Neuroscience Informatics 2 (2022) 100087

Contents lists available at ScienceDirect

Neuroscience Informatics

www.elsevier.com/locate/neuri

Original article

# Performance analysis of VEP signal discrimination using CNN and RNN algorithms

Zineb Cheker [a,*], Saad Chakkor [a], Ahmed EL Oualkadi [a], Mostafa Baghouri [b], Rachid Belfkih [c], Jalil Abdelkader El Hangouche [d], Jawhar Laameche [e]

[a] *Laboratory of Information and Communication Technologies (LabTIC), Ecole Nationale des Sciences Appliquées de Tanger, University of Abdelmalek Essaadi [1], Morocco*
[b] *Smart Materials and Artificial intelligence Team, LCCPS, École Nationale Supérieure des Arts et Métiers de Casablanca, Casablanca, Morocco*
[c] *Faculty of Medicine and Pharmacy, Laboratory of Neurology, Tangier, Morocco*
[d] *Faculty of Medicine and Pharmacy, Laboratory of Physiology, Tangier, Morocco*
[e] *Faculty of Medicine and Pharmacy, Laboratory of Chemistry and Biochemistry, Tangier, Morocco*

A R T I C L E I N F O

*Article history:*
Received 31 January 2022
Received in revised form 2 June 2022
Accepted 8 June 2022

*Keywords:*
Visual evoked potential
Latency P100
Deep learning
Recurrent neural network
Convolutional neural network

A B S T R A C T

The visual evoked potential as an electrophysiological signal is mainly used in the neurophysiological exploration of the optic nerves. Traditionally, medical doctors base their diagnosis of specific pathologies related to the time delay of the nerve flow on the time scale. In this context, the VEP latency P100 that reflects a temporal notion is considered the main characteristic on which human interpretation is based. However, its value is influenced by different factors and remains a limited method. This insufficiency triggers our interest instead in deep learning architectures, taking into consideration and adapting to the specificity of each particularity related to the laboratory of the neurophysiological exploration unit in the hospital. The comparison between the results obtained from Matlab by the application of the CNN as well as the RNN, based on the evaluation parameters calculated after k-fold cross-validation, confirms that the CNN-1D architecture can be considered powerful in terms of reliability of classification between signals that are related to pathological subjects and normal ones, which privileges the use of this architecture compared with recurrent neural networks that are less reliable and require more time for execution, subsequently the use of the CNN will allow us to avoid even the extraction of attributes for the discrimination between the two classes object of classification, with the possibility to progressively improve the performance of the solution over time based on the new signals acquired in the VEP analysis laboratory.



## 1. Introduction

This paper discusses the classification of VEP examinations that are considered primarily useful to measure the duration of conduction relative to the neuronal activity from the retina to the occipital cortex [1]. Clinically, this measure gives the clinician information that confirms or not the good integrity of the neuronal function [1] following the use of a visual stimulus (checkerboard or flash), by using electrodes to acquire the electrophysiological signal in the form of a trace in the function of the time. We note that the P100 component called latency, which occurs approximately at 100 ms [2] is considered a discriminating characterization of pathological and normal cases (see Fig. 1).

Since VEP signals can confirm some important information about the integrity of the visual pathways, this test is particularly used to diagnose different types of pathologies, based on the notion of P100 latency for its different clinical indications [3]; Fig. 2 presents different examples of pathologies that are reflected or related to a change or variance of the P100 wave in terms of lengthening/prolongation or based on its marked delay on the time axis accompanied or not by a reduction in amplitude, there may also be bilateral delays of the latency [3].

It is noted that the reliability of the interpretation of VEP signals is influenced in one way or another by different technical factors [3] [4], as well as factors related to the subject who will be diagnosed by the clinical examination as shown in the Fig. 3, and subsequently, this interpretation has triggered a need for standards that must be adopted by each unit of analysis of VEP signals

* Corresponding author.
*E-mail address:* zineb.cheker@gmail.com (Z. Cheker).
[1] Full postal address: Riad Tanger-Alliance Darna-A app 50, Route Rabat, Gzenaya, Tangier, Morocco.

https://doi.org/10.1016/j.neuri.2022.100087

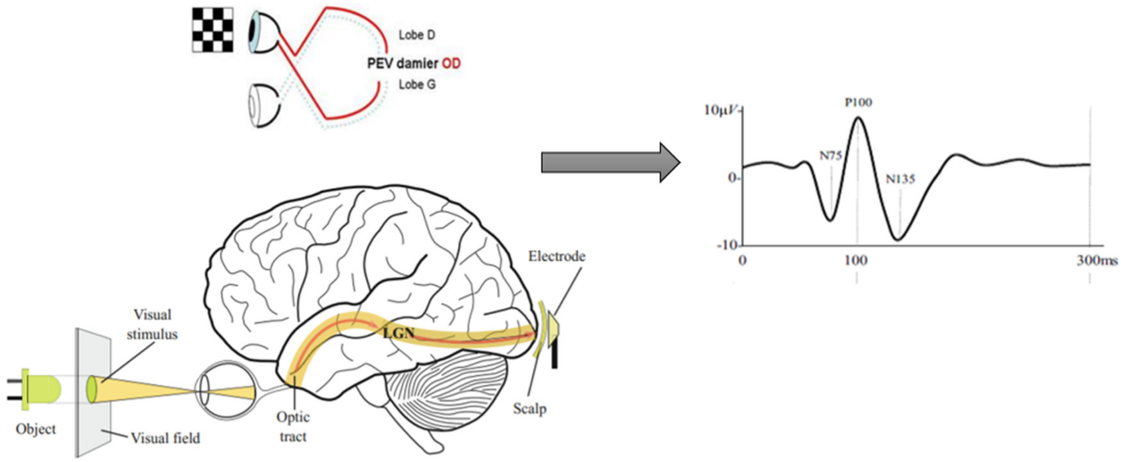


**Fig. 1.** Representation of the stimulation of the visual system and the trace of the signal resulting from the acquisition [1].

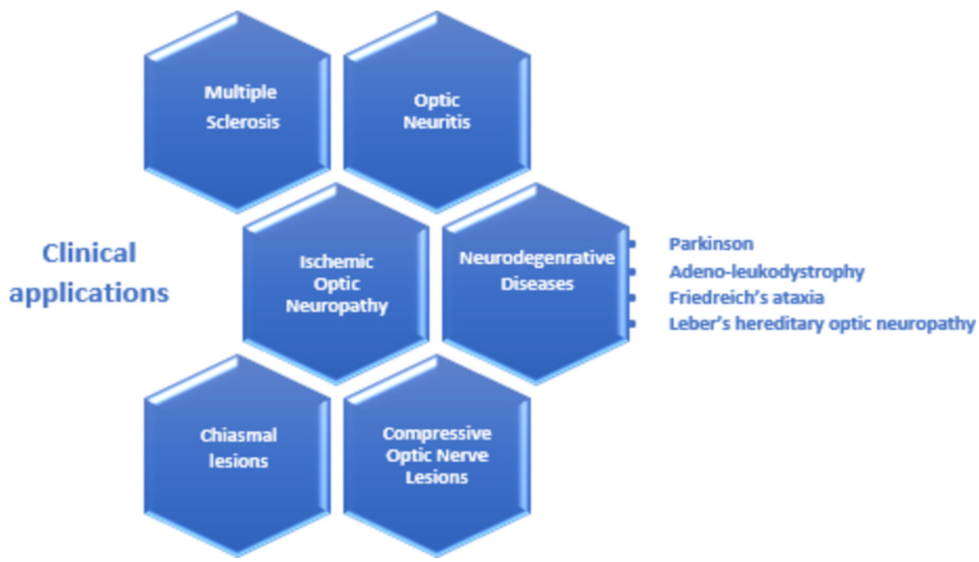


**Fig. 2.** The different categories of pathologies that can be diagnosed by the VEP examination.

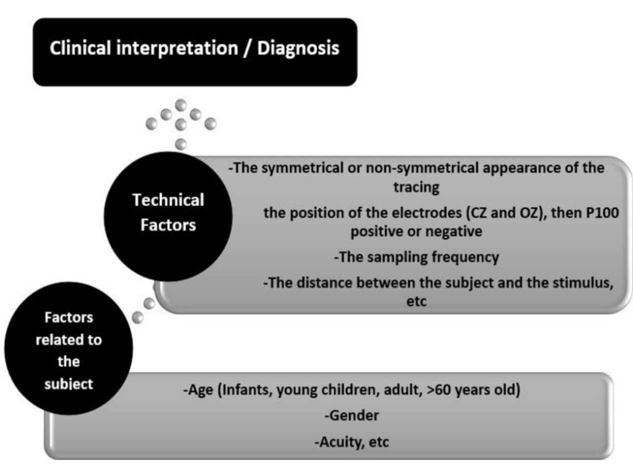


**Fig. 3.** The different factors that influence the clinical interpretation based on the P100 latency.

[5] in order to not have subjective results, [6] confirms that the clinical interpretation based on P100 has an unsatisfactory rate of about 67%.

In our work, we aim to provide another method of classification based on artificial intelligence, comparing the use of two types of neural network architectures CNN and RNN, which allows us to adopt a powerful tool that can be used in analyzing the informational content of the VEP signals, and subsequently the classification of signals based on the entire electrophysiological signal VEP and not only the latency P100.

Which can offer a base of automatic learning that will be at the disposal of the laboratory for analysis of the VEP signals in a hospital structure, by the possibility of learning the program progressively to distinguish the normal and pathological cases, all this aims to ensure an adaptable system with the different factors that influence on the decision making and even the subjectivity of the clinical interpretation.

## 2. Related work

There are several studies in the literature, that deal with the classification of electrophysiological signals based on signal processing and artificial intelligence. In [6], the classification of VEP signals was based on the spectral analysis to high resolution with the approach of Pisarenko, used with classifiers containing neural networks, the result shows that this kind of classification could reach 95%. In the article [4] the authors made an analysis based on the CWT wavelet transform in order to have a discriminating spectral representation between normal and pathological cases on a

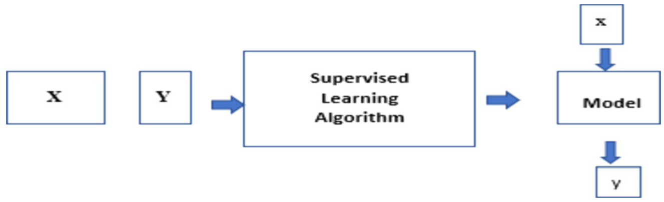


**Fig. 4.** General structure of supervised learning.

dataset that gathers 30 normal subjects and 20 pathological cases, the results that are based on this method have outperformed the latency-based method.

In neurophysiology, the VEP is considered as a modification of the electrical potential produced by the nervous system using a stimulus, it is recorded through a technique similar to the EEG, in this sense, several works have been applied on this type of biosignal. [7] Was the first study to apply CNN for EEG diagnosis, in order to detect seizure crisis; the classification system achieved an accuracy equal to 88,67%. Furthermore, the authors in [8] discuss the effectiveness of using CNN, RNN, and DBN in deep learning classification applied to EEG tasks: such as motor imagery, sleep stage scoring, etc. the authors of [9], investigate the use of the RNN to monitor mental depression from EEG signals, which has given good results.

Since the ECG is an electrophysiological signal like the EEG and the VEP, we have also benefited from works that are focused on the ECG signal; the authors in [7] propose a system for automatic diagnosis of the myocardial infraction using; the results show that the CNN can have a good accuracy equal to 93,22% with a signal without noise.

While in [10] authors showed the results when we apply the RNN architecture to classify the ECG arrhythmia; this approach investigates the different accuracies of three RNN models, and the maximum accuracy was 88,1%.

## 3. Proposed methodology

### 3.1. Supervised learning

We have based our work mainly on supervised learning, the signals included in the database where each signal has a well-defined class [11], in our case we are dealing with a binary classification (0 and 1), as shown in Fig. 4.

According to [12], supervised learning is a very powerful method in terms of signal classification, especially when we talk about deep learning. Given the fact that our dataset is quite limited in the number of signals available to evaluate the two architectures, supervised learning may be more appropriate in this case.

### 3.2. Deep learning

Deep learning is well used for the classification of 1D signals and proved to be of remarkable interest [13], we are talking precisely about artificial neural networks as illustrated in Fig. 5 and characterized by their activation specificity [14] which will be used as a basis for classification in our topic, the CNN and the RNN will be compared proving that deep learning architectures can remedy the insufficiencies of the human and classical interpretation.

The purpose of our subject is to use the VEP signals as inputs for the two types of algorithms based on CNN and RNN, then compare the two types of results in order to have a reliable classification for the two classes, and adopt the most discriminating architecture and try to adapt and optimize it with the new VEP signals from neurological functional exploration units as shown in Fig. 6.

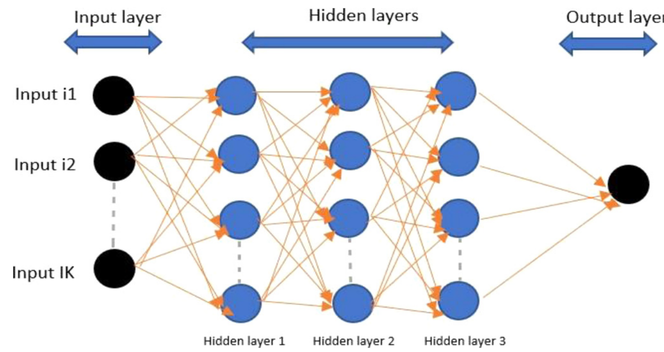


**Fig. 5.** Illustration of artificial neural with 3 hidden layers.

### 3.3. CNN: convolutional neural networks

The basic configuration of the CNN is structured on the following steps [7] illustrated in Fig. 7: (1) Convolution that contains filters to highlight the important attributes in the input, (2) ReLU which is responsible for converting values below zero to 0, (3) Pooling function that simplifies data by performing non-linear subsampling, (4) Fully connected layer which applies a linear combination, and softmax to present probabilities of each category in our classification [7].

### 3.4. RNN: recurrent neural networks

It is a powerful type of neural network that is mainly based on the previous information [15]. Fig. 8 shows the details of a recurrent layer with the inputs (I) and outputs (O) at a time (t), the recurrent connections are noted with a red color.

### 3.5. Evaluation metrics

Our comparative analysis can be based on what are called evaluation and differentiation parameters, calculated after the simulation (training and test), which can clarify the advantages and the inconveniences of each method to have in the end a good justification on which the comparative analysis will be based on.

In our subject, we have based our comparison mainly on 4 parameters: accuracy, recall, precision, and F1 score, [16] the equations that define these metrics are described below:

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \times 100\% \tag{1}$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \tag{2}$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \tag{3}$$

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100\% \tag{4}$$

## 4. Simulation

### 4.1. Dataset

We use a dataset that gathers signals acquired in the 'University of medicine and dentistry of New Jersy, Department Directory | University Hospital, Newark', recorded under visual stimulation by the checkerboard pattern that is in front of the subject with a distance of one meter as presented in Fig. 9. The recordings were based on the CZ and OZ electrodes placed on the scalp in relation to the third reference electrode, the recorded signals were amplified, and converted through an A/D converter with a sampling frequency of 1 KHz [4].

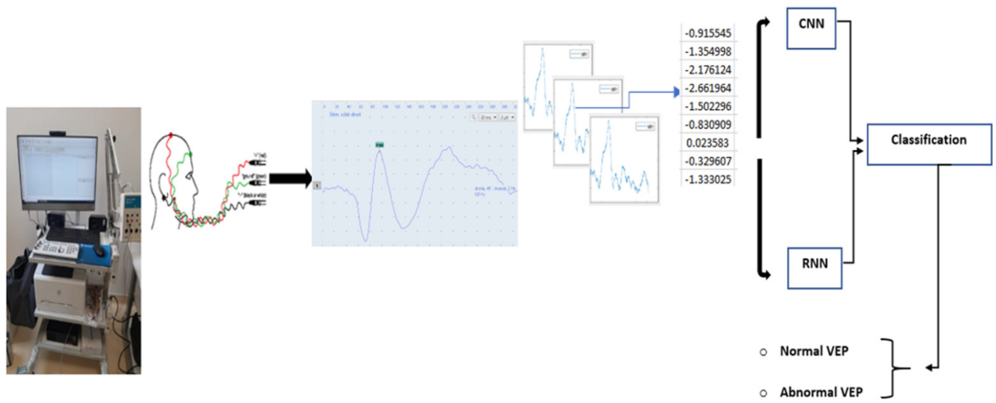


**Fig. 6.** The studied approach for the classification of VEP signals.

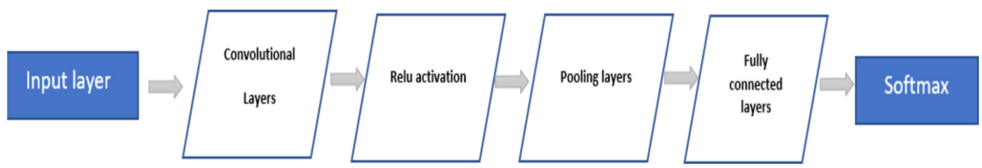


**Fig. 7.** The architecture of the convolutional neural networks.

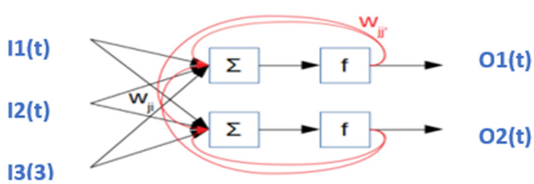


**Fig. 8.** Example of a simple RNN layer with three inputs and two outputs.

The dataset used consists of 14 normal signals, the rest is related to subjects that represent the pathological category, with 512 temporal values.

The number of data available for our subject is considered insufficient [17], which can lead to relatively unreliable evaluations, in order to remedy this problem, we have integrated in our evaluation signals that have a normal P100 but that are related to pathological subjects in order to train our system and ensure a good classification that exceeds the one based on classical interpretation, in this sense, it is also necessary to choose the method on which our evaluation will be based:

- Train-Test Split:

We can divide our database into two classes such as 80%-20% [18] or 60%-40%, which can give us an idea about the difference between the different performances of the learning model adopted on data that are not used during training. Fig. 10 illustrated a dataset divided into 2 classes a% and b%.

- K-fold cross-validation:

In our study, we divided the dataset into 25 parts. In this analysis, we will use cross-validation 25 times. In this procedure, we use one signal to test, and the other $k - 1 = 24$ to train our model in each interaction [19]. By doing this, each time we start a new interaction, we use a different fold for testing. Figs. 11 and 12 present the procedure adapted with $k = 5$.

### 4.2. Architecture CNN-1D

Convolutional neural network (CNN) is the most effective type of deep learning architecture for 2D image classification [20], Several works confirm that the CNN gives very good results, among the interesting work, we find the detection of tumors from MRI images [21], and the multiclass classification of skin tumors [22].

In order to use a CNN model with electrophysiological signals, some papers have converted 1D signals into 2D data [23], [24]. However, [25] has shown that the architecture used must be simple with a minimum of hidden layers, in our case, we will use a simple CNN architecture with only 5 hidden layers to obtain an optimal classification as shown in Fig. 13. We note that the data used for training and for testing must be grouped in a 4D array [26], [27].

### 4.3. Architecture RNN

Regarding the RNN architecture, it should be pointed out that networks that are too small are unable to learn the problem correctly, while networks that are too large tend to overfit the training data and consequently to low performance [28], in our work the optimal number of layers is developed through an elaborate trial and error procedure. However, there are not enough studies that have used RNNs, it is necessary to investigate further the problematic looking for the best data that will be used as input for the good exploitation of the RNN.

The RNN algorithm adopted in our work consists of 6 layers, we are based on the results of the training confusion matrices for less than 6 layers that were not satisfactory.

## 5. Result – overview of a practical experiment

### 5.1. Comparison between CNN and RNN classification results

Since we have limited training samples, we used the exclusion principle to train the two architectures CNN and RNN; in fact, the training is done on the 24 signals, and the test is done with the remaining signal, it is a type of k-fold cross-validation. Fig. 14 illustrates the procedure followed to divide our dataset into 25 signals, and we apply the cross-validation 25 times.

We note that: Red is for pathological signals – Green for normal signals – Blue for test signals.

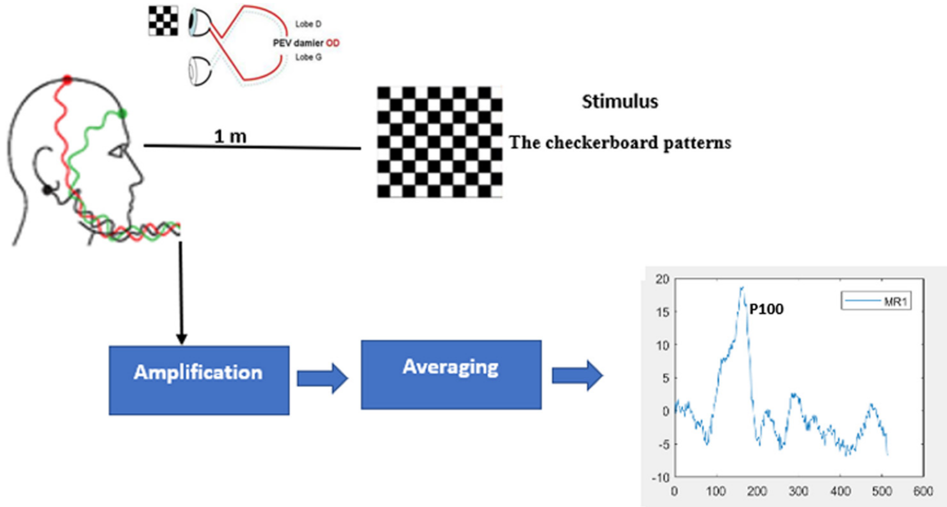


**Fig. 9.** Acquisition of VEP signals by the checkerboard method.


**Fig. 10.** Division of the database into two parts: training and testing.

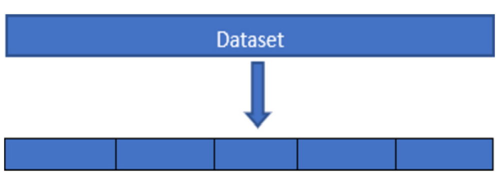


**Fig. 11.** Division of the database into $K = 5$ parts.

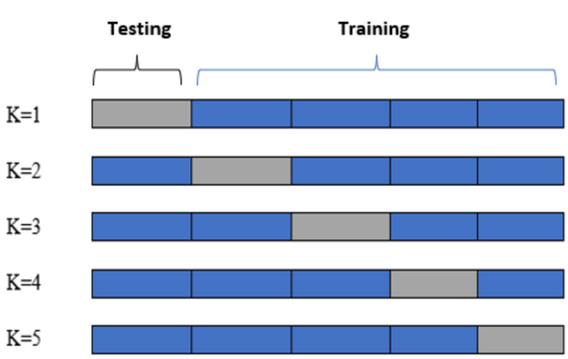


**Fig. 12.** Test procedure by exclusion for $k = 5$.

**Table 1**
Test results by exclusion.

| Class - Signal | CNN | RNN |
|---|---|---|
| AS | F | F |
| BS | F | T |
| CS | T | F |
| DS | F | F |
| ES | F | F |
| FS | F | F |
| GS | F | F |
| HS | F | F |
| IS | T | T |
| JS | T | T |
| KS | T | T |
| LS | T | T |
| MS | T | T |
| NS | T | T |
| OS | T | T |
| PS | F | F |
| QS | F | F |
| RS | T | T |
| SS | T | F |
| TS | F | F |
| US | F | F |
| VS | F | F |
| WS | F | F |
| XS | T | T |
| YS | T | F |

**Table 2**
Comparison between CNN and RNN in terms of accuracy and sensitivity.

| Performance evaluation metrics | CNN | RNN |
|---|---|---|
| Accuracy | 96% | 88% |
| Precision | 91% | 90% |
| Recall | 100% | 81% |
| F1 Score | 95% | 85% |

The output values resulting from the two algorithms are presented in Table 1, for each signal (AS, BS, ..., YS), F represents false cases, and T represents pathological signals. We note that the correct results are colored green and the wrong ones are colored red.

The results of the values from the 4 evaluation parameters are presented in Table 2, we can clearly notice the difference between the two algorithms, which gives an important consideration for the CNN rather than the RNN.

### 5.2. Practical experiment – VEP classification

In order to ensure the results confirm that CNN is better placed for the classification of VEP signals, a practical test was carried out on a group of people in the neurological functional exploration unit of the University Hospital of Tangier, it consists in acquiring a number of VEP signals in order to test the ability of CNN and RNN to determine the class of each signal (see Fig. 15).

- Acquisition conditions:

We tried to make the VEP examination in almost similar conditions to those related to the dataset used in the previous section for training and testing, we note that the sampling frequency is equal to 1 KHz, as well as the distance between the checkerboard

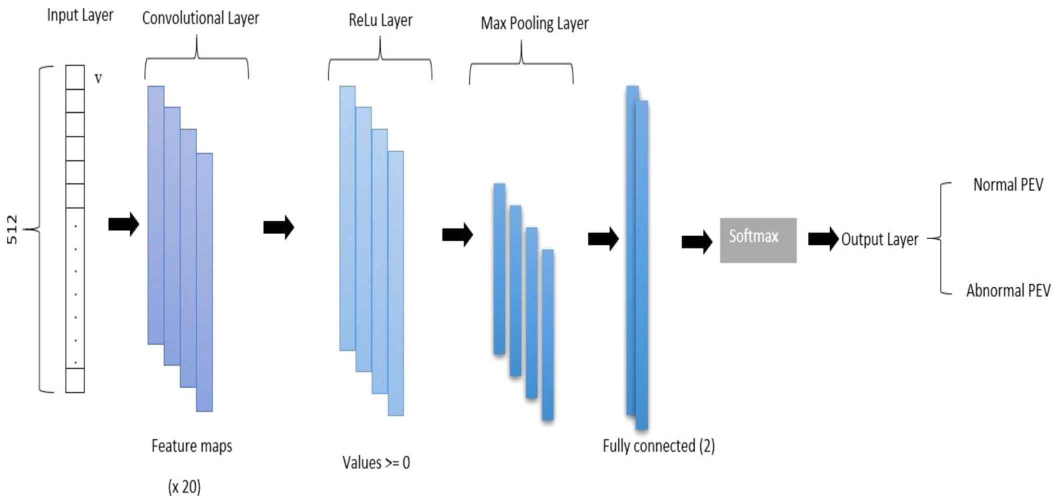


**Fig. 13.** The proposed CNN architecture for VEP signal classification.

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 23 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |

**Fig. 14.** Test procedure by exclusion of our dataset with k= 25.

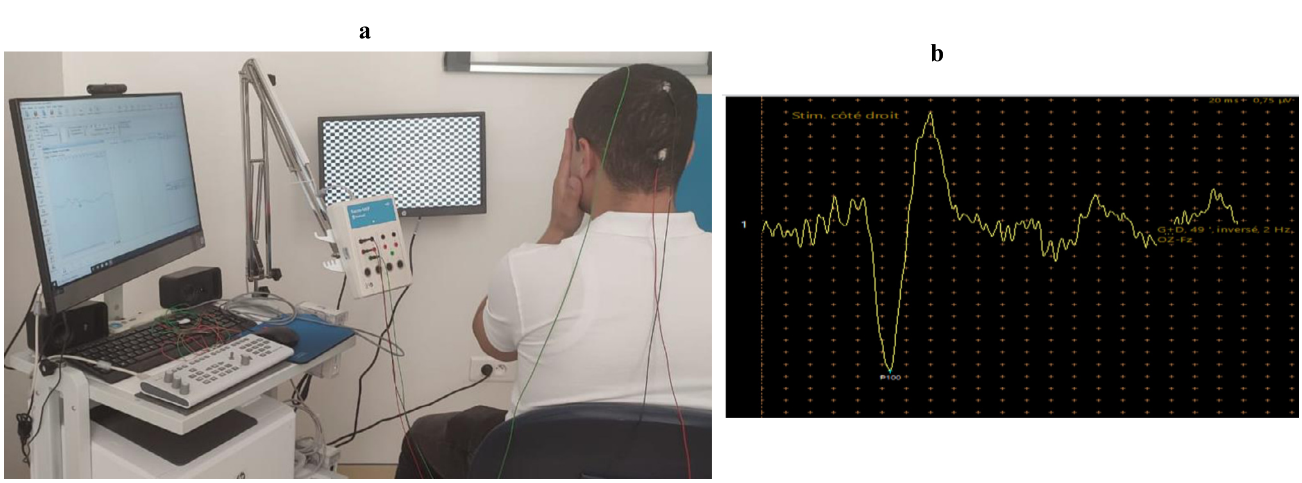


**Fig. 15.** Illustration of the subject's position during the acquisition connected to the three electrodes (a), and the VEP trace after the acquisition (b).

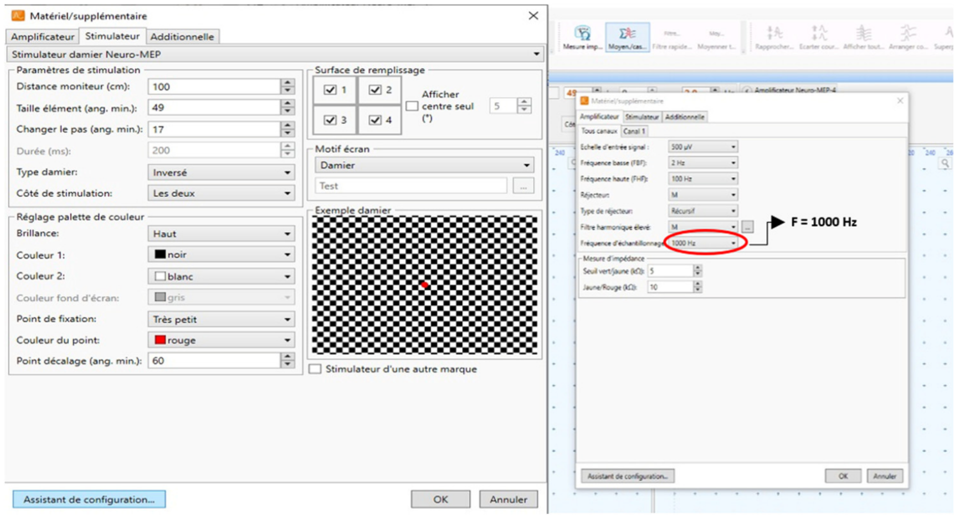


**Fig. 16.** Options for setting values before the start of the acquisition: sampling frequency, checkerboard parameter.

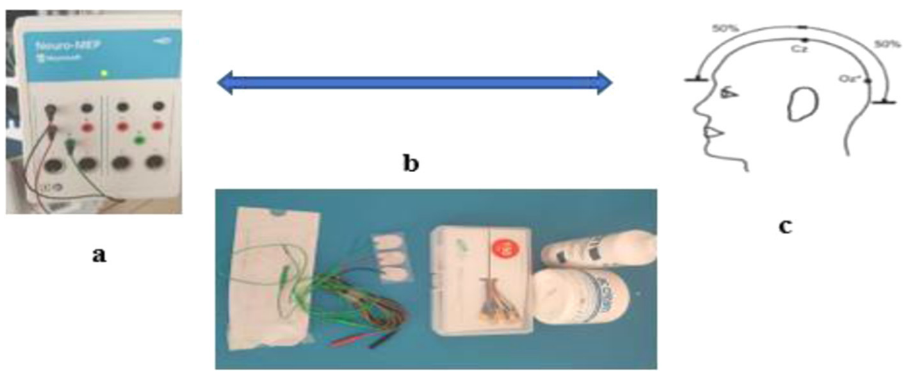


**Fig. 17.** Connection material used (b) for the link between the integrated amplifier (a) in the VEP device and the patient (c).

monitor and the subject is equal to 1 m, the subjects have to concentrate on the small red fixation point placed in the middle of the monitor screen as shown in Fig. 16.

We note in this context that several types of electrodes can be used to capture the signal as shown in Fig. 17, in our test we will use the cups EEG electrodes (Ag/AgCl) with preparation in advance of the area of contact with the abrasive gel and the application of a conductive paste to decrease the impedance electrode-skin, three electrodes were used: OZ, CZ, and the neutral.

- Subjects:

A group of 17 individuals representing the non-pathological cases that have passed the VEP test, this dataset of people includes women (F) and men (H) of different ages as indicated in Table 3.

After the acquisition procedure, the displayed signals have been divided into two parts: the first part gathers 3 signals representative of the male and female VEPs and also with a positive and negative P100, which will be added to the database related to the training so that our program can learn new specifications of the signals related to this examination session, the other 14 signals will all be used to test the two architectures in order to confirm or not the result found in the previous section, Fig. 18 illustrates 4 signals acquired from 4 different subjects.

- Result:

The program based on the CNN was able to classify all the signals, the result of the whole test gave the value 0, in other words,

**Table 3**
Information about the subjects participating in the VEP examination.

| Subject | Age | Gender | P100 |
|---|---|---|---|
| 1 | 29 | H | 107 |
| 2 | 25 | H | 98 |
| 3 | 26 | H | 113 |
| 4 | 26 | F | 99 |
| 5 | 30 | H | 101 |
| 6 | 35 | H | 102 |
| 7 | 28 | H | 106 |
| 8 | 36 | H | 101 |
| 9 | 27 | H | 99 |
| 10 | 26 | F | 95 |
| 11 | 25 | H | 102 |
| 12 | 30 | F | 94 |
| 13 | 34 | F | 94 |
| 14 | 42 | F | 110 |
| 15 | 28 | F | 97 |
| 16 | 26 | F | 97 |
| 17 | 21 | F | 95 |

the signals tested are related to subjects who do not present the pathological cases.

On the other hand, the RNN gave values almost equal to 0 for the 14 signals, reflecting the non-similarity of the ability to decide on the class of non-pathological signals, but in return, the resulting value reflects the class 0.

The comparison must take into consideration the execution time, since we aim to make the program in continuous training using the new acquisitions to make it more powerful and discriminating, and subsequently that means the increase of dataset

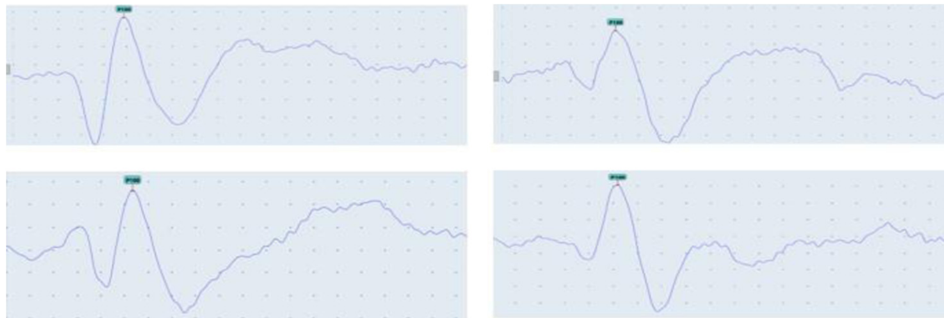

**Fig. 18.** VEP signals collected from 4 different subjects (2 women and 2 men).

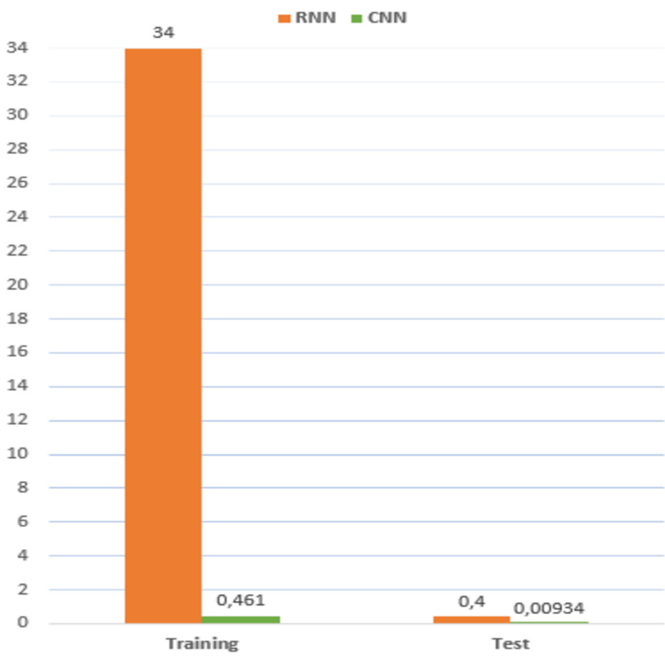


**Fig. 19.** Comparison between CNN and RNN in terms of time needed for training and testing.

implies a change relative to the execution time. Fig. 19 represents a comparison between the average time related to the two architectures.

## 6. Discussion

It is clear that the traditional method for the classification of VEP signals based on the P100 latency is less reliable, with the confirmation of the existence of signals that have latencies included in the margin of normality, but that are related to pathological cases [4], which requires avoiding the interpretation based on temporal information as the P100 by triggering trials based on signal processing; the authors of [4] have applied the continuous wavelet transform CWT, allowing good discrimination between the two classes, and subsequently offering the possibility of minimizing the acquisition time of VEP signals.

On the other hand, several works confirm that for an efficient classification, an accurate feature method is imperative to extract distinctive information from the original signals, in this context, the authors of [6] applied a high-resolution spectral analysis considering the non-stationarity of the signal and then applying a classifier based on neural networks, which gave a classification rate equal to 92%. However, in the case where we aim to use attribute extraction, it should be pointed out that classification performance can be degraded if features are not extracted in an appropriate way.

In our work, the classification provided by the CNN appears to be efficient without resorting to the extraction step, which is already integrated by default in its process, exceeding the classification based on RNN architecture. Subsequently, the use of CNN will allow us to adapt it to each laboratory by training our program, ensuring a more reliable diagnosis and subsequently the appropriate treatment for the patients.

## 7. Conclusion

It goes without saying that the new techniques resulting from artificial intelligence, and especially from deep learning have brought a remarkable added value to the analysis of electrophysiology signals, this is proven in our research work, given the fact that the CNN has shown its powerful character in terms of classification of VEP signals, which will allow considerably to good support of the patients who have problems related to this type of signal.

The comparative analysis between the two algorithms considerably favored the classification based on the CNN with a rate of 95% for the F1 score, which directs our interest in the investigation of this type of algorithm to ensure a good classification of VEP signals; this result could be explained by the fact that CNN architectures have the particularity to group together the two main phases of the traditional electrophysiological signal classification into a simple learning structure: determination of the attributes and then give the label-class corresponding to each case. Our proposed solution can assist doctors in their clinical diagnosis related to VEP signals. Among the perspectives of this work, the experimentation of CNN in the classification of VEP signals in the neurological exploration service of the University Hospital of Tangier and its application in the diagnosis and monitoring of pathologies related to the VEP signal as neuropathy and multiple sclerosis.

In future research, it will be interesting to investigate other comparisons such as the one between the use of CNN- 1D and CNN-2D using a large VEP dataset.

## Informed consent and patient details

The authors declare that they obtained a written informed consent from the patients and/or volunteers included in the article and that this report does not contain any personal information that could lead to their identification.

## Funding

This work has been supported by: University of Abdelmalek Essaâdi, Tetouan, Morocco.

## Declaration of competing interest

The authors declare that they have no known competing financial or personal relationships that could be viewed as influencing the work reported in this paper.

## Acknowledgements

This work is funded and supported by the University of Abdelmalek Essaâdi, Tetouan, Morocco.

Our special thanks to the "University Hospital Center of Tangier".

## References

[1] Anibal Cotrina, Toward Brain-Computer Interaction in Paralysis, Human–Computer Interaction Series, Springer International Publishing, Cham, 2017.

[2] Maelenn Cabon, Nouvelle méthode de traitement des potentiels évoqués visuels pour l'estimation de l'acuité visuelle chez le jeune enfant, n.d., 275, submitted for publication.

[3] Frank W. Drislane, Visual evoked potentials, in: A.S. Blum, S.B. Rutkove (Eds.), The Clinical Neurophysiology Primer, Humana Press Inc., Totowa, NJ, 2007.

[4] H. Drissi, F. Regragui, J.P. Antoine, M. Bennouna, Wavelet transform analysis of visual evoked potentials: some preliminary results, ITBM-RBM 21 (2) (April 2000) 84–91, https://doi.org/10.1016/S1297-9562(00)90010-X.

[5] Ruby Sharma, Sandeep Joshi, K.D. Singh, Avnish Kumar, Visual evoked potentials: normative values and gender differences, J. Clin. Diagn. Res. 9 (7) (July 2015) CC12–CC15, https://doi.org/10.7860/JCDR/2015/12764.6181.

[6] Mohammed Majidi Himmi, Outils d'aide au diagnostic basés sur l'analyse spectrale haute résolution du potentiel évoqué visuel, s. d., 125 (2007).

[7] U. Rajendra Acharya, Hamido Fujita, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, Muhammad Adam, Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals, Inf. Sci. 415–416 (November 2017) 190–198, https://doi.org/10.1016/j.ins.2017.06.027.

[8] Alexander Craik, Yongtian He, Jose L. Contreras-Vidal, Deep learning for electroencephalogram (EEG) classification tasks: a review, J. Neural Eng. 16 (3) (June 2019) 031001, https://doi.org/10.1088/1741-2552/ab0ab5.

[9] Avik Sarkar, Ankita Singh, Rakhi Chakraborty, A deep learning-based comparative study to track mental depression from EEG data, Neurosci. Inform. 2 (4) (December 2022) 100039, https://doi.org/10.1016/j.neuri.2022.100039.

[10] Shraddha Singh, Saroj Kumar Pandey, Urja Pawar, Rekh Ram Janghel, Classification of ECG arrhythmia using recurrent neural networks, Proc. Comput. Sci. 132 (2018) 1290–1297, https://doi.org/10.1016/j.procs.2018.05.045.

[11] Manuel Lopez-Martin, Belen Carro, Antonio SanchezEsguevillas, Application of deep reinforcement learning to intrusion detection for supervised problems, Expert Syst. Appl. 141 (March 2020) 112963, September 2019.

[12] Jason Brownlee, Impact of dataset size on deep learning model skill and performance estimates, Machine Learning Mastery (blog), https://machinelearningmastery.com/, January 2019.

[13] Laith Alzubaidi, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie, Laith Farhan, Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, J. Big Data 8 (1) (December 2021) 53.

[14] Michael Zgurovsky, Victor Sineglazov, Elena Chumachenko, Artificial Intelligence Systems Based on Hybrid Neural Networks: Theory and Applications, Studies in Computational Intelligence., vol. 904, Springer International Publishing, Cham, 2021.

[15] R. Vinayakumar, K.P. Soman, Prabaharan Poornachandran, Evaluation of recurrent neural network and its variants for intrusion detection system (IDS), in: Deep Learning and Neural Networks, Information Resources Management Association, IGI Global, 2020, pp. 295–316.

[16] Amir Hossein Ansari, Ofelie De Wel, Mario Lavanga, Alexander Caicedo, Anneleen Dereymaeker, Katrien Jansen, Jan Vervisch, Maarten De Vos, Gunnar Naulaers, Sabine Van Huffel, Quiet sleep detection in preterm infants using deep convolutional neural networks, J. Neural Eng. 15 (6) (December 2018) 066006.

[17] Pankaj Mehta, Marin Bukov, Ching-Hao Wang, Alexandre G.R. Day, Clint Richardson, Charles K. Fisher, David J. Schwab, A HighBias, low-variance introduction to machine learning for physicists, Phys. Rep. 810 (May 2019) 1–124, https://doi.org/10.1016/j.physrep.2019.03.001.

[18] Dima Shulga, 5 reasons why you should use cross-validation in your data science projects, Medium (September 2018), https://towardsdatascience.com/5-reasons-why-you-should-use-crossvalidation-in-your-data-science-project-8163311a1e79.

[19] Zahra Nematzadeh, Roliana Ibrahim, Ali Selamat, Comparative studies on breast cancer classifications with k-fold cross validations using machine learning techniques, in: 2015 10th Asian Control Conference (ASCC), IEEE, Kota Kinabalu, 2015, pp. 1–6.

[20] Grega Vrbancic, Vili Podgorelec, Automatic classification of motor impairment neural disorders from EEG signals using deep convolutional neural networks, Elektron. Elektrotech. 24 (4) (August 2018) 3–7, https://doi.org/10.5755/j01.eie.24.4.21469.

[21] Arkapravo Chattopadhyay, Mausumi Maitra, MRI-based brain tumour image detection using CNN based deep learning method, Neurosci. Inform. 2 (4) (December 2022) 100060, https://doi.org/10.1016/j.neuri.2022.100060.

[22] Karar Ali, Zaffar Ahmed Shaikh, Abdullah Ayub Khan, Asif Ali Laghari, Multiclass skin cancer classification using EfficientNets – a first step towards preventing skin cancer, Neurosci. Inform. 2 (4) (December 2022) 100034, https://doi.org/10.1016/j.neuri.2021.100034.

[23] Yunan Wu, Feng Yang, Ying Liu, Xuefan Zha, Shaofeng Yuan, A comparison of 1-D and 2-D deep convolutional neural networks in ECG classification, s. d., 4.

[24] Serkan Kiranyaz, Turker Ince, Moncef Gabbouj, Real-time patient-specific ECG classification by 1-D convolutional neural networks, IEEE Trans. Biomed. Eng. 63 (3) (2016) 664–675, https://doi.org/10.1109/TBME.2015.2468589.

[25] Imran Ullah Khan, Sitara Afzal, Jong Weon Lee, Human activity recognition via hybrid deep learning based model, Sensors 22 (1) (January 2022) 323.

[26] Vadim Lebedev, Algorithms for speeding up convolutional neural networks, s. d., p. 106.

[27] R. Ramadevi, V. Prakash, B. Sheela Rani, Role of hidden neurons in an Elman recurrent neural network in classification of cavitation signals, Int. J. Comput. Appl. 37 (7) (January 2012) 9–13, https://doi.org/10.5120/4618-6626.

[28] Meysam Vakili, Mohammad Ghamsari, Masoumeh Rezaei, Performance analysis and comparison of machine and deep learning algorithms for IoT data classification, n.d., 13.