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

Asma Baghdadi1,2,3 · Rahma Fourati2,4 · Yassine Aribi5 · Sawsen Daoud6,7 · Mariem Dammak6,7 · Chokri Mhiri6,7 · Habib Chabchoub8 · Patrick Siarry3,9 · Adel Alimi2,10

Received: 11 January 2022 / Accepted: 4 April 2023 / Published online: 16 May 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

Epilepsy affects almost 1% of the worldwide population. An early diagnosis of seizure types is a crucial patient-dependent step for the treatment selection process. The selection of the proper treatment relies on the correct identification of the seizure 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.

Keywords Seizure type classification · EEG feature learning · Patient-independent · Attention mechanism · Deep LSTM · Explainable model

Asma Baghdadi
asma.baghdadi@ieee.org

Rahma Fourati
rahma.fourati@ieee.org

Yassine Aribi
yaliraibi@tu.edu.sa

1 Esprit School of Business, Computer Science and Applied Mathematics Department, Tunis, Tunisia
2 Research Groups in Intelligent Machines, National Engineering School of Sfax (ENIS), University of Sfax, BP 1173, 3038 Sfax, Tunisia
3 Laboratoire Images, Signaux et Systèmes Intelligents, Université Paris Est Créteil, 120 Rue Paul Armangot, 94400 Vitry sur Seine, France
4 Faculty of Law, Economics and Management Sciences of Jendouba (FSIEGJ), University of Jendouba, Jendouba, Tunisia
5 Sciences and Technology Department, University College of Ranyah Taif University, Taif, Saudi Arabia
6 Department of Neurology, Habib Bourguiba University Hospital, Sfax, Tunisia
7 Neuroscience Laboratory, Sfax university, LR-12-SP-19, Sfax, Tunisia
8 College of Business, Al Ain University, Abu Dhabi, United Arab Emirates
9 LiSSi, Université de Paris 12, 94010 Créteil, France
10 Department of Electrical and Electronic Engineering Science, Faculty of Engineering and the Built Environment, University of Johannesburg, Johannesburg, South Africa

Springer
1 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 seizure type. Seizure types are classified by the International League Against Epilepsy based on the symptoms manifestation. 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 (Goldenberg 2010).

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 (Panayiotopoulos 2005a). Recently, several researches have demonstrated that for some cases an experienced neurologist can have troubles in recognizing the correct type of seizure (Panayiotopoulos 2005b). 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.

When working with EEG signals in general, there are two types of experiments: patient-dependent and patient independent. The formal consists of designing and training a model per patient. So, we will have a number of models equals to the number of patients. Each model is trained and tested with EEG signals of the same patient.

In contrast, only one model is conceived in patient-independent classification tasks. It is trained with data from different patients and then tested with data from different patients as well. Due to the high variability of EEG signals from one person to another, the patient-independent classification remains challenging in comparison to the patient-dependent one.

In fact, most of the work done in EEG-based seizure detection or classification type is in the patient-dependent scheme such as Gao et al. (2022). It is not surprising, since the difficulty to design a model that generalizes well, prevents some researchers from achieving good performance, thereby choosing to proceed in this type of experiment. For instance, recent works on CHB-MIT dataset that involves the detection of seizure and normal states in patient-dependent scheme achieved 96.38% (Shen et al. 2022) and even reached 96.69% (Cimr et al. 2022), while another work (Jiang et al. 2023) on the same dataset in patient-independent scheme achieved only 83.36%.

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 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 three 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 endorsed with an explainable case study where a correlation between the seizure type and the localization of the highest active channels is established. (3) An extensive validation of the proposed method on seizure detection and seizure type classification.

As already mentioned above, we address two major classification problems in our work, which are seizure detection and seizure types classification. The first aims to classify two epileptic states for patients suffering from epilepsy. Ictal state, which is the period of time when the onset occurs, and non-ictal states, which are periods free of onsets. This task was evaluated on two datasets: CHB-MIT and TUHZ since both include ictal and non-ictal periods in their recording.

The second one aims to classify eight types of seizures, and this by classifying only ictal state of eight seizure types available in the TUSZ recording. Since CHB-MIT dataset don’t offer any information about each seizure type, we couldn’t use it for the evaluation phase.
A channel-wise attention-based representation learning method for epileptic seizure detection…

2 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 (Shoeb 2009) and TUSZ (Harati et al. 2014) are widely used free benchmarks. We first describe these datasets and next overview the related work on them.

2.1 Existing epileptic benchmarks

The scalp long-term EEG dataset CHB-MIT (Shoeb 2009) contains data gathered from 23 pediatric patients admitted in the Boston Children’s Hospital suffering from intractable seizures. EEG Recordings consist of 983 h 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 (Table 1). 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 (Shah et al. 2018) 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 2 indicating the total number of events per type. The used version consists of EEG recordings resampled to 250 Hz.

2.2 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 Yuan et al. (2018b) and providing a model interpretability (Yuan and Jia 2019). Most of the recent seizure detection and type classification methods 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 (Zhang et al. 2020) 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 (Lerche et al. 2013).

Yuan et al. (2018b, 2018a) 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 a 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 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 Yuan et al. (2018a) 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 1 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 classes        | 8    | 2       |

Table 2 Details of all eight types of seizure included in the TUSZ dataset

| Seizure code | Seizure type        | Total recording (variable duration) | Nb. of samples |
|--------------|---------------------|--------------------------------------|----------------|
| FNSZ         | Focal non-specific  | 1536                                 | 17,021         |
| 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 s segments can change the data distribution over classes.
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 datasets with a variable amount of data across classes. To meet the aims of seizure type classification based on deep learning approaches, the authors in Asif et al. (2020) 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 Asif et al. (2020) consists of 45.94 M trainable parameters, while the hybrid bilinear structure proposed in the study of Liu et al. (2020) contains only 1.2 M trainable parameters. In the latter, myoclonic seizures was discarded from the classification task due to the insufficient number of samples on this class. Both Ahmedt-Aristizabal et al. (2019) and Sriraam et al. (2019) proposed solutions for a multi-class classification problem evaluated on the TUSZ dataset, achieving 84.06% and 94.05%, respectively. As observable from Table 3, the bilinear model proposed by Liu et al. (2020) achieved better performance on the 8 seizure type classification problem. Machine learning methods, including the k-nearest neighbor (KNN) proposed in Roy et al. (2019) and the support vector machine (SVM) proposed in Saputro et al. (2019), exhibit acceptable performance (90.70% and 91.40%), nevertheless, achieved through undesirable extensive feature engineering.

| References | Method | Dataset | Data | (#) Seizure types | Performance |
|------------|--------|---------|------|-------------------|-------------|
| Yuan et al. (2018b) | $\text{Attloc}$ and $\text{Attglob}$ attention mechanisms | CHB-MIT | EEG Spectrograms | Seizure detection: Ictal Vs Interictal | Acc: 96.61 |
| Saputro et al. (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% |
| Roy et al. (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% |
| Ahmedt-Aristizabal et al. (2019) | CNN | TUSZ | multi-spectral feature from EEG spectograms | (8) focal non-specific, generalized non-specific, simple-partial, complex-partial, absence, tonic, tonic-clonic, myoclonic | Acc: 84.06% |
| Liu et al. (2020) | Neural memory network | TUSZ | Intensity values of STFT | (8) focal non-specific, generalized non-specific, simple-partial, complex-partial, absence, tonic, tonic-clonic, myoclonic | F1: 94.50% |
| | Hybrid bilinear | TUSZ | As above | (8) focal non-specific, generalized non-specific, simple-partial, complex-partial, absence, tonic, tonic-clonic, myoclonic | F1: 97.40% |
| Asif et al. (2020) | Extreme gradient boosting | TUSZ | multi-spectral feature from EEG spectograms | (8) focal non-specific, generalized non-specific, simple-partial, complex-partial, absence, tonic, tonic-clonic, myoclonic | F1: 70.70% |
| | Ensemble CNN’s | TUSZ | multi-spectral feature from EEG spectograms | (8) focal non-specific, generalized non-specific, simple-partial, complex-partial, absence, tonic, tonic-clonic, myoclonic | F1: 98.40% |
| Zhang et al. (2020) | Adversarial learning using CNNs | TUSZ | Raw EEG | Seizure detection: Normal Vs Seizure | Acc: 80.50 |
| Tang et al. (2020) | Unified multi-level spectral–temporal feature extraction | CHB-MIT | PCA-CSP and MMSE | Onset detection | Acc: 97.80 |
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 Cisotto et al. (2020), the author has discussed the difference between 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 (Harati et al. 2014). The comparison invoked in the work (Nahmias et al. 2020) shows that a small difference exists dated on TUSZ dataset since it is the largest available dataset as between the two anticonvulsants. The method was validated on TUSZ dataset since it is the largest available dataset (Harati et al. 2014). The comparison invoked in the work (Nahmias et al. 2020) shows that a small difference exists

The 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 it’s 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 \( x = (x_1, x_2, \ldots, x_T) \) to the output sequence \( y = (y_1, y_2, \ldots, y_T) \) by figuring out the network unit activations using the following equations:

\[
\begin{align*}
i_t & = \sigma(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_i) \\
f_t & = \sigma(W_{fx}x_t + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f) \\
c_t & = f_t \odot c_{t-1} + i_t \odot g(W_{cx}x_t + W_{cm}m_{t-1} + b_c) \\
o_t & = \sigma(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_{t-1} + b_o) \\
m_t & = o_t \odot hc_t \\
y_t & = \phi(W_{ym}m_t + b_y)
\end{align*}
\]

In the above equations, \( W \) represents the weight and \( W_{ic} \) is the maximum weight of the input gate to the input. \( W_{fc} \) 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 compromised mostly features extracted from EEG signals. However, the signal itself and EEG images (spectrograms) were also used (Craik et al. 2019).

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 (Kaushik et al. 2018; Tsiouris et al. 2018; Abbasi et al. 2019). These studies reported an accuracy rate of 50.00% ± 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.

### 3.2 Attention mechanism for multi-channel epileptic signals

Attention mechanism was first introduced by Bahdanau et al. (2015) 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 Chen et al. (2017), 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 (Temko et al. 2011). Thus, an attention mechanism is introduced for channel importance learning and to pay different attention to various brain lobes. As above mentioned, the attention mechanism allows modeling of dependencies among EEG channels (Zhang et al. 2020) and has shown success in some research topics (Cisotto et al. 2020; Hu et al. 2020; Eom et al. 2020).

Meanwhile, the excellent temporal feature learning ability of recurrent neural networks (RNNs) has been extensively used in research areas such as speech recognition (Miao et al. 2015), language modeling (Yin et al. 2017), diseases prediction (Tsiouris et al. 2018) 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 attention mechanism takes into consideration the information of all channels and assigns weights to different channels based on their importance. The attention applied in our model is Soft, since weights are generated for each $x_k$ and then used to produce a weighted temporal features by multiplying weights and the LSTM output. This mechanism will allow to explain the contribution of each channel to the final decision. Consider that, $X = [X_1, X_2, \ldots, X_n]$ represents EEG samples, and $X_i = [x_1, x_2, \ldots, x_k]$ 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 after normalisation. Notice that the applied normalization happens across the channels within each trial, rather than across different trials of a given patient. 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

$$X = [X_1, X_2, \ldots, X_n]$$

$$X_i = [x_1, x_2, \ldots, x_k]$$

$$x_k$$ 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 after normalisation. Notice that the applied normalization happens across the channels within each trial, rather than across different trials of a given patient. 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:

$$Y_1 = f_{nor}(X_0)$$  \hspace{1cm} (7)

$$Y_2 = w_{al} \times Y_1 + b_{al}$$  \hspace{1cm} (8)

$$Y_3 = f_{tens}(Y_2)$$  \hspace{1cm} (9)

$$Y_4 = \sigma(Y_3)$$  \hspace{1cm} (10)

Here, $X_0$ denotes an input of size $(N_{\text{samples}}, N_{\text{timesteps}}, N_{\text{channels}})$. Symbols ($N_{\text{samples}}, N_{\text{timesteps}}$ and $N_{\text{channels}}$) 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 $(N_{\text{channels}}, N_{\text{channels}})$, a bias $b_{al}$ of $(N_{\text{samples}}, N_{\text{timesteps}})$, and $Y_2$ have the same size as $Y_1$. A symbol $\sigma(.)$ represents a nonlinear activation function, which transforms the importance of channels to a probability distribution, like softmax(.) and sigmoid(.). $Y_3$ is a tensorized matrix of $Y_2$.

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

$$\text{attscores} = Y_3 = [att_1; att_2; \ldots; att_k]$$  \hspace{1cm} (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 (Hochreiter and Schmidhuber 1997). The predicted seizure class $Y_5$ is related to the last dense layer (D2) and the learnt attention scores:

$$Y_5 = \text{Dense}_\text{out} \odot \text{attscores}$$  \hspace{1cm} (12)

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

### 4.2 Training and evaluation

In order to extensively evaluate the proposed model performance, a stratified five-fold cross validation is used. 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 2). The minor class, absence seizures, has 6 min 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.

| 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 prob | [0.0, 0.2, 0.5] | None |
| Dense activation | ['softmax', 'relu', 'tanh', 'sigmoid', 'linear'] | Softmax |
| Optimizer | None       | Adam       |
| Epochs | [20, 50, 100] | 100        |
| Batch size | [10, 20, 50] | 20         |
4.3 Hyper-parameters fitting

In our work, a grid search for all parameters was adopted as depicted in Table 4. 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 $[N_{\text{timesteps}}; N_{\text{channels}}]$. The number of units on the first and second LSTM layers was tuned to be between $[0.0, 0.2, 0.5]$. The two FC layers have $N_{\text{features}}$ and $N_{\text{classes}}$ hidden neurons respectively, and several activation functions were tested ['softmax', 'relu', 'tanh', 'sigmoid', 'linear']. The attention layer has $N_{\text{channels}}$ 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 960 M. The average Accuracy, Precision, Recall and F1-scores were reported. AUC scores are reported only for seizure detection on CHB-MIT dataset.

4.4 Optimal architecture details

The model architecture contains 3 main blocks where each one has its specificity with respect to the input and output shapes and their parameters. Table 5 depicts the sequencing of data processing between all the three blocks. Attention block along with LSTM take the same input. Since both attention and LSTM blocks are applied directly to the raw EEG data, the input for these blocks is a time series with a shape of (1250, 22).

In the attention block, two vectors are initialized randomly (W and b) based on the input. In order to calculate the weighted channel attention vector, Weights are normalized using the Softmax function to obtain the contribution probability of each channel in a 2D tensor of a shape equal to (1, 22). The LSTM block is composed of a first LSTM layer with the attribute return_sequences = True to get the activation of each cell. These activations are then fed to a second LSTM layer with the attribute return_sequences = False to get the activation of the last output cell. Then it is compressed by two Dense layers to obtain 1D tensor with 176 elements. In order to value the relevance of each channel data, a Reshape layer is added at the bottom of the LSTM block to obtain a 2D tensor with a shape of (8, 22).

The Attention block output and the LSTM block output are then passed through a multiplication layer. In our work, it is implemented as a customized layer.

Finally, the classification block gets the attention-based LSTM representation. It is composed on one dense layer followed by a sofmax layer with 8 nodes.

5 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 Sect. 3.2 and the basic deep LSTM model Sect. 3.1. 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.

The classification problems are:

- Seizure type classification on TUSZ: aims to classify 8 types of seizures. In this task, we only used Ictal segments related to the 8 seizure types existing in the TUSZ dataset
- Seizure detection on TUSZ: aims to classify two states that are Ictal (during the seizure) and Non-ictal (before and after the seizure) using signals from the TUSZ dataset
- Seizure detection on CHB-MIT: also aims to classify two states that are Ictal and Non-ictal but using signals from the CHB-MIT dataset.

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.

5.1 Evaluation on TUSZ for seizure detection and type classification

In seizure detection problem, the LSTM-att achieves an accuracy of $96.78\pm0.21\%$ which outperforms the basic LSTM model of approximately $9.46\%$ for the TUSZ dataset as illustrated in Table 6. 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
Table 5 Details of the optimal model including blocks parameters

| Variable/layer | Input shape   | Parameters   | Layer   | Input shape | Units | Parameters   |
|----------------|---------------|--------------|---------|-------------|-------|--------------|
| Input          | (1250, 22)    | –            | Input   | (1250, 22)  | –     | –            |
| W              | (1, 1250)     | mean=0, stddev=0.1 | LSTM1   | (1250, 22)  | 250   | return_sequences=True, rate=0.0 |
| b              | (1, 22)       | mean=0, stddev=0.1 | Dropout | –           | –     | –            |
| Matmul (Input, W) | (1, 1250), (1250, 22) | –           | Dense   | –           | 300   | –            |
| Add (attention_W, b) | (1, 22), (1, 22) | –           | Dense   | –           | 176   | –            |
| Softmax        | (1, 22)       | –            | Reshape | 176         | –     | target_shape=(8, 22) |

Classification block

| Layer          | Units | Input shape |
|----------------|-------|-------------|
| Multiplication layer | –     | (8, 22), (22, 1) |
| Flatten        | –     | (8, 1)      |
| Dense          | 8     | (8)         |
| Softmax        | 8     | (8)         |
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 7.

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 is achieved for the Tonic seizure (TNSZ) class. However, the lowest

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

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

According to the aforementioned results, it was shown that the model trained using learnt feature in combination with 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 8, 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 (Zhang et al. 2020). 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 (Ahmedt-Aristizabal et al. 2019) (Asif et al. 2020) is handled, where our model greatly outperforms them.

5.2 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 (Baghdadi et al. 2020b).

According to Table 9, the LSTM model achieves an accuracy of 88.40 ± 1.31%, while our LSTM-att model reaches 96.48 ± 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 contributes 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.

Table 10 illustrates the comparison with raw data-based works for seizure detection. While bidirectional parsing of EEG signals tends to collect richer information, our LSTM-att model outperforms the BiLSTM-att model (Yao et al. 2021) 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%.

6 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.

6.1 Seizure detection on CHB-MIT

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 samples 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.
6.2 Seizures classification and detection on TUSZ

6.2.1 Seizure detection analysis

For the case of seizure detection task elaborated on the TUSZ dataset and as shown in Fig. 6, the right side of the frontal, temporal, parietal and occipital areas of the brain are more informative. The band-wise topographies show that the value of F4, C4, T7 and O2 channels are higher than others for all bands. The topographies for the O1 channel show that its activation is only related to high frequencies (Gamma).

The activity at the Pz channel appears for the Theta, Alpha and Gamma bands. Frontal (Fp2 and F4), temporal (T7 and T8) and parietal (P7, Pz and P8) areas are more activated and thus are more informative, which correlates with the attention weights illustrated by the Heatmap in Fig. 5.

Comparing the topographical maps for a FNSZ seizure from TUHZ dataset, the images in the Fig. 8 show that the activity for all frequency bands is a bilateral temporop-occipital, which implicates the Fz, Cz, T5, Pz, T6 and O2 channels. The activity of Fp1 and Fp2 channels are higher than other channels for the Gamma band compared with other bands, while our model does not assign largest weights for the aforementioned channels as shown in the Fig. 7. This fact was explained by the presence of muscular artifacts that enhance the activity around these channels only for the Gamma band, knowing that muscle artifacts occur around 50–60 Hz. However, it can be related to the power-line interference artefact that needs to be removed for more investigation of this hypothesis. A pre-processing step based on the application of a notch filter has proven its effectiveness as reported by the authors in Leske and Dalal (2019).

6.2.2 Seizure type classification analysis

By analysing several Heatmaps related to seizure type classification, we observed that for different timestamps, the 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 neurologist reported a Nonconvulsive status epilepticus in a patient with a drug resistant epilepsy. The EEG plotted on the top of the Fig. 7 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. 9.

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. 9, a medium level of scores is attributed to Fz, Cz and Pz.

The brain topography maps for different EEG frequency bands are depicted in Figs. 6 and 8. We have not applied the normalization to guaranty the visualization of the smallest brain activity variation.

7 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.

Acknowledgements 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 LR11IES48.

References

Abbasi MU, Rashad A, Basalamah A, Tariq M (2019) Detection of epilepsy seizures in neo-natal eeg using lstm architecture. IEEE Access 7:179074

Ahmedt-Aristirizabal D, Fernando T, Denman S, Peterssson L, Aburn MJ, Fookes C (2019) Neural memory networks for robust classification of seizure type. arXiv:1912.04968

Asif U, Roy S, Tang J, Harrer S (2020) Seizurenet: multi-spectral deep feature learning for seizure type classification. In: Machine learning in clinical neuroimaging and radiogenomics in neuro-encephalography

Baghdadi A, Fourati R, Aribi Y, Chaabouni S, Mhamdi M, Chabchoub H, Siarry P, Alimi A (2023) A channel-wise attention-based representation learning method for epileptic seizure detection and type classification. Preprint on TechRxiv, 2023. https://doi.org/10.36277/techrxiv.17087147.v1

Baghdadi A, Aribi Y, Fourati R, Halouani N, Siarry P, Alimi A (2020a) Psychological stimulation for anxious states detection based on eeg-related features. J Ambient Intell Humaniz Comput 1

Baghdadi A, Fourati R, Aribi Y, Siarry P, Alimi AM (2020b) Robust feature learning method for epileptic seizures prediction based on long-term eeg signals. In: 2020 international joint conference on neural networks (IJCNN), p 1. IEEE

Bahdanau D, Cho K, Bengio Y (2015) Neural machine translation by jointly learning to align and translate. In: Bengio Y, LeCun Y (eds) 3rd international conference on learning representations, ICLR 2015, San Diego, May 7–9, 2015, Conference Track Proceedings

Chen L, Zhang H, Xiao J, Nie L, Shao J, Liu W, Chua T (2017) Scann: spatial and channel-wise attention in convolutional networks for image captioning, p 6298

Cimr D, Fujita H, Tomaskova H, Cimler R, Selamat A (2022) Automatic seizure detection by convolutional neural networks with computational complexity analysis. Comput Methods Progr Biomed. https://doi.org/10.1016/j.cmpb.2022.107277

Cisotto G, Zanga A, Chlebus J, Zoppis I, Manzoni S, Markowska-Kaczmar U (2020) Comparison of attention-based deep learning models for eeg classification. arXiv:2012.01074

Craik A, He Y, Contreras-Vidal JI (2019) Deep learning for electroencephalogram (eeg) classification tasks: a review. J Neural Eng 16(3):031001

Eom H, Lee D, Han S, Hariyani YS, Lim Y, Sohn I, Park K, Park C (2020) End-to-end deep learning architecture for continuous blood pressure estimation using attention mechanism. Sensors 20(8):2338

Fourati R, Ammar B, Aouiti C, Sanchez-Medina J, Alimi AM (2017) Optimized echo state network with intrinsic plasticity for eeg-based emotion recognition. In: International conference on neural information processing, p 718. Springer

Fourati R, Ammar B, Jin Y, Alimi AM (2020a) Eeg feature learning with intrinsic plasticity based deep echo state network. In: 2020 international joint conference on neural networks (IJCNN), p 1. IEEE

Fourati R, Ammar B, Sanchez-Medina J, Alimi AM (2020b) Unsupervised learning in reservoir computing for eeg-based emotion recognition. IEEE Trans Affect Comput. https://doi.org/10.1109/TAFFC.2020.2982143

Gao X, Zhu Y, Yang Y, Zhang F, Zhou F, Tian X, Xu K, Chen Y (2022) A seizure detection method based on hypergraph features and machine learning. Biomed Signal Process Control 77:103769. https://doi.org/10.1016/j.bspc.2022.103769

Goldenberg MM (2010) Overview of drugs used for epilepsy and seizures: etiology, diagnosis, and treatment. Pharm Ther 35(7):392

Harati A, Lopez S, Obeid I, Picone J, Jacobson M, Tobochnik S (2014) The tuh eeg corpus: a big data resource for automated eeg interpretation. In: 2014 IEEE signal processing in medicine and biology symposium (SPMB), p 1. IEEE

Hocheirer S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735

Hu H, Li Q, Zhao Y, Zhang Y (2020) Parallel deep learning algorithms with hybrid attention mechanism for image segmentation of lung tumors. IEEE Trans Ind Inf 17(4):2880

Jiang L, He J, Pan H, Wu D, Jiang T, Liu J (2023) Seizure detection algorithm based on improved functional brain network structure feature extraction. Biomed Signal Process Control 79:104053. https://doi.org/10.1016/j.bspc.2022.104053

Kaushik P, Gupta A, Roy PP, Dogra DP (2018) Eeg-based age and gender prediction using deep lstm-lstm network model. IEEE Sens J 19(7):2634

Lerche H, Shah M, Beck H, Noebels J, Johnston D, Vincent A (2013) Ion channels in genetic and acquired forms of epilepsy. J Physiol 591(4):753

Leske S, Dalal SS (2019) Reducing power line noise in eeg and meg data via spectrum interpolation. Neuroimage 189:763

Li Y, Zheng W, Wang L, Zong Y, Cui Z (2019) From regional to global brain: a novel hierarchical spatial-temporal neural network model for eeg emotion recognition. IEEE Trans Affect Comput. https://doi.org/10.1109/TAFFC.2019.2922912

Liu T, Truong ND, Nikpour A, Zhou L, Kavehei O (2020) Epileptic seizure classification with symmetric and hybrid bilinear models. IEEE J Biomed Health Inform 24(10):2844

Miao Y, Gowayyed M, Metze F (2015) Eesenn: end-to-end speech recognition using deep rnn models and wst-based decoding. In: 2015 IEEE workshop on automatic speech recognition and understanding (ASRU), p 167. IEEE

Namias DO, Civlicio EF, Koutson KL (2020) Deep learning and feature-based medication classifications from eeg in a large clinical data set. Sci Rep 10(1):1

Panayiotopoulos C (2005a) Clinical aspects of the diagnosis of epileptic seizures and epileptic syndromes. In: The epilepsies: seizures, syndromes and management. Bladon Medical Publishing

Panayiotopoulos C (2005b) The tuh eeg corpus: a big data resource for automated seizure detection and treatment. PhD thesis, Massachusetts Institute of Technology.
Sriraam N, Temel Y, Rao SV, Kubben PL (2019) 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), p 2547. IEEE

Tang F-G, Liu Y, Li Y, Peng Z-W (2020) A unified multi-level spectral-temporal feature learning framework for patient-specific seizure onset detection in EEG signals. Knowl Based Syst 205:106152

Temko A, Lightbody G, Thomas EM, Boylan GB, Marnane W (2011) Instantaneous measure of EEG channel importance for improved patient-adaptive neonatal seizure detection. IEEE Trans Biomed Eng 59(3):717

Tsiouris KM, Pezoulas VC, Zervakis M, Konitsiotis S, Koutsouris DD, Fotiadis DI (2018) A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals. Comput Biol Med 99:24

Yao X, Li X, Ye Q, Huang Y, Cheng Q, Zhang G-Q (2021) A robust deep learning approach for automatic classification of seizures against non-seizures. Biomed Signal Process Control 64:102215

Yin W, Kann K, Yu M, Schütze H (2017) Comparative study of CNN and RNN for natural language processing. arXiv:1702.01923

Yuan Y, Jia K (2019) Fusionatt: deep fusional attention networks for multi-channel biomedical signals. Sensors 19(11):2429

Yuan Y, Xun G, Jia K, Zhang A (2018a) A multi-view deep learning framework for EEG seizure detection. IEEE J Biomed Health Inform 23(1):83

Yuan Y, Xun G, Ma F, Suo Q, Xue H, Jia K, Zhang A (2018b) A novel channel-aware attention framework for multi-channel EEG seizure detection via multi-view deep learning. In: 2018 IEEE EMBS international conference on biomedical & health informatics (BHI), p 206. IEEE

Zhang X, Yao L, Dong M, Liu Z, Zhang Y, Li Y (2020) Adversarial representation learning for robust patient-independent epileptic seizure detection. IEEE J Biomed Health Inform 24(10):2852–9

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.