



HAL
open science

A self-attention model for cross-subject seizure detection

Tala Abdallah, Nisrine Jrad, Fahed Abdallah, Anne Humeau-Heurtier, Patrick van Bogaert

► **To cite this version:**

Tala Abdallah, Nisrine Jrad, Fahed Abdallah, Anne Humeau-Heurtier, Patrick van Bogaert. A self-attention model for cross-subject seizure detection. *Computers in Biology and Medicine*, 2023, 165, pp.107427. <10.1016/j.combiomed.2023.107427>. <hal-04192901>

HAL Id: hal-04192901

<https://univ-angers.hal.science/hal-04192901v1>

Submitted on 1 Oct 2025

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

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



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

A Self-Attention model for cross-subject seizure detection

Tala Abdallah^{a,*}, Nisrine Jrad^{a,b}, Fahed Abdallah^c, Anne Humeau-Heurtier^a, Patrick Van Bogaert^{a,d}

^a*Univ Angers, LARIS, SFR MATHSTIC, F-49000 Angers, France, 62 avenue Notre-Dame du Lac, Angers, 49000, France*

^b*University of Catholique de l'Ouest, Angers-Nantes, 49000, France*

^c*University of Lebanon, Beirut, Lebanon*

^d*The Department of Pediatric Neurology, CHU, Angers, 49000, France*

Abstract

Epilepsy is a neurological disorder characterized by recurring seizures, detected by electroencephalography (EEG). EEG signals can be detected by manual time-consuming analysis and recently by automatic detection. The latter poses a significant challenge due to the high dimensional and non-stationary nature of EEG signals. Recently, deep learning (DL) techniques have emerged as valuable tools for seizure detection. In this study, a novel data-driven model based on DL, incorporating a self-attention mechanism (SAT), is proposed. One notable advantage of the proposed method is its simplicity in application, as the raw signal data is directly fed into the suggested network without requiring expertise in signal processing. The model leverages a one-dimensional convolutional neural network (CNN) to extract

*Corresponding author: Tala Abdallah, email: tala.abdallah@etud.univ-angers.fr, phone number: +33656685857

relevant features from EEG signals. These features are then passed through a long short-term memory (LSTM) module to benefit from its memory capabilities, along with a SAT mechanism. The key contribution of this paper lies in the addition of the SAT layer to the LSTM encoder, enabling enhanced exploration of the latent mapping during the encoding step. Cross-subject experiments revealed good performance of this approach with F1-score of 97.8% and 92.7% for binary and five-class epileptic seizure recognition tasks, respectively, on the public UCI dataset, and 97.9% on the CHB-MIT database, surpassing state-of-the-art DL performance. Besides, the proposed method exhibits robustness to inter-subject variability.

Keywords: convolutional neural network (CNN), deep learning (DL), electroencephalography (EEG), epileptic seizure recognition, long short-term memory (LSTM), self-attention (SAT)

1. Introduction

Epilepsy is a long-standing neurological condition characterized by the recurrence of seizures. It is the fourth most common neurological disorder after migraine, stroke and Alzheimer’s disease [1]. Based on the World Health Organization, the disorder affects 50 million individuals globally, with 70 percent of them being able to live seizure-free with proper diagnosis and treatment [2], [3]. However, 30 percent of patients continue to have unpredictable recurrent seizures [4]. This explains the need for new technologies that can detect or even predict seizure occurrences is crucial [5]. Usually,

epileptologists rely on their expertise in order to diagnose seizures by visually examining electroencephalography (EEG). One of the major drawbacks is time and labor intensive because EEG signals are non-stationary and complex data [3]. Hence, automatic seizure detection has become essential.

Recent developments in machine learning techniques (ML), especially deep learning techniques (DL) advancements, have revealed their superiority for automatically learning very powerful features that exceed human-engineered features in a variety of domains, including diagnostic testing, speech recognition, natural language processing, and computer vision [6], [7]. The benefit of DL models is that, unlike the ML approaches, they can automatically extract features from the signal without the need for feature engineering and experts. The Convolutional Neural Networks (CNNs) are used by most of the studies dedicated to epilepsy since they have revealed outstanding efficiency in tackling the classification problem of EEG data [8], [9]. Unfortunately, CNN has a drawback: it cannot remember prior time series patterns that can be significant to the classifier.

To overcome this issue, some researchers proposed to use recurrent neural network (RNN) and its variations (Gated recurrent unit (GRU) and long short-term memory (LSTM)) [10], [11]. These networks are commonly used in time series processing. GRU requires less memory and is faster than LSTM, but LSTM is more accurate when working with datasets with longer sequences, explaining why most researchers used it over GRU. It notably addresses the gradient vanishing problem in the simple RNN and aids in the

learning of long-term relationships, allowing it to more effectively learn the temporal properties of sequential data. Many studies have used hybrid approaches, which combine CNN and RNN or GRU or LSTM to make benefit of the advantages of these methods [12], [13].

The Encoder-Decoder LSTM is a technique that has shown to be quite successful to time series prediction. This architecture is composed of two models: one receives the input sequence and encodes it into a constant length vector, and the other one produces the predicted result after decoding the fixed-length vector. The fact that the encoding step of LSTM must represent the full input sequence as a single vector will lead to loss of information because all information will be compressed. Furthermore, the decoder must decode the given data using only this single vector, which is a difficult problem. Another issue is that it is not possible to prioritize some subsets of the input over others. Attention mechanisms (AT) were proposed to overcome these problems, and were combined with neural networks in the detection of seizures [14], [15].

This paper proposes a model that incorporates three mechanisms: CNN to benefit from its powerful capability of extracting EEG feature, LSTM to benefit from its memory by extracting prior temporal information, and a self-attention (SAT) mechanism. The main difference of the proposed model, when compared to other studies that used LTSM combined to a SAT layer [16], [17], [18], lies in its inclusion of the SAT layer alongside the LSTM encoder and its FC decoder. This SAT layer liberates the latent mapping of

the encoder step by computing an attention map over the entire input sequence, effectively deciding which part to focus on. The AT is incorporated within the encoder/decoder of an LSTM, the proposed model gains the ability to attend to various segments of the input sequence simultaneously. This parallel attention mechanism empowers the model to capture and utilize relevant information from multiple time steps concurrently. Consequently, the proposed model achieves a more comprehensive understanding of the input sequence, leading to more precise and accurate predictions or decisions. The parallel nature of the attention mechanism also enhances the flexibility and the adaptability of the proposed model during processing. It dynamically allocates attention to different parts of the input sequence based on their significance to the specific task at hand. This adaptability allows the model to adapt its focus and optimize its analysis. This will contribute to improve the performance across a wide range of tasks and data variations.

The proposed method is validated on an inter-subject experiment to show its robustness to inter-subject variability.

The paper is organized as follows: Section II goes through the latest state-of-the-art approaches used for epileptic seizures detection. In Section III, the data sets used in this article are presented. The new proposed method is then described in Section IV. Section V shows the outcomes of the EEG data analysis. Finally, Section VI brings the article to a conclusion and sets the stage for future studies.

2. Related Work

The majority of automated seizure detection systems use ML methods with two main stages: feature engineering and classification [19, 20, 21, 22, 23, 24]. Due to the non-stationary nature of EEG signals, the feature extraction step often requires a substantial amount of effort and domain expertise to investigate and evaluate the signals [25]. A main idea to consider is to learn feature automatically and to uncover features straight from data without relying on human expertise.

DL advancements may hold the key to answer this issue. DL is a subclass of ML that uses a signal representation with several hierarchical layers and advanced training algorithms to imitate the structures and activities of the human brain. DL differs from typical ML in that the spatial, spectral, and temporal feature representations are learned immediately from the training data rather than by human presumption, leading to more realistic and efficient signal representation and better results [26]. First, researchers used a single DL method for both extraction and classification. CNN is among the most well-known DL models. CNN can capture a significant amount of features by utilizing multiple filters in the convolutional layers, pooling layers, and fully connected layers, consequently boosting task operational performances [27]. 1D-CNNs are naturally suited to process biological inputs such as EEG for the identification of seizures. Due to fewer parameters, these architectures have a clearer structure, and one complete sweep of them is quicker than CNNs with 2D architecture. The most significant advantage of

1D architectures over 2D systems is the ability to use bigger pooling and convolutional layers [28]. Furthermore, signals are of 1D origin, and converting them to 2D via pre-processing methods may not only result in information loss, but also it could add errors related to filters. 1D designs such as 1D-LeNet, AlexNet, VGGnet, ResNet, and DenseNet are used in the context of 1D space in [29]. Authors in [30] have proposed a paradigm for epileptic seizure classification that combines the capacity to comprehend probabilistic graphical models (PGMs) with improvements in DL. Moreover, a CNN-BP (standing for CNN bipolar) defined by a 1D-CNN architecture was used in [31]. The algorithm is trained to detect epileptic signals from impulses using only one bipolar intracranial channel. Outputs from several bipolar channel inputs are averaged to generate the final detector output.

Since EEG signals are sequential data that have big variations and lengths [32], RNNs and their derivatives are employed in order to remember prior time series patterns. They are commonly used for physiological signals. For instance, authors of [10] have used three-layer LSTMs to extract and classify EEG features for the detection of epileptic seizures. In another study, Tahereh et al. [33] extracted and selected handcraft features then fed them to an LSTM model to detect epilepsy.

Different researchers are currently using hybrid deep learning by integrating multiple deep learning algorithms [34]. Hybrid approaches combine the advantages of different DL methods while minimizing their drawbacks. This will lead to more accurate results and less computationally costly solutions.

Xu et al. [12] proposed a CNN-LSTM model to extract features from signals and automatically recognize and identify epileptic seizures. They obtained high recognition results on the public UCI epileptic seizure recognition dataset [35]. In [36], researchers introduced a RNN-BiLSTM model to identify epileptic seizures, utilizing the RNN for feature extraction and the BiLSTM for classification. In a different study [37], authors proposed a model called 1D-DSCNN-LSTMs for seizure detection. Their model incorporates a depthwise separable convolutional neural network and long short-term memory networks, for autonomously extracting temporal features from raw EEG data. Both approaches were applied to the same dataset, showcasing different techniques for seizure detection and highlighting the diversity of methodologies used in the field. A hybrid DL model that integrates a DenseNet with an LSTM was used on CHB-MIT scalp EEG dataset [13]. Other studies used CNN to extract spatial information followed by LSTM to extract temporal information and classify EEG data [38]. In [39], researchers utilized wavelet transforms to filter EEG signals, followed by a Bi-GRU model for seizure detection on patient specific study applied to the CHB-MIT dataset. In another study [40], a novel approach merged a 2D-CNN and an LSTM to automatically classify epileptic seizures using signal images from the CHB-MIT dataset. More recently, authors of [41] explored three data augmentation (DA) techniques to increase the percentage of seizures within the dataset. Subsequently, they employed the widely used CNN-LSTM model to analyze the augmented data. Results are very promising.

All these methods show good results but none of them have the ability to prioritize some subsets of the input over others. Besides, all information is treated equally.

Recently, some researchers have included the mechanism of AT in their pipeline. The use of virtual sample generation with an AT mechanism to detect very high frequency rhythms (ripples) is presented in [14]. Another study examined the spatial and time-frequency domain to extract EEG features using a 3D residual AT model [15]. However, to the best of our knowledge, none of these studies have employed the SAT model to address the encoder-decoder LSTM’s constant length problem. As a result, the hypothesis emerged that adding the SAT mechanism to the encoder of the LSTM could be more efficient in detecting epileptic seizures. This would allow the model to focus on a large portion of the inputs over other inputs by releasing the internal fixed-length encoder-decoder architecture representation.

3. Materials and Method

3.1. Datasets Description

Two datasets were employed in this study: the public UCI epileptic seizure recognition [35] and the Children’s Hospital Boston–Massachusetts Institute of Technology (CHB–MIT) [42].

3.1.1. The public UCI epileptic seizure recognition dataset

There are five separate folders in the original epileptic seizure recognition dataset, each with 100 files. Each file is a recorded sample of one subject’s

brain activity, and contains 4097 data points recorded for 23.6 seconds. Electrodes were placed on the scalp and the EEG signals were captured using the 10–20 International system.

In total, this dataset contains the recordings of 10 participants (5 normal controls and 5 patients with temporal lobe epilepsy related to hippocampal sclerosis). Each of the 4097 data points in each sample is separated into 23 data chunks, each comprising 178 data points of 1 second. So the sampling frequency is 178 Hz here. The 23 data pieces are then jumbled. In total, 11500 samples were acquired. The number of electrodes used is not mentioned.

The UCI dataset contains five conditions: one epileptic seizure condition and four non-epileptic conditions in which patients do not have epileptic seizures. Each condition has 2300 data points. The five conditions are:

- (1) Epileptic seizure condition: recordings of EEG signal obtained during a seizure of an epileptic patient.
- (2) First non-epileptic condition: authors pinpointed the epileptic focus of the epileptic patients and recorded EEG activity from the healthy brain regions.
- (3) Second non-epileptic condition: recordings of EEG signals collected from the location of the epileptic focus of the epileptic patients.
- (4) Third non-epileptic condition: open eyes EEG signals from normal subjects.
- (5) Fourth non-epileptic condition: closed eyes EEG signals from normal

subjects.

The dataset is detailed in Table 1.

Table 1: The detailed of the public UCI epileptic seizure recognition data set.

Data set	The public UCI
Type	Scalp EEG
Subjects	10 (5 epileptic and 5 normal)
Sampling Rate	256 Hz
EDF File For Each Patient	50
Electrode Positions	The international 10-20 method

3.1.2. CHB-MIT dataset

This dataset contains EEG recordings from pediatric patients with intractable seizures that were gathered at Children’s Hospital Boston [42]. Following the cessation of anti-seizure medication, subjects were recorded for up to several days to characterize their seizures and determine their suitability for surgical intervention. 22 patients (5 boys, ages 3–22; and 17 girls, ages 1.5–19) provided recordings, which were divided into 23 cases (case chb21 was taken from the same patient 1.5 years after case chb01). The chb24 case was added to this collection later in December 2010, his/her age and gender are not precise. Each patient has between 9 to 42 continuous .edf files. Most of EDF files contain precisely one hour of signals. All signals were sampled at a rate of 256 Hz, with a resolution of 16 bits. Expert opinions and judgments were used to designate the beginning and end times of epileptic seizures. The majority of the files contain 23 EEG electrodes. The International 10-20

method of EEG electrode positions using longitudinal bipolar montages was used [43].

On average, each patient had 5.75 seizures of different durations. The dataset is detailed in Table 2.

Table 2: The detailed of the CHB-MIT dataset.

Data set	CHB-MIT
Type	Scalp EEG
Subjects	23
Male	5
Female	17
Age	between 1.5-22
Electrodes	23
Sampling Rate	256 HZ
EDF File For Each Patient	between 9-42
Average number of seizure per Patient recording	5.75
Electrode Positions	644 Hours
	The international 10-20 method

3.2. The Proposed Method Architecture

The proposed model comprises an input layer, four 1D-CNN layers, one pooling layer, one LSTM layer mixed with one SAT layer, four FC layers, and a softmax output layer (see Fig. 1). In a simplified Encoder-Decoder design, the conventional decoder relies solely on the final output of the encoder stage, which contains compressed information, to initiate the generation of predictions. To enhance this process, the Self-Attention (SAT) mechanism is introduced, allowing the decoder to consider all hidden states from the encoder sequence for making predictions. In this approach, the LSTM and

SAT layers were combined in the encoding/decoding stage. Specifically, each SAT neuron is interconnected with all neurons in the LSTM layer, facilitating the reception of information from all LSTM neurons and enabling the computation of corresponding attention weights.

The key concept behind the SAT layer is to enable the model to selectively focus on different segments of the input sequence by assigning distinct attention weights to the LSTM neurons. These attention weights are calculated based on the individual characteristics of each LSTM neuron, typically represented by learned weights within the model. Consequently, the SAT mechanism empowers the model to dynamically adjust its attention, giving prominence to relevant parts of the input sequence during the encoding process.

In this architecture, the decoding stage is performed using a fully connected (FC) layer, where the final predictions are generated based on the accumulated knowledge and attention weights computed during the encoding phase. This integration of LSTM, SAT, and FC layers enhances the model’s ability to comprehend and process input sequences effectively, contributing to improved performance and accuracy in our proposed method. The hyperparameters and the parameters of the model are represented in tables 3 and 4. Hyperparameter tuning involves systematically exploring different combinations of hyperparameter values and evaluating their impact on the model performance metrics. By adjusting the values of the hyperparameters, the model performance is improved and its generalization ability

is enhanced. Grid search is used to learn parameters and hyperparameters.

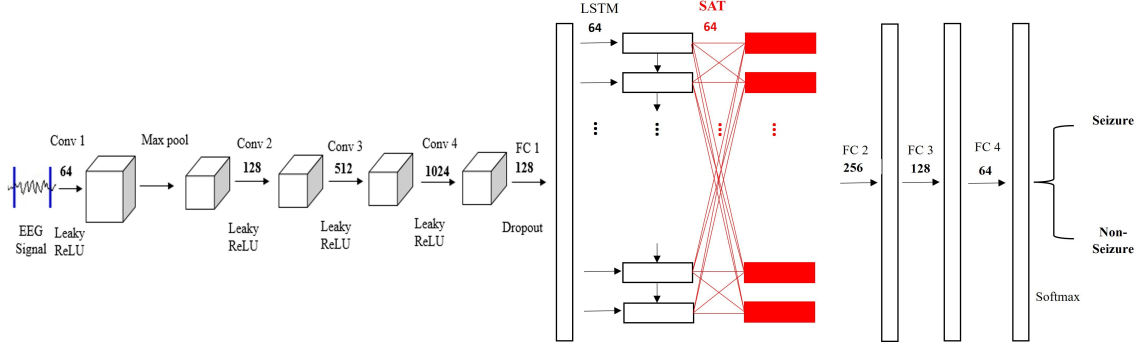


Figure 1: The Architecture of the CNN-LSTM-SAT model.

Table 3: The hyperparameters of the 1D CNN-LSTM-SAT model.

Hyperparameters	Hyperparameter tuning	Best values
CNN Layers	1,2,3,4	4
LSTM Layers	1,2,3,4	1
Number of neurons in the first CNN Layer	64,128,1024	64
Number of neurons in the first LSTM Layer	64,128	64
Kernel size	$2 \times 1, 3 \times 1, 5 \times 1$	3×1
Activation function	ReLu, LeakyReLu	LeakyReLu
Optimizer	Adam, RAdam [44]	Adam
Dropout	0.2, 0.3, 0.5	0.5

Table 4: The parameters of the 1D CNN-LSTM-SAT model.

Layer Type	Number of Neurons	Kernel Size	Dimensions in the public UCI data	Dimensions in the CHB-MIT data
Input	-	-	(178, 1)	(256, 1)
CNN+ Leaky Relu	64	3×1	(None, 176, 64)	(None, 254, 64)
Max Pooling	-	2×1	(None, 88, 64)	(None, 127, 64)
CNN+ Leaky Relu	128	3×1	(None, 86, 128)	(None, 125, 128)
CNN + Leaky Relu	512	3×1	(None, 84, 512)	(None, 123, 512)
CNN+ Leaky Relu	1024	3×1	(None, 82, 1024)	(None, 121, 1024)
FC	128	-	(None, 82, 128)	(None, 121, 128)
LSTM	64	-	(None, 82, 64)	(None, 121, 64)
AT	64	-	(None, 64)	(None, 64)
FC	256	-	(None, 256)	(None, 256)
FC	128	-	(None, 128)	(None, 128)
FC	64	-	(None, 64)	(None, 64)
Softmax	-	-	(None, 2)	(None, 2)

4. Theoretical Foundations

4.1. 1D-CNN

The 1D-CNN can extract the pertinent and useful features from 1D time-series sequence data by performing 1D convolutional filters. In this study, the convolutional filters and the feature maps of the 1D-CNN are fully one-dimensional, matching the one-dimensional characteristic of raw EEG signal data without losing any information. The CNN can progressively extract higher-level features that are resilient and discriminating for tasks involving classifying epileptic seizures. To reduce the dimensionality, pooling is commonly performed after a convolution procedure. This limits the number of parameters, hence shortening the training period and avoiding overfitting. The 1D convolutional operation is described mathematically as follows:

$$\sigma \left(\sum_{i=1}^{N_{l-1}} \text{conv1D}(w_{i,j}^{l-1}, s_i^{l-1}) + b_j^l \right) = x_j^l \quad (1)$$

where s_i^{l-1} is the i th feature map of a signal in the $(l-1)$ th layer, x_j^l is the j th feature map in the l th layer, $w_{i,j}^{l-1}$ is the trainable convolutional kernel, N_{l-1} is the number of feature mappings in the $(l-1)$ th layer, and the conv1D is the 1D convolution operation. The j th feature map bias in the l th layer is represented by b_j^l , and σ is an activation function.

4.2. LSTM Architecture

LSTMs were created specifically to address the issue of long-term dependency. At step t in time, the input, hidden state, and cell state are x_t , h_t , and C_t , respectively. The LSTM computes the h-sequence (h_1, h_2, \dots, h_m) and the C-sequence (C_1, C_2, \dots, C_m) from a sequence of inputs (x_1, x_2, \dots, x_m), where m is the number of units in the LSTM layer. The composition of the LSTM is as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (2)$$

$$z_i = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (3)$$

$$c_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \quad (4)$$

$$C_t = f_t \odot C_{t-1} + z_i \odot c_t, \quad (5)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (6)$$

$$h_t = o_t \odot \tanh(C_t), \quad (7)$$

where the function σ is a logistic sigmoid, \odot is an element-wise multiplication, C_t represents the current state of the cell, while W_f , W_i , W_c , W_o , W' and b_f , b_i , b_C , b_o are respectively the weights and bias model parameters that may be learned. Furthermore, f_t , z_i , c_t , and o_t are also known as the forget gate, input gate, cell state gate, and output gate, respectively. Each LSTM unit contains a memory cell with the state C_t at time t that is controlled by

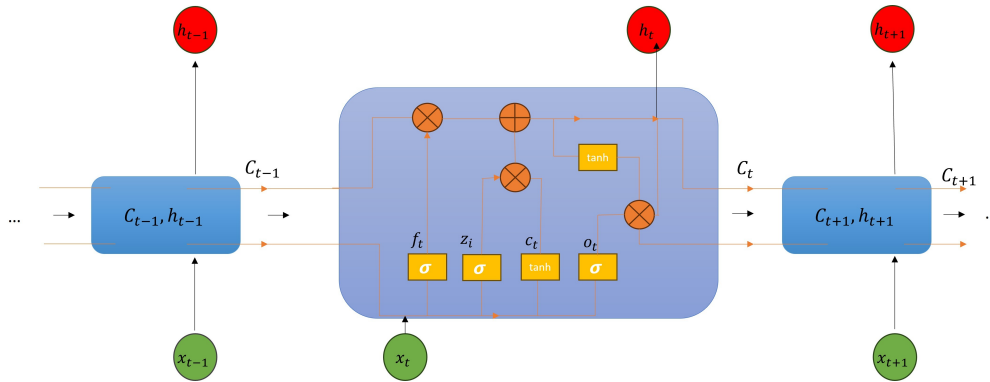


Figure 2: The architecture of the LSTM model.

the gates mentioned above. The architecture of the LSTM model is illustrated in Fig. 2.

A potential drawback of the LSTM encoder is that a neural network compresses all of the essential information from an input sequence into a fixed-length vector. As a result, long series may be difficult for the neural network to deal with. Another issue is that there is no ability to prioritize certain of the inputs over others.

4.3. The Self-Attention Mechanism

In natural language processing (NLP), AT mechanism outperforms the encoder decoder-based neural machine translation system. This approach and its adaptations were recently applied in numerous applications such as computer vision, signal processing, and others [45].

To overcome the shortcoming of LSTM, a SAT layer is implemented in the encoding stage of the model. At each time step, the SAT architecture pays attention to every hidden state from each encoder node and then generates

predictions depending on which one is more informative. Using a neural network, it learns which hidden states to attend to and by how much using a neural network. In other words, the SAT mechanism takes into account various inputs as the LSTM analyzes each one and determines which ones are significant by giving those inputs various weights. The encoded input and the decoder will receive the weights provided by the SAT mechanism. Technically, the output h_t of the LSTM layer is the input of the SAT layer. The SAT layer first maps h_t to h'_t in the first component, as follows:

$$h'_t = f'(h_t), \quad (8)$$

where h'_t is the hidden state at time t and f' is a scoring function that maps the LSTM output to a scalar score.

In the second part, the attention scores, denoted as e_t , are calculated by combining the h'_t vectors using a weight vector. These scores represent the relative importance or relevance of each h'_t vector:

$$e_t = v^T \cdot \tanh(h'_t \cdot W' + b'), \quad (9)$$

where the vector v is a learnable parameter of the model, W' is the weight matrix and b' is the bias vector of the SAT layer.

Softmax is used to normalize these elements as follows:

$$\beta_t = \text{softmax}(e_t) = \frac{\exp(e_t)}{\sum_{i=1}^m \exp(e_i)}, \quad (10)$$

where β_t is a SAT weight, which is a score that indicates how much focus must be given to the latent encoder mapping. The SAT model final output at time t , i.e. the weighted input feature sequence Z_t , may alternatively be obtained as follows:

$$Z_t = (\beta_1 h_1, \beta_2 h_2, \dots, \beta_m h_m)^T. \quad (11)$$

Thus the job of transforming raw time series to SAT-based time series is ultimately completed. The architecture of the SAT model is illustrated in Fig. 3.

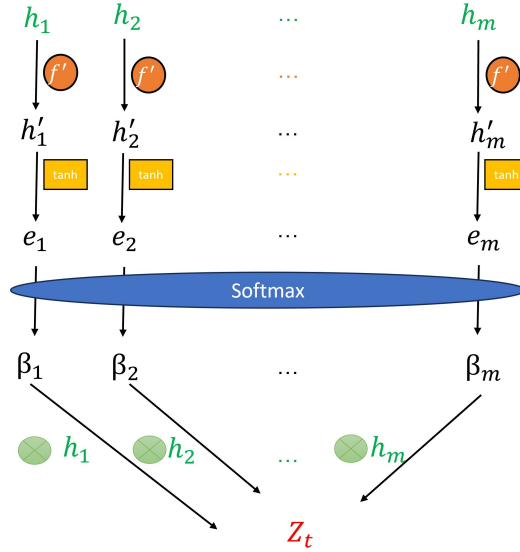


Figure 3: The architecture of the self-attention model.

5. Methodology

5.1. Datasets Preprocessing

5.1.1. The public UCI epileptic seizure recognition dataset

The original dataset has been reorganized and processed by a UCI official as described above. Thus, no preprocessing step was required.

5.1.2. CHB-MIT dataset

The preprocessing step of the CHB-MIT dataset was performed using the Matlab programming language.

The same patients as [46] were chosen for this study. Because seizures are rare, the number of positive observations should be raised. To avoid an extremely imbalanced dataset, for each file, a length window of 1 second was moved along the time axis for each EEG signal, with 30% overlap when there was no seizures and 90% overlap during seizures to avoid an extremely imbalanced dataset as illustrated in Fig. 4. The use of a one second sliding window on EEG data was approved by the state-of-the-art studies. Indeed, one second seems to be an appropriate time interval for epileptic seizures since epileptic seizures last from seconds to minutes. Therefore, the average percentage of seizures time relative to the total seizure and non-seizure time increases from 2% to 15.65%. Finally, a total of 7,703,262 segmented signals were obtained, with 6,497,064 labeled 0 for the seizure-free segment and 1,206,198 labeled 1 for seizure.

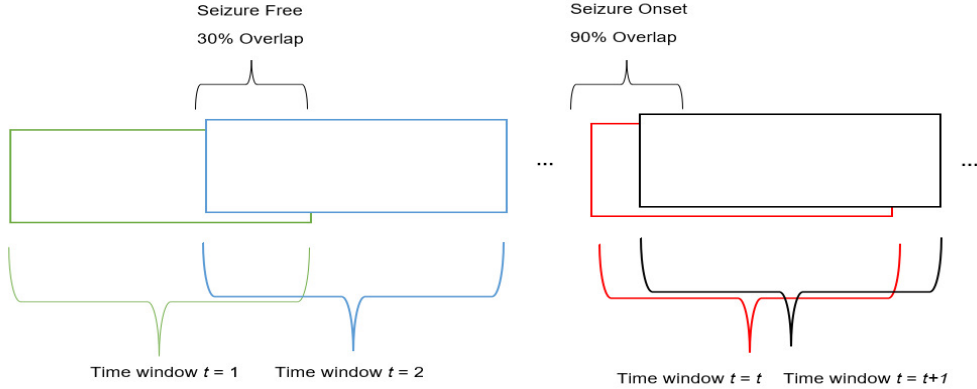


Figure 4: Segmenting the signal of EEG along the time axis with 30% overlap for non-seizure and 90% for seizure.

5.2. Experimental Setup

Throughout the experiment, each dataset was split per patient into training, validation and testing sets with percentages of 80%, 10% and 10% respectively. The number of training epochs was 100, and the batch size was 32. Adam’s optimizer with a 0.001 learning rate and the categorical cross entropy loss function were used here. Therefore the data was converted into a categorical format before being used for multi-class classification. The data was re-sampled 5 times according to the percentage cited above, and the training-test process of each model was repeated 5 times to assess the generalization ability. The mean and the standard deviation of the performance metrics were reported and compared for the different models. The models were created using the Python, as well as a variety of supporting libraries, including the Keras DL API from the Tensorflow ML toolbox.

This research looks at two forms of epileptic seizure recognition challenges: binary and five-class recognition tasks. The binary task consists of discriminating the epileptic seizure conditions from the four normal condition of the public UCI epileptic seizure recognition data and to discriminate seizure and free seizure segments of the CHB–MIT dataset. The five-class task consists of discriminating the five conditions cited in section 3.1 of the public UCI epileptic seizure dataset.

5.3. Evaluation Metrics

In this section, the accuracy, precision, recall, and F1-score metrics are derived to further analyze the seizure classification performance of our model.

The formulas of these metrics are as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad (12)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (13)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (14)$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (15)$$

where TP and FN correspond to the count of accurately and inaccurately classified seizures for a particular seizure recognition task, respectively. TN represents the number of correctly classified non-seizure recognition tasks for

a specific class. FP indicates the instances when a non-seizure recognition task is wrongly classified as belonging to seizure class. The F1-score, serving as a consolidated metric of precision and recall, is mathematically defined as the harmonic average of these two scores.

5.4. Statistics

Non-parametric Wilcoxon-signed rank tests are also applied to ensure the performance of the findings. It measures the differences between the results of two classifiers. Tests with p-values less than 0.05 were considered significantly different and are marked with ”*”.

5.5. Patient Specific

The performances of the proposed model were tested on each of the nine patients of the CHB–MIT dataset, in order to see if the model is able to give good performance on patients with a low percentage of seizures. The seizure rate is the ratio between the total seizure and non-seizure segments after the overlapping sliding window (see section 5.1.2) for each patient. It therefore shows how unbalanced the dataset is.

6. Results and Discussion

6.1. Experimental Results

To measure the generalization ability of the proposed algorithm, each dataset was re-sampled 5 times into a training, validation and testing sets. The mean and standard deviation of the results on the testing set were reported (as cited above in section 5.2). Comparisons of the model with other DL methods (Deep Neural Network (DNN), CNN, LSTM, CNN-LSTM, and CNN-GRU) are presented in Tables 5, 6 and 7. A DNN is a form of neural network consisting of at least two layers. The GRU is similar to an LSTM with a forget gate, but it misses an output gate, hence it has lower variables. GRUs have been shown to work better on some small and less common datasets. The proposed model outperforms the results obtained using all five methods in the binary classification as well as in the five class recognition. The obtained p-values are all less than 0.05. This shows the tremendous promise of our findings in the area of epileptic seizure research using EEG. The average training time per epoch for each model was also calculated, taking into account the specific hardware configuration of the experimental setup. The experiments were conducted on a processor of 11th Gen Intel(R) Core(TM) i9-11950H @ 2.60GHz 2.61 GHz with 64 GB of RAM. As noticed, the CNN-LSTM-AT model has significantly reduced the training time compared to the CNN-LSTM model.

Finally, as for the patient specific analysis, according to Table 8, the best results refer to patient 11 who has the highest seizure rate. However,

Table 5: Comparison of the performances between state of the art models and the proposed 1D CNN-LSTM-SAT model for the binary classification task on the public UCI dataset. Wilcoxon sign-rank tests with p-values less than 0.05 are considered significantly different and are marked with ”*”.

Methods	Accuracy	Precision	Recall	F1-Score	Average time per epoch (seconds)
DNN	90.3%±0.44*	89.8%±0.40*	89.5%±0.44*	89.2% ±0.46*	1.3
CNN	91.8%±0.50*	91.3%±0.58*	91.2%±0.56*	90.4%±0.52*	3.6
LSTM	90.2%±0.38*	92.4%±0.30*	92.5%±0.30*	89.0%±0.36*	9.8
CNN-GRU	95.5%±0.03*	94.5%±0.04*	94.5%±0.03*	95.5%±0.04*	7.3
CNN-LSTM	95.0%±0.05*	94.3%±0.05*	94.2%±0.04*	95.0%±0.04*	8
The CNN-LSTM-SAT model	98.5%±0.03	98.0%±0.02	97.7%±0.03	97.8%±0.03	6

Table 6: Comparison of the performances between state of the art models and the proposed 1D CNN-LSTM-SAT model for the five-class classification task on the public UCI dataset. Wilcoxon sign-rank tests with p-values less than 0.05 are considered significantly different and are marked with ”*”.

Methods	Accuracy	Precision	Recall	F1-Score	Average time per epoch (seconds)
DNN	76.9%±0.66*	78.7%±0.69*	77.6%±0.65*	77.7%±0.060*	1.3
CNN	77.0%±0.44*	77.3%±0.40*	77.7%±0.40*	76.7%±0.42*	3.6
LSTM	77.8%±0.50*	78.2%±0.56*	77.0%±0.51*	77.2%±0.53*	9.8
CNN-GRU	85.4%±0.05*	84.4%±0.06*	84.4%±0.05*	84.5%±0.05*	7.3
CNN-LSTM	85.2%±0.06*	83.6%±0.07*	83.7%±0.07*	84.2%±0.06*	8
The CNN-LSTM-SAT model	93.0%±0.03	92.6%±0.03	92.9%±0.03	92.7%±0.03	6

Table 7: Comparison of the performances between state of the art models and the proposed 1D CNN-LSTM-SAT model for the binary classification task on the CHB-MIT dataset. Wilcoxon sign-rank tests with p-values less than 0.05 are considered significantly different and are marked with ”*”.

Methods	Accuracy	Precision	Recall	F1-Score	Average time per epoch (seconds)
DNN	81.9%±0.50*	80.2%±0.55*	80.3%±0.50*	80.3% ±0.53*	864.5
CNN	84.4%±0.32*	82.5%±0.48*	82.5%±0.40*	82.5%±0.36*	1776
LSTM	78.3%±0.02*	76.6%±0.03*	77.0%±0.02*	76.8%±0.02*	3680
CNN-GRU	79.9%±0.03*	78.5%±0.02*	78.4%±0.02*	78.4%±0.02*	2998
CNN-LSTM	94.8%±0.03*	93.6%±0.04*	93.6%±0.03*	93.6%±0.04*	3320
The CNN-LSTM-SAT model	98.9%±0.03	97.1%±0.03	98.7%±0.03	97.9%±0.3	2353

it is worth noting that the model still achieved good performance on the other patients, even when they had a low seizure rate (4.24% for patient 14). Additionally, the mean, standard deviation, as well as the first, second, and third quartiles, were calculated for all the results obtained from the nine patients.

Table 8: The performances of the proposed 1D CNN-LSTM-SAT model on patient specific approach on the CHB-MIT Hospital dataset.

Patients	Accuracy	Precision	Recall	F1-Score	Seizure Rate
P5	98.4%	98.00%	98.00%	98.00%	18.00%
P8	98.53%	98.50%	98.00%	98.24%	26.77%
P11	98.96%	98.56%	98.57%	98.56%	26.92%
P12	96.23%	95.50%	95.55%	95.52%	15.55%
P13	94.75%	93.53%	93.50%	93.51%	10.83%
P14	85.36%	84.50%	84.00%	84.24%	4.24%
P15	98.11%	97.88%	97.85%	97.86%	20.92%
P23	90.33%	88.00%	87.50%	87.74%	8.07%
P24	90.46%	86.50%	88.00%	87.24%	7.32%
Mean	94.57%	93.44%	93.44%	93.43%	15.40%
Standard Deviation	4.80%	5.64%	5.54%	5.58%	8.39%
Q1	90.46%	88.00%	88.00%	87.74%	8.07%
Q2	96.23%	95.50%	95.55%	95.52%	15.55%
Q3	98.40%	98.00%	98.00%	98.00%	20.92%

6.2. Discussion

This work introduces CNN-LSTM-SAT, an original hybrid DL approach using EEG data, to identify epileptic seizures.

This model is able to learn consistent and appropriate EEG features from the beginning to the end with no need of customized feature engineering. By incorporating a SAT layer into the architecture, the fixed-length encoder-decoder representation is transformed into a more flexible and dynamic one. This addition allowed the model to selectively emphasize on certain parts of the input sequence, giving it the capability to prioritize and focus on the most relevant information while downplaying less significant aspects.

The model is tested on two datasets. The first is the public UCI epileptic seizure recognition. In the latter, data has already been preprocessed, thus the model was applied directly to it, addressing both binary and five-class

recognition tasks for epileptic seizure identification. According to Tables 5 and 6, the approach yielded significantly better results compared to other existing advanced methods. The DNN had the lowest results, explaining why it is the least used generally in this field compared to the others.

The CNN-LSTM-SAT model exhibited comparatively lower performance in the five-class classification task as opposed to the binary classification task. The confusion matrices in tables 9 and 10 provide a comprehensive analysis of this behavior. This addition serves to enhance the understanding of the model performance and its limitations in accurately distinguishing between the aforementioned normal classes. The disparity can be attributed to its occasional confusion between the first two non-epileptic cases, case 2: “recordings from EEG activity from the healthy brain regions of the patient” and case 3 “recordings of EEG signals from the location of the epileptic focus of the epileptic patient”, and the last two normal cases, case 4 “eyes opened” and case 5 “eyes closed”. The reason behind this confusion is that the EEG patterns observed during instances of these normal cases exhibit notable similarities, posing a challenge for the model to accurately discern them. Furthermore, the limited availability of the training data might affect the capacity of the model to effectively learn the distinguishing patterns and make it more difficult. It is worth noting that these errors do not occur between seizure and seizure-free cases or between epileptic and normal patients. Therefore, their impact may be considered relatively minor in terms of clinical perspectives. Moreover, Fig. 5 displays the accuracy and loss

curves, providing an insightful representation of the model performance during training. The accuracy curve visually demonstrates the model’s evolving accuracy for each task as training progresses, indicating its improvement over time. On the other hand, the loss curve illustrates the value of the loss function at different stages of training. Typically, as the model learns from the data and refines its parameters, the accuracy curve tends to ascend, reflecting enhanced performance, while the loss curve tends to descend, indicating a reduction in the model’s error over successive epochs. This consistent trend of increasing accuracy and decreasing loss underscores the effectiveness of the training process, highlighting the model’s ability to optimize and better fit the data throughout its training duration.

Table 9: The confusion matrix for the binary classification on the public UCI dataset.

		Actual	
		Non-Epileptic case	Epileptic case
Predicted	Non-Epileptic case	901	8
	Epileptic case	9	232

Table 10: The confusion matrix for the five class recognition task on the public UCI dataset.

		Actual				
		Epileptic case: 1	Non-Epileptic case: 2	Non-Epileptic case: 3	Non-Epileptic case: 4	Non-Epileptic case: 5
Predicted	Epileptic case: 1	229	2	1	0	1
	Non-Epileptic case: 2	3	259	23	1	3
	Non-Epileptic case: 3	0	18	151	1	2
	Non-Epileptic case: 4	1	0	1	169	17
	Non-Epileptic case: 5	0	3	2	2	251

Although the public UCI dataset has been used in several studies [47], [12], there is still some missing information (age of patients, their sex, the

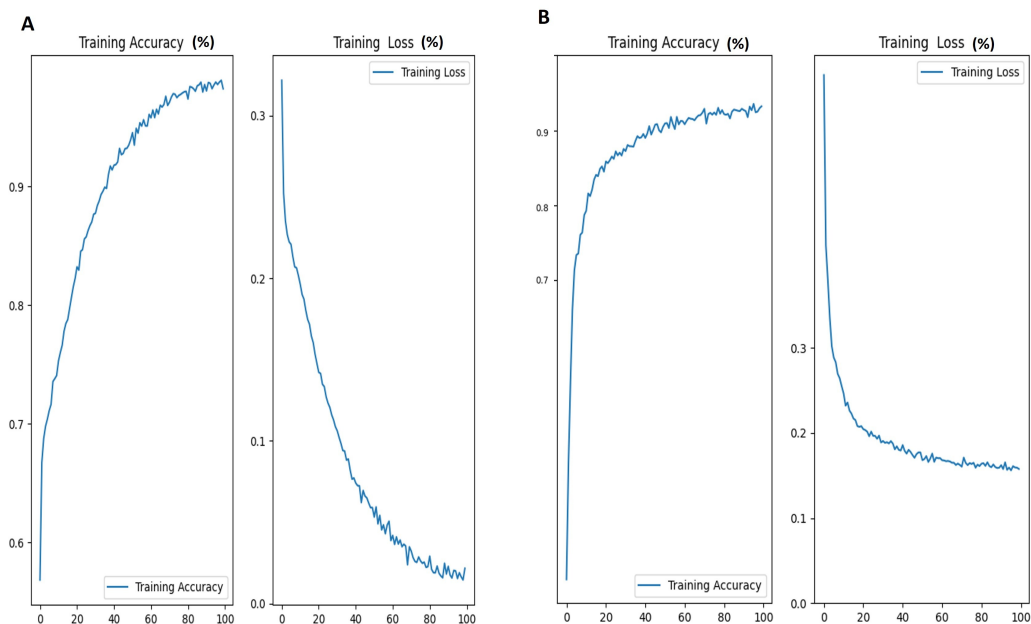


Figure 5: The training accuracy and loss of the proposed model: (A) for the binary classification and (B) for the five-class recognition on the public UCI dataset.

electrodes used...). Additionally, the dataset is not very imbalanced between seizure and non-seizure segments, which does not reflect real life conditions of long-term EEG recording. This is why a second database was used: the CHB-MIT dataset.

As shown in Table 7, the superiority of the proposed model compared to other models on this dataset for the binary classification was confirmed. Here the lowest result refers to the CNN-GRU model. This corroborates with the state of the art findings that show that GRU gives bad results when working on large dataset.

Accuracy, precision, recall and F1-score were used to evaluate the proposed model and they all gave significantly higher values than the other models.

Because of the highly imbalanced numbers of seizure and non-seizure frames, the accuracy is not a reliable performance metric in studies of seizure detection. However it was calculated to enable comparison with existing results. Since the accurate classification of seizure frames is of utmost importance, the recall or sensitivity rate becomes the primary performance metric. The proposed model exhibited a high level of accuracy as it achieved a consistently high recall rate, indicating a low number of false negatives and a strong ability to correctly identify positive instances in all cases.

The confusion matrix in Table 11 provides a comprehensive assessment of the CNN-LSTM-SAT model’s predictive ability by comparing its predictions with the assigned labels in the CHB-MIT dataset. Additionally, the accuracy and loss curves (shown in Fig. 6) illustrate the model’s improved classification performance over time, demonstrating a decreasing trend in error as the model undergoes training. Moreover, the proposed model is robust to

Table 11: The confusion matrix on the CHB-MIT dataset.

		Actual	
		non-epileptic case	Epileptic case
Predicted	Non-Epileptic case	647,382	6808
	Epileptic case	1809	114328

inter-subject variability, since it is easily adaptable to several subjects at the same time, making it usable in the service of medicine. On top of all this, the strength of the proposed model appears not only over all patients but also on each patient of the CHB-MIT dataset individually as shown in Table 8. The model gives good results even on patient 14 who has the lowest percentage

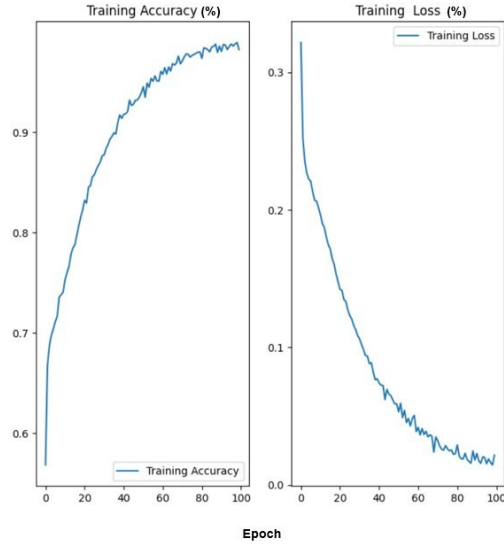


Figure 6: The training accuracy and loss of the proposed model on the CHB-MIT dataset.

of seizures (4.24%). It is also crucial to keep in mind that that the model performed well even though the datasets used were quite imbalanced.

In Table 12, a comprehensive comparison of the proposed CNN-LSTM-SAT model with other DL models used on the public UCI dataset is conducted for both binary and five-class recognition tasks. Notably, the CNN-LSTM-SAT model consistently exhibited superior performance in the five-class recognition tasks, outperforming alternative approaches. In the binary task, the model demonstrated comparable performance to existing works, showcasing its competitiveness. The effectiveness of the CNN-LSTM-SAT model is further validated by its low standard deviation across multiple runs. This observation indicates that the model’s performance remains highly sta-

ble and reproducible, adding to its reliability and robustness. The reliable and consistent outcomes reinforce the credibility of the proposed model and its potential for practical applications in seizure recognition.

Table 12: Comparison of existing deep learning approaches results and the CNN-LSTM-SAT results on the public UCI dataset.

Authors	Year	Method	Task	Accuracy	Precision	Recall	F1-Score
Gaowei et al.[12]	2020	1D-CNN-LSTM	5 class recognition	82.0%	81.7%	81.7%	81.5%
Zhentaο et al.[37]	2022	DSCNN-2LSTMs	5 class recognition	81.3%	79.2%	79.5%	79.5%
Nagwan et al.[36]	2022	RNN-BiLSTM	Binary	98.90%	98.2%	98.5%	-
This work	2023	CNN-LSTM-SAT	5 class recognition	93.0%±0.03	92.6%±0.03	92.9%±0.03	92.7%±0.03
This work	2023	CNN-LSTM-SAT	Binary	98.5%±0.03	98.0%±0.02	97.7%±0.03	97.8%±0.03

Furthermore, when compared to other existing DL methods tested on the CHB-MIT dataset, the performances of the CNN-LSTM-SAT model are better (see Table 13).

Table 13: Comparison of existing deep learning approaches results and the CNN-LSTM-SAT results on the CHB-MIT scalp EEG dataset.

Authors	Year	Method	Accuracy	Precision	Recall	F1-Score
Truong et al.[48]	2018	CNN	-	-	87.8%	-
Ozcan et al.[49]	2019	3D CNN	-	-	81.2%	-
Syed Muhammad Usman[50]	2020	CNN-SVM	-	-	92.7%	-
Sanguk Ryu[13]	2021	Dense-LSTM	93.3%	-	92.9%	92.3%
Yanli et al.[39]	2022	Bi-GRU	98.5%	-	93.9%	-
Muhammet Varli[40]	2023	2D CNN-LSTM	96.2%	94.8%	98.2%	96.5%
Tala et al.[41]	2023	DA-CNN-LSTM	98.7%	97.0%	96.8%	96.9%
This Work	2023	CNN-LSTM-SAT	98.9%	97.1%	98.7%	97.9%

Finally, Table 14 shows that the model not only outperforms other existing DL models but also traditional ML methods in the detection of epileptic seizures.

Table 14: Comparison of existing machine learning approaches results and the CNN-LSTM-SAT results on the CHB-MIT scalp EEG dataset.

Authors	Year	Method	Accuracy	Precision	Recall	F1-Score
Birjandtalab et al.[51]	2017	Random Forest + KNN	-	-	80.8%	56.2
Syed Muhammad Usman[52]	2017	SVM + EDM	92.2%	-	-	-
Jacobs[53]	2018	Random Forest	82.4%	-	87.9%	-
Mohammed Khubeb[25]	2020	SVM, ANN and KNN	98.3%	-	-	-
Zarqa et al.[24]	2023	Ensemble learning-based classifiers	91.0%	85.9%	98.0%	91.5%
This Work	2023	CNN-LSTM-SAT	98.9%	97.1%	98.7%	97.9%

7. Conclusion and Future Work

DL has recently been used to identify epileptic seizures. In this research, a new approach was implemented for the classification of epileptic seizures using EEG data analysis. Based on the achieved outcomes, the suggested method outperforms standard DL methods. The inclusion of a SAT layer addresses the flaw in the previous models and shows that the proposed model is one of the most effective seizure classification methods on the datasets processed. The model exhibited a low error rate as it showed a high true positive rate (sensitivity) and a high precision across all cases. Its robustness to inter-subject variability is also confirmed by its ability to perform well on multiple subjects simultaneously. Once the model is trained, the testing procedure becomes highly efficient and quick making the model an essential tool in the field of medicine.

In the future, the CNN-LSTM-SAT model will be used on a home-made dataset of EEG signals from long-term recordings of epileptic children with various types of epileptic seizures.

8. Data & Code Availability:

The public UCI epileptic seizure recognition dataset [35] and the CHB-MIT dataset [42] are both publicly available and can be downloaded online. Every source code will be available online after acceptance and publication of the article.

9. Competing interests

None declared

10. Funding

None

11. Ethical approval

Not required

References

- [1] M. Ikram, S. Rasheed, Genetics of dravet syndrome and its targeted therapy by nanomedicine: A roadmap for future treatment of drug resistant seizures., *Current Molecular Pharmacology* (2022).
- [2] Y. Yuan, G. Xun, K. Jia, A. Zhang, A multi-view deep learning method for epileptic seizure detection using short-time fourier transform, in: *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, 2017, pp. 213–222.

- [3] I. Ullah, M. Hussain, H. Aboalsamh, et al., An automated system for epilepsy detection using eeg brain signals based on deep learning approach, *Expert Systems with Applications* 107 (2018) 61–71.
- [4] V. Gabeff, T. Teijeiro, M. Zapater, L. Cammoun, S. Rheims, P. Ryvlin, D. Atienza, Interpreting deep learning models for epileptic seizure detection on eeg signals, *Artificial Intelligence in Medicine* 117 (2021) 102084.
- [5] I. Hubbard, S. Beniczky, P. Ryvlin, The challenging path to developing a mobile health device for epilepsy: The current landscape and where we go from here, *Frontiers in Neurology* 12 (2021).
- [6] X. Wang, Y. Zhao, F. Pourpanah, Recent advances in deep learning, *International Journal of Machine Learning and Cybernetics* 11 (4) (2020) 747–750.
- [7] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, Y. Zhang, Short-term residential load forecasting based on lstm recurrent neural network, *IEEE Transactions on Smart Grid* 10 (1) (2017) 841–851.
- [8] A. R. Johansen, J. Jin, T. Maszczyk, J. Dauwels, S. S. Cash, M. B. Westover, Epileptiform spike detection via convolutional neural networks, in: *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2016, pp. 754–758.
- [9] R. Zuo, J. Wei, X. Li, C. Li, C. Zhao, Z. Ren, Y. Liang, X. Geng, C. Jiang, X. Yang, et al., Automated detection of high-frequency oscil-

- lations in epilepsy based on a convolutional neural network, *Frontiers in computational neuroscience* 13 (2019) 6.
- [10] K. S. Shekokar, S. Dour, Automatic epileptic seizure detection using lstm networks, *World Journal of Engineering* 19 (2) (2021) 224–229.
- [11] S. Chauhan, L. Vig, Anomaly detection in ecg time signals via deep long short-term memory networks, in: *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, IEEE, 2015, pp. 1–7.
- [12] G. Xu, T. Ren, Y. Chen, W. Che, A one-dimensional cnn-lstm model for epileptic seizure recognition using eeg signal analysis, *Frontiers in Neuroscience* 14 (2020) 1253.
- [13] S. Ryu, I. Joe, A hybrid densenet-lstm model for epileptic seizure prediction, *Applied Sciences* 11 (16) (2021) 7661.
- [14] J. Guo, H. Li, Y. Pan, Y. Gao, J. Sun, T. Wu, J. Xiang, X. Luo, Automatic and accurate epilepsy ripple and fast ripple detection via virtual sample generation and attention neural networks, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28 (8) (2020) 1710–1719.
- [15] Y. Feng, R. Zheng, X. Cui, T. Wang, T. Jiang, F. Gao, J. Cao, 3d residual-attention-deep-network-based childhood epilepsy syndrome classification, *Knowledge-Based Systems* 248 (2022) 108856.

- [16] L. Mou, C. Zhou, P. Zhao, B. Nakisa, M. N. Rastgoo, R. Jain, W. Gao, Driver stress detection via multimodal fusion using attention-based cnn-lstm, *Expert Systems with Applications* 173 (2021) 114693.
- [17] M. A. Khatun, M. A. Yousuf, S. Ahmed, M. Z. Uddin, S. A. Alyami, S. Al-Ashhab, H. F. Akhdar, A. Khan, A. Azad, M. A. Moni, Deep cnn-lstm with self-attention model for human activity recognition using wearable sensor, *IEEE Journal of Translational Engineering in Health and Medicine* 10 (2022) 1–16.
- [18] S. P. Singh, M. K. Sharma, A. Lay-Ekuakille, D. Gangwar, S. Gupta, Deep convlstm with self-attention for human activity decoding using wearable sensors, *IEEE Sensors Journal* 21 (6) (2020) 8575–8582.
- [19] P. Boonyakitanont, A. Lek-Uthai, K. Chomtho, J. Songsiri, A review of feature extraction and performance evaluation in epileptic seizure detection using eeg, *Biomedical Signal Processing and Control* 57 (2020) 101702.
- [20] K. Anusha, M. T. Mathews, S. D. Puthankattil, Classification of normal and epileptic eeg signal using time & frequency domain features through artificial neural network, in: *2012 international conference on advances in computing and communications*, IEEE, 2012, pp. 98–101.
- [21] Z.-K. Gao, Q. Cai, Y.-X. Yang, N. Dong, S.-S. Zhang, Visibility graph from adaptive optimal kernel time-frequency representation for classi-

- fication of epileptiform eeg, *International Journal of Neural Systems* 27 (04) (2017) 1750005.
- [22] H. Adeli, Z. Zhou, N. Dadmehr, Analysis of eeg records in an epileptic patient using wavelet transform, *Journal of neuroscience methods* 123 (1) (2003) 69–87.
- [23] H. Adeli, S. Ghosh-Dastidar, N. Dadmehr, A wavelet-chaos methodology for analysis of eegs and eeg subbands to detect seizure and epilepsy, *IEEE Transactions on Biomedical Engineering* 54 (2) (2007) 205–211.
- [24] Z. Altaf, M. A. Unar, S. Narejo, M. A. Zaki, et al., Generalized epileptic seizure prediction using machine learning method, *International Journal of Advanced Computer Science and Applications* 14 (1) (2023).
- [25] M. K. Siddiqui, R. Morales-Menendez, X. Huang, N. Hussain, A review of epileptic seizure detection using machine learning classifiers, *Brain informatics* 7 (1) (2020) 1–18.
- [26] Y. LeCun, Y. Bengio, G. Hinton, et al., Deep learning. *nature*, Google Scholar Google Scholar Cross Ref Cross Ref 521 (2015) 436–444.
- [27] F. Radenović, G. Toliás, O. Chum, Fine-tuning cnn image retrieval with no human annotation, *IEEE transactions on pattern analysis and machine intelligence* 41 (7) (2018) 1655–1668.
- [28] A. Shoeibi, N. Ghassemi, R. Alizadehsani, M. Rouhani, H. Hosseini-Nejad, A. Khosravi, M. Panahiazar, S. Nahavandi, A comprehensive

- comparison of handcrafted features and convolutional autoencoders for epileptic seizures detection in eeg signals, *Expert Systems with Applications* 163 (2021) 113788.
- [29] P. Bizopoulos, G. I. Lambrou, D. Koutsouris, Signal2image modules in deep neural networks for eeg classification, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2019, pp. 702–705.
- [30] J. Craley, E. Johnson, A. Venkataraman, Integrating convolutional neural networks and probabilistic graphical modeling for epileptic seizure detection in multichannel eeg, in: *International Conference on Information Processing in Medical Imaging*, Springer, 2019, pp. 291–303.
- [31] M. Abou Jaoude, J. Jing, H. Sun, C. S. Jacobs, K. R. Pellerin, M. B. Westover, S. S. Cash, A. D. Lam, Detection of mesial temporal lobe epileptiform discharges on intracranial electrodes using deep learning, *Clinical Neurophysiology* 131 (1) (2020) 133–141.
- [32] C. Angermueller, T. Pärnamaa, L. Parts, O. Stegle, Deep learning for computational biology, *Molecular systems biology* 12 (7) (2016) 878.
- [33] T. Najafi, R. Jaafar, R. Remli, W. A. Wan Zaidi, A classification model of eeg signals based on rnn-lstm for diagnosing focal and generalized epilepsy, *Sensors* 22 (19) (2022) 7269.

- [34] R. Yang, S. K. Singh, M. Tavakkoli, N. Amiri, Y. Yang, M. A. Karami, R. Rai, Cnn-lstm deep learning architecture for computer vision-based modal frequency detection, *Mechanical Systems and signal processing* 144 (2020) 106885.
- [35] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, C. E. Elger, Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state, *Physical Review E* 64 (6) (2001) 061907.
- [36] N. A. Samee, N. F. Mahmoud, E. A. Aldhahri, A. Rafiq, M. S. A. Muthanna, I. Ahmad, Rnn and bilstm fusion for accurate automatic epileptic seizure diagnosis using eeg signals, *Life* 12 (12) (2022) 1946.
- [37] Z. Huang, Y. Ma, R. Wang, B. Yuan, R. Jiang, Q. Yang, W. Li, J. Sun, Dscnn-lstms: A lightweight and efficient model for epilepsy recognition, *Brain Sciences* 12 (12) (2022) 1672.
- [38] X. Liu, J. Jia, R. Zhang, Automatic detection of epilepsy eeg based on cnn-lstm network combination model, in: *2020 4th International Conference on Computer Science and Artificial Intelligence, 2020*, pp. 225–232.
- [39] Y. Zhang, S. Yao, R. Yang, X. Liu, W. Qiu, L. Han, W. Zhou, W. Shang, Epileptic seizure detection based on bidirectional gated recurrent unit

- network, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 30 (2022) 135–145.
- [40] M. Varlı, H. Yılmaz, Multiple classification of eeg signals and epileptic seizure diagnosis with combined deep learning, *Journal of Computational Science* 67 (2023) 101943.
- [41] T. Abdallah, N. Jrad, F. Abdallah, A. Humeau-Heurtier, P. Van Bogaert, Cross-site generalization for imbalanced epileptic classification, in: *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2023, pp. 1–5.
- [42] A. H. Shoeb, Application of machine learning to epileptic seizure onset detection and treatment, Ph.D. thesis, Massachusetts Institute of Technology (2009).
- [43] T. D. Lagerlund, Manipulating the magic of digital eeg: montage reformatting and filtering, *American journal of electroneurodiagnostic technology* 40 (2) (2000) 121–136.
- [44] Radam explained — papers with code.
URL <https://paperswithcode.com/method/radam>
- [45] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, *Advances in neural information processing systems* 30 (2017).

- [46] A. Abdelhameed, M. Bayoumi, A deep learning approach for automatic seizure detection in children with epilepsy, *Frontiers in Computational Neuroscience* 15 (2021) 650050.
- [47] K. Nanthini, A. Tamilarasi, M. Pyingkodi, M. Dishanthi, S. Kaviya, P. A. Mohideen, Epileptic seizure detection and prediction using deep learning technique, in: *2022 International Conference on Computer Communication and Informatics (ICCCI)*, IEEE, 2022, pp. 1–7.
- [48] N. D. Truong, A. D. Nguyen, L. Kuhlmann, M. R. Bonyadi, J. Yang, S. Ippolito, O. Kavehei, Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram, *Neural Networks* 105 (2018) 104–111.
- [49] A. R. Ozcan, S. Erturk, Seizure prediction in scalp eeg using 3d convolutional neural networks with an image-based approach, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27 (11) (2019) 2284–2293.
- [50] S. M. Usman, S. Khalid, M. H. Aslam, Epileptic seizures prediction using deep learning techniques, *Ieee Access* 8 (2020) 39998–40007.
- [51] J. Birjandtalab, M. B. Pouyan, D. Cogan, M. Nourani, J. Harvey, Automated seizure detection using limited-channel eeg and non-linear dimension reduction, *Computers in biology and medicine* 82 (2017) 49–58.

- [52] S. M. Usman, M. Usman, S. Fong, Epileptic seizures prediction using machine learning methods, *Computational and mathematical methods in medicine* 2017 (2017).
- [53] D. Jacobs, T. Hilton, M. Del Campo, P. L. Carlen, B. L. Bardakjian, Classification of pre-clinical seizure states using scalp eeg cross-frequency coupling features, *IEEE Transactions on Biomedical Engineering* 65 (11) (2018) 2440–2449.