# Drowsiness Detection Using Electroencephalography Signals: A Deep Learning Based Model

**Luis Guarda[a], Nicolás Astorga[a], Enrique López Droguett[a], Marcio Moura[b], Marcelo Ramos[c]**

[a] Mechanical Engineering Department, University of Chile, Santiago, Chile
[b] Production Engineering, Federal University of Pernambuco, Recife, Brazil
[c] Naval Engineering, University of Sao Paulo, Sao Paulo, Brazil

**Abstract:** Due to the automation and technological development in the work environment, it is increasingly common the execution of monotonous tasks requiring high levels of attention. When these tasks play an important role in maintaining and guaranteeing safety in work environments such as control rooms, it is imperative for the operators to retain adequate level of alertness. Early detection of drowsiness has become of vital importance to ensure the correct and safe development of the aforementioned tasks. Due to the transient mental state of a human subject between alertness and drowsiness, automated drowsiness detection is a complex problem to tackle. Using electroencephalogram signals (captured from various electrodes positioned in the scalp), it is possible to record variations in the electrical potential of an individual's brain where each of them gives specific information about a subject's mental state. In particular, the frontal and parietal channels provide information about the level of drowsiness of an individual, which is represented in variations of the alpha and theta waves. Currently, Convolutional Neural Networks or ConvNets are one of the most effective deep learning techniques for image recognition due to its high capacity to extract essential hierarchical characteristics and patterns from images. Therefore, this paper presents a deep learning-based model for drowsiness detection via ConvNets. The proposed method is applied to the "ULg Multimodality Drowsiness Database" through the use of spectrograms from electroencephalography signals' channels. The preliminary results are encouraging with the proposed model achieving an average accuracy superior to 86% for the classification of subjects between baseline (alert) and drowsy states.

**Keywords:** Deep Learning, Drowsiness, Electroencephalography, Convolutional Neural Network.

## 1. INTRODUCTION

Reliability in every engineering process is fundamental to ensure its correct and safe development. For this reason, it becomes imperative to study of reliability of the equipment's operator, who usually performs monotonous tasks that require a high level of attention, and in many cases they are linked to guarantee the safety of the process.

In this context, human reliability is defined as the probability that a person (1) correctly performs an action required for the time required and (2) that he does not perform any strange activity that may degrade the system [1]. The problem is that it is insufficient to determine if an operator can perform a critical task only by analyzing its apparent level of performance since it can perform a compensatory mental process so that its performance in the task might seem normal.

In this context, drowsiness is defined as a state of consciousness with oscillations between sleep and wakefulness, and an irresistible desire to sleep, accompanied by heaviness and / or clumsiness. This can cause slow reaction times and a reduction in vigilance. Drowsiness is also used interchangeably with the term of fatigue [1].

Using electroencephalography signals (captured from several electrodes located on the scalp), it is possible to record variations in the electrical potential of an individual's brain, where each of them provides specific information about the mental state of the subject. In particular, the frontal and parietal channels provide information about the level of drowsiness of an individual, which is represented in variations of the alpha and theta waves [2]. Also, it has been found that there is a significant relationship between the variation of alpha and theta waves and the Karolinska sleepiness scale [3].

In general, the use of electroencephalography signals for the detection of drowsiness has been accomplished by the manual extraction and selection of features. Alternatively, eye tracking devices has also been employed [4]. Although good results have been obtained, the use of eye tracking devices makes this technique too invasive for the operator.

Convolutional Neural Networks (ConvNets or CNNs) are one of the most effective deep learning techniques for image recognition due to their capacity to automatically extract hierarchical characteristics and essential patterns to perform classification tasks [5]. In this context, Hung [7] used electroencephalography signal spectrograms to perform the detection of drowsiness in conductors.

In this paper we present a Convolutional Neural Network based model that processes electroencephalography signals to detect whether a subject is under drowsiness. In addition, we compare the CNN model against shallow Neural Networks, Support Vector Machines and Random Forest.

## 2. BACKGROUND

### 2.1 Fatigue/Drowsiness

Mental fatigue is described as a change in both psychological and physiological states that an individual experiments during the development of a cognitive activity that demands high concentration for a prolonged period [8]. These changes are manifested mainly as a decline in the subject's cognitive and psychomotor performance, that is, a deterioration in efficiency during the execution of a task. Because of this, drowsiness is linked to a state of diminished alertness, where fatigue and lack of energy gradually predominate.

From this perspective, fatigue is considered as an indicator of an implicit problem: the loss of an individual's basic resources such as reacting in a timely and appropriate manner to an emergency or unforeseen event [9].

As a person suffers from mental fatigue, the following symptoms appear:

- Slow or clumsy movements
- Decrease in the motor speed of reaction
- Appearance of blurred or double vision
- Difficulty concentrating or staying alert
- Difficulty remembering

This implies that the individual who, under the effects of fatigue, works in areas where a high level of attention is necessary, can generate accidents.

### 2.2 Electroencephalography

Electroencephalography (EEG) is an electrophysiological method capable of recording the bioelectrical activity generated by the cortical neurons of the brain and their variation over time. The EEG is performed by placing small disc electrodes on the scalp of the head or using an electrode cap. The electrodes are connected to an electroencephalogram, which amplifies the potentials of brain waves and their activity [10].

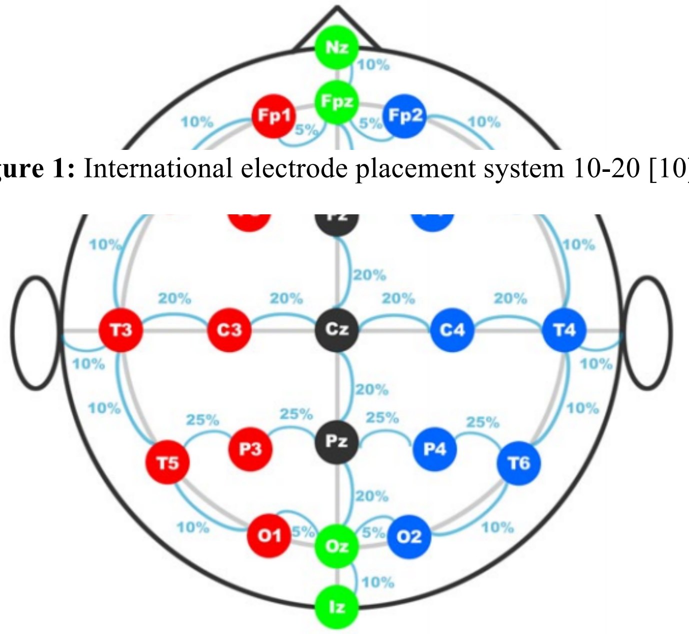



**Figure 1:** International electrode placement system 10-20 [10].

To ensure the standardized reproducibility of the EEG signal capture, the international system 10-20 [10] is used. This system is based on the relationship between the location of an electrode and the underlying area of the cerebral cortex. In Figure 1, the conventional disposition of the international system 10-20 is shown.

EEG signals can be classified into four main segments depending on their frequency band. In addition, each frequency band is related to different states of consciousness, such as alertness, intense concentration, drowsiness, etc. The frequency bands can be classified as [10]:

- **Delta (0.5-4 Hz):** its appearance occurs mainly during deep sleep
- **Theta (4-8 Hz):** this type of wave is characterized by a state of drowsiness of the individual with reduced consciousness
- **Alpha (8-13 Hz):** alpha waves represent a state of low brain activity and relaxation both physical and mental, although aware of the environment
- **Beta (13-20 Hz):** waves of this type are emitted when we are in a conscious or alert state. Beta waves denote intense mental activity

### 2.3 Karolinska Sleepiness Scale

In order to establish a relationship between the actual level of drowsiness perceived by an individual and the physiological data that can be extracted from it to study their level of sleepiness, and as a way to avoid the detection of fatigue events where the subject is self-perceive alert, there is the Karolinska Sleepiness Scale (KSS). This scale measures the subjective level of drowsiness of an individual at a certain time of the day. On this scale, which ranges from 1 to 9, the subjects indicate which level best reflects the psychophysical state experienced in the last 10 minutes [11] [3]. The description of each level of the scale is shown in the Table 1.

**Table 1:** Karolinska Sleepiness Scale [11].

| Scale | Psychophysical state |
|---|---|
| 1. | Extremely alert |
| 2. | Very alert |
| 3. | Alert |

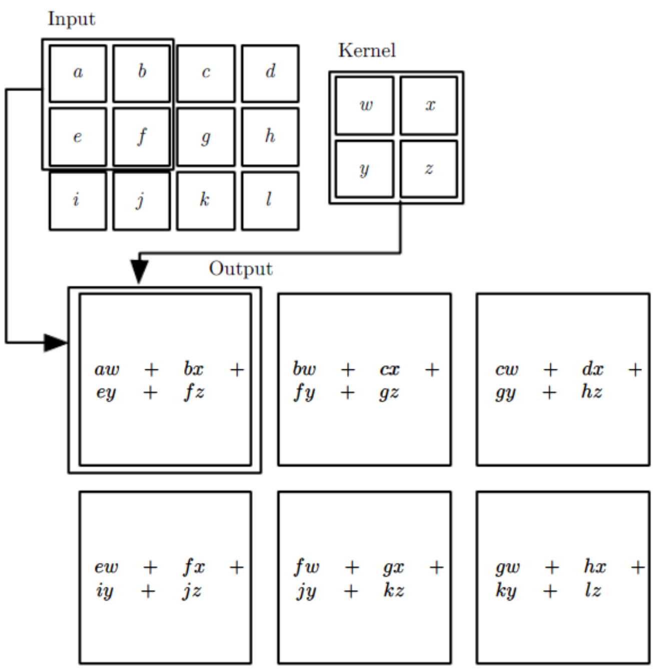
| | |
|---|---|
| 4. | Rather alert |
| 5. | Neither alert nor sleepy |
| 6. | Some sighs of sleepiness |
| 7. | Sleepy, but no difficult remaining awake |
| 8. | Sleepy, some effort to keep alert |
| 9. | Extremely sleepy, fighting sleep |

### 2.4 Convolutional Neural Network

Convolutional Neural Networks are a deep learning algorithm that explicitly assumes that the input data corresponds to images. Under this assumption, the CNNs are composed of three types of layers, each of which fulfills a specific role. These layers are:

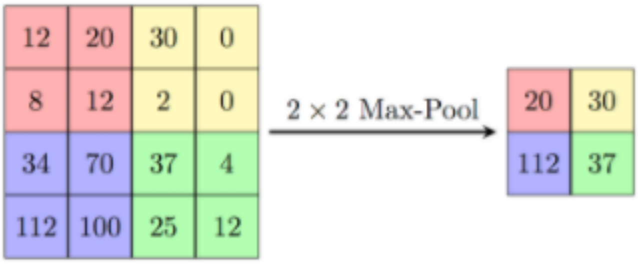
- **Convolutional layer**: In this first layer, a convolution is made between the input image (s) and a filter function (also called kernel), which seeks to extract specific characteristics of the image (s). This technique allows certain characteristics to become more dominant in the output image, because they have a higher weight in the pixels that represent them. The convolution performed between the aforementioned functions is addressed numerically as a matrix product, which allows to work with entries of variable size [12]. The operation of a convolutional layer by way of example is illustrated in the Figure 2.

**Figure 2:** Example of 2D convolution [12].




- **Pooling layer:** Generally the pooling layer is located after the convolution layer, and the function of this is to perform a sampling reduction, which implies a loss of information at the same time. However, loss of information might be beneficial for the network, because it leads to a lower calculation overhead for the successor layers of the network. Usually the max-pooling function is used in this layer. This function finds the maximum value between a sample window, saving only this value for the next layer [12]. In the Figure 3, the operation performed by the max-pooling function can be observed graphically.

**Figure 3:** Example of 2x2 max-pooling [12].

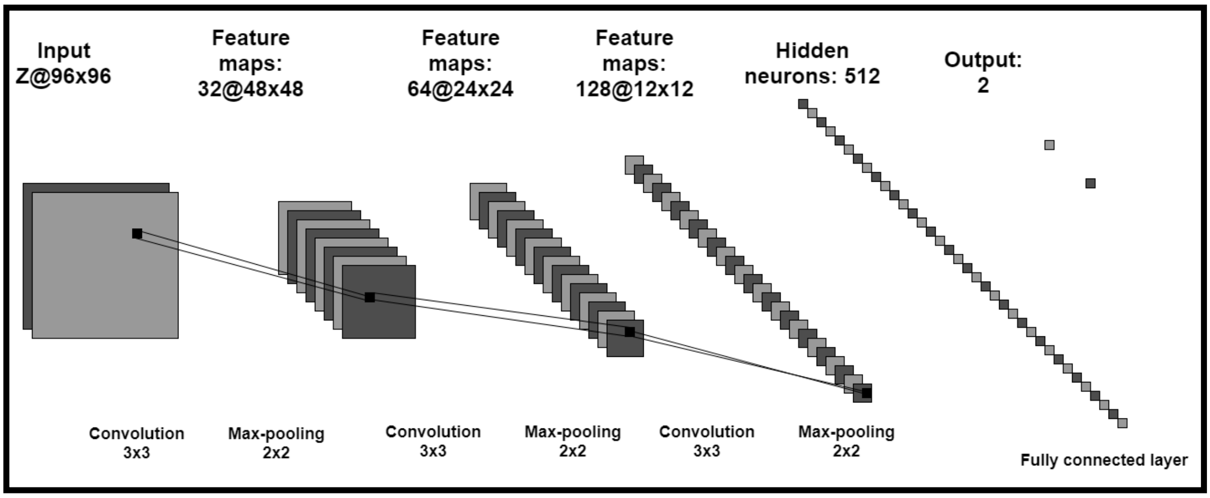



- **Classification layer**: After the convolution and pooling layers, a fully connected layer of neurons is used, in which each pixel of the image is considered as a separate neuron. Finally, the fully connected layer performed the desired classification.

## 3. PROPOSED MODEL

The proposed architecture consists of an input of one or two 96x96 images (depending on how many EEG sensors are used), followed by a 3x3 convolution with 32 feature maps, which gives way to a layer of max-pooling with a size of 2x2 and 32 feature maps. After this first stage, another 3x3 convolution layer is opened, but with 64 feature maps, followed by a 2x2 max-pooling layer with 64 feature maps. Then there is a last 3x3 convolution layer with 128 feature maps which is followed by a last 2x2 max-pooling layer of 128 feature maps. Finally, there is a single fully connected layer consisting of 512 hidden neurons, which is followed by the output of the network, which gives the classification of subjects in a state of drowsiness or base state (without drowsiness). In the Figure 4 a graphic representation of the architecture used is shown.

**Figure 4:** Used CNN architecture.




The summary of the architecture used is shown below, where C corresponds to the convolutional layers, P corresponds to the pooling layers and FC corresponds to the fully connected layer. The preceding number to each of the letters corresponds to the number of filters applied in each layer, while the numbers in brackets correspond to the size of the filters or the number of neurons in each layer, as appropriate. Moreover, Z indicates whether the input layer has a value of one or two depending on the data set used (one or two EEG sensors).

**Input[96x96xZ]-32C[3x3]-32P[2x2]-64C[3x3]-64P[2x2]-128C[3x3]-128P[2x2]-FC[512]-Output[2]**

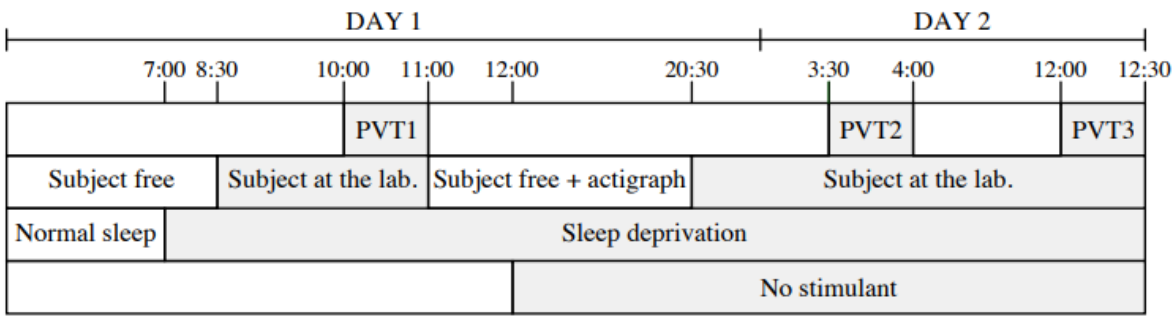
The main difference between the architecture in Hung [7] and the one here proposed resides in the format of the used images. While Hung uses the devices that have the information of the signals of thirty-two channels, the proposed model uses the spectrograms of the channels chosen independently (one per image), thus the proposed architecture has two input channels for the case in which both channels are used simultaneously.

## 4. APPLICATION

### 4.1 Experiment

In order to illustrate the applicability of the proposed model, "ULg Multimodality Drowsiness Database", also called DROZY was used [13]. This database was obtained by the Laboratory of Exploitation of Signals and Images (INTELSIG) from 14 healthy young subjects (3 men and 11 women), where each of them performed three psychomotor vigilance tests (PVT) of 10 minutes duration under conditions of sleep deprivation induced by a prolonged waking state, which was preceded by a normal night's sleep. The schedule used to carry out the data collection is shown in the Figure 5.

**Figure 5:** Summary of data collection schedule [13].




For each subject and psychomotor vigilance test the DROZY database contains: Karolinska sleepiness scale scores (KSS), stimulus and reaction times, polysomnography signals (including five EEG channels, two electrooculography channels (EOG), electrocardiogram (ECG) and electromyography (EMG)) taken in full at 512 Hz, Kinect v2 sensor images, Kinect v2 sensor videos and face signals. All the database described is perfectly synchronized over time [13].

### 4.2 Dataset

For the development of the proposed classifier, only the EEG signals and the ratings of the Karolinska sleepiness scale were used. Of the five EEG channels available in the database, only the *Fz* and *Pz* channels were used to train the model, due to the relationship between mental fatigue and the amplitude increase of theta waves in the frontal lobe of the brain and the amplitude increase of alpha waves in the parietal lobe of the brain under this condition [2].

To generate the classes of the database, the Karolinska Sleepiness Scale scores indicated by each subject before each PVT were used. In this way and following the description of each index of the scale, the subjects who presented a KSS index less than or equal to 4 were defined as a base state (without drowsiness), and as a drowsy state the subjects who presented a KSS index greater than or equal to 7 [11].

The base state was defined with all subjects of PVT 1 who presented a KSS less than or equal to 4 and the state of drowsiness with all subjects of PVT 3 who presented a KSS greater than or equal to 7. It should be noted that in no case exist subjects with a KSS index corresponding to the state of drowsiness in the PVT 1 or vice versa. The data corresponding to the PVT 2 were not considered because the average of the KSS scores of PVT 2 (5.6 ± 1.7) did not correspond to any of the previously established classes.

Then we proceeded to segment the signals of the channels *Fz* and *Pz*, with the aim of increasing the number of available data. First, each channel was segmented into files of thirteen seconds, since that is the minimum duration that an EEG signal must have to contain all the information referring to the individual's mental fatigue [2]. After making this segmentation, a total of 460 files were obtained for each class, constituting a total of 920 files for the data bank of thirteen seconds.

In parallel, the signals were segmented into files of five seconds, in order to evaluate the relevance of the signal length with the number of files available for classification. With this signal length, a total of 1200 files were obtained for each class, constituting a total of 2,400 files for the data bank of the five-second signals.

Subsequently, we proceeded to generate the spectrograms of each of the files (segmented signals) for each of the data banks previously generated. The spectrograms were obtained using a window of 512 points, an overlay 50% and only frequencies from 0 to 20 Hz were considered, as the brain waves relevant to the problem at hand are in this range [2]. In addition, each image was obtained in 96x96 pixels dimension.

Then each of the generated spectrograms was converted from RGB to gray scale format, so the representative matrix of each image has a dimension of two. In Figure 6, a spectrogram after being converted to gray scale is shown.

**Figure 6:** Example of EEG spectrogram in grayscale.

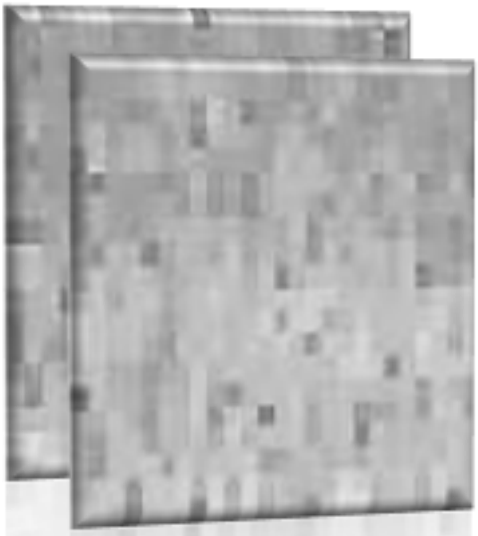
Finally, after pre-processing the entire database and generating the corresponding labels, we obtained six sets of spectrograms (all in gray scale) to be used to achieve the detection of subjects under drowsiness, which are described next:

1. **Spectrograms channel *Fz* / 13 seconds duration:** set of spectrograms of the channel *Fz* with duration of thirteen seconds and a range of frequencies 0-20 Hz, each of them with a dimension of 96x96x1 pixels.
2. **Spectrograms channel *Pz* / 13 seconds duration:** set of spectrograms of the channel *Pz* with duration of thirteen seconds and a range of frequencies 0-20 Hz, each of them with a dimension of 96x96x1 pixels.
3. **Spectrograms channels *Fz-Pz* / 13 seconds duration:** set of spectrograms of the channels *Fz* and *Pz* with duration of thirteen seconds and a range of frequencies 0-20 Hz. For this dataset, we have with images in three dimensions, where there are two levels of depth, each of which corresponded to the spectrograms of the channels *Fz* and *Pz* respectively, having a dimension of 96x96x2 pixels.

4. **Spectrograms channel *Fz* / 5 seconds duration:** set of spectrograms of the channel *Fz* with duration of five seconds and a range of frequencies 0-20 Hz, each of them with a dimension of 96x96x1 pixels.
5. **Spectrograms channel *Pz* / 5 seconds duration:** set of spectrograms of the channel *Pz* with duration of five seconds and a range of frequencies 0-20 Hz, each of them with a dimension of 96x96x1 pixels.
6. **Spectrograms channels *Fz-Pz* / 5 seconds duration:** set of spectrograms of the channels *Fz* and *Pz* with duration of five seconds and a range of frequencies 0-20 Hz. For this dataset, we have with images in three dimensions, where there are two levels of depth, each of which corresponded to the spectrograms of the channels *Fz* and *Pz* respectively, having a dimension of 96x96x2 pixels.

Figure 7 shows an example of a 3D image where the spectrogram of the front corresponds to the channel *Fz* and the one of back to the *Pz* channel.

**Figure 7:** Example of 3D image used, containing the spectrograms of the channels *Fz* and *Pz*.



### 4.3 Result and Discussion

All the results presented in this chapter were obtained through cross validation. **Table 2** shows the average values obtained for each performance metric in the detection of subjects under drowsiness for each of the six datasets.

**Table 2:** Summary of performance metrics for the different datasets using the proposed CNN model.

| Signal duration | Channel | Accuracy | F1 Score | Sensitivity |
|---|---|---|---|---|
| Thirteen seconds | *Fz* | 79.66% ± 1.12% | 79.48% ± 2.90% | 78.38% ± 5.38% |
| | *Pz* | 64.78% ± 3.67% | 66.22% ± 3.33% | 69.08% ± 4.42% |
| | *Fz-Pz* | 86.74% ± 1.81% | 87.09% ± 1.99% | 88.42% ± 5.06% |
| Five seconds | *Fz* | 73.37% ± 0.85% | 73.12% ± 1.64% | 71.85% ± 6.16% |
| | *Pz* | 63.81% ± 3.13% | 73.12% ± 4.25% | 64.42% ± 5.81% |
| | *Fz-Pz* | 77.98% ± 0.27% | 76.47% ± 2.01% | 75.32% ± 5.88% |

| Signal duration | Channel | Specificity | Precision |
|---|---|---|---|
| Thirteen seconds | ***Fz*** | 80.67% ± 4.67% | 79.11% ± 3.06% |
| | ***Pz*** | 60.69% ± 7.64% | 63.95% ± 5.92% |
| | ***Fz-Pz*** | 84.88% ± 3.19% | 85.96% ± 1.41% |
| Five seconds | ***Fz*** | 74.77% ± 7.34% | 74.95% ± 3.35% |
| | ***Pz*** | 62.98% ± 5.54% | 62.24% ± 3.59% |
| | ***Fz-Pz*** | 80.19% ± 5.16% | 78.09% ± 2.96% |

Table 3 presents the mean values obtained for each performance metric in the detection of subjects under somnolence state for the proposed model, Neural Network (NN), Support Vector Machines (SVM) and Random Forest (RF) using the thirteen-second spectrogram dataset corresponding to the *Fz-Pz* channels, dataset with which the highest performance was achieved.

**Table 3:** Comparative summary of results for each of the models for drowsiness detection.

| Model | Accuracy | F1 Score | Sensitivity | Specificity | Precision |
|---|---|---|---|---|---|
| **CNN** | 86.74% ± 1.81% | 87.08% ± 1.99% | 88.42% ± 5.06% | 84.88% ± 3.19% | 85.96% ± 1.41% |
| **NN** | 74.45% ± 2.55% | 75.77% ± 3.02% | 79.35% ± 6.23% | 69.49% ± 5.67% | 72.78% ± 3.44% |
| **SVM** | 73.03% ± 2.06% | 72.18% ± 1.50% | 73.10% ± 2.68% | 72.79% ± 4.31% | 71.41% ± 71.42% |
| **RF** | 76.19% ± 2.56% | 77.28% ± 3.76% | 80.56% ± 3.76% | 72.14% ± 4.83% | 74.55% ± 5.19% |

Analyzing the results presented in Table 2, it can be seen that for both thirteen and five-second signals, the joint use of the *Fz-Pz* channels gives the best performance for all the metrics evaluated. Also, channel *Fz* provides better performance than the *Pz* channel. From this, it can be argued that the presence of drowsiness in an individual causes more relevant changes in theta waves than in the alpha waves of the brain due to the relationship between the alpha and theta waves with the *Fz* and *Pz* channels, respectively. However, the changes existing in the presence of drowsiness in the *Pz* channel also provide information about the mental state of the individual, which is evident when evaluating the results obtained when using both channels.

Moreover, it was found that it is necessary that the signals to be used have a minimum duration of thirteen seconds because signals with a duration shorter than this do not have all the information regarding the level of fatigue of an individual [2]. This factor is evident not only in the metrics shown in Table 2, where it is observed that for thirteen second signals, better results are obtained per channel compared to five-second signals, but also when evaluating the number of available spectrograms images for each dataset. Datasets with thirteen-second signals have a total of 920 images, while datasets with five second signals have a total of 2,400 images. Having more than double the data, models trained on the five-second signals could have a greater generalization capacity, because the

CNNs usually present better performance with larger databases [5]. However, this is clearly not the case here.

When analyzing the results presented in Table 3, we can see that the model that delivers the best performance in all the metrics corresponds to proposed CNN based model, followed by RF, NN and SVM. With these results, we can verify that the reduction of dimensionality made by the pooling layers of CNN is essential to be able to generate a classification based on images.

In order to quantify the effect of the pooling and convolution layers, we can compare the performance of the CNN with that of the NN, given that the latter has the same architecture of the fully connected layers in the CNN model. In this way, it can be observed that the proposed CNN model obtains results 11% higher than those of the NN on average, a difference that is explained by the presence of the convolutional and pooling layers.

## 5. CONCLUSIONS

Detecting when an individual has a drowsy state, which can put at risk their own safety and of the others, is a need not satisfactorily met at present. The present work presents a deep CNN based model that allows detecting the presence of drowsiness in an operator by using non-invasive electroencephalography technique, and the use of the most important signal patterns, delivering up to 86% accuracy.

The obtained results indicate that the proposed Convolutional Neural Network model represents the most suitable drowsiness detection approach using electroencephalography spectrograms in comparison to Neural Networks, Support Vector Machines and Random Forest.

## REFERENCES


[1] A. D. Swain y H. E. Guttmann, Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications, 1983.

[2] L. J Trejo, K. Kubitz y L. D. Montgomery, «EEG-Based Estimation and Classification of Mental Fatigue,» *Psychology,* 2015.

[3] K. Kaida, M. Takahashi, T. Äkerstedt, A. Nakata, Y. Otsuka, T. Haratani y K. Fukasawa, «Validation of the Karolinska sleepiness scale against performance and EEG variables,» *Clinical Neurophysiology,* vol. 117, 2006.

[4] Mervyn V.M. Yeo, Xiaoping Li, Kaiquan Shen y Einar P.V. Wilder-Smith, «Can SVM be used for automatic EEG detection of drowsiness during car driving?,» *Safety Science,* 2009.

[5] Y. LeCun, Y. Bengio y G. Hinton, «Deep Learning,» *Nature,* 2015.

[6] D. Verstraete, A. Ferrada, E. López Droguett, V. Meruane y M. Modarres, «Deep Learning Enabled Fault Diagnosis Using Time-Frequency Image Analysis of Rolling Element Bearings,» *Shock and Vibration,* 2017.

[7] Y.-C. Hung, Y.-K. Wang, M. Prasad y C.-T. Lin, «Brain Dynamic States Analysis based on 3D Convolutional Neural Network,» de *International Conference on Systems, Man, and Cybernetics (SMC)*, Banff, 2017.

[8] T. Marek, C. Noworol y W. Karwowski, «Mental fatigue at work and pain perception,» *Work and Stress,* vol. 2, pp. 133-137, 1988.

[9] A. C. d. Seguridad, «Boletín Técnico de Ergonomía, Antecedentes sobre fatiga en la conducción de vehiculos,» Santiago, 2007.

[10] T. C. Technologies, Cortical Functions, Hong Kong, 2012.

[11] T. Äkerstedt y M. Gillberg, «Subjective and Objective Sleepiness in the Active Individual,»

*International Journal of Neuroscience,* vol. 52, pp. 29-37, 1990.

[12] I. Goodfellow, Y. Bengio y A. Courville, Deep Learning, 2016.

[13] Q. Massoz, T. Langohr, . C. Francois y J. G. Verly, «The ULg Multimodality Drowsiness Database (called DROZY) and Examples of Use,» 2016.