Epilepsy Seizure Prediction from EEG Signal Using Machine Learning Techniques

Boutkhil SIDAOUI[1], Kaddour SADOUNI[2]
[1]*University Center SALHI Ahmed, BP 66, 45000 Naâma, Algeria*
[2]*University of Science and Technology Mohamed Boudiaf, BP 1505, El M'naouer, Oran, Algeria*
b.sidaoui@cuniv-naama.dz

***Abstract*—Automatic seizure prediction is an important task to help epilepsy patients and epilepsy specialists. In addition, measuring electrical activity in different brain parts is an important step before any prediction. The best tool for recording electrical activity is electroencephalography (EEG), which uses electrodes placed on the head. This paper examines the performance of the convolutional neural network (CNN) architectures and support vector machine (SVM) method for predicting epileptic seizure activity using rich information recorded in the signal of EEG segments. The proposed approach is based on 22 features extracted from different EEG segments to produce a representative dataset. SVM classification models and two CNN architectures are proposed to predict ongoing seizures and different states of epilepsy patients. Two CNN architectures are presented: the first is trained with a dataset of features extracted from the EEG signal, and the second is trained with a dataset of Scalogram images from the EEG signal, whose purpose is to predict the imminence of an epileptic seizure in patients. A dataset of 6 patients is used to predict all states of epilepsy patients. Both CNN architectures and binary SVM classifiers achieve a classification rate above 98%.**

***Index Terms*—epilepsy seizure, EEG, prediction, Convolutional Neural Network, SVM.**

I. Introduction

Epilepsy is a common neurological disease that affects the central nervous system and causes seizures. Epilepsy is one of the most common neurological diseases in the world. During an epileptic episode, normal neural activity is disrupted, causing strange sensations, emotions, or sometimes seizures, muscle spasms, and loss of consciousness. Hans Berger made the most important invention in understanding epilepsy in 1924. He developed an electroencephalography (EEG) device, which records the brain's electrical activity. EEG is a multi-channel recording of the electrical activity generated by several neurons in the brain using several electrodes placed at different locations in the brain. EEG has been a fundamental tool, especially for diagnosing epilepsy. The analysis is performed by the epileptologist that observes the EEG continuously.

However, visual examination is time-consuming and can be inefficient. Automatic EEG signal detection is a basic tool that facilitates the diagnosis of epilepsy and improves the exploitation of EEG signals. Since 1970, considerable attention has been paid to automatic seizure detection. Feature extraction describes the behavior of EEG signals, an important phase for seizure detection. The feature extraction step should reduce the original signal to a dimension that is as small as possible and contains the maximum amount of relevant information to obtain the best performance of the classifier. Several types of feature extraction from epilepsy EEG data have been proposed, such as parametric and non-parametric methods and time-frequency methods. The features extracted are used for distinguishing between classes of epilepsy.

In general, two tasks are proposed for EEG signal classification: seizure detection and seizure prediction, using machine learning tools such as neural network (NN), support vector machine (SVM), and decision trees. Many approaches have been proposed to improve the performance of automatic prediction from EEG signals. This work presents two machine learning methods for predicting epileptic seizures, deep learning architectures named convolutional neural networks and support vector machines. In addition, efficient heuristics for selecting electrodes are introduced based on their frequencies and positions, in the feature extraction step, from the EEG signal.

The rest of the paper is organized as follows: Section 2 discusses some successful, related methods for epilepsy detection and prediction; Section 3 introduces the convolutional neural network, followed by support vector machines in Section 4; Section 5 introduces the seizure prediction and electrode selection process in the feature extraction stage; Sections 6 and 7 analyze the proposed SVM models and CNN architectures, while Section 8 presents the results of experiments on a dataset, using a variety of parameters; Section 9 concludes the paper with some work perspectives.

II. Related Works

EEG classification consists of predicting or detecting the seizure class given a set of appropriate features extracted from the EEG signal. Many works are proposed for automatic detection and prediction in the EEG signal. Alexandros et al. in [1] demonstrate the relevance of time-frequency analysis to classify EEG segments for seizures. However, the authors in [2] extract temporal and frequency features from EEG signals using the Empirical mode decomposition algorithm, before creating a prediction model based on machine learning methods. The authors in [3] present a novel approach for automatic epilepsy seizure detection based on two contributions: the use of non-linear classifiers through the so-called kernel trick and the proposal of a bag-of-words model for extracting a non-linear feature representation of the input data. In another work [4], the authors present a novel EEG classification method for epilepsy seizure detection. The authors employ a continuous wavelet transform method for obtaining the time-frequency

Digital Object Identifier 10.4316/AECE.2023.02006

images of EEG signals. The processed EEG signals are then decomposed into five sub-band frequency components. Gaussian mixture model features and gray-level co-occurrence matrix descriptors are extracted from these sub-band time-frequency images. Authors in [5] propose an epilepsy seizure-detecting method that can be implemented in a hardware device to help epilepsy patients. Their study utilized an EEG dataset used in various researches regarding epilepsy detection. The researchers process the EEG signal in both time and frequency domains, apply a Chebyshev filter to pre-process the signal, and then use wavelet analysis. In [6], a novel deep-learning approach is proposed for detecting seizures in pediatric patients based on the classification of raw multichannel EEG signal recordings that are minimally pre-processed. The new approach takes advantage of the automatic feature learning capabilities of a two-dimensional deep convolution auto-encoder linked to a neural network-based classifier to form a unified system that is trained in a supervised way to achieve the best classification between the ictal and interictal brain states.

In [7], the authors propose a novel and efficient algorithm to detect the presence of seizures in heart rate variability. This algorithm includes feature extraction and classification. The ten features include time and frequency domain analysis and nonlinear features extracted from one-lead electrocardiogram signals of epilepsy patients. The extracted features were used as input to multilayer perceptual neural networks. In [8], authors present convolutional neural network architecture to develop seizure detection models for three patients with nocturnal frontal lobe epilepsy. The authors state that the performance, in terms of accuracy, sensitivity, and specificity, exceeds that of the most recent literature by several percentage points. The authors of [9] propose an end-to-end automatic seizure detection system based on deep learning, which does not require heavy preprocessing of EEG data or feature engineering. The fully convolutional network with three convolution blocks is used to learn the expressive features of seizures from the EEG data. In addition, the EEG features relevant to seizures are presented as an input to the nested long short-term memory (NLSTM) model to explore the inherent temporal dependencies of EEG signals. Finally, the high-level features obtained from the NLSTM model are fed into the Softmax layer to produce predicted labels. In [10], a review is proposed. The authors discuss recent advances in seizure detection, signal processing, time or frequency domain analysis, and classification algorithms to detect and classify seizure stages. Then, they show a strong potential for applying for recent advances in non-invasive brain stimulation technology to treat seizures.

III. Convolutional Neural Networks

Artificial neural networks are a set of architectures that successfully solve a large class of machine learning tasks such as recognition, optimization, and prediction [26-28]. Recently, a new efficient architecture called convolutional neural network (CNN) has been proposed. The convolutional neural network (CNN) is a deep learning algorithm originally developed for image recognition. The pre-processing required in a CNN is much lower than in other classification algorithms. CNN, like artificial neural networks, is a model of the connectivity of neurons in the human brain and was inspired by the organization of the visual cortex. The CNN is a feed-forward architecture composed of blocks of layers, such as an input layer, hidden layers, and an output layer. Each layer comprises nodes connected to others in the next layer and has an associated weight and activation function. The output of any individual node sends data to the next layer of the network.

Originally, CNN consists of two parts, the first part is dedicated to extracting features from the input data, and the second is reserved for the classification stage. In practice, the first part comprises several construction layers, such as convolution, normalization, and pooling. The second part comprises fully connected layers, followed by the output layer. The CNN uses the backpropagation algorithm to optimize the weights. A standard architecture consists of repetitions of several convolution layers and a pooling layer, followed by one or more fully connected layers (Figure 1). The CNN architecture consists of repetitions of blocks used for feature learning (multiple convolution layers and pooling layers), followed by one or more fully connected layers.

A. Input Layer

It represents the source of training data. At first, CNN is developed for image recognition; the input layer takes the structure of an image of one or more channels, mathematically a matrix of n rows and m colonies of binary or color image. However, it is possible to define the input layer as a signal data vector or a feature matrix.

B. Convolutional Layer

The convolutional layer is the core of a CNN. It requires a few components, which are input data (image, signal, or features), and a set of filters or kernels. The convolution operation is applied between the input and filters. Moreover, after convolution, an activation function is calculated for each neuron. The Rectified Linear Activation function (ReLU) is the most used, and it is given by the formula (1):

$$\mathrm{Re}LU(z) = \begin{cases} z & z > 0 \\ 0 & z \leq 0 \end{cases} \quad (1)$$

However, some hyper-parameters need a prior optimization to select the best values for them, before the neural network training begins. These include the number and size of the filters.

C. Pooling Layer

Pooling layers are used to reduce the dimensionality of the previous convolutional layer. Like the convolutional layer, pooling sweeps a kernel through the entire input. The difference is that the kernel has no weight. There are two main types of pooling: max pooling and average pooling.

D. Flatten and Fully Connected Layer

All neurons are connected in one layer in this step. As a nonlinear combination of the best features, this layer is considered an output of the convolutional layers. The fully connected layer is fed to a feed-forward neural network with a backpropagation algorithm.

E. Output Layer

The output layer is the terminal layer in the neural network that produces the final prediction. Usually, in CNN, a Softmax function is used to activate neurons in the output layer. The Softmax function is given as follows:

$$Soft\max(z) = \frac{e}{\sum_j e^{zj}} \quad (2)$$

In recent years, the convolutional neural network (CNN) has been one of the most widely used architectures in several fields, especially for image processing. For a more detailed description of CNN, see [11-14].

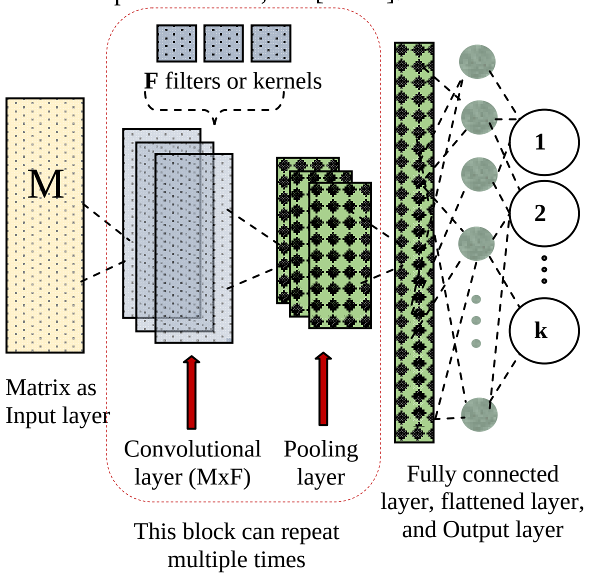

Figure 1. Architecture of convolutional neural network

IV. SUPPORT VECTOR MACHINE

Support vector machines (SVM) [15-16], introduced during the last decade in the context of statistical learning, have been successfully used for the solution of a large class of machine learning tasks [17-18] such as categorization, prediction, novelty detection, ranking, and clustering. In their general form, Binary SVM extends an optimal linear hypothesis, in terms of an upper bound on the expected risk that can be interpreted as the geometrical margin, to nonlinear ones using kernels k (.,.). Kernels can be interpreted as dissimilarity measures of pairs of objects in the training set. In standard SVM formulation, the optimal hypothesis sought is of the form (1).

$$f(x) = \sum \alpha_i k(x, x_i) \quad (3)$$

where, α_i are the components of the unique solution of a linearly constrained quadratic programming problem (QP). Usually, QP has a size equal to the number of training examples. The solution vector obtained is generally sparse, and the non-zeroes α_i are called support vectors (SVs). The number of SVs determines the query time, which is the time it takes to predict novel observations. The training process is implicitly performed in a reproducing kernel Hilbert space (RKHS) in which k (x, y) is the inner product of the images of two examples: x et y. Moreover, the optimal hypothesis can be expressed in terms of the kernel that can be defined for non-Euclidean data, such as biological sequences and speech utterances. Popular positive kernels include Polynomial (4) and Gaussian (5).

$$k(x_i, x_j) = \left(\gamma x_i^T x_j + r\right)^d, \gamma > 0 \quad (4)$$

$$k(x_i, x_j) = \exp\left(-\gamma \left\|x_i - x_j\right\|^2\right), \gamma > 0 \quad (5)$$

Formally, SVM is defined as given training vectors, $x_i \in R^n, i = 1,..,m$ belongs to two classes labeled by a vector $y \in R^m$ such that $y \in \{1, -1\}$. Support vector classifiers [15-16] solve the following linearly constrained convex quadratic programming problem:

$$\max \quad W(\alpha) = \frac{1}{2}\sum_{i,j} \alpha_i \alpha_j y_i y_j k(x_i, x_j) - \sum_{i=1}^{m} \alpha_i$$
$$under\_constraints: \quad \forall i \ 0 \le \alpha_i \le C \quad (6)$$
$$\sum_{i=1}^{m} \alpha_i y_i = 0$$

The optimal hypothesis is:

$$f(x) = \sum_{i=1}^{m} \alpha_i y_i k(x, x_i) + b \quad (7)$$

where, the term b can be computed separately and $k(x_i, x_j)$ is a kernel that can be defined for nonlinear data such as positive Polynomial and Gaussian kernels.

The optimal hypothesis f depends only on the non-null coefficients α_i whose corresponding patterns are called Support vectors (SV). The quadratic programming objective function involves the problem gram matrix whose entries are the similarities $k(x_i, x_j)$ between the examples x_i and x_j . It is important to note that the pattern input dimension d, in the above formulation, is implicit and does not affect to some extent the complexity of training, and the representation of the patterns is not needed, and only pair-wise between objects must be specified. This feature makes SVM very attractive for high input dimensional recognition problems and for those where patterns cannot be represented as fixed dimensional real vectors such as text, strings, and DNA. Several algorithms are used to solve quadratic programming based on decomposition techniques, such as sequential minimal optimization SMO implemented in the LibSVM library [23].

A. *SVM Multiclass*

Support Vector Machines are inherently binary classifiers. Several frameworks have been introduced to extend SVM to multi-class contexts and a detailed account of the literature, such as [22]. Typically, multi-class classifiers are built by combining several binary classifiers. The first method of this type is the “one against all” method [19] which constructs L classifiers, where L is the number of classes. The L^{th} classifier is trained by labeling all the examples in the L^{th} class as positive and the remainder as negative. The final hypothesis is given by the following formula:

$$f_{OVA}(x) = \arg\max_{i=1,\dots,L}\left(f_i(x)\right) \quad (8)$$

Another popular approach called one-against-one, proceeds by training *((L-1)L)/2* binary classifiers corresponding to all the pairs of classes. The hypothesis consists of choosing the class with the most votes [20-21].

V. EPILEPSY SEIZURE PREDICTION

Seizure prediction systems can detect ongoing seizures and provide clinicians with detailed data. Closed-loop systems built around seizure detection might also provide rapid therapy in response to seizures early in their clinical onset, thereby limiting the complications or potentially arresting the spread of seizures. Automatic seizure prediction machines must be able to detect the presence of epilepsy seizures and predict the coming epilepsy seizures before there are happening. Various artificial intelligence algorithms predict epilepsy seizures from different biometric signals. All these algorithms involve two main steps: extraction of characteristics from biometric signals and creation of classification models.

A. EEG Signals

Electroencephalography (EEG) is a method of exploring the brain that measures electrical activity in the brain through electrodes placed on the scalp, often shown as a trace called an electroencephalogram. Acquisition of brain electrical activity with EEG has the most valuable and rich source of information for epilepsy research. EEG is essential for diagnosing and classifying epilepsy seizures because it provides real-time information and excellent temporal resolution. EEG is one of the main diagnostic tests for epilepsy. EEG provides an important dataset that can be analyzed and interpreted via several methods. It can be difficult for patients to wear the EEG electrodes for longer. Several features are extracted from the EEG signals. These features collect all important information about seizures without noise and redundancy information. In this study, we are using characteristics extracted from EEG signals captured by electrodes placed on the scalp of the head.

B. Feature Extraction from EEG Signal

Several types of characteristic extraction from epilepsy EEG data have been proposed in the literature, such as parametric and non-parametric, and time-frequency methods. The pertinent pieces of information extracted are used for distinguishing between classes of epilepsy. In this work, the 22 characteristics extracted from each EEG segment of 5 minutes are: {Time_Offset, AR_Coeff, Decorr_Time, LT_Energy, Hjorth_M, Hjorth_C, Rel_Pow_Delta, Rel_Pow_Theta, Rel_Pow_Alpha, Rel_Pow_Beta, RelPow_Gamma, Spectr_Edge_Freq, Spectr_Edge_Pow, Mean, Variance, Skewness, Kurtosis, Wavelet_db4_EnergyBand1, Wavelet_db4_EnergyBand2, Wavelet_db4_EnergyBand3, Wavelet_db4_EnergyBand4, Wavelet_db4_EnergyBand5}. The 10–20 electrode system is used to recording EEG signals, with bandwidth from 0.1 Hz to 2000 Hz according to the type of bands (delta, theta, alpha, beta, and gamma) as defined in EPILAB project of Center for Informatics and Systems [24], [29].

In practice, three parameters are requisite in this step of characteristic (features) extraction step:

1) The window (segment) size: each EEG signal is segmented on a window of 5 min;

2) The prevision time (period) or pre-ictal time before instance seizure: 30 min and 10 min values are chosen;

3) Selection of the appropriate electrodes: three heuristics are proposed and discussed in the next subsection.

C. Electrodes Selection

The selection of electrodes is an important phase for feature selection. In this study, we propose three heuristics of electrodes selection:

1) In the first heuristic, we select all electrodes of the highest frequency for each patient. Generally, high-frequency signals can indicate the presence of seizures. The instances of high frequency in each signal represent the instances of seizures. Table I shows electrodes selected for patients;

2) In the second heuristic, we select all electrodes placed in the left part of the head for each patient, i.e., all electrodes of even numbers, without considering the frequency of electrodes. In this case, we find if there is any relationship between the left part and seizures;

3) In the third heuristic, we select all electrodes placed in the right part of the head for each patient, i.e., all electrodes of odd numbers, without considering the nature of the signals. In this case, we find if there is any relationship between the right part and seizures.

It should be noted that the data used in this study contain signals recorded for 6 epilepsy patients. These data belong to the EPILAB project of the Center for Computer Science and Systems (CISUC), Department of Computer Engineering, University of Coimbra, Portugal [25].

TABLE I. HIGHEST FREQUENCY ELECTRODES SELECTED FOR EACH PATIENT

Patient	Highest frequency electrodes
P 1200	AF7, F7, T7, P7, F9, T9, FT7, FT9, TP7
P 1300	AF7, F7, T7, P7, F9, T9, FT7
P 1400	AF7, F7, T7, P7, F9, T9, FT7
P 1500	FP2, F8, T8, P8, AF8, F10, T10, FT8
P 1600	AF7, F7, T7, F9, T9, FT7, FT9
P 1700	P8, TP8
P 2100	FP1, AF7, F7, F3, F8, AF8, F9, F10, T9, T10, FT9
P 2300	F7, F3, T7, C3, P7, P3, F9, T9, FT7

D. Automatic Seizure Prediction Problems

Automatic seizure prediction can be presented as two types of problems: predict the coming seizures as a binary classification problem, and expect all seizures state of patients as a multi-class problem. For this, four classes are defined in epilepsy seizures problem: pre-ictal, ictal, inter-ictal, and post-ictal. These classes represent the different states of epilepsy patients (Table II). For the binary problem, two classes are used to predict the ongoing seizures by a supervised binary classifier. The pre-ictal is used as the positive class, and all others (ictal, interictal, post-ictal) as the negative class. In addition, the four classes, pre-ictal, ictal, interictal, and post-ictal, were used to create a multi-class classifier to predict the four states of epilepsy patients.

TABLE II. CLASSES DESCRIPTIONS

Classes	Description
pre-ictal	before epilepsy seizure instance
Ictal	Instance of epilepsy seizure
inter-ictal	normal brain state
post-ictal	After epilepsy seizure

E. Models for Seizures Prediction

In this stage, classifier models are created to predict ongoing seizures. These models are derived from machine learning methods and trained on sets of features extracted in the first step (dataset) or from the EEG signals directly (in the case of CNN). Before creating models, the dataset is divided into two sets: the training dataset and the testing dataset. During the training step, the classification models are constructed using the training dataset, i.e., parameters that indicate criteria for the presence or absence of seizures are determined. After that, a testing dataset is used to illustrate the performance of the proposed models. Two machine learning methods (SVM and CNN) are used to build classification models.

VI. SVM Models Proposed

This section presents the first contribution based on the SVM method. It is represented by two classifiers: binary SVM and multi-class SVM (Figure 2).

The first classifier is developed to predict the ongoing seizure as a binary classification problem, and the second classifier is constructed to predict four patient states as a multi-class problem.

In the Training stage, both binary and multi-class SVM classifiers are constructed with the Gaussian kernel, the best values of the C parameter. The best values of C and the parameter gamma of the Gaussian kernel are optimized by a cross-validation algorithm with 5 folders. Moreover, the one-versus-all (OVA) strategy is applied in the multi-class cases.

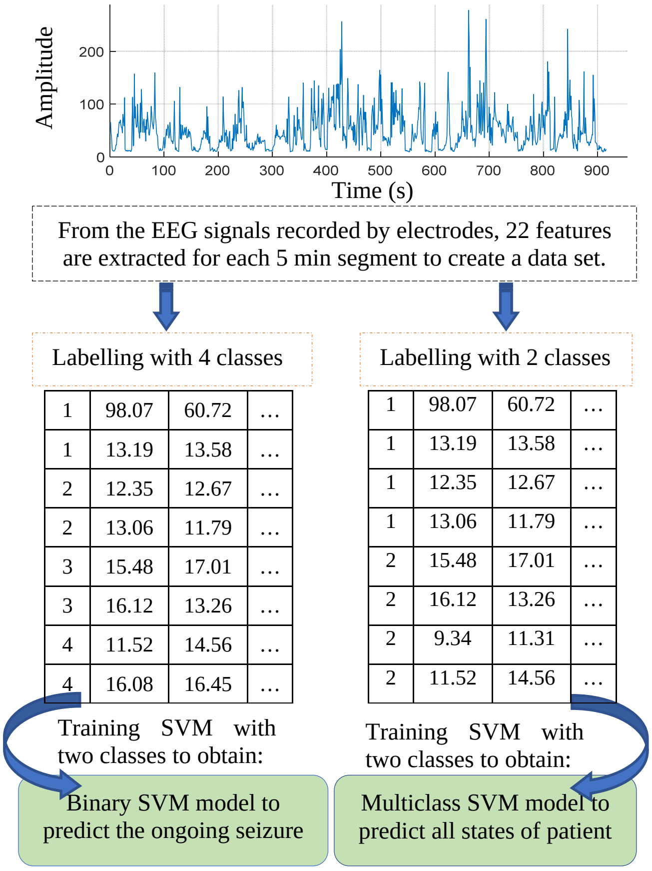

Figure 2. Creation of SVM models using EEG signals

VII. CNN Architectures Proposed

This section presents the second contribution based on CNN and represented by two architectures for epilepsy seizure prediction: the first uses the data of EEG signals, and the second uses the Scalogram images of EEG signals. Each of these CNN architectures is developed to predict epilepsy seizures as two problems: The prediction of the ongoing seizure as a binary classification problem, i.e., pre-ictal state (class) versus the rest of the patient states {ictal, inter-ictal, post-ictal}, and the prediction of four patient states as a multi-class problem. i.e., classification with four classes {pre-ictal, ictal, inter-ictal, post-ictal}.

A. CNN Trained on EEG Signal (CNN_EEG)

In this case, CNN trained on EEG signal to construct an architecture powered by the EEG signals data (Figure 3), and consists of:

1) Input layer as a vector of 22 (features number) times the number of channels selected for each patient;

2) Two convolutional layers: the first uses 30 kernels of dimension 5, and the second uses 100 kernels of 5 dimensions. Each convolutional layer is followed by a batch normalization layer, a ReLU Layer, and the highest pooling layer of 2 dimensions with a stride of 2;

3) Another convolutional layer uses 50 kernels of dimension 5, followed by a batch normalization layer and a ReLU Layer without pooling;

4) A fully connected layer regroups all flats of the previous layer on one layer as a column fully connected to the output layer with a Softmax function.

Finally, the output layer contains k neurons corresponding to the number of classes, i.e., k=2 in the binary case (pre-ictal versus others), and k=4 in the multi-class case with four classes (pre-ictal, ictal, inter-ictal, post-ictal).

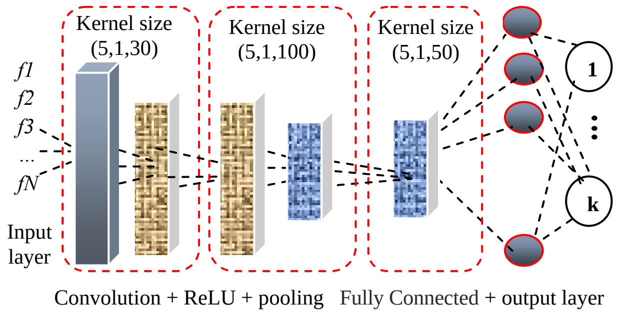

Figure 3. CNN trained on EEG signal of dimension N (CNN_EEG)

B. CNN Trained on Scalogram (CNN_Scalogram)

In this case, the CNN architecture needs to transform EEG signals into Scalogram images, as shown in the next figure (Figure 4). The CNN architecture (Figure 5) in this case, consists of the following layers:

1) Input layer as an image color of size 28 x 28;

2) The first convolutional layer uses 20 kernels of dimension 5x5, and the second uses 50 kernels of 5x5 dimensions. Each convolutional layer is followed by a batch normalization layer, a ReLU Layer, and the highest pooling layer of 2 dimensions with a stride equal to 2;

3) Another convolutional layer uses 30 kernels of dimension 5x5, followed by a batch normalization layer, a ReLU Layer without pooling;

4) A fully connected layer regroups all flats of the previous layer on one layer as a column fully connected to the output layer with a Softmax function;

5) Finally, the output layer contains k neurons corresponding to the number of classes, i.e., k=2 in the binary case (pre-ictal versus others), and k=4 in the multi-class case with four classes (pre-ictal, ictal, inter-ictal, post-ictal).

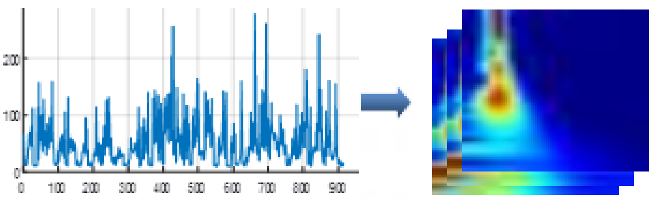

Figure 4. Scalogram images from EEG signals

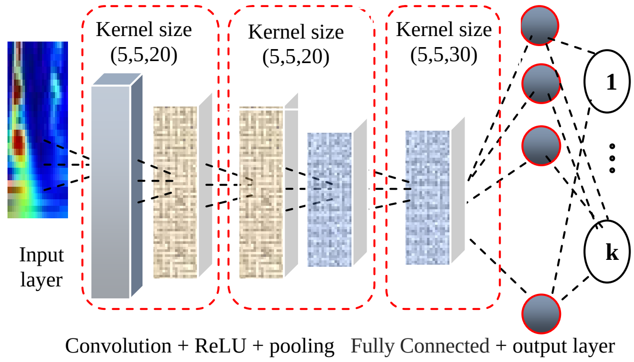

Figure 5. CNN trained on Scalogram EEG images (CNN_Scalogram)

VIII. Implementation and Results

In our experiments, we used a dataset consisting of 6 patients belong the EEG dataset of Coimbra University [24]. After the feature's extraction step, six models are constructed: four architectures based on CNN and two SVM classifiers, as follows:

1) CNN trained on an EEG signal dataset with an output layer of two classes (pre-ictal versus others);

2) CNN trained on the EEG signal dataset with an output layer of four classes (pre-ictal, ictal, inter-ictal, post-ictal);

3) CNN trained on Scalogram images with an output layer of two classes (pre-ictal versus others);

4) CNN trained on Scalogram images with an output layer of all classes;

5) Binary SVM classifier (pre-ictal versus others);

6) Multi-class SVM classifier with four classes.

On the other hand, we note that MATLAB 2020 version is used in the implementation step. In the experiments reported below, we performed 60% of the data for the training step, 10% of the data for validation, and 30% for the testing step. All experiments were performed on a standard Intel i3 3.40 processor with 4 GB memory running the Windows 7 operating system.

A. Evaluation Criteria

For the evaluation, three criteria are used: precision, recall, and accuracy. Precision and recall are very used to evaluate binary classification, they are given by the next formulas:

$$Precision = TP/(TP+FP) \tag{9}$$

$$Recall = TP/(TP+FN) \tag{10}$$

where, TP is true positive, FP is false positive, and FN is false negative.

The accuracy is used to measure how well a classification test correctly identifies. It is the proportion of true results (both true positives (TP) and true negatives (TN)) among the total number of cases examined. The popular formula of accuracy is given by:

$$Accuracy = (TP+TN)/\,Total\ number \tag{11}$$

B. Experimentations Results

Tables III and IV summarize the best-obtained results by SVM classifiers for each patient with the highest frequency electrodes. The results presented below in Table III indicate the performance of the SVM machine in predicting the ongoing seizure, especially for a short time of 30 min and 10 min, before the seizure, with 98% and 99% of accuracy. According to the results of the multi-class problem presented in Table IV, good results are obtained with an accuracy of 97% for some patients. Otherwise, kernel Gaussian gives the best projection to the description space for SVM, with parameter C=100.

Table III. Prediction of Pre-ictal State

Patient	Period (min)	Accuracy (%)	Precision (%)	Recall (%)
P 1200	30	79.91	79.91	100
	10	92.58	92.58	100
P 1400	30	94.04	94.04	100
	10	98.02	98.02	100
P 1500	30	88.57	88.57	100
	10	96.19	96.19	100
P 1600	30	98.77	98.77	100
	10	99.59	99.59	100
P 2100	30	98.10	98.10	100
	10	99.07	99.07	100
P 2300	30	96.50	96.50	100
	10	98.83	98.83	100

Table IV. Prediction of Pre-ictal, Inter-ictal, Post-ictal and Ictal

Patient	Period (min)	Accuracy (%)
P 1200	30	60.34
	10	73.01
P 1400	30	90.58
	10	94.56
P 1500	30	80.13
	10	87.74
P 1600	30	97.84
	10	98.66
P 2100	30	96.08
	10	97.05
P 2300	30	93.97
	10	96.30

Tables III and IV above present the best results obtained by the binary classifier SVM and multi-classifier SVM trained with: the gaussian kernel, gamma parameter =0.5, parameter C=100, highest frequency electrodes, and One-Versus-All multi-class strategy for multi-class case. The column 'Period' is the duration time before the ongoing seizures.

On the other hand, Tables V, VI, VII and VIII summarize the best results obtained by CNN architectures for each patient with the highest frequency electrodes. Tables V and VI below present the best results obtained by the two architectures, CNN_EEG and CNN_Scalogram, to predict the ongoing seizure with previous times of 30 min and 10 min and a number of epochs equal to 4. Accuracies of more than 95 are obtained. In contrast, Tables VII and VIII list the best results obtained by CNN_EEG and CNN_Scalogram to predict all states of patients with times of 30 min and 10 min before seizure instance.

TABLE V. PREDICTION OF PRE-ICTAL WITH PREVIOUS TIME OF 30 MIN

4 Epochs	CNN_EEG		CNN_Scalogram	
	Validation (%)	Test (%)	Validation (%)	Test (%)
P 1200	85.78	85.30	87.92	88.24
P 1400	96.80	97.11	97.14	97.13
P 1500	94.05	94.08	94.13	94.14
P 1600	96.44	96.67	96.55	96.55
P 2100	95.78	95.34	96.33	96.34
P 2300	95.35	95.28	95.41	95.42

TABLE VI. PREDICTION OF PRE-ICTAL WITH PREVIOUS TIME OF 10 MIN

4 Epochs	CNN_EEG		CNN_Scalogram	
	Validation (%)	Test (%)	Validation (%)	Test (%)
P 1200	94.28	94.65	94.73	94.79
P 1400	98.94	99.14	99.04	99.04
P 1500	97.97	98.10	98.05	98.04
P 1600	98.74	98.76	98.85	98.85
P 2100	97.94	98.25	98.41	98.41
P 2300	98.23	98.31	98.35	98.35

TABLE VII. PREDICTION OF ALL CLASSES (PRE-ICTAL, ICTAL, POST-ICTAL AND INTER-ICTAL) WITH PREVIOUS TIME = 30 MIN

4 Epochs	CNN_EEG		CNN_Scalogram	
	Validation (%)	Test (%)	Validation (%)	Test (%)
P 1200	71.65	71.21	75.27	74.66
P 1400	94.80	95.06	95.09	95.07
P 1500	90.03	89.27	89.95	89.95
P 1600	94.02	93.87	94.03	94.05
P 2100	90.51	91.41	92.91	92.90
P 2300	91.18	91.10	91.23	91.17

TABLE VIII. RESULTS OF PREDICTION OF ALL CLASSES (PRE-ICTAL, ICTAL, POST-ICTAL AND INTER-ICTAL) WITH PREVIOUS TIME = 10 MIN

4 Epochs	CNN_EEG		CNN_Scalogram	
	Validation (%)	Test (%)	Validation (%)	Test (%)
P 1200	80.60	80.45	81.75	81.43
P 1400	96.95	96.88	97.01	96.97
P 1500	92.98	93.59	93.53	93.68
P 1600	96.29	96.43	96.33	96.33
P 2100	93.80	93.75	94.60	94.53
P 2300	94.00	94.14	94.16	94.53

C. Graphics and Discussion

The results obtained by convolutional neural networks proposed in this work are better than those obtained by the support vector machines. Additionally, while we are close to the seizure instance, the accuracy increases, and we can also see that the results obtained by Scalogram images are better. Selecting the electrodes with the highest frequency gives better results than the other approaches used to select electrodes. Nevertheless, the position of the electrodes is not essential to predict seizures from EEG signals. Figures 6 and 7 compare the CNN architectures and the SVM in terms of accuracy. These graphics show that CNN and SVM give similar results; nevertheless, CNN proposed architectures present some positive points, such as stable results for all patients and for different previous times. Moreover, in terms of CPU time, CNN is clearly faster than SVM as Figure 11 shown. On the other hand, the two curves of comparison presented in Figure 8 and the histogram in Figure 9 indicate that the number of epochs does not significantly influence the obtained results for all CNN models. Figure 10 shows that CNN with Scalogram images gives better results than CNN with signal data. Also, the selection of the electrodes of the highest frequency gives better results. Globally, the proposed systems of prediction based on SVM and CNN give encouraging results, compared to the literature, as in [24].

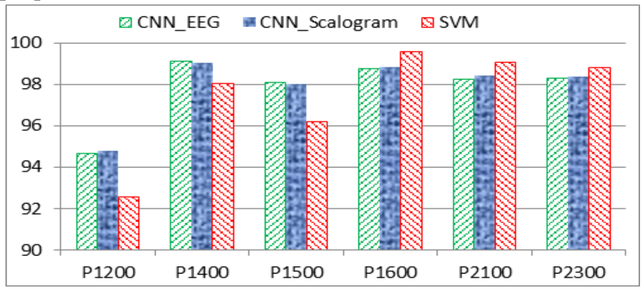

Figure 6. Accuracies prediction of pre-ictal obtained by binary classifiers CNN and SVM, with a previous time equal to 10 minutes

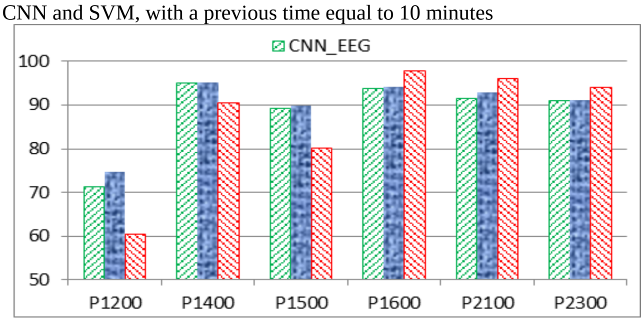

Figure 7. Comparison of accuracies obtained by multi-classifiers CNN and SVM, with a previous time equal to 10 minutes

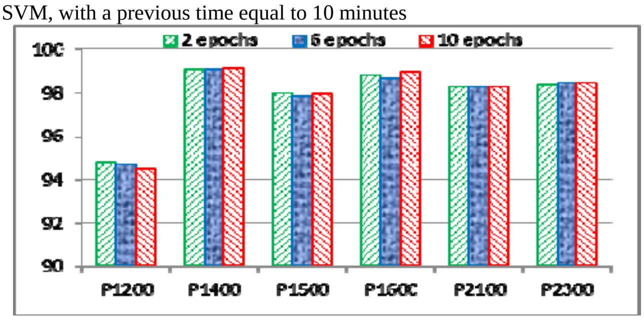

Figure 8. Variation of accuracy for 6 patients using CNN with EEG features as input layer and previous time equal to 10 minutes, for three epochs

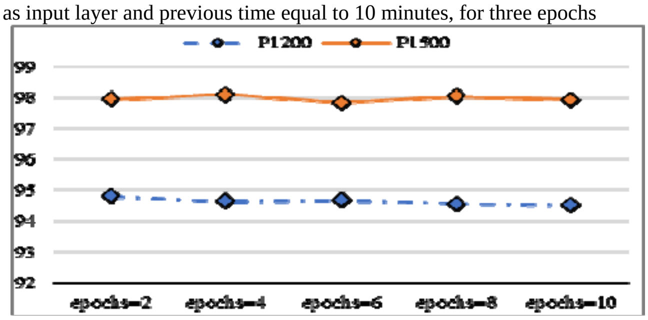

Figure 9. Variation of accuracy obtained by CNN with different epochs, and a previous time equal to 10 minutes

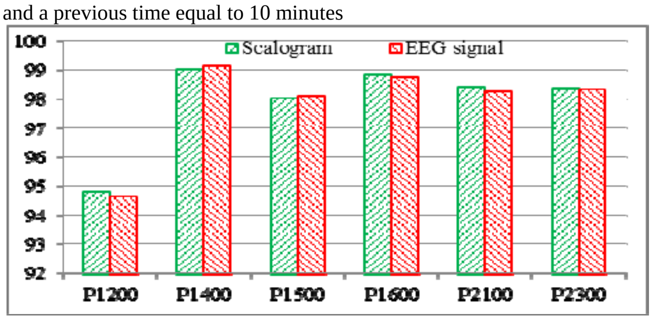

Figure 10. Comparison between CNN with EEG signal and CNN with Scalogram, with previous time equal to 10 minutes, and 4 epochs

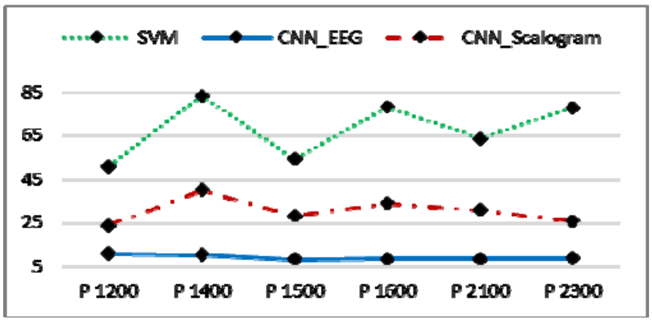

Figure 11. Comparison of CPU time of testing step, obtained by multi-classifiers CNN and SVM, with previous time equal to 30 minutes

IX. Conclusion

In this work, two machine-learning tools were implemented and compared to predict ongoing seizures. The first is composed of two architectures based on a convolutional neural network. The second one is based on support vector machines. However, three heuristics are proposed to select the electrodes before extracting the best features: right side, left side, and highest frequency electrodes. Based on the results obtained, it can be concluded that there is no difference between the left and right electrode positions. The selection of the electrodes with the highest frequency is sufficient. The proposed tools, SVM and CNN give encouraging results in predicting pre-ictal time and predicting all patient states, compared to the literature. The precision, recall, and accuracy give a good picture of the SVM predictor's ability to give better predictions for each patient, with a simple advantage for CNN over SVM, especially for multi-class problems. However, we believe that to improve the technology for automatic seizure capture. More research should be invested in developing efficient machine learning paradigms capable of handling multiple patients to develop a common model that is efficient in terms of accuracy and generality.

References

[1] Y. Kumar, M. L. Dewal, R. S. Anand, “Epileptic seizure detection using DWT based fuzzy approximate entropy and support vector machine,” Neurocomputing, col. 133, pp. 271-279, 2014. doi:10.1016/j.neucom.2013.11.009

[2] S. M. Usman, M. Usman, S. Fong, “Epileptic seizures prediction using machine learning methods,” Computational and Mathematical Methods in Medicine, vol. 2017, ID 9074759, 2017. doi:10.1155/2017/9074759

[3] Y. Li, W. Cui, M. Luo, K. Li, and L. Wang, “Epileptic seizure detection based on time-frequency images of EEG signals using gaussian mixture model and gray level co-occurrence matrix features,” International Journal of Neural Systems, vol. 28, no. 07, 2018. doi:10.1142/S012906571850003X

[4] Z. Lasefr, S. Shiva, V. N. R. Ayyalasomayajula, K. Elleithy, “Epilepsy seizure detection using EEG signals,” In 8th IEEE Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, New York City, NY, USA, October 19-21, pp. 162-167, 2017. doi:10.1109/UEMCON.2017.8249018

[5] S. Vani, G. R. Suresh, T. Balakumaran, T. Ashawise Cross, “EEG signal analysis for automated epilepsy seizure detection using wavelet transform and artificial neural network,” Journal of Medical Imaging and Health Informatics, vol. 9, no. 6, pp. 1301-1306, 2019. doi:10.1166/jmihi.2019.2713

[6] M. A. Ahmed and M. A. Bayoumi, “A deep learning approach for automatic seizure detection in children with epilepsy,” Frontiers in Computational Neuroscience, vol. 15, 2021. doi:10.3389/fncom.2021.650050

[7] S. Bani, J. Nader, M. Ali, A. Cesar and D. Antonio, “A new algorithm for detection of epileptic seizures based on HRV signal,” Journal of Experimental & Theoretical Artificial Intelligence, vol. 26, no. 2, pp. 251-265, 2014. doi:10.1080/0952813X.2013.861874

[8] F. Pisano, et al., “Convolutional neural network for seizure detection of nocturnal frontal lobe epilepsy,” Complexity, vol. 2020, ID 4825767, 10 pages, 2020. doi:10.1155/2020/4825767

[9] M. Zhou, C.Tian, R. Cao, B. Wang, Y. Niu, T. Hu, H. Guo and J. Xiang, “Epileptic seizure detection based on eeg signals and CNN,” Frontiers in Neuroinformatics, vol. 12, 2018. doi:10.3389/fninf.2018.00095

[10] Y. Li, Z. Yu, Y. Chen, C. Yang, Y. Li, X. Allen Li, and B. Li, “Automatic seizure detection using fully convolutional nested LSTM,” International Journal of Neural Systems, vol. 30, no. 4, 2020. doi:10.1142/S0129065720500197

[11] M. A. Nielsen, “Neural networks and deep learning”, pp. 167- 205, Determination Press, 2015

[12] B. Liquet, S. Moka and Y. Nazarathy, “The mathematical engineering of deep learning,” pp. 195-224, 2022

[13] R. Yamashita, et al., “Convolutional neural networks: an overview and application in radiology,” Insights Imaging, vol. 9, pp. 611–629, 2018. doi:10.1007/s13244-018-0639-9

[14] M. Emec, M. H. Ozcanhan, “A hybrid deep learning approach for intrusion detection in IoT networks,” Advances in Electrical and Computer Engineering, vol.22, no.1, pp.3-12, 2022. doi:10.4316/AECE.2022.01001

[15] A. Mariette and K. Rahul, “Efficient learning machines: Theories, concepts, and applications for engineers and system designers,”, pp. 1-62, Apress Berkeley, 2015. doi:10.1007/978-1-4302-5990-9

[16] A. R. Webb, “Statistical pattern recognition,”, pp. 169-200, John Wiley & Sons Ltd, 2002. doi:10.1002/9781119952954

[17] E. Osuna, R. Freund, and F. Girosit, “Training support vector machines: an application to face detection,” Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, USA, pp. 130-136, 1997. doi:10.1109/CVPR.1997.609310

[18] T. Joachims, “Making large scale SVM learning practical,” Advances in Kernel Methods: Support Vector Learning book, pp. 169–184, MIT Press, 1999. doi:10.17877/DE290R-5098

[19] Y. Liu and Y. F. Zheng, “One-against-all multi-class SVM classification using reliability measures,” Proceedings. IEEE International Joint Conference on Neural Networks, Montreal, Que., vol. 2, pp. 849-854, 2005. doi:10.1109/IJCNN.2005.1555963

[20] P. Chen and S. Liu, “An improved DAG-SVM for multi-class classification,” Fifth International Conference on Natural Computation, Tianjian, China, pp. 460-462, 2009. doi:10.1109/ICNC.2009.275

[21] H. Chih-Wei, L. Chih-Jen, “A comparison of methods for multiclass support vector machines,” IEEE Trans. on Neural Networks, vol. 13, no. 2, pp. 415-425. 2002. doi:10.1109/72.991427

[22] B. Sidaoui, and K. Sadouni, “Binary tree multi-class SVM based on OVA approach and variable neighbourhood search algorithm,” Int. J. Computer Applications in Technology, vol. 55, no. 3, pp. 183–190, 2017. doi:10.1504/IJCAT.2017.084772

[23] C. Chang and C. Lin, “LIBSVM: a library for support vector machines,” 2021. https://www.csie.ntu.edu.tw/~cjlin/libsvm/

[24] S. Ramgopal et al., “Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy,” Epilepsy & Behavior, vol. 37, pp. 291–307, 2014. doi:10.1016/j.yebeh.2014.06.023

[25] J. Klatt, et al., “The Epilepsiae database: An extensive electroencephalography database of epilepsy patients,” Epilepsia, vol. 53, no. 9, pp. 1669–1676, 2012. doi:10.1111/j.1528-1167.2012.03564.x

[26] A. Mohd Aftar, M. Noratiqah, S. Mohd, L. Ong, A. Fatin, “Prediction of multivariate air quality time series data using long short-term memory network,” Malaysian Journal of Fundamental and Applied Sciences, vol. 18, no. 1, 2022. doi:10.11113/mjfas.v18n1.2393

[27] Z. Muralimohanbabu, K. Radhika, ‘‘Multi spectral image classification based on deep feature extraction using deep learning technique,” Int. J. of Bioinformatics Research and Applications, vol. 17, no. 3, pp 250-261, 2021. doi:10.1504/IJBRA.2021.117169

[28] O. Israa, A. Mouhammad, and A. Mohammad, “Diabetic retinopathy detection method using artificial neural network,” Int. J. of Bioinformatics Research and Applications, vol. 18, no. 4, pp 300-317, 2022. doi:10.1504/IJBRA.2022.126740

[29] C. A. Teixeira et al., “EPILAB: A software package for studies on the prediction of epileptic seizures,” Journal of Neuroscience Methods, vol. 200, no 2, pp.257-271, 2011. doi:10.1016/j.jneumeth.2011.07.002