whole data in CHB-MIT, making it only usable for seizure detection or for seizure prediction as done in our previous paper [37].

According to Table VIII, the LSTM model achieves an accuracy of $88.40 \pm 1.31\%$, while our LSTM-att model reaches $96.48 \pm 1.16\%$. In terms of AUC and F1-score, the LSTM-att model reached 97.60% and 96.50% with an improvement

of approximately 6% and 10%, respectively. This can be explained by the fact that not all channels contribute equally to the decision of the presence or not of the seizure. The attention mechanism endows the LSTM model with a capability of weighting the channels such that they contribute differently and individually in each EEG sample to make the final decision.

To further understand the LSTM-att behavior, the confusion matrix is plotted as shown in Fig. 4. Actually, the model misclassified 2.75% of the seizure samples as non-ictal and 4.28% of the non-ictal samples as ictal. In general, the proposed model is able to classify 97.25% correctly of ictal cases and 95.72% of non-ictal cases. The achieved results are encouraging.

Fig. 4. Confusion matrix of seizure detection for the LSTM-att model with CHB-MIT dataset

Table IX illustrates the comparison with raw data-based works for seizure detection. While bidirectional parsing of

TABLE VII
COMPARISON WITH STATE-OF-THE-ART METHODS VALIDATED ON TUSZ DATASET

Ref	Year	Input	Classifier	Results		
				F1-score	AUC	Accuracy
Seizure detection						
[12]	2020	Raw EEG	CNN	—	94.92	80.50
Ours	2021	Raw EEG	LSTM-att	96.60	97.60	96.78
Seizure type classification						
[9]	2019	EEG spectrograms	CNN	—	—	84.06
[10]	2019	Intensity values of STFT	Neural Memory Network	94.50	—	97.40
[11]	2020	EEG spectrograms	Extreme Gradient Boosting	70.70	—	—
Ours	2021	Raw EEG	LSTM-att	96.87	—	98.41

TABLE VIII
RESULTS ON CHB-MIT DATASET

Class	LSTM				LSTM-att			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
Ictal	87.47	88.00	84.15	86.03	97.25	97.25	95.88	96.56
Non-ictal	89.33	85.24	89.66	87.39	95.72	95.72	97.65	96.67
Mean	88.40	86.50	87.35	86.44	96.48	96.88	96.28	96.50
Std	1.31	0.86	0.78	0.25	1.16	0.40	0.45	0.17
	AUC=0.916				AUC=0.976			

TABLE IX
COMPARISON WITH STATE-OF-THE-ART METHODS ON SEIZURE DETECTION WITH CHB-MIT DATASET

Ref	Year	Input	Classifier	Results		
				F1-score	AUC	ACC
[14]	2019	Raw EEG	FusionAtt	89.53	96.22	97.01
[24]	2021	Raw EEG	BiLSTM-att	84.15	84.15	91.51
Ours	2021	Raw EEG	LSTM-att	96.50	97.60	96.48

EEG signals tends to collect richer information, our LSTM-att model outperforms the BiLSTM-att model [24] with an improvement of 12.35%, 12.45% and 4.97% for F1-score, AUC, and accuracy metrics, respectively. Another work known as FusionAtt achieves similar results on AUC and accuracy metrics compared to our LSTM-att model, but it degrades in term of F1-score with a percentage of 6.97%.

VI. EXPLAINABILITY ANALYSIS

We conceptualized our attention mechanism to recognize different brain region signals and to produce various weights across channels. A single patient may experience seizures in different types from various brain regions. Accordingly, it is more reasonable to adaptively calculate channel weights in our attention mechanism. In our method, a kernel matrix and a bias matrix are trainable parameters, which undergo transformations by combining them with data segments. The transformation outputs represent the segment attention weights. If a channel weight is close to 0, it indicates that the corresponding signal characteristics are comparably weak to characterize a seizure type. This does not entail a lack of contribution of the corresponding channel to this seizure type. EEG signal manifestations vary between the seizure-free segment and the

onset segment according to the brain region contribution. In our seizure detection experiments, we observed that channels having great differences between ictal and non-ictal signals were assigned rather large weights

An example of attention weights of 18 channels for a set of seizure segments is shown in Fig. 5. The channels P7, C3, P3, FP2, F4, C4, O2, F8 and Cz have the large weights compared to other channels.

Since epileptic seizures have a patient-dependent characteristics, the plot of randomly selected ictal can't provide a relevant interpretation about the channels contribution on seizure detection for each patient. For this purpose, the model should be trained with patient-related data separately in order to plot and interpret the attention weights learnt by the model. Based on the results shown in Fig.5, we can only deduce that for the the randomly selected set of seizures, the aforementioned channels having the highest scores contributes the most to detecting ictal vs non-ictal segments.

Fig. 5. Mean attention weights on channels for a set of seizure samples from CHB-MIT

By analysing several Heatmaps related to seizure type classification, we observed that for different timestamps, the

Fig. 6. Left: Visualization of a multi-channel signal containing a Focal Non-specific seizure from TUSZ, Right: Correspondent Heatmap of calculated weights

Fig. 7. Position of high-weighted channels: Fz, Cz, Pz, (O1,O2) and (P5,P6) according to the 10-20 electrode system placement.

attention scores learnt by the attention mechanism are different. Specifically, when the seizure is generalized or begin as a focal and ends generalized, the distribution of attention scores are relatively uniform. This is because no such ictal pattern related to the seizure type is found within the whole channel views and hence the attention scores make even contribution to the seizure type classification. For some seizure types, we remarked that the attentional representations have the same view, then they depend on the seizure type.

According to the session-wise description attached to the data, there are a significant correlation between the neurological comments and our results. The larger attention score means the probability of seizure onset on this area is higher. In summary, the case study indicates that we can learn accurate attention scores with interpretable representations by our channel-wise Attention-based model, which not only improve the detection performance, but also identify the influential clinical concepts of seizure onset in healthcare. In our experiments, it was observed that relatively large weights were assigned to channels that contribute to characterize a specific seizure type. For this example of a Focal Non-Specific seizure, the neuro-

logist reported a Nonconvulsive status epilepticus in a patient with a drug resistant epilepsy. The EEG plotted on the top of the Fig. 6 demonstrates continuous 3 Hz bilateral temporo-occipital seizure activities, which coincided with the attention weights of 21 channels for correspondent seizure segments. The channels (O1,O2) and (T5,T6) have the large weights compared to other channels at the seizure start. The comments of the neurologist when reading the correspondent EEG signal correlate with our findings. Regarding a reported temporo-occipital origin, we demonstrated that the high weights are assigned to channels in this brain location as shown in the Fig. 7.

Since there is no specific lobe that includes the median line channels (Fz,Cz and Pz), these ones are almost affected by the seizure, which explains the implication of these channels in the seizure type classification. As shown in the Heatmap of Fig. 7, a medium level of scores is attributed to Fz,Cz and Pz.

VII. CONCLUSION

Classification of epileptic seizures has been a challenge for neurologists diagnosing epilepsy, prescribing treatment and arriving to a prognosis. The automated seizure classification method proposed in this paper can assist clinical professionals in diagnosing the disease, reducing time and potentially improves accuracy and reliability.

This paper proposes a novel channel-wise attention-based deep LSTM model which demonstrates the capability of the attention layer in enhancing the classification performance. An explainability analysis of our model showed that a high correlation exists between neurological interpretation and the reading of the Heatmaps of the learnt features. The LSTM-attention model achieves a significant improvement in classification accuracy up to 98.41% on the TUSZ dataset for 8 types of epileptic seizures and 96.78% for a binary classification (ictal vs non-ictal). The model also generalizes well across different minority seizure types.

Future works can focus on improving the performance of this study for the minor classes, by using multi-modal data, primarily EEG videos.

ACKNOWLEDGMENT

The research leading to these results has received funding from the Ministry of Higher Education and Scientific Research of Tunisia under the grant agreement number LR11ES48.

REFERENCES

- [1] M. M. Goldenberg, "Overview of drugs used for epilepsy and seizures: etiology, diagnosis, and treatment," *Pharmacy and Therapeutics*, vol. 35, no. 7, p. 392, 2010.
- [2] C. Panayiotopoulos, "Clinical aspects of the diagnosis of epileptic seizures and epileptic syndromes," in *The epilepsies: Seizures, syndromes and management*. Bladon Medical Publishing, 2005.
- [3] —, "Optimal use of the eeg in the diagnosis and management of epilepsies," in *The epilepsies: seizures, syndromes and management*. Bladon Medical Publishing, 2005.
- [4] (2010) Chb-mit scalp eeg database. [Online]. Available: <http://www.physionet.org/pn6/chbmit/>
- [5] (2020) Temple university hospital (tuh) seizure corpus dataset. [Online]. Available: https://www.isip.piconepress.com/projects/tuh_eeg/html/downloads.shtml
- [6] Y. Yuan, G. Xun, F. Ma, Q. Suo, H. Xue, K. Jia, and A. Zhang, "A novel channel-aware attention framework for multi-channel eeg seizure detection via multi-view deep learning," in *Biomedical & Health Informatics (BHI), 2018 IEEE EMBS International Conference on*. IEEE, 2018, pp. 206–209.
- [7] I. R. D. Saputro, N. D. Maryati, S. R. Solihati, I. Wijayanto, S. Hadiyoso, and R. Patmasari, "Seizure type classification on eeg signal using support vector machine," in *Journal of Physics: Conference Series*, vol. 1201, no. 1. IOP Publishing, 2019, p. 012065.
- [8] S. Roy, U. Asif, J. Tang, and S. Harrer, "Machine learning for seizure type classification: setting the benchmark," *arXiv preprint arXiv:1902.01012*, 2019.
- [9] D. Ahmedt-Aristizabal, T. Fernando, S. Denman, L. Petersson, M. J. Aburn, and C. Fookes, "Neural memory networks for robust classification of seizure type," *arXiv preprint arXiv:1912.04968*, 2019.
- [10] T. Liu, N. D. Truong, A. Nikpour, L. Zhou, and O. Kavehei, "Epileptic seizure classification with symmetric and hybrid bilinear models," *IEEE journal of biomedical and health informatics*, vol. 24, no. 10, pp. 2844–2851, 2020.
- [11] U. Asif, S. Roy, J. Tang, and S. Harrer, "SeizureNet: Multi-spectral deep feature learning for seizure type classification," in *Machine Learning in Clinical Neuroimaging and Radiogenomics in Neuro-oncology*. Springer, 2020, pp. 77–87.
- [12] X. Zhang, L. Yao, M. Dong, Z. Liu, Y. Zhang, and Y. Li, "Adversarial representation learning for robust patient-independent epileptic seizure detection," *IEEE Journal of Biomedical and Health Informatics*, 2020.
- [13] F.-G. Tang, Y. Liu, Y. Li, and Z.-W. Peng, "A unified multi-level spectral-temporal feature learning framework for patient-specific seizure onset detection in eeg signals," *Knowledge-Based Systems*, vol. 205, p. 106152, 2020.
- [14] Y. Yuan and K. Jia, "Fusionatt: deep fusional attention networks for multi-channel biomedical signals," *Sensors*, vol. 19, no. 11, p. 2429, 2019.
- [15] Y. Yuan, G. Xun, K. Jia, and A. Zhang, "A multi-view deep learning framework for eeg seizure detection," *IEEE journal of biomedical and health informatics*, vol. 23, no. 1, pp. 83–94, 2018.
- [16] N. Sriraam, Y. Temel, S. V. Rao, P. L. Kubben *et al.*, "A convolutional neural network based framework for classification of seizure types," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 2547–2550.
- [17] G. Cisotto, A. Zanga, J. Chlebus, I. Zoppis, S. Manzoni, and U. Markowska-Kaczmar, "Comparison of attention-based deep learning models for eeg classification," *arXiv preprint arXiv:2012.01074*, 2020.
- [18] D. O. Nahmias, E. F. Civillico, and K. L. Kontson, "Deep learning and feature based medication classifications from eeg in a large clinical data set," *Scientific Reports*, vol. 10, no. 1, pp. 1–11, 2020.
- [19] Y. Li, W. Zheng, L. Wang, Y. Zong, and Z. Cui, "From regional to global brain: A novel hierarchical spatial-temporal neural network model for eeg emotion recognition," *IEEE Transactions on Affective Computing, to be published*, 2019, doi:10.1109/TAFFC.2019.2922912.
- [20] R. Fourati, B. Ammar, J. Sanchez-Medina, and A. M. Alimi, "Unsupervised learning in reservoir computing for eeg-based emotion recognition," *IEEE Transactions on Affective Computing, to be published*, 2020, doi:10.1109/TAFFC.2020.2982143.
- [21] R. Fourati, B. Ammar, Y. Jin, and A. M. Alimi, "Eeg feature learning with intrinsic plasticity based deep echo state network," in *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020, pp. 1–8.
- [22] R. Fourati, B. Ammar, C. Aouiti, J. Sanchez-Medina, and A. M. Alimi, "Optimized echo state network with intrinsic plasticity for eeg-based emotion recognition," in *International Conference on Neural Information Processing*. Springer, 2017, pp. 718–727.
- [23] A. Baghdadi, Y. Aribi, R. Fourati, N. Halouani, P. Siarry, and A. Alimi, "Psychological stimulation for anxious states detection based on eeg-related features," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–15, 2020.
- [24] X. Yao, X. Li, Q. Ye, Y. Huang, Q. Cheng, and G.-Q. Zhang, "A robust deep learning approach for automatic classification of seizures against non-seizures," *Biomedical Signal Processing and Control*, vol. 64, p. 102215, 2021.
- [25] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (eeg) classification tasks: a review," *Journal of neural engineering*, vol. 16, no. 3, p. 031001, 2019.
- [26] P. Kaushik, A. Gupta, P. P. Roy, and D. P. Dogra, "Eeg-based age and gender prediction using deep blstm-lstm network model," *IEEE Sensors Journal*, vol. 19, no. 7, pp. 2634–2641, 2018.
- [27] K. M. Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, "A long short-term memory deep learning network for the prediction of epileptic seizures using eeg signals," *Computers in biology and medicine*, vol. 99, pp. 24–37, 2018.
- [28] M. U. Abbasi, A. Rashad, A. Basalamah, and M. Tariq, "Detection of epilepsy seizures in neo-natal eeg using lstm architecture," *IEEE Access*, vol. 7, pp. 179 074–179 085, 2019.
- [29] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, Y. Bengio and Y. LeCun, Eds., 2015.
- [30] L. Chen, H. Zhang, J. Xiao, L. Nie, J. Shao, W. Liu, and T. Chua, "Scan-n: Spatial and channel-wise attention in convolutional networks for image captioning," pp. 6298–6306, 2017.
- [31] A. Temko, G. Lightbody, E. M. Thomas, G. B. Boylan, and W. Marnane, "Instantaneous measure of eeg channel importance for improved patient-adaptive neonatal seizure detection," *IEEE transactions on biomedical engineering*, vol. 59, no. 3, pp. 717–727, 2011.
- [32] H. Hu, Q. Li, Y. Zhao, and Y. Zhang, "Parallel deep learning algorithms with hybrid attention mechanism for image segmentation of lung tumors," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2880–2889, 2020.
- [33] H. Eom, D. Lee, S. Han, Y. S. Hariyani, Y. Lim, I. Sohn, K. Park, and C. Park, "End-to-end deep learning architecture for continuous blood pressure estimation using attention mechanism," *Sensors*, vol. 20, no. 8, p. 2338, 2020.
- [34] Y. Miao, M. Gowayyed, and F. Metze, "Eesen: End-to-end speech recognition using deep rnn models and wfst-based decoding," in *2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*. IEEE, 2015, pp. 167–174.
- [35] W. Yin, K. Kann, M. Yu, and H. Schütze, "Comparative study of cnn and rnn for natural language processing," *arXiv preprint arXiv:1702.01923*, 2017.
- [36] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [37] A. Baghdadi, R. Fourati, Y. Aribi, P. Siarry, and A. M. Alimi, "Robust feature learning method for epileptic seizures prediction based on long-term eeg signals," in *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020, pp. 1–7.