Automatic Drowsiness Detection based on a Single Channel of EEG Signals using a Hybrid Analysis and Decision Tree Classification Method under Python

Elmouzoun Elidrissi Mouad[1][a], Ben Taleb Lhoucine[1], Lyazidi Aissam[1,2], Rattal Mourad[1,2], Mouhsen Azeddine[1], Harmouchi Mohammed[1] and Essoukaki Elmaati[1,2]

[1]*Laboratory of Radiation-Matter and Instrumentation, Faculty of Science and Technology, Hassan First University of Settat, Morocco*

[2]*Higher Institute of Health Sciences, Hassan First University of Settat, Morocco*

Keywords: Driver drowsiness, EEG signals, road accident, hybrid system, automatic detection, real-time, machine learning.

Abstract: The human brain generates millions of signals because they translate all our movements and thoughts, our physical and psychological state. During driving, all these signals are generated simultaneously. Vigilance while being behind the wheel is necessary. However, when the roads are monotonous, especially when taking highways, that alertness state decreases, and the drowsiness state shows off. In Morocco, 1/3 of fatal accidents on the highway are caused by drowsiness and sleeping while driving. Wherefore, we proposed the idea of developing an automatic system based on electroencephalogram signals (EEG) that can predict the state of drowsiness in real-time using several features extracted from the EEG records when this state occurs to drivers while driving. The proposed work is based on time-frequency analysis of EEG signals based on one single-channel (FP1-Ref), and drowsiness is predicted using machine learning techniques under Python. The results are significant where we could reach an accuracy of prediction with an average of 95.7% within 0.062 seconds using the decision tree classification method.

1 INTRODUCTION

In Morocco, 33.3% is the rate of fatal/deadly accidents on highways caused by drowsiness and sleep while driving based on the latest statistics of the Ministry of Equipment, Transport, Logistics and Water, Directorate of Roads (Des et al., 2016) (Statistiques, 2017). These statistics gave us the idea of developing an automatic system that can predict drowsiness when occurring and before the situation becomes worst leading to dangerous accidents.

Therefore, the idea of our system is not new, but it came to improve the performance and solve the limitations of the existing ones by using the latest processing software 'Python', also by providing the best processing techniques 'time and frequency' and machine learning algorithms to perform the most performant hybrid and automatic method of detecting drowsiness.

Though, our work aims to develop a method that analyses the EEG signals according to time and frequency domains (Fourier and Spectrum analysis) based on single channel records before being fed to our machine learning classifier chosen (DT: Decision Tree classifier) for the training and testing process.

As a result, our model shows high performance compared to the existing models, especially the works that used the same database (Physionet) with an accuracy of 95.7% and a time consuming of 0.062 seconds.

2 RELATED WORKS

To make our system more efficient, and based on a heavy analysis, a detailed study was carried on the existing systems to precise and overcome their limitations.

[a] https://orcid.org/0000-0002-3784-8967

Elmouzoun Elidrissi, M., Ben Taleb, L., Aissam, L., Rattal, M., Mouhsen, A., Harmouchi, M. and Elmaati, E.
Automatic Drowsiness Detection based on a Single Channel of EEG Signals using a Hybrid Analysis and Decision Tree Classification Method under Python.
DOI: 10.5220/0010732900003101
In *Proceedings of the 2nd International Conference on Big Data, Modelling and Machine Learning (BML 2021)*, pages 301-305
ISBN: 978-989-758-559-3

Detecting drowsiness for drivers was the primary goal of many researchers. The processing techniques were wildly variant, where some were based on sensors only, based on physiological signals like electroencephalogram, electrocardiogram, electro-oculogram (EEG, ECG, EOG) or even a mix of these two techniques. Chang et al, proposed at the year of 2018 a system based on smart glasses that detect drowsiness based on signals generated by accelerometers and gyroscopes by capturing the micro-falls of the head in addition to an infra-red transceiver that captures the frequency of the blinking and the degree of closure of the eyes (Chang et al., 2018).
Other works used algorithms of detecting drowsiness based on face or eyes recognition like (Ouabida et al., 2020) & (Dhanalakshmi et al., 2016) or even using thermal imaging (Kiashari et al., 2020).

However, when we use just signals and information issued from sensors and no physiological signals, it is not evident to confirm that the detection is efficient because blinking or eyes closure or even the head's movement are some standard and random actions that the driver may do while driving, so the goal is not reached. So, the solution is to use some techniques based on signals recorded from EEG, ECG, EOG and others.

To situate our work, the following works used a single channel study in addition to using the same dataset of EEG signals available at the Physionet database (eegmat 1.0.0 and sleep-edfx) to compare our results and show the improvement added by our hybrid method. (Belakhdar et al., 2018) proposed a method of analysing the spectral domain of the EEG signals under MATLAB by applying Fourier transform and a classification by artificial neural network. With this method, their work reached an accuracy of 88.8%. (Bajaj et al., 2020) reached an accuracy of 91% after extracting the features using tunable Q-factor wavelet transform (TQWT) from the EEG signals, and classifying them using the extreme machine learning classifier (ELM). The highest accuracy of 94.45% was reached by (B et al., 2021) after using a processing method based on wavelet packet transform (WPT) fed to the extra-tree classifier.

The proposed work aims to improve the algorithm's efficiency of detecting drowsiness of drivers based on a single channel of EEG signals in the term of rapidity and accuracy. We will explain the method with more details in the next section.

3 PROPOSED WORK

Our method proposed in this paper aims to provide a new hybrid algorithm of detecting drowsiness of drivers based on temporal and frequential domains by processing a single channel of EEG records (FP1). Many researchers confirmed that the FP1 position is one of the most accurate positions to be based on to study drowsiness (Strijkstra et al., 2003). The proposed method is shown in the figure below:

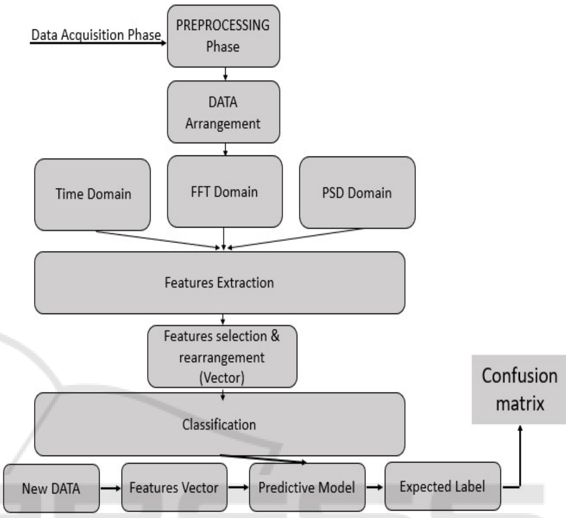

Figure 1: Flowchart of the proposed method.

3.1 Pre-processing Phase

The databases used (eegmat 1.0.0 and sleep-edfx) are two open databases provided by the Physionet to use them for researchers works.

All the EEG records were artefact-free. All the noise was filtered right after the acquisition phase of these signals due to a 30 Hz low-pass filter and a 50 Hz notch filter to eliminate the noise of the alimentation source.

The selected subjects were males and females with a different marge of ages (an average of 18 years old), and the recording was done under the 10-20 international system.

3.2 Time Segmentation Phase

A segmentation of 3 seconds was applied on the EEG records in the current work to ensure stationarity of spectral analysis (Fourier and Power Spectral Density analysis).

3.3 Features Extraction Phase

This step aims to extract the most significant features from the single channel of EEG recording following three domains (temporal, Fourier and spectral).
We designed a function that can extract all the features one by one, and it scales all of them in the right shape for the classification procedure.

3.3.1 Temporal Domain Analysis

In the time domain, eight parameters were calculated to differentiate between the alert/awake and drowsy states. After calculating the minima, the maxima, the amplitude peaks and the mean of amplitude peaks, these parameters are also processed.

The median:

$$P(y \leq x) = P(z \geq x) \quad (1)$$

The mean:

$$\bar{X} = \frac{\Sigma xi}{n} \quad (2)$$

The variance:

$$\text{Var} = \frac{\Sigma(xi - \bar{x})^2}{n} \quad (3)$$

The standard deviation:

$$\text{Std} = \sqrt{\frac{\Sigma(xi - \bar{x})^2}{n}} \quad (4)$$

The root-mean-square:

$$\text{RMS} = \sqrt{\frac{\Sigma xi^2}{n}} \quad (5)$$

3.3.2 Fourier Domain Analysis

In the present phase, we proposed a frequency analysis of the EEG signals recorded using Fast Fourier Transform. After extracting the same feature as the time domain, we calculated the modulus of these features to get rid of the imaginary part and to have just a real significant part.

For $0 \leq k \leq N-1$:

$$X_k = \sum_{n=0}^{n-1} x_n e^{-2\Pi i} \frac{kn}{N} \quad (6)$$

3.3.3 Power Spectral Density Analysis

The spectrum of the signal is calculated using the burg algorithm to compare the power of each band of the brain frequencies to have a good differentiation between the two states of drivers (awake and drowsy).

3.4 Features Selection & Classification

A total of eight machine learning classification methods were tested in our study to compare the efficiency and get the best classifier, and secondly, to select the most appropriate features. As a result, our hybrid model showed the best accuracies and time performances.

The classifiers used in our method are Support Vector Machine (with its four kernels), Gaussian Process (GP), K-Nearest-Neighbors (KNN), Multilayer Perceptron (MLP), and Decision Tree (DT).

3.5 Results

After extracting the features, all the calculated parameters were scaled and processed using machine learning classifiers. These classifiers depend on four parameters which are:

i. True-positive (TP): The subject is drowsy, and the prediction is positive (Drowsy state predicted)

ii. True-negative (TN): The subject is awake, and the prediction is negative (Awake state predicted)

iii. False-positive (FP): The subject is awake, and the prediction is positive (Drowsy state predicted).

iv. False-negative (FN): The subject is drowsy, and the prediction is negative (Awake state predicted).

Based on these parameters, the classifier's accuracy is calculated according to:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (8)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (9)$$

$$\text{F-1 score} = 2 * \frac{precision * recall}{precision + recall} \quad (10)$$

Table 1: Performance comparison between different classifiers applied on our selected features

Classifier	First method results	Second method results	Third method results	Hybrid method results
SVM (Linear kernel)	49.7%	49.9%	49.4%	49.5%
SVM (Polynomial kernel)	54.6%	85.5%	93.2%	83.6%
SVM (Sigmoid kernel)	35.3%	66.8%	88.7%	66.0%
SVM (RBF kernel)	71.9%	86.5%	93.3%	87.8%
MLP	49.8%	74.1%	48.9%	75.6%
KNN	90.6%	92.9%	94.1%	93.1%
GP	49.1%	68.9%	49%	56%
DT	49.3%	93.6%	94.3%	**95.7%**

As we can see, our hybrid model based on the time, Fourier and PSD domains achieved the best accuracy compared to all the other selection of features and classifiers, and here is a comparison in terms of the executing time, in addition to the confusion matrix result of the proposed machine learning model including the processing time.

Classifier	Accuracy	Time (s)
SVM (Linear kernel)	49.5%	0.983
SVM (Polynomial kernel)	83.6%	0.722
SVM (Sigmoid kernel)	66.0%	1.170
SVM (RBF kernel)	87.8%	0.985
MLP	75.6%	5.144
KNN	93.1%	0.142
GP	56%	12.57
DT	95.7%	0.062

Our proposed method is the best in terms of both accuracy of detecting the drowsiness state and the processing time.

```
--- Execution time is : 0.06298518180847168 seconds ---
Train Accuracy : 100.0%
Test Accuracy : 95.7%
Test precision : 95.6%
Classifier's Accuracy : 95.7%
Recall : 95.7%
              precision    recall  f1-score   support

           0       0.96      0.96      0.96      1438
           1       0.96      0.96      0.96      1427

    accuracy                           0.96      2865
   macro avg       0.96      0.96      0.96      2865
weighted avg       0.96      0.96      0.96      2865
```

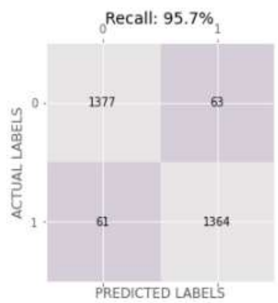

Figure 2: Confusion matrix of the proposed method.

4 CONCLUSIONS

The present work proposed a new hybrid method of detecting drivers' drowsiness based on time-frequency analysis of EEG signals from a single channel (FP1).

We extracted a total of eight features from the three domains, the time, Fourier and PSD. After that, we trained eight machine learning models, support vector machine (with its four kernels), Gaussian process (GP), K-Nearest-Neighbors (KNN), Multi-layer Perceptron (MLP), and Decision Tree (DT).

We compared our work to previous ones based on the same dataset available at the Physionet database, and the use of a single channel of EEG records. The added value of our model is the improvement of the detection's performance in the term of accuracy, which achieved 95.7% and the processing time 0.062 seconds.

REFERENCES

B, V. P., & Chinara, S. (2021). Automatic classification methods for detecting drowsiness using wavelet packet transform extracted time-domain features from single-channel EEG signal. *Journal of Neuroscience Methods*, *347*, 108927. doi: 10.1016/j.jneumeth.2020.108927

Bajaj, V., Taran, S., Khare, S. K., & Sengur, A. (2020). Feature extraction method for classification of alertness and drowsiness states EEG signals. *Applied Acoustics*, *163*, 107224. doi: 10.1016/j.apacoust.2020.107224

Belakhdar, I., Kaaniche, W., Djemal, R., & Ouni, B. (2018). Single-channel-based automatic drowsiness detection architecture with a reduced number of EEG features. *Microprocessors and Microsystems*, *58*, 13–23. doi: 10.1016/j.micpro.2018.02.004

Budak, U., Bajaj, V., Akbulut, Y., Atilla, O., & Sengur, A. (2019). *An Effective Hybrid Model for EEG-Based*. *19*(17), 7624–7631.

Chang, W., Chen, L., Member, S., & Chiou, Y. (2018). Design and Implementation of a Drowsiness -Fatigue-Detection System Based on Wearable Smart Glasses to Increase Road Safety. *IEEE Transactions on Consumer Electronics*, *PP*(c), 1. doi: 10.1109/TCE.2018.2872162

Des, R., & Des, S. (2016). *Recueil des statistiques des accidents corporels de la circulation routière 2016*.

Dhanalakshmi, S., Rosepet, J. J., Rosy, G. L., & Philominal, M. (2016). *Drowsy Driver Identification Using MATLAB*. *4*(Iv), 198–205.

Kiashari, S. E. H., Nahvi, A., Bakhoda, H., Homayounfard, A., & Tashakori, M. (2020). Evaluation of driver drowsiness using respiration analysis by thermal imaging on a driving simulator. *Multimedia Tools and Applications*. doi: 10.1007/s11042-020-08696-x

Ouabida, E., Essadike, A., & Bouzid, A. (2020). Optik Optical correlator based algorithm for driver drowsiness detection. *Optik - International Journal for Light and Electron Optics*, *204*(December 2019), 164102. doi: 10.1016/j.ijleo.2019.164102

Statistiques, R. D. E. S. (2017). *Recueil des statistiques des accidents corporels de la circulation routière 2017*. 1–129.

Strijkstra, A. M., Beersma, D. G. M., Drayer, B., Halbesma, N., & Daan, S. (2003). *Subjective sleepiness correlates negatively with global alpha (8 – 12 Hz) and positively with central frontal theta (4 – 8 Hz) frequencies in the human resting awake electroencephalogram*. *340*, 17–20. doi: 10.1016/S0304-3940(03)00033-8