Mobile EEG Based Drowsiness Detection using K-Nearest Neighbor

Conference Paper · October 2019

DOI: 10.1109/ICAwST.2019.8923161

4 authors, including:

Prima Dewi Purnamasari
University of Indonesia

Anak Agung Putri Ratna
University of Indonesia

Dodi Sudiana
University of Indonesia

Mobile EEG Based Drowsiness Detection using K-Nearest Neighbor

Prima Dewi Purnamasari
Department of Electrical Engineering
Faculty of Engineering, Universitas Indonesia
Depok, Indonesia
prima.dp@ui.ac.id

Pratiwi Yustiana
Department of Electrical Engineering
Faculty of Engineering, Universitas Indonesia
Depok, Indonesia
pratiwi.yustiana@ui.ac.id

Anak Agung Putri Ratna
Department of Electrical Engineering
Faculty of Engineering, Universitas Indonesia
Depok, Indonesia
ratna@eng.ui.ac.id

Dodi Sudiana
Department of Electrical Engineering
Faculty of Engineering, Universitas Indonesia
Depok, Indonesia
dodi.sudiana@ui.ac.id

Abstract— **In this research, a drowsiness detection system, named Drowsiver, was developed for a mobile electroencephalograph (EEG) and a mobile phone. The system is expected to minimize the causes of accidents caused by drowsy drivers. By using Electroencephalogram (EEG), the condition of drowsiness is detected by recording the electrical activity that occurs in the human brain and is represented as a frequency signal. The signal is transmitted to the Android mobile application via Bluetooth and will give an alarm notification if the drowsiness is detected. The brainwave from the mobile EEG is processed using Fast Fourier Transform (FFT) to extract its features. These features are classified using K-Nearest Neighbor (KNN) classifier. The system produces the best performance with the highest accuracy of 95.24% using the value of *k*=3 and four brain waves as features, namely Delta, Theta, Alpha, and Beta waves.**

Keywords— Drowsiness, Electroencephalogram (EEG), Brain Wave, Android Application

I. Introduction

Usually, humans need 7 hours to sleep. However, they can have fewer sleeping time because of their activity or work, which take a lot of time that cause sleepless or make them drowsy. This condition marked when eyes feel sore, slowly blinking, yawns or the body feels trembling that affect people activities; one of them is driving. Some people who drive even often ignore their physical condition. When the physical condition of the driver is not feeling well or even just drowsy, it can endanger themselves, passengers, or other drivers.

Based on data from the Indonesian National Police, the number of traffics accidents up to June 2018 reached 26,729 cases which caused a large number of victim and substantial losses. Of the total accidents throughout Indonesia, there were 7,505 victims died, 3,203 victims were seriously injured, 32,294 people had minor injuries and caused a loss of Rp. 54,883,131,888 or about 3.9 Million US Dollars. One of the causes of the accidents was that the driver felt tired and drowsy [1].

Drowsiness can cause microsleep or someone experiencing unconsciousness for a few seconds and cannot respond to the situation. To find out whether someone is experiencing microsleep can be seen from several aspects, one of them through brain waves. The human brain has a continuous electrical activity that represent frequency signal known as Electroencephalogram (EEG). Brain waves are divided into several frequency regions, namely: Delta (0.5-4 Hz), Theta (4-8 Hz), Alpha (8-13 Hz), Beta (13-30 Hz), and Gamma (30-45 Hz) [2].

In this research, a drowsiness detection system will be created to prevent driver negligence that can cause accidents. So, when the driver has given signs that she/he is drowsy, the system can warn the driver to be more careful. The system used EEG device, Neurosky Mindwave Mobile 2 to detect human brain waves and analyze whether the waves show the drowsiness or not.

There are several studies to detect drowsiness using NeuroSky MindWave Mobile. Song et al. has made a drowsiness detection application using five brain waves. The signal is classified using Support Vector Machine (SVM) [3]. He et al. proposed a fatigue detection application using attention and meditation signals from NeuroSky features, and the signal is classified using KNN with values of *k*=3 [4]. In this paper, we proposed a mobile system to detect (and warn) driver drowsiness using a fast classifier, the KNN. The system compares several features and shows which combination of features that provide the best recognition rate or accuracy.

This paper is structured as follow. Section II explains theories that support this research include the EEG, human brainwaves, FFT, and KNN. The methodology implemented in this research and the design of both hardware and software are described in Section III. The implementation and simulation results are presented in Section IV. Section V concludes the paper.

II. Literature Review

A. Electroencephalogram (EEG)

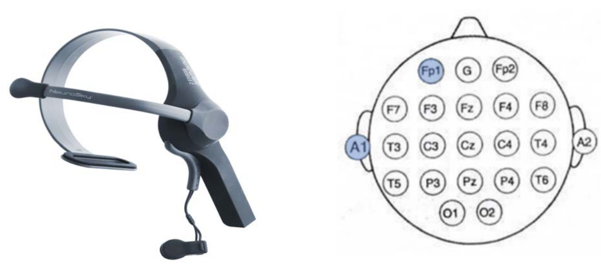
Electroencephalogram (EEG) is a device used to record electrical activity produced by the human brain using several electrodes placed on the scalp. While Electroencephalography is the process of recording from the results of recording using the EEG device. The EEG signal is a representation of the electric current that flows in the human brain. Electrical activity that occurs in the brain produces different levels for each individual [5].

Conventional EEG devices that are usually used in clinics or hospitals record the electrical activity of the human brain by placing electrodes on the scalp with a conductive gel or paste. Most EEG recording systems use electrodes, each of which is connected by a cable. Besides, there are also EEG systems that use nets or caps with electrodes in them.

The placement of electrodes with their names is listed on the International 10-20 system, which ensures consistent electrode naming throughout the laboratory [5]. The system uses 21 electrodes placed at 10% and 20% of the head circumference. The electrode is named according to its position. The first part refers to the region, and the second part refers to the area. For the first part, each electrode is named as follows: F = frontal, C = central, P = pariental, T = temporal, O = occipital, A = auricular (ear), and Fp = fronto-polar.

B. NeuroSky MindWave Mobile 2

NeuroSky MindWave Mobile 2 is a commercial and consumer-grade brain-computer interface (BCI) device. To acquire human brain wave, it provides a 1-channel EEG in Fp1 position or forehead above left eyebrow and ear-clip as ground in position A1, by following International 10-20 System [5]. The illustration of this BCI device and the electrode's position are shown in Fig. 1.



(a) (b)

Fig. 1. Single channel EEG a) Neurosky Mindwive 2 BCI device and b) Fp1 electrode's position in 10-20 System

The MindWave Mobile 2 EEG device can be connected to several devices, such as computers with Windows or Mac operating systems, or Android or iOS-based smartphones using a Bluetooth connection. The MindWave Mobile 2 device requires 1 AAA size alkaline battery as power that can work for 8 hours. NeuroSky has several separate technologies to facilitate operation in its products. These technologies are ThinkGear and eSense. ThinkGear is a technology on NeuroSky products that allows this device to connect with other devices. This technology is used to represent brain waves received from the electrode into the connected device.

C. Brainwaves during drowsy conditions

The human brain has nerve cells that produce a very small amount of electricity (mV). These electrical activities can be recorded and displayed using an EEG device and translated into frequency waves in Hertz (Hz). Brain waves have a frequency range that can be grouped into Delta, Theta, Alpha, Beta, and Gamma waves respectively, according to their respective frequency regions. At one time, the brain will produce these waves at the same time, but with different levels. The brainwave, along with the activities related to the production of waves, is presented in Table I.

Sleep has several stages that can be distinguished from each other based on the pattern of brain waves that occur at each level [6]. This change in brain wave activity can be described using EEG and can be distinguished from the amplitude and frequency of each brain wave. Sleep can be divided into two phases: REM (Rapid Eye Movement) and Non-REM (NREM). The REM phase can be identified by eye movements when the eyes are closed. Brain waves during the REM phase are very similar to brain waves when the condition is awake (awake). The NREM phase is divided into four stages that can be distinguished through brain wave patterns. Stages 1-4 are the NREM phases, and the fifth stage or the final stage is the REM phase.

TABLE I. Human Brainwaves [2]

Brainwaves	Activity
Delta (0.1-4 Hz)	Sleep
Theta (4-8 Hz)	Drowsy, Meditation
Alpha (8-13 Hz)	Relaxation
Beta (13-30 Hz)	Focus, Concentration
Gamma (30-100 Hz)	Panic, Fear

The first stage of the NREM phase is the transition phase that occurs between wakefulness and sleep. During this stage, there is a slowdown in respiration and heart rate. In terms of brain waves, stage 1 sleep is associated with Alpha waves and Theta. This stage produces Alpha waves (8-13 Hz) with a high amplitude pattern. At this stage, one can say that the person is awake but still calm. When someone continues this step 1, it will produce a Theta wave with a lower frequency (4-8 Hz). This wave will generate a higher amplitude than the Alpha wave.

In the second stage, the body will enter the deep relaxation stage. Theta waves still dominate brain wave activity. In stages 3 and 4, it is usually called deep sleep which enters low frequencies (up to 4 Hz), high amplitude in Delta waves. The last stage is the REM phase, which is the period of sleep when someone dreams.

To determine whether a person is drowsy or cannot be seen from eye movements, the heart rate and breathing are slowing, and the brain waves are produced. A person who is sleepy can be seen from the change in amplitude, namely, the increase in the Theta wave and the decrease in the Alpha wave [7].

D. Fast Fourier Transform (FFT)

Fast Fourier Transform (FFT) is a signal processing algorithm that is a development of Discrete Fourier Transform (DFT) [8]. FFT uses algorithms that are better and work like DFT but with a shorter time. DFT itself is important in frequency analysis because DFT takes discrete time-domain signals and transforms these signals into discrete representations in the frequency domain. DFT only evaluates several frequency components from a segment being analyzed.

DFT is a Fourier transform for analysis of discrete-time signals with limited domains. Inputs from DFT are sequences of real or complex numbers, making them ideal for digital signal processing. The DFT formula is defined in (1).

$$X_k = \sum_{n=0}^{N-1} x_n e^{-2\pi ink/N} \tag{1}$$

for k=0, …, N-1.

Where i is an imaginary unit and is the root of n-th union, the DFT algorithm requires an arithmetic operation $O(n^2)$. The FFT algorithm can get the same and better results (faster) with O(n log n) operations. Optimization of this transformation equation converts the signal to a simpler one until it gets a transformation from two elements where n is 0 and 1. After being transformed it becomes simpler, FFT must group it at the top level and must be completed again and again to get the highest level. At the end of the process, the results of the transformation must then be rearranged [9].

E. K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) is a non-parametric algorithm and lazy algorithms are used as algorithms for data classification [10]. The purpose of KNN is to be used in a dataset where data points are separated into several classes to predict the classification of new sample points. What is meant by non-parametric is KNN does not assume on data distribution, but the structure of the model will be determined from the data.

KNN is also a lazy algorithm; this indicates that KNN does not use training data to generalize. In other words, the training process is very minimal and very fast. By not generalizing, KNN will store most of the training data to be used in the testing process. KNN is based on the similarity of features (feature similarity), which is evaluated from how closely a feature of the new sample data resembles training data to classify the given data points.

KNN classifies the projected training data in many dimensions. This space represents the criteria for each training data which are points k. The new data to be classified will be projected on many dimensions of space with learning data points. The classification process is done by finding the nearest k point from the new data or called the nearest neighbor. The nearest neighbor search technique or distance search generally uses the Euclidean distance formula as defined in (2).

$$D_E(a, b) = \sum_{i=1}^{N} \sqrt{(a_i - b_i)^2} \tag{2}$$

The stages of classification using the KNN algorithm are as follows:

1. Determine the parameter k (number of closest neighbors).
2. Calculate the distance using the Euclidean formula between new data on the training data provided.
3. Sort the results of the ascending distance calculation sequentially from the highest to the lowest value.
4. Collect the Y category or the nearest neighbor classification based on the value k.
5. The most frequent nearest neighbor categories are the results of new data categories.

In determining the k value, if the number of classes to be classified is even, then k value should be set with an odd number and vice versa. This is done to make it easier to obtain the classification results from the number of closest neighbors. If there is one class that has the most neighbors, then the test data will get the results of that class [10].

III. METHODOLOGY

The system requirement is the stage to set a minimum limit on the success of the system. This stage has a role in completing the design of the sleep detection system using human brain waves. In this sleep detection system, there is one user, the driver. The driver will use this system against himself. Some of the requirements for brainwave-based sleep detection systems include:

- The system can perform a series of processes to detect drowsiness in the driver who uses an EEG device while driving his vehicle.
- The sleep detection system will take brain wave signals using an EEG device that does not interfere with driver activity.
- Driver drowsiness will be analyzed using the application on the driver's smartphone to sound an alarm if sleepiness conditions have been detected.
- The system is real-time that produces output within a certain period so that it can prevent accidents directly.

A. Hardware Design

This sleep detection system device uses one EEG device and an Android-based smartphone. EEG devices will take the driver's brain wave data. Smartphones are used as a tool to process signals and produce an alarm notification output. The block diagram of the hardware system is shown in Fig. 2.

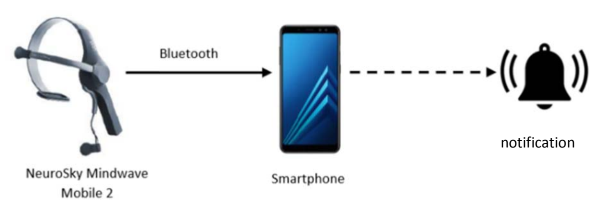


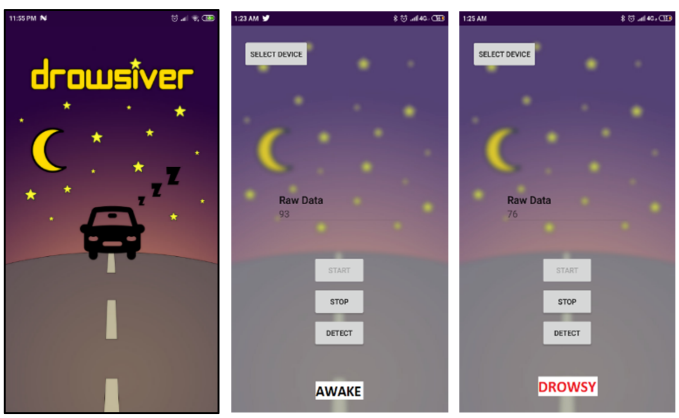
Fig. 2. The block diagram of the hardware system

The EEG device used is the NeuroSky Mindwave Mobile 2. EEG devices are used to retrieve data in the form of brain waves from the driver who drives the vehicle. Then the signal will be sent to the driver's smartphone via Bluetooth. The smartphone will receive data obtained from the EEG device and processed in the application to detect drowsiness.

B. Software Design

Fig. 3. Block diagram of the software part of the developed system

The block diagram of the sleep detection system software is shown in Fig. 3. First of all, the user EEG signals are acquired using NeurosSky Mindwave. Then, the EEG signals are pre-processed to remove the noise and downsampling. The next step is the essential step, which is the feature extraction stage. We use FFT to transform the time domain EEG signal into the frequency domain signal. By using Power Spectrum Density (PSD), the frequency signal that has been obtained will be taken the power of each frequency point. Furthermore, these frequency signals are categorized into five bandpass categories, namely: Delta waves (0.1 – 4 Hz), Theta (4 – 8 Hz), Alpha (8 – 13 Hz), Beta (13 – 30 Hz), and Gamma (30 – 100 Hz). It must be ensured that the amount of data to be calculated must be even before FFT processing. If it is of odd value, a value of 0 is added to the last line of data. Then the average value of each power frequency range will be taken according to the categories mentioned. After EEG data has been transformed into data in the frequency domain, the next step is to classify the data to find out the drowsy condition of the driver. Data classification uses the KNN algorithm.



(a) (b) (c)

Fig. 4. GUI of the Drowsiver Andoid application: a) splash screen, b) user awake, andc) user drowsy

The Graphical User Interface (GUI) of the developed system is shown in Fig. 4. The application was developed using Java Language on Android Studio. Once the application is started, and the EEG device is connected, the application will process the EEG data and classify whether the driver is awake or drowsy. The GUI of the status is presented in Fig. 3a and 3b, respectively. If the driver's drowsy condition is detected, the application will send a notification on the driver's smartphone, and a bell is sounded to warn the driver.

IV. Experiment and Analysis

EEG based drowsiness detection system aims to determine the relationship between someone's drowsy condition with brainwaves. Based on the theory mentioned above, the drowsy condition of a person will affect the EEG signal. The experiments were aimed to test the accuracy of the algorithm used in the drowsiness detection system, mainly the feature extraction method using FFT and the classification method using KNN.

To test the developed drowsiness detection system, we have collected data from 7 participants aged 17 – 30 years, when the participant drowsy, and also in an awake condition. The experiments were conducted during the bedtime, starting from the participants are still awake and continue until they are drowsy and at the end, sleeping. The EEG signals are collected, and for the experiment purpose, the data were then grouped into two categories, namely awake and drowsy. From the data collected, 50% of the data were used as the training data, and the rest 50% were used as testing data.

A. Feature Type

The first experiment is to test the best number and type of features used. As explained earlier, someone who is in an awake state or drowsy state has a difference in the brain signals produced. When a person is in the first phase of sleep (someone who is still conscious but still calm), he will produce Alpha waves with high amplitude. Whereas someone who continues the first phase sleep or someone in a drowsy state will produce a Theta wave with a higher amplitude than the Alpha wave.

In this experiment, five combinations of features are compared. As mentioned before, there are five frequency bands of brainwaves, Delta, Theta, Alpha, Beta, and Gamma. Based on theory, only 2 of these frequencies are closely connected to the drowsy state, Alpha and Theta. On the other hand, in machine learning theory, it is usually good to have more features if we do not know exactly which features are the best. Therefore, in this experiment, we tested several combinations of brainwave frequencies to see if there is a combination better then only Theta and Alpha, as stated in theory.

As the benchmark, the first features tested are Theta and Alpha, which according to theory, these two frequencies related to drowsiness. Next, we increase the number of features into 3. Here, Delta and Beta are compared, which one is better to accompany Alpha and Theta. In theory, Delta waves are produced when a person is asleep, while Beta waves are produced when someone is concentrating. Then, we also compared if both Delta and Beta added to the feature vector. And the last one, we use all the five frequency bands as the features.

The KNN classifier with k=3 was used in this experiment. The k value was chosen based on similar work [4], [11]. The result of the test is presented in Table II. In this result, it can be seen that four features (Delta, Theta, Alpha, Beta) produced the best accuracy, which is 95.24% recognition rate. It can be analyzed that when a person is drowsy, he will have less concentration (changes in Beta) and sleep frequency also arise (changes in Delta). The addition of Gamma frequency (related to panic or fear) will decrease the recognition as it does not have a relation to drowsy or sleeping condition.

TABLE II. TEST RESULTS BASED ON NUMBER OF FEATURES USED.

Number of Features	Accuracy (%)
2 (Theta, Alpha)	92.86
3 (Delta, Theta, Alpha)	90.48
3 (Theta, Alpha, Beta)	90.48
4 (Delta, Theta, Alpha, Beta)	**95.24**
5 (Delta, Theta, Alpha, Beta, Gamma)	90.48

B. K Values in k-Nearest Neighbor

The second experiment is testing the system performance based on various *k* value of the KNN classifier. The *k* value in question is the number of closest neighbors with the dataset under test. There is no exact value of *k* to be set in theory; thus, the value of *k* has to be found. In this experiment, the *k* value was varied from 3 to 10. All combination of features used in the first experiment were also used in this second experiment. The result is depicted in Table III.

TABLE III. TEST RESULTS BASED ON K VALUES VARIATION

K values	2 features	3 features (Delta)	3 features (Beta)	4 features	5 features
3	92.86	90.48	90.48	**95.42**	90.48
4	35.71	69.05	71.43	54.76	71.43
5	92.86	90.48	90.48	90.48	90.48
6	35.71	61.90	59.52	61.90	66.67
7	90.48	88.10	90.48	90.48	88.10
8	69.05	54.76	57.14	66.67	54.76
9	90.48	88.10	88.10	85.71	85.71
10	50.00	78.57	66.67	66.67	40.48

From Table III, it can be seen that the higher accuracy is achieved when the *k* value or the number of closest neighbors for classification is an odd number; namely when *k*=3, *k*=5, *k*=7, and *k*=9. This result proves that if the class for the classification of using KNN is even-numbered, in this case, "Alert" and "Drowsy" then the value of *k* should be an odd number.

Based on the first and second experiment, the best setting for the drowsiness detection system is by using four features and with the value of k for KNN is *k*=3.

V. CONCLUSION

The drowsiness detection system has been successfully developed using FFT to extract brain signal features and classification using KNN. This drowsiness detection system has been tested using a variation of the K value in the KNN classification from numbers 3 – 10 and uses several number of features. The best results of all tests conducted were an accuracy rate of 95.24% using four wave features, namely Delta, Theta, Alpha, and Beta waves and *k* = 3.

ACKNOWLEDGMENT

The authors would like to thank the Ministry of Research, Technology and Higher Education of the Republic of Indonesia to partly fund this research under PDUPT 2019 grant no No. NKB-1658/UN2.R3.1/HKP.05.00/2019.

REFERENCES

[1] Indonesian National Police, “Accident Statistics (*Statistik Laka*).” [Online]. Available: http://korlantas.polri.go.id/statistik-2/. [Accessed: 22-Oct-2018].

[2] Brainwave Entrainment, “Different Types of Brain Waves: Delta, Theta, Alpha, Beta, Gamma,” Itsusync.Com. 2016.

[3] C.X. Song, S-C. Yoon, E. Rex, J. Nieves, and C. Moretz, “Driver Drowsiness Detection Using Single-Channel Dry EEG,” Proc. 2017 IEEE Signal Proc. in Med. and Biol. Symp. (SPMB), 2017.

[4] J. He, D. Liu, Z. Wan, and C. Hu, “A noninvasive real-time driving fatigue detection technology based on left prefrontal Attention and Meditation EEG,” Process. 2014 Int. Conf. Multisens. Fusion Inf. Integr. Intell. Syst. MFI 2014, 2014.

[5] B.H. Paudel, R. Panta, N. Limbu, and B. Shrestha, “Electroencephalography: Neurophysiology Application Notes,” January, 2012.

[6] Lumen Learning, “Stages of Sleep | Introduction to Psychology.” [Online]. Available: https://courses.lumenlearning.com/intropsychmaster/chapter/stages-of-sleep/ [Accessed: 09-Jan-2019].

[7] S.K.L. Lal, A. Craig, P. Boord, L. Kirkup, and H. Nguyen, “Development of an algorithm for an EEG-based driver fatigue countermeasure,” J. Safety Res., vol. 34, no. 3, pp. 321–328, 2003.

[8] University of Rhode Island Departement of Electrical and Computer Engineering, “FFT Tutorial,” 2013.

[9] M. Rivera, L. Salas, S. By, S. Begum, and P. Funk, “Monitoring of Micro-sleep and Sleepiness for the Drivers Using EEG Signal,” Thesis School of Innovation, Design and Engineering (IDT) Mälardalen University Västerås, Sweden 2013.

[10] A. Bronshtein, “A Quick Introduction to K-Nearest Neighbors Algorithm,” 2017. [Online]. Available: https://medium.com/@adi.bronshtein/a-quick-introduction-to-k-nearest-neighbors-algorithm-62214cea29c7. [Accessed: 09-Jan-2019].

[11] A. Baseer, S. J. Weddell, and R. D. Jones, “Prediction of microsleeps using pairwise joint entropy and mutual information between EEG channels,” Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, pp. 4495–4498, 2017.