

Evaluation of recurrent neural networks as epileptic seizure predictor

Luciano Bongiorni, Alexandre Balbinot *

Universidade Federal do Rio Grande do Sul -UFRGS, Graduate Program in Electrical Engineering - PPGEE, Brazil

ARTICLE INFO

Keywords:

Epilepsy
Categorical prediction
Recurrent neural networks (RNN)
LSTM
Electroencephalography(EEG)
Intelligent diagnosis

ABSTRACT

The occurrence of epileptic seizures is a problem that makes everyday life difficult for many people who suffer from this disorder, which may cause risks to the subjects and people around them. The possibility of predicting epileptic seizures could create a way to minimize risks associated with these attacks. Through computational analysis of Electroencephalography (EEG) signals, a variation of Recurrent Artificial Neural Network (RNN), of Long Short Term Memory (LSTM) type, was developed in an attempt to classify segments of EEG signals that occur before the onset of epileptic seizures, segments which evince the pre-ictal state. Using an EEG database of pediatric subjects with intractable seizures, groups of patients were created, then redundant EEG channels were selected and reduced by only 25%. Temporal and spectral features were extracted to improve the evaluation of data patterns related to seizures. Besides using the unprocessed data, two techniques were proposed, first using the superposition of inputs, then anticipating targets in the training data. A third technique was also proposed by joining the first two. Finally, two LSTM structures were created. The proposed Neural Network achieved high accuracies above 99%; however, the results showed that it was not possible to classify pre-ictal regions by using this combination of architecture and individuals. Furthermore, it was observed that the Neural Network was able to classify ictal regions with up to 61% sensitivity and 99% specificity, confirming the capacity of RNNs of LSTM type to assimilate temporal patterns in EEG data.

1. Introduction

Epilepsy is a brain disorder caused by abnormal activity, which is observable through neural electrical biosignals. Epilepsy can incur in epileptic seizures, events in which the epileptic individual feels sensory disturbances, commonly accompanied by strong pain and involuntary muscular contraction [1,2]. Due to the bioelectrical nature of seizures, Electroencephalography (EEG) is a direct means of analyzing it. Through the acquisition of EEG signals, useful information can be processed and used for monitoring individuals who suffer from this disorder.

There are several ways of predicting seizures, ranging from their own clinical analysis to EMG [3] and the monitoring of diverse electrical and bioelectrical variables [4]. Among all these techniques, the Artificial Neural Network (ANN) is an intelligent computational technique, usually fed and trained with EEG data with the final purpose of extracting significant features from the data and learning how to generalize it to new inputs of the same kind.

A special class of ANN is the Recurrent Neural Network (RNN), which holds an internal structure with a feedback loop. This kind of ANN is able to store information regarding previously analyzed inputs, establishing

relations which consider not only the current inputs, but also a sequence of past inputs. In this way the RNNs can make predictions considering sequences of data, making it suitable to analyze temporal signals. Due to the vanishing or exploding gradient seen in RNNs training when presented with long sequences [5], the Long Short Term Memory (LSTM) [6, 7] variation of RNNs was proposed in order to alternately forget or consider new data into the memory, making it a better structure to deal with long sequences of data, like EEG signals.

Through the employment of intelligent computation techniques, it is possible to minimize human error in epileptic signals analysis and implement an evaluation process of neurological signals, presenting generalization in the characterization of epileptic events. Thus, in case the algorithms here proposed are able to abstract the stochastic variation of cerebral signals [8], obtained through EEG, and extract features that sufficiently describe epileptic events, it would be possible to predict those events, allowing the own epileptic individual, relatives or professionals of the medical area to take appropriate actions with a certain antecedence to the epileptic event.

This work first presents the study of the data obtained from the Physionet database [9], following the selection and labeling process of

* Corresponding author.

E-mail address: alexandre.balbinot@ufrgs.br (A. Balbinot).

several cases within the database. Then clusterization of cases in groups is proposed in order to improve the training and evaluation of the ANN, alongside a channel reduction aimed to decrease the computational load of the algorithm. After the data sets are structured, feature extraction techniques are proposed, using temporal and spectral analysis like Standard Deviation and Short Time Fourier Transform (STFT). The pre-processing of data includes the proposal of two techniques to improve the ANN accuracy, being those of the data superposition technique and the anticipation of target data, both for the training phase of the algorithm. Finally, the LSTM network is developed, using two structures due to the techniques proposed, and its performance is analyzed by the training accuracy, and the sensitivity (true-positive rate) and specificity (true-negative rate) of the predictions.

Given the previous discussions presented in a summarized manner, this work focuses on the implementation and investigation of the efficiency of specific algorithm techniques of neural networks in the prediction of epileptic events. In a more specific fashion, the objectives are:

- analyze an EEG database of individuals with the occurrence of recurrent epileptic events;
- analyze the extraction of signal patterns presenting significant correlations with pre-ictal segments;
- develop an algorithm with LSTM ANN for early prediction of epileptic seizures;
- analyze the latency of the algorithm in determining epileptic events.

2. Methodology

Given the objective of developing an RNN for pre-ictal segments categorization, the network structure and feature extraction aim to capture the most significant mathematical patterns that can yield high classification accuracy. The proposal for this method is presented in Fig. 1.

The method starts by selecting and grouping patients based on the number of epileptic events presented in the recordings, then only EEG channels common to all recordings are maintained and the data entries containing epileptic events are added in a target column vector. The data pre-processing step starts with channel reduction to less than a quarter of the channels selected, followed by the extraction of 5 temporal and spectral features, just as data normalization and target encoding. The last proposed technique consists of the proposal of two data manipulation techniques - input signals superposition and anticipation of target values, both for training data. Finally, considering the two techniques proposed, two RNN structures are created to test the accuracy, sensitivity, and specificity in identifying pre-ictal regions.

For measuring the accuracy of this proposed method, the metrics used were ANN training accuracy, classification specificity, and classification sensitivity. The sensitivity, as shown in Equation (1), stands for the rate between true-positives and the sum of it with false-negative predictions, and the sensitivity stands for the rate between true-negative predictions and the sum of it with false-positive predictions, just as shown in Equation (2):

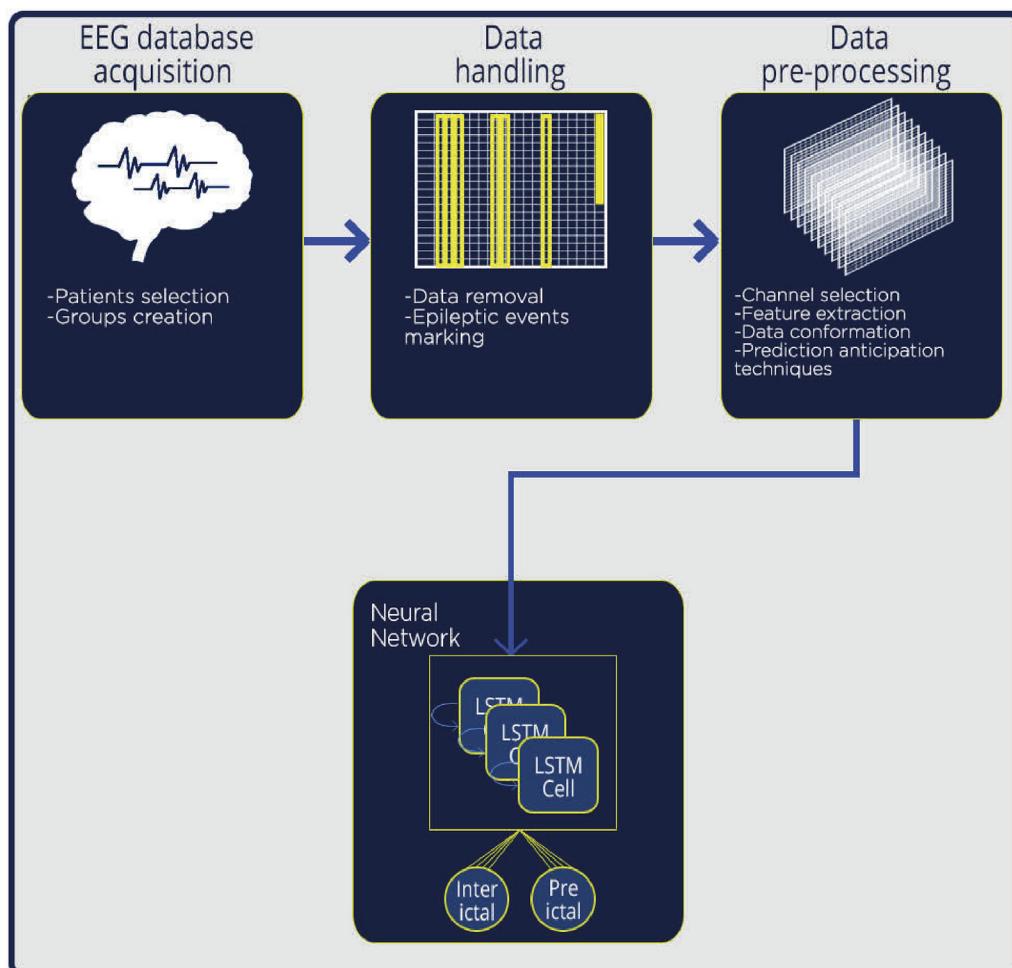


Fig. 1. Solution overview presented as a block diagram.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (1)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (2)$$

2.1. Data selection

Focusing on the analysis of pre-ictal signal pattern classification, instead of the acquisition of such signals, an open database was chosen for this work. The database selected was obtained from Ref. [9,10], which consists of scalp EEG data of 22 children with intractable epileptic seizures, gathered by a group of Massachusetts Institute of Technology and Children's Hospital Boston professionals. Each record was obtained using the 10–20 International System for electrodes, using 256 samples per second with a 16 bits ADC, being the values recorded in microvolts. These records were labeled using information about the time of occurrence of epileptic events, in such a way that each entry had a new column indicating if a seizure was occurring at the time of data capturing or not.

The recordings displayed a different series of channels, so the more frequent channels present in the recordings were selected, resulting in FP1-F7, F7-T7, T7-P7, P7-O1, FP1-F3, F3-C3, C3-P3, P3-O1, FZ-CZ, CZ-PZ, FP2-F4, F4-C4, C4-P4, P4-O2, FP2-F8, F8-T8, T8-P8.

Patients in this database displayed an unbalanced number of epileptic occurrences, and due to this fact they were grouped in a first patient group with 3 patients showing 10 or more events, a second group of 9 patients with 6–9 events and a third group of 12 patients with 5 or fewer events in the set of recordings.

Considering that the increase in specificity rate can have a correlation with inter-ictal regions, and also considering that the ANN can be prone to learn much better only one kind of region if there is a data unbalance (ex: these data sets signals are almost entirely inter-ictal), it is proposed a combination of stratified and random sampling methods. This sampling strategy aims to better evaluate the ANN performance, taking into consideration the density of the epileptic events to improve representation of ictal signals.

The sampling is accomplished by creating four training groups, as seen in Table 1, based on the three groups of patients discussed earlier in this chapter. The groups and selection criteria are: Group 1) All patients from the patient group with more epileptic events; Group 2) The first patient from each of the three patient groups; Group 3) Three patients from full randomization including all patients groups; Group 4) One random case from each group.

The training groups created includes 10 subjects, using 104 h of EEG data.

Table 1

Random grouping of patients for unbiased training. The groups enumerate patients with different epileptic events density in a same time span, evincing possible differences in cerebral activities patterns.

Group	Patient selection criteria	Patients	Number of files with epileptic events
1	All from the group with higher number of events	chb15	14
		chb12	13
		chb24	12
2	Biggest number of events from each group	chb15	14
		chb13	8
		chb05	5
3	Complete randomization	chb14	7
		chb17	3
		chb19	3
4	Randomized block	chb15	14
		chb16	6
		chb08	5

2.2. Data pre-processing

Aiming to obtain features with a higher probability of correlation with the pre-ictal regions (based on preliminary tests), pre-processing is made with the data before training and testing the ANN. The pre-processing steps are shown in Fig. 2:

At first, a channel reduction is proposed, due to the excessive volume of data, then feature extraction is made, followed by data normalization to better analyze each channel disregarding the difference in intensity of the electric field captured by the EEG electrodes in different brain regions, and finally, two data manipulation techniques are proposed in an attempt to better assess pre-ictal regions and its characteristics.

2.2.1. Channel reduction

The use of an excessive number of channels incurs huge computational load, in addition to undesired factors like greater data acquisition complexity, related costs, energy consumption and read time for portable acquisition equipment, patient discomfort, and especially data overfitting due to redundant data [11].

Among previous works [12], obtained less than 10% of the variation in sensitivity and specificity when using only the C3-C4 channel, in comparison with the results when using all 11 available EEG channels. [13], which used the same database as the present work, applied the Decision Trees algorithm to identify EEG channels with a greater contribution in correct seizure classification, identifying the channels FZ-CZ, F3-C3, and P3-O1 as the most relevant channels for classification, although pointing out that it was a patient-specific result. Also using the Physionet Database [14], evaluated the performance of using from 2 to 6 out of 22 channels. Initially, a 60% accuracy was obtained when using 22 channels; with 5–6 channels, the accuracy increased to 70%; finally, with 3–4 channels, the accuracy was the same as with 22 channels.

It is proposed to select the top 4 most frequent channels present in the combinations that provided the higher accuracy in Ref. [14] work, being them F7-T7, F8-T8, FP2-F8, and T8-P8, with 87%, 67%, 60% and 53% participation, respectively, in the best combinations. It was considered that, even though there are channels that may not present the best results when used alone, they might produce good results when used in conjunction.

2.2.2. Feature extraction and data preparation

Epileptic seizure EEG data presents a variety of patterns, which are not easily identified. Feature extraction is applied as an attempt to obtain features of the EEG signals by transforming raw data.

To obtain preliminary understandings, the Short Time Fourier Transform (STFT) of an entire record is presented in the Fig. 3, showing the spectrum throughout the temporal samples, with superposing rectangles identifying the occurrence of epileptic seizures.

As seen in Fig. 3, there are abrupt temporal and spectral variations in imminence and the occurrence of epileptic events (highlighted by white rectangles superimposed to the graph). Similar variations occur in other sections without epileptic events, such as right after 500 and 3000 s, indicating at first that there is no coherent pattern variations in epileptic events, but it should be considered that not only the frequency pattern characterizes the regions of interest, but rather a set of characteristics considered simultaneously in ANN training.

To better evaluate the spectral behavior for differentiating inter-ictal, pre-ictal and ictal regions, a more detailed view of the spectral components in such regions is shown in Fig. 4, using the same case as Fig. 3, regarding the analysis of the F7-T7 channel along with the untreated EEG signal of the four channels selected for analysis.

By analyzing Fig. 4, it can be seen that the untreated EEG signals for the ictal and inter-ictal regions are very similar, only visually distinct ictal signals presented wide amplitude variations between the channels. However the spectral components have more distinct patterns for each region, having the pre-ictal regions time-varying spectral components, which are somewhat similar to the components of the ictal regions.

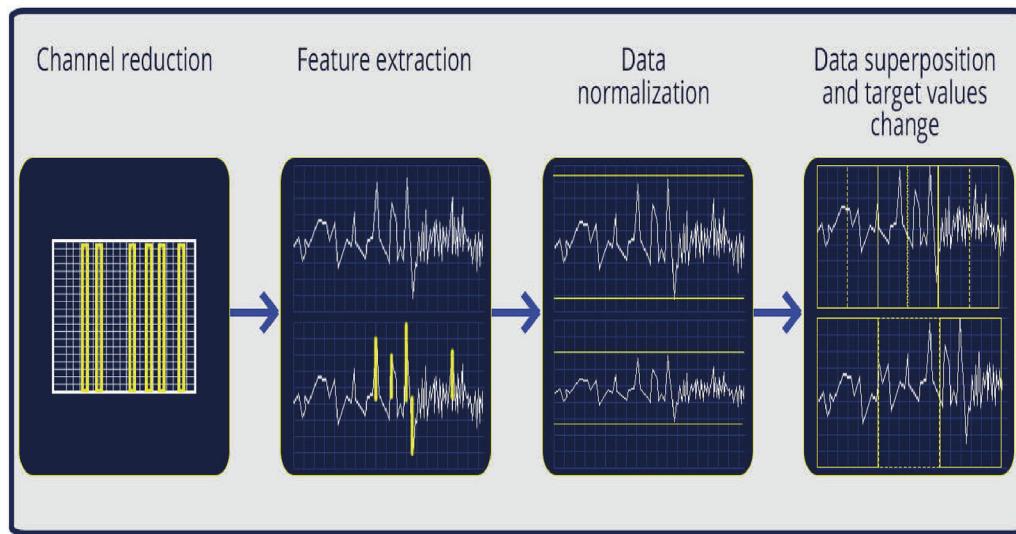


Fig. 2. Pre-processing steps.

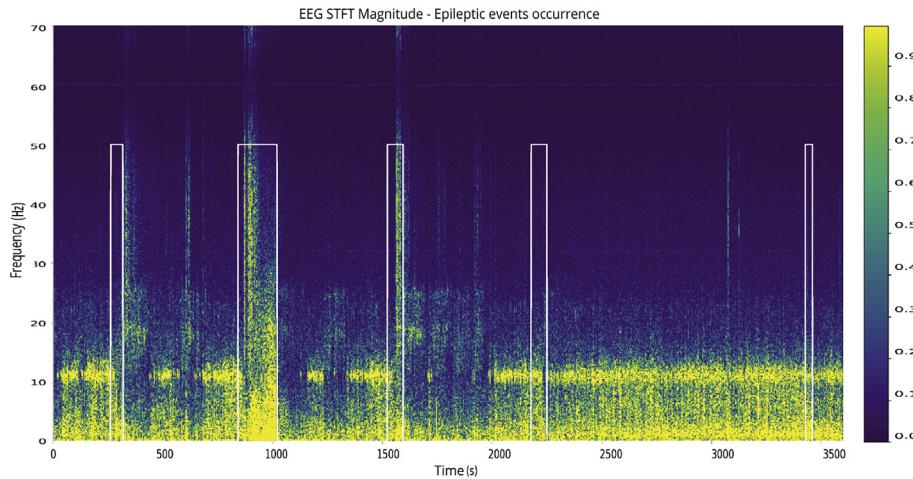


Fig. 3. STFT of an EEG signal along with marking of recorded epileptic events occurrence. Illustration of temporal and spectral characteristic variations of epileptic events.

Analyzing the EEG signals shape, it's possible to notice that there are many sparse peaks over time, with these peaks correlating with specific cognitive activities as well as with epileptic events, as analyzed by Ref. [15]. Given the characteristic of constant variation in amplitude and presence of peaks, in the present work is taken into consideration temporal characteristics such as Root Mean Square (RMS), standard deviation, and the number of peaks observed above a certain threshold.

The threshold used for peak counting was 80% of the maximum signal amplitude for each EEG channel used, considering the entire patient record being analyzed. This threshold was obtained in preliminary tests by varying the threshold from 50 to 90%, looking for a value that could better represent visually distinctive peaks, but minimizing the detection of local maxima. All observations in a time frame of 1 s were summed, then this sum was replicated for all samples in the second in question, being 256 samples due to the sample rate of 256 Hz.

Especially in EEG signals of epileptic individuals, low-frequency signal peaks, such as alpha and theta waves, are observed in both inter-ictal and pre-ictal periods; however, in the occurrence of an epileptic event, a sudden change of frequencies and the emergence of polymorphic waves of prolonged duration is observed, with some decrease in frequency, but with an increase in amplitude [1,16]. Given the nature of this type of biosignal, the spectral characteristic Fast Fourier

Transform (FFT) was extracted, then followed by Power Spectral Density (PSD) extraction, obtaining, this way, power information about the spectral feature. The PSD indicates the energy density verified in the spectral components, showing how significant a spectral component is for a given segment of data.

To obtain spectral features, the Short Time Fourier Transform was applied, as shown in Equation (3), by using a Hanning Window in the $w[n-m]$ term, 512 samples, 256 Hz sample rate and 50% overlapping between the windows. The choice of sample and sample rate values is a product of the best adjustment between time and frequency resolution obtained after several preliminary tests, in which either temporal or spectral resolutions significantly decreased in detriment of one another, reiterating Heisenberg's Uncertainty Principle [17], which says that frequency resolution is directly proportional to the window size applied, while time resolution is inversely proportional.

In addition to contributing with a fair balance between time and frequency resolutions, the length of 512 samples for STFT processing, in combination with the other described parameters, results in a matrix with temporal dimension equals to the input length, allowing it to be used as a feature column without further dimension adjustments. With 50% overlapping, all EEG inputs are evaluated at least once to create the spectral feature.

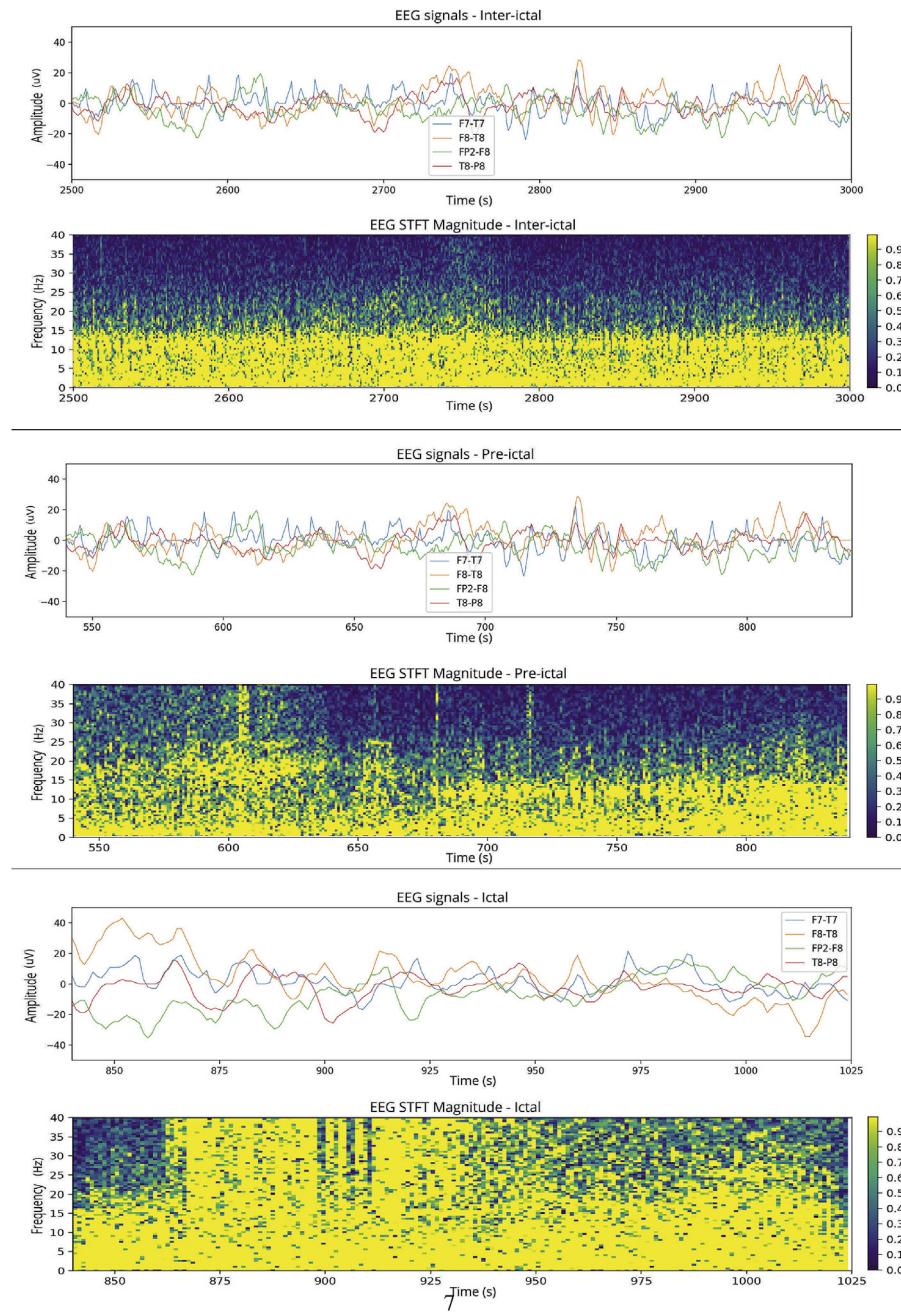


Fig. 4. Spectral analysis of inter-ictal, pre-ictal and ictal regions of the F7-T7 channel using STFT. Graphical display of EEG reading of the four channels selected for this work.

$$\text{STFT}\{x[N_{\text{seg}}]\} = X(m, \omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n} \quad (3)$$

Although overlapping proportion improves the FFT consideration of each data point, it also increases overlapping correlation, which is undesired due to data redundancy associated. Considering this, the value of 50% was chosen since this is the lowest value for which Hanning Window amplitude flatness is 1, that is, the sum of all weights applied to each data point sums to one [18].

By obtaining the spectral characteristics, it is possible to select certain frequencies for analysis. It was chosen to use one frequency in each of delta, theta, alpha and beta brainwaves, since, as analyzed by many works [2,19–22], these brainwaves hold a relation with seizures. Therefore the frequencies 4 Hz, 8 Hz, 13 Hz and 18 Hz were selected, representing each of the brainwaves mentioned.

2.2.3. Data normalization and target categorization

The features extracted, being EEG electrical fields or magnitude of spectral components, differ not only in nature but also in order of magnitude and eventually will interact in the training process and subsequent prediction of ANN, so it is recommended that they be normalized to have similar bases of comparison, with equal lengths or numerical limits [23]. In the present work the method used for normalization was Min-Max, as shown in Equation (4):

$$\text{MinMax}(x[n]) = \frac{x[n] - \min(x)}{\max(x) - \min(x)} (max - min) + min \quad (4)$$

This type of normalization preserves the relationship between the original data [24], which is assumed in the present work as a relevant criterion to be maintained. A scale with a minimum of -1 and a maximum of 1 is used in order to preserve the original scale characteristic

and to include negative and positive values.

The normalization process is done by analyzing the values of a random file among the current patient records, minimizing any selection bias, and then applying to all records.

After the data normalization, the targets are encoded in such a way that the ANN can easily process the data. This is done by categorizing values that are originally zeros and ones into an array, by using the one-hot encoding process. The outcome of the encoding process is that each set of data entry have a target represented by a two columns array, specifying if an epileptic is occurring or not in that sample.

After the manipulations described, the data is resized in batches, as seen in Fig. 5. A set of N inputs is divided in B batches, having each batch T observations or inputs, being B the rate between the total number on inputs N of the data set and the number of observations T , which can also be seen as time steps. Finally, the target A is divided in T_a time steps in the same fashion as N inputs to T time steps.

The main purpose of batch division is to reduce the computational cost, which is generated due to the large data size. When long data sequences are used in a single training simultaneously, it demands computational loads that sometimes are not well fitted for computers without dedicated setups for such task.

In this work it was chosen to use 1-s batches, that is, 256 observations per batch, since the EEG acquisition rate is 256 Hz, totaling 3600 batches per EEG file. Lower batches of 0.5 and 0.25 s, as well as bigger batches as 2–5 s, were tested; the outcome showed that batch size presents a trade-off between computational load and generalization error, in such a way that higher batches significantly increased the computational load with some improvement in accuracy. The value of 1 s showed a reasonable memory consumption while achieving good results in comparison with larger batches.

2.2.4. Input data superposition and target values change

In an attempt to increase the antecedence in prediction of epileptic event occurrences, two preprocessing techniques are used, acting only on the training values, so that ANN is trained with this preprocessed data to later generate early predictions when exposed to data that has not undergone such processes.

One approach to try to anticipate predictions of epileptic events is to overlay the data by replicating it with the segmentation technique named sliding windows. The technique registers a window of data W , then move it forward by S steps/inputs and adds this new data window to the first one, and so on. It is noted in the Fig. 6 the superposed signal (lower half) from three windows applied in the original signal (upper half), in which the windows height, just as the vertical displacement of the signals, are altered for better visualization.

It is empirically defined in the present work the use of 1 s windows (256 entries) and 0.5 s steps (128 inputs), i.e. 50 % overlap, expanding the training set to double of its original extension. Thus it is intended that the ANN be exposed more than once to various data. A consequence of

the sliding windows application is that there is a possibility that ANN can better comprise data sequences preceding epileptic events, especially in input data segments where targets correspond to epileptic events, as overlap creates sequences with data repetition, and consequently expands (through temporal repetition) the excerpts referring to epileptic events.

In conjunction with data overlapping, it is applied an approach where the target values are anticipated, based on [25], which applies to a partial data set the target verified at the end of the set. The anticipation is made by shifting the target values to earlier positions in time in such a way to move the positive targets (which indicate the incidence of an epileptic event) to the pre-ictal sections. The target offset can be seen in Fig. 7, where Fig. 7a shows the original signal and Fig. 7b shows the result of the offset.

It is used in the present work the displacement of the targets in 5 s, based on preliminary tests. Thus, it is intended that ANN attempts to classify pre-ictal events by directly indicating the targets of those segments with the category 1 value, and partially indicating the ictal segments with category 0 targets.

2.3. Recurrent LSTM network

The goal of the proposed network is to classify data sequences between pre-ictal and inter-ictal, indicating the high or low probability of being each class. This is done by modeling the ANN to accept temporal sequences and retain relevant information - a task for which the LSTM topology is especially efficient.

The LSTM network was modeled to compute batches of data, in such a way to generate, at the end of each batch, an output. It was chosen to use batches due to the high size of the data set being processed (several Gigabytes of data for each patient after data processing).

The developed network processes such input data using several LSTM cell layers. In the present work, the configuration developed has three LSTM network layers and one output layer. The first two LSTM layers have 30 cells, returning one output at each input, configuring an architecture of multiple inputs to multiple outputs. In addition, these first two layers pass to each subsequent batch the internal states computed in the last batch. The third layer has 4 cells, being this a random value chosen by the error minimization criterion. This layer returns an output only after analyzing all observations in a batch, having a structure of multiple inputs to one output, and maintains the procedure of passing internal states to subsequent batches.

Finally, there is a layer with 2 neurons and the Softmax activation function, which actually calculates the probability that the processed sequence belongs to each of the present classes. The diagram of the developed LSTM network can be seen in Fig. 8a, where the data stream of a batch is represented. Because ANN performs the classification of which class should belong to the analyzed sequence, the loss function used is Categorical Cross Entropy.

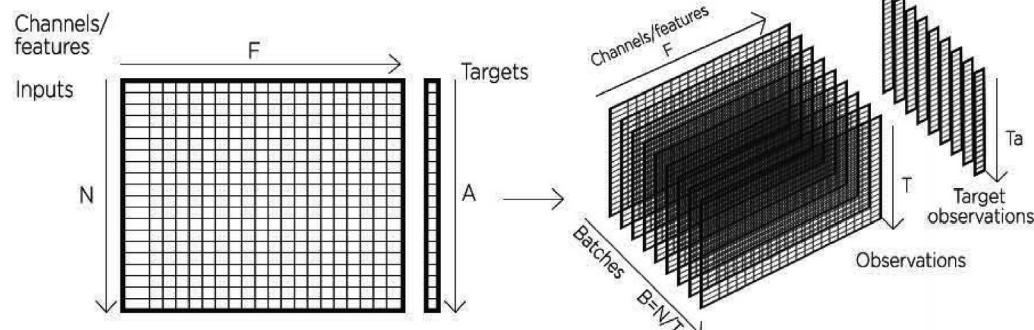


Fig. 5. Division of data sets in batches and observations.

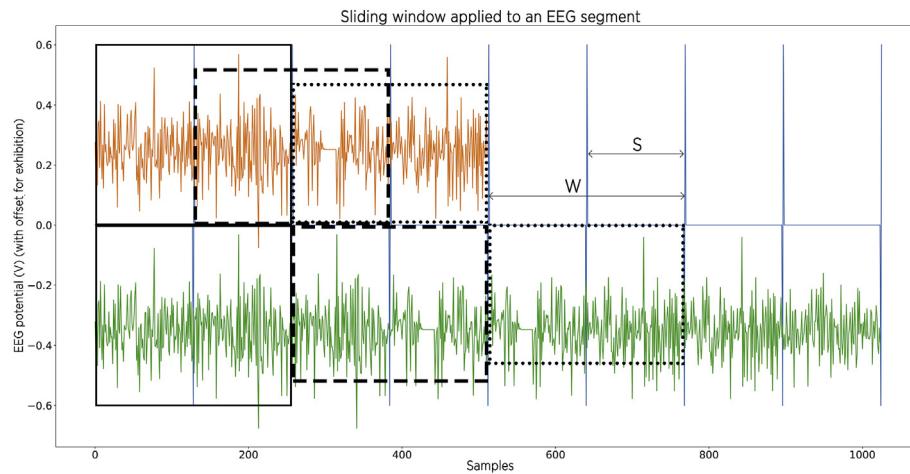


Fig. 6. Input data overlay (upper signal) achieved by applying sliding windows of size W and step S , obtaining an overlapped signal (lower signal). Different dashes indicate different windows. The procedure is used in training sets to anticipate predictions of epileptic events.

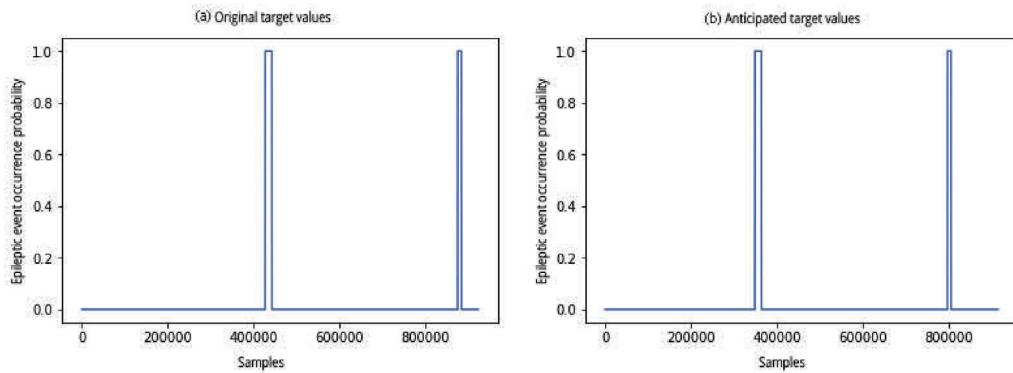


Fig. 7. Target displacement applied to anticipate epileptic seizure predictions: (a) original targets and (b) displaced targets.

The diagram shown in Fig. 8a corresponds to the general architecture used when there is no data overlap. However, when using an overlay, the input format of the first LSTM layer is changed to receive any number of batches so that the data length is longer in training and of normal size in tests. For this, the second layer of 30 neurons that received sequences of defined size was removed, because the matrix sequence has a variable size, as seen in Fig. 8b. In addition, the internal states computed in the last batch are no longer passed to each subsequent batch due to the same change in network architecture to accept any number of batches.

It was used 30 epochs for the network training, a value for which the decrease of training error and increase of the accuracy have very small variations when increasing the number of epochs (according to the analysis presented in the Results chapter). In addition, the data set partition was selected between 80% for training and 20% for testing, thus preventing performance evaluation from inducing results that apply only to the training set, so evaluation also tests ANN's generalizability when analyzing performance in the test set.

Once the ANN is trained, it is necessary to evaluate the results generated by it against new data. ANN's performance evaluation metric is based on detecting signals such as pulses or detection windows between 5 and 60 s before an epileptic event occurs, being these values are arbitrated as a near period to the epileptic event but with certain antecedence, so that it is possible to take the necessary actions by epileptic patients or related individuals. Other detection time values are also evaluated in a summarized manner (as per analysis presented in the Results chapter).

For evaluating the ANN performance, sensitivity and sensibility metrics are used. In the eventuality of failure in early prediction of

epileptic events, the analysis is changed in a way to verify the ability of the ANN to act as a classifier of ictal events, that is, the ability to generate predictions with zero anticipation time.

3. Results

In order to evaluate the techniques employed, the performance evaluation of the developed ANN is separated into three tests with variation in the use of the techniques, the variations being: basic structure without data superposition and target anticipation, use of data superposition, and finally the use of superposition in conjunction with target anticipation. For each combination, it is evaluated the early classification performance of epileptic events in the test set with each of the training groups shown in Table 1.

By applying all the techniques described above, no predictions were obtained as intended, but the ANN LSTM was able to classify epileptic events into processing time during testing. Therefore, the correct prediction is evaluated every second during the occurrence of epileptic event windows at each prediction, that is, it is evaluated at every second (when a whole lot is evaluated and an output is generated) if the prediction of an epileptic seizure matches the prediction target of the given test lot.

3.1. Performance evaluation and neural network predictions

The LSTM network used in this work has many variable parameters, being the epochs number among them. It was observed that the network's performance is directly affected by the number of epochs. By using the training group 1, it was generated the data relating the ANN's

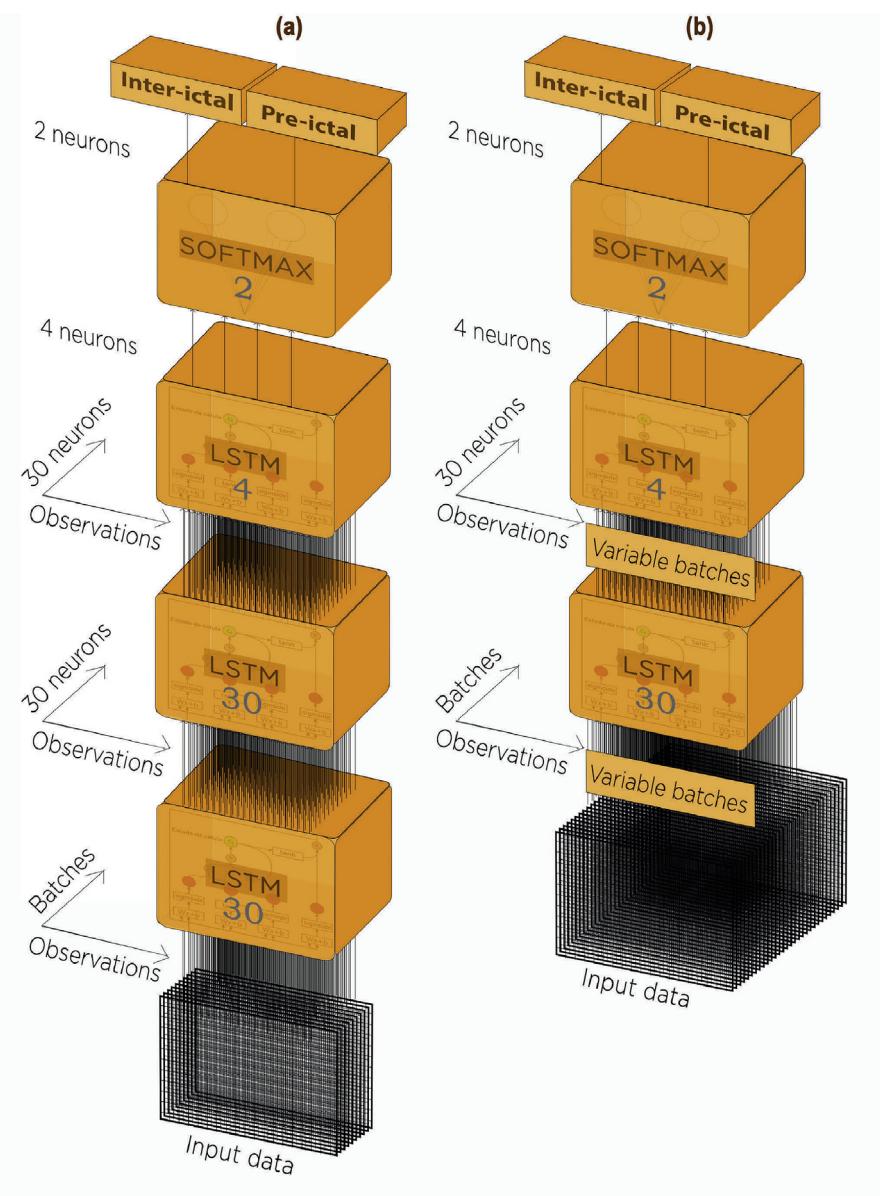


Fig. 8. Diagram of the ANN LSTM developed for categorical classification of EEG temporal signals between inter-ictal and pre-ictal. Datastream representation for a single batch. The number in each layer indicates the number of cells or neurons: (a) general architecture diagram and (b) architecture diagram for data superposition.

performance and its loss (prediction error) given the number of epochs. It is summarized in the [Table 2](#).

Considering the results of [Table 2](#), it can be seen that the performance has a direct relation with the accuracy and even a more strong relation with the loss, but it is not linear. The accuracy only varies by the least significant digit, while the loss has a significant decay, specially between 15 and 30 epochs, when it gets 50% better. Considering that the training time is directly proportional to the number of epochs, it was chosen to use 30 epochs, given its good balance between accuracy, loss and training

time.

Once the network performance is evaluated, the results of epileptic events detection are evaluated. The results obtained with the developed ANN LSTM did not present in any of the tests the detection of anticipated epileptic events with consistent patterns between files, or without the presence of several other similar detections preceding the supposed early detection, even when using several different anticipation periods. It was considered for this analysis the detection metric established within 5–60 s prior to an epileptic event, that is, the period considered pre-ictal.

In order to briefly analyze the application of different metrics, expanding the analysis window of anticipated epileptic events to even longer periods, [Table 3](#) shows a comparison of the percentage of early epileptic events detected with all previously proposed techniques by varying the window time for analysis of epileptic event detection, ranging from 60 s from the original metric to 240 s in 60 s steps.

The best individual result, being this the group 1 with upper limit detection windows ranging from 120 to 240 s, reached a maximum of 9.37% of expected detections. Therefore, it can be concluded that the analysis of ANN's ability to detect epileptic events in an early stage is not

Table 2

ANN performance according to the number of training seasons used. Tests done with ANN basic structure and training group 1.

Number of epochs	Accuracy (%)	Loss
10	99.5	0.04
15	99.5	0.04
30	99.6	0.02
60	99.6	0.02

Table 3

Percentage of early detection of epileptic events considering all proposed techniques and varying the size of the pre-ictal period. Results obtained for each group of patients and for the whole set of groups.

Period pre-ictal(s)	Anticipated detections Group 1 (%)	Anticipated detections Group 2 (%)	Anticipated detections Group 3 (%)	Anticipated detections Group 4 (%)	Anticipated detections of epileptic events Total for the period (%)
5–60	3.12	0.00	8.33	5.00	3.33
5–120	9.37	0.00	8.33	5.00	5.00
5–180	9.37	0.00	8.33	5.00	5.00
5–240	9.37	0.00	8.33	7.5	5.83

being strongly affected by the established detection metric.

Although early detection of epileptic events did not achieve significant results through the tests, the predictions obtained demonstrated that the LSTM structure acted as a classifier, analyzing data sequences and predicting for a set of tests the precise periods of occurrence of epileptic events. Examples of unanticipated prediction can be seen in Fig. 9.

In Fig. 9a, from group 1 of patients, the correct prediction is observed in practically every window of occurrence of the epileptic event to the right of the figure. Earlier in the same figure one can also observe erroneous predictions. Fig. 9b shows that the same result is observed in other training and testing groups, which is an example of group 3. In both figures, the ANN's accuracy in classifying epileptic events is evident, even if there are small sections within the event window where the rating fluctuates.

The results demonstrate that the techniques discussed in the present study were not sufficient to anticipate predictions, considering the pre-ictal period between 5 and 60 s, but in fact generated correct

predictions, allowing the network to assimilate temporal patterns of EEG and classify the occurrence of epileptic seizures in the presence of such patterns.

It is important to consider, for comparison purposes, that an extensive database and combinations of data from different patients were used in the present development, with the intention of maximizing the generalization of the possible scenarios to be observed. Ideally, this combination of factors would make ANN learn the generalization of pre-ictal segments regardless of possible physiological or psychic factors unrelated to epileptic events, but what possibly occurred was the difficulty in training so many scenarios and for so many patients with distinct EEG patterns.

Given the discussion about the absence of anticipated predictions, in order to explore the results obtained by ANN LSTM with the different types of structures proposed in the present work, only the performance of the network for classification of ictal regions is evaluated next, that is, the unanticipated prediction of epileptic events.

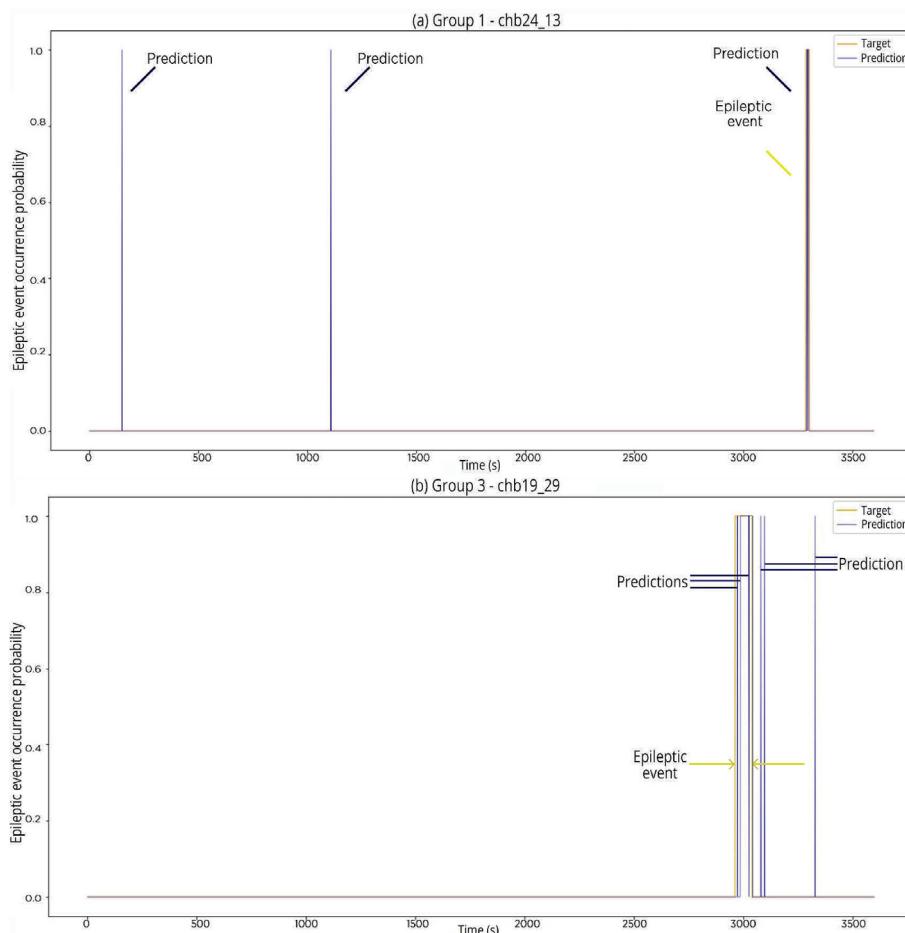


Fig. 9. Examples of unanticipated predictions, both files analyzed present epileptic events at the end. Although predictions weren't classified in advance, they correctly indicate the occurrence of epileptic events. Presence of false-positives: (a) at the beginning of the file belonging to group 1 and (b) at the end of the file belonging to group 3, immediately after the epileptic event.

3.2. Basic structure training

Using the ANN shown in Fig. 8, using the basic structure, the confusion matrices shown in Fig. 10 were obtained, relating the sensitivities and specificities in the classification of epileptic events for each of the four training groups, that is, determination of ictal periods.

The results show that the classification of ictal periods did not present a good rate of true-positives, resulting in an average sensitivity of 14.3% for the 1494 ictal segments for all four groups, and a maximum sensitivity of 45.5% for group 3. There are partial detections of the periods for half of the groups, except for groups 2 and 4 that did not present any correct answers. However, it should be considered that there were detections within peak event windows in the form of spikes or smaller windows. An example of the detections can be seen in Fig. 11, which shows an example of detection for group 1.

By evaluating the accuracy, it is possible to verify that it does not necessarily match the specificity and sensitivity verified in the test set. This may be because part of the training is done with data from more than one patient and the test involves patients who have had partial participation in the training because only files where epileptic events were checked.

By analyzing Table 1 and the results of Fig. 11 it can be stated that there are indications that the proposed structures present results that are patient-dependent or dependent of distinct EEG patterns, and that ANN performance is negatively affected for patients with more discrepant

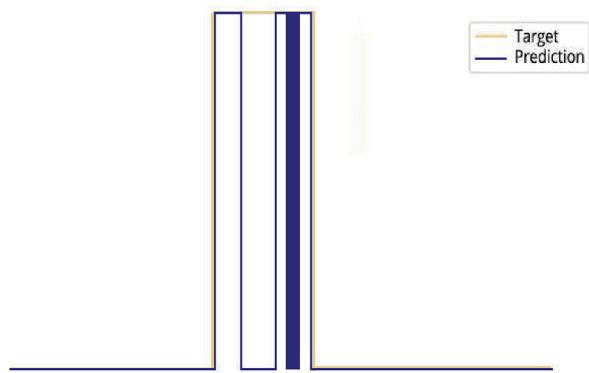


Fig. 11. Detection peaks within a window of occurrence of an epileptic event. Undetected sections reflect network sensitivity.

seizure incidence rates.

Groups 2 and 4 show extreme scenarios where there are no true-positive detections. These groups show the most discrepancy in number of epileptic events per patient, ranging from 5 to 14 in both groups, while group 1 ranges from 12 to 14 events and group 3 ranges from 3 to 7 cases, although this last group has two patients with the same number of epileptic events (chb 17 and chb 19). This could lead to an indication that

		(a) Group 1			(b) Group 2		
		Inter ictal	Ictal	Target	Inter ictal	Ictal	Target
Prediction	Inter ictal	28468 98.8%	212 0.74%	28680 99.3% 0.74%	20759 96.1%	841 3.89%	21600 96.1% 3.89%
	Ictal	19 0.07%	101 0.35%	120 84.2% 15.8%	0 0.00%	0 0.00%	0 0.00%
		28487 99.9% 0.07%	313 32.3% 67.7%	28800 99.2% 0.80%	20759 100% 0.00%	841 0.00% 100%	21600 96.1% 3.89%
		Inter ictal	Ictal	Target	Inter ictal	Ictal	Target
		(c) Group 3			(d) Group 4		
Prediction	Inter ictal	10453 99.8%	134 1.24%	10587 98.7% 1.27%	21504 99.6%	94 0.44%	21598 99.6% 0.44%
	Ictal	101 0.94%	112 1.04%	213 52.6% 47.4%	2 0.01%	0 0.00%	2 100%
		10554 99.0% 0.96%	246 45.5% 54.47%	10800 97.8% 2.18%	21506 100% 0.01%	94 0.00% 100%	21600 99.6% 0.44%
		Inter ictal	Ictal	Target	Inter ictal	Ictal	Target

Fig. 10. Confounding matrices of ictal period determination for all training groups using basic training structure.

the frequency, or severity of epilepsy clinical presentation, have a strong relation to the internal group of EEG signals patterns present in different epilepsy patterns among patients. This behavior is seen for all approaches in this work.

3.3. Data superposition

By overlapping the data, as shown in the example in Fig. 6, the results arranged in Table 4 were obtained.

Analyzing the results shown in Table 4, it can be seen that the sensitivity of ANN increased with the use of data superposition technique when compared to the basic structure, resulting in an average sensitivity of 19.9% for the total of 1494 ictal segments for all four groups, and a maximum of 61.3% for group 3. This shows that this technique attenuates the occurrence of false-negatives. However, the false-positive rate increases too, as seen in Fig. 12, which shows the results for groups 1 and 3.

In Fig. 12a there is an epileptic event between 3000 and 3500 s, overlapped by several detection peaks, in which it can be seen that ANN correctly identifies this segment, resulting in a higher sensitivity compared to the basic training structure. Fig. 12b shows a similar result for group 1, with a high concentration of detection peaks during the epileptic event.

In addition to the test performed with the previously described methodology, a brief test was also performed to compare with a different overlapping window and step value from those originally proposed. Instead of 1-s overlap windows with half-second steps, the parameters used for data overlap were halved then doubled, that is, a 0.5-s window with 0.25-s steps and 2-s windows with 1-s steps. Using training group 1 only, the results shown in Table 5 were obtained.

The results show that the data superposition parameters have an optimal value, instead of improvement trends directly proportional to the size of these parameters. Therefore it can be stated that the parameters used for data superposition in the present methodology are correctly adjusted.

It was seen that for [22] there were methodologies for improving sensitivity with the disadvantage of worsening the number of false-positives, highlighting the possibility that certain techniques could not improve the classification of temporal signal regions without significantly increasing the amount of erroneous classification in other regions. Therefore, if an RNN LSTM application designed for EEG signals wishes to improve sensitivity and can tolerate worsening false-positive rates, overlapping training data may prove to be an effective technique.

3.4. Target data anticipation

The usage of target data anticipation resulted in the accuracy, sensitivities and specificities summarized in Table 6:

Although training targets were anticipated by 5 s before an epileptic event, with a value of 1 being set for targets, sensitivity improved by 1.60% for group 1 and worsened by 5.30% for group 3, that is, there was no consistent improvement in results. Other than that, ANN's hit rate and specificity worsened for virtually all training groups. This result demonstrates that for the ANN structure and configurations proposed in the present work the anticipation of targets tends to show a slight worsening of results, or at best a small increase in sensitivity.

In order to briefly explore the results obtained for the other target anticipation values, it is shown in the Table 7 the sensitivity and specificity values obtained for the patient group 1 when varying the target anticipation time between several values going from 1 to 40 s.

The results show that there is no consistent variation trend in prediction performance as the anticipation time varies, as, for example, the best sensitivity result was obtained for 2-s target anticipation, whereas the worst result was for the 7-s anticipation; however, the second-best result is for the 9-s anticipation.

It is interesting to note that the training target anticipation time

Table 4

Results obtained with sliding window data superposition technique.

Training group	ANN accuracy (%)	Sensitivity (%)	Specificity (%)
1	99.5	47.0	99.7
2	93.2	0.00	100
3	98.1	61.3	97.3
4	100	0.00	98.2

arbitrarily chosen in the present methodology was not optimal and that this value does not appear to have a local maximum either. Using a similar target anticipation technique [25], was able to anticipate facial movement predictions, obtained by monitoring with a camera. However, although the present ANN architecture differs from the architecture proposed by Ref. [25], the principle of data anticipation training is the same.

Combining the aspect that random results were observed by varying the anticipation time of targets, with the inability to anticipate prediction for the present methodology, it is suggested that the difference in results may be related to the nature of the data, that is, there is a possibility that there is an inherent characteristic of EEG signal patterns that renders target anticipation ineffective in the goal of anticipating predictions.

3.5. Target data anticipation and data superposition

The last proposed technique applies superposition to the training set input and target data and anticipates targets of the same training set. This combination of techniques resulted in the accuracy, sensitivity and specificity summarized in Table 8.

As with target anticipation techniques without overlapping data, there was no consistent increase in sensitivity. Despite the improvement of sensitivity of 16.9% for group 1, there was a worsening of 16.3% for group 3, characterizing a strong dependence on performance with the patient analyzed.

Although data overlap resulted in higher sensitivity for groups 1 and 3, remembering that for the other groups the sensitivity was null, the results followed the pattern observed in the anticipation of targets. Compared to the basic structure, sensitivities are better for group 1 and worse for group 3. This shows that for the ANN structure and techniques proposed in this paper, the use of target anticipation is not good. Consistent results for prediction sensitivity also show the same behaviors when used in conjunction with other techniques.

3.6. Comparison with similar works

This work reached a maximum percentage of 8.33% and 61.3% of anticipated seizures onset detection and epileptic events classification, respectively. It is important to consider for this comparison that the present development used an extensive database and a combination of data from several patients, to generalize all scenarios present in the database as much as possible. Ideally, this combination of factors would make the ANN apprehend the generalization of pre-ictal segments; however, results show a model difficulty in learning the variety of EEG patterns for so many patients.

When comparing the results with works such as [19], which showed sensitivity around 98% in the classification of ictal regions, inter-ictal and pre-ictal with SVM classifier, it can be noticed that the result obtained with the present ANN LSTM is less satisfactory as it reaches a maximum of 61.3% sensitivity (using the data superposition). However, it must be considered that, as cited by Refs. [19], the database of such previous work was relatively small, being about 30 times smaller in time of recordings in comparison to the database used in this work, with 3.28 h of data versus 101 h of the files used in the present work. When analyzing a shorter data set, there is less probability of facing specific seizure scenarios with particular patterns that differ from others. Even with the present considerations, it is interesting to notice the use of

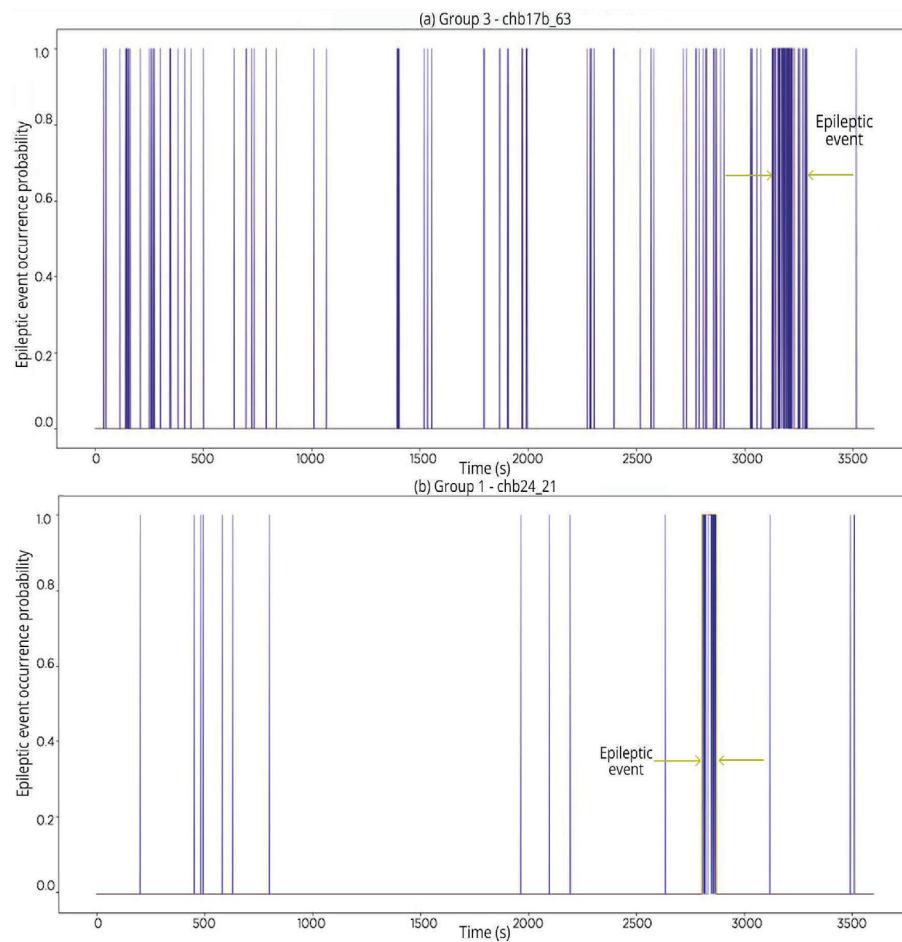


Fig. 12. Example of data overlapping results for groups 1 and 3. There is an epileptic event occurring at the end of each file, where there are many detection peaks. Sensitivity has good results over specificity, a result evinced by many false-positives. Epileptic events are marked by arrows for better identification. (a) group 3 file, with the epileptic event between 3000 and 3500 s, covered by several detection peaks. (b) group 1 file, with a clear classification of the ictal period between 2500 and 3000 s.

Table 5

Comparison of results obtained from varying window size and overlapping steps applied to training data. The test was performed with group 1.

Window size (s)	Step size (s)	Sensitivity (%)	Specificity (%)
0.50	0.25	32.6	99.9
1.00	0.50	47.0	99.7
2.00	1.00	34.8	99.9

Table 6

Results obtained with target data anticipation technique in the training set.

Training set	Accuracy (%)	Sensitivity (%)	Specificity (%)
1	99.5	33.9	99.7
2	92.9	0.00	100
3	97.6	40.2	99.1
4	100	0.00	99.8

Autoregressive Moving Average (ARMA) to get autocorrelation with moving averages and identify trends.

The work of [20] adopted the same database as in this development, aiming to classify epileptic events by using the Extreme Learning technique. The temporal and spectral characteristics have similarities, although the final results differ widely, with the work cited reaching 95% of correct classifications. The authors used only one EEG channel from the 3 patients selected, contrasting with the present research, which used 4 EEG channels from 10 different patients, resulting in 5 times more data. Therefore, it is again considered that the referred work may have been

Table 7

Results obtained by varying the anticipation time of training targets for patient group 1.

Anticipation time (s)	Sensitivity (%)	Specificity (%)
1	39.9	99.8
2	50.2	99.8
3	36.4	99.9
4	39.3	99.9
5	33.9	99.7
6	42.5	99.8
7	13.1	99.9
8	36.7	99.9
9	47.3	99.7
10	36.1	99.8
20	44.7	99.7
30	29.4	99.9
40	38.0	99.8

Table 8

Results obtained with data superposition and target anticipation technique in the training set.

Training set	ANN's Accuracy (%)	Sensitivity (%)	Specificity (%)
1	99.6	49.2	99.2
2	93.1	0.00	100
3	96.9	29.3	99.6
4	100	0.00	100

exposed to a limited number of scenarios, reducing the number of patterns to be learned, or there is still the possibility that the ELM technique was more appropriate for this scenario.

In [26], the authors used the same database as this work to implement seizure prediction using LSTM with patient-specific analysis; the work achieved over 99% sensitivity and specificity, and 0.11 false positives per hour when evaluating pre-ictal segments of 15 min before seizures. Some significant differences between the present work and [26] remain in the signal segmentation, class imbalance approach and time window used to assess the performance of anticipated predictions.

The authors in Ref. [26] divided the data into segments of 5 s each, instead of 1 s, resulting in less segmentation and possibly leading to better feature extraction at the cost of higher computational load. The class imbalance was solved by extracting random inter-ictal results with the same size of pre-ictal segments, creating balanced segments. The main difference remains in the pre-ictal window size used for metrics. The current work evaluated early predictions from 5 to 60 to 5–240 s, while [27] windows ranged from 15 to 120 min, a window eight times larger. This larger window drastically changes the results; therefore, a strict direct performance comparison is not possible. The authors compared their work with other studies that used windows of at least 5 min, being all windows at least two times wider than the ones used in this work.

4. Conclusions

By analyzing the results obtained, it is possible to state that with the adopted methodology it is not possible to determine pre-ictal periods in EEG signals with LSTM networks. Early determination of epileptic events may require more characteristics, different combinations of these or different techniques for manipulating training data.

It was possible to verify a strong dependence on the proposed technique with the patient and their EEG patterns, suggesting that it may be not possible to obtain EEG patterns generalization for several patients in a single ANN model. This strong dependency was especially visible in sensitivity results. Considering this, it is suggested that such models be first trained and tested with only one patient, then generalized with the combination of other patients. It was possible to verify a strong dependence on the proposed technique with the patient and their EEG patterns, suggesting that it may be not possible to obtain EEG patterns generalization for several patients in a single ANN model. This strong dependency was especially visible in sensitivity results. Considering this, it is suggested that such models be first trained and tested with only one patient, then generalized with the combination of other patients.

The high accuracy found in all tests was consistent only with the detection of ictal regions - and not for pre-ictal - showing that the pre-processing techniques or configurations of the ANN LSTM topology were not sufficiently adjusted to categorize the desired pre-ictal temporal sequences.

While detection of epileptic events did not occur early, the ANN was able to learn temporal patterns for categorical prediction of ictal regions, generating peaks or detection windows during the occurrence of epileptic events. This demonstrates that the RNN LSTM topology presents learning and memory capacity, just as data correlation for long time sequences, specifically in today's application for EEG signals.

Although the abundance of data increases the chances for the network to learn the desired patterns, it was also verified that the excess of records of periods without epileptic events (inter-ictal periods), configuring a data imbalance, favored the tendency for a better classification of inter-ictal periods in detriment of ictal and pre-ictal periods. This was evinced in the evolution of accumulated training error when encountering data with various epileptic events, significantly increasing the error at first contact but significantly decreasing with the execution of more epochs.

Considering that target data anticipation did not present good results and caused an overall decrease in performance, even when used in conjunction with another technique, it is suggested to avoid using this

technique when performing seizure prediction using RNN LSTM and EEG data, or at least modify the RNN structure to attempt a different computation of this modified data before providing it to the RNN. It was also verified that there is no clear value or pattern for which the anticipation performs better, reinforcing the attention and further analysis required when using this technique.

When evaluating data processing techniques, reducing channels to less than 25% of selected channels demonstrated positive results in temporal sequence classification, although extensive tests with different combinations of the 18 initially selected channels could better evaluate the correlation between channel composition and classification of ictal and pre-ictal regions. Therefore, it is encouraged to consider channel reduction techniques in similar situations, for improvement of training performance and reduction of related computational and equipment costs.

The directions of future research would be to experiment with the combination of other features, using Multivariate Discriminant Analysis to find resulting features that have a greater correlation with pre-ictal periods. The combination of different features can result in dimensionality reduction and eventual elimination of redundant features for training. Experimentation with larger databases and Deep Learning structures, using only raw EEG data, could result in a model with a low number of EEG channels, in case significant features are successfully extracted by the deep learning model. It is also suggested to test similar EEG databases, compare results, and further analyze the ANN generalization.

Credit author statement

Luciano Bongiorni: performed the signal processing and wrote and revised the paper. Alexandre Balbinot: coordinated the Project, wrote and revised the paper.

Declaration of competing interest

I declare that I have no significant competing financial, professional or personal interests that might have influenced the performance or presentation of the work described in this manuscript.

References

- [1] Siuly S, Li Y, Zhang Y. EEG signal analysis and classification: techniques and applications. Melbourne: Springer International Publishing AG; 2016. p. 2016.
- [2] Shorvon SD, Guerrini R, Cook M, Lhatoo SD. Oxford textbook of epilepsy and epileptic seizures. Oxford: Oxford University Press; 2013. p. 2013.
- [3] Vandecasteele K, De Cooman T, Gu Y, Cleeren E, Claes K, Paesschen W, Huffel S, Hunyadi B. Automated epileptic seizure detection based on wearable ECG and PPG in a hospital environment. Sensors 2017;17(10):2338 (Oct 2017).
- [4] Jory C, Shankar R, Coker D, McLean B, Hanna J, Newman C. Safe and sound? a systematic literature review of seizure detection methods for personal use. Seizure 2016;36:4–15 (Nov. 2016).
- [5] Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult. IEEE Trans Neural Network 1994;5(2):157–66 (Mar. 1994).
- [6] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9(8): 1735–80 (Nov. 1997).
- [7] Gers F, Schmidhuber J. Recurrent nets that time and count. Como, Italy: IEEE; 2000. p. 189–94 (2000).
- [8] Klonowski W. From conformons to human brains: an informal overview of nonlinear dynamics and its applications in biomedicine. Nonlinear Biomed Phys 2007;1(5):5 (Jul. 2007).
- [9] Physionet. Chb-mit scalp eeg database. available at: <https://www.physionet.org/pn6/chbmit/>. [Accessed 1 June 2018].
- [10] Goldberger AL, Amaral LA, Glass L, Hausdorff J, Ivanov P, Mark R, Mietus J, Moody G, Peng CK, Stanley H. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. Circulation 2000;101(23):215–20 (Jun. 2000).
- [11] Alotaibi T, El-Samie FEA, Alshebili SA, Ahmad I. A review of channel selection algorithms for eeg signal processing. EURASIP Journal on Advances in Signal Processing 2015 2015;1:66 (Aug. 2015).
- [12] Greene B, Boylan G, Marnane W, Lightbody G, Connolly S. Automated single channel seizure detection in the neonate, no. 30. Vancouver: Canada; 2008. p. 915–8 (Feb. 2008).

- [13] Birjandtalab J, baran pouyan M, Cogan D, Nourani M, Harvey J. Automated seizure detection using limited-channel eeg and non-linear dimension reduction. *Comput Biol Med* 2017;82(1):49–58 (Jan. 2017).
- [14] Chang N, Chen T, Chiang C, Chen L. Channel selection for epilepsy seizure prediction method based on machine learning. San Diego, USA: IEEE; 2012. p. 5162–5 (2012).
- [15] Nunez PL, Srinivasan R. Electric fields of the brain: the neurophysics of EEG. second ed. New York: Oxford University Press; 2006. p. 2006.
- [16] Tzallas AT, Tsipouras MG, Tsalikakis DG, Karvounis EC, Astrakas L, Konitsiotis S, Tzaphlidou M. Automated epileptic seizure detection methods : a review study. In: Stevanovic D, editor. Epilepsy - histological, electroencephalographic and psychological aspects. first ed. InTech, [S.l.]; 2012. p. 75–98. Ch. 4 (2012).
- [17] Busch P, Heinonen T, Lahti P. Heisenberg's uncertainty principle. *Phys Rep* 2007; (452):155–76 (Oct. 2007).
- [18] Heinzel G, Rüdiger A, Schilling R. Spectrum and spectral density estimation by the discrete fourier transform (DFT), including a comprehensive list of window functions and some new flat-top windows. *Max Plank Inst* 2002;(12):18–9 (Feb. 2002).
- [19] Gupta A, Singh P, Karlekar M. A novel signal modeling approach for classification of seizure and seizure-free eeg signals. *IEEE Trans Neural Syst Rehabil Eng* 2018. 1–1 (Mar. 2018).
- [20] Ammar S, Senouci M. Seizure detection with single-channel eeg using extreme learning machine. Sousse, Tunisia: IEEE; 2016. p. 776–9 (2016).
- [21] Larmuseau M, Dambre J. Epileptic seizure prediction using deep learning. 2016.
- [22] Dadgar-Kiani E, Alkan C, Shameli A. Applying machine learning for human seizure prediction. Stanford, California: Tech. rep.; 2016.
- [23] Guyon I, Gunn S, Nikravesh M, Zadeh LA. Feature extraction: foundations and applications. New York: Springer-Verlag New York Inc; 2006. p. 2006.
- [24] Ogasawara E, Martinez LC, de Oliveira D, Zimbrão G, Pappa GL, Mattoso M. Adaptive normalization: a novel data normalization approach for non-stationary time series, no. 21. Barcelona; 2010. p. 1–8 (Jul. 2010).
- [25] Jain A, Singh A, Koppula HS, Soh S, Saxena A. Recurrent neural networks for driver activity anticipation via sensory-fusion architecture. *CoRR* 2015;1:1–8 (Sep. 2015).
- [26] Tsioris KM, Pezoulas VC, Zervakis M, Konitsiotis S, Koutsouris DD, Fotiadis DI. A long short-term memory deep learning network for the prediction of epileptic seizures using eeg signals. *Comput Biol Med* 2018;99(1):24–37 (May 2018).