RESEARCH

Open Access

Classifying blinking and winking EOG signals using statistical analysis and LSTM algorithm

Ahmed M. D. E. Hassanein^{1*}, Ahmed G. M. A. Mohamed² and Mohamed A. H. M. Abdullah²

*Correspondence: ahmed.diaa.hassanein@gmail.com

 ¹ Systems and Information Department, Engineering and Renewable Energy Research Institute, National Research Centre (NRC), Dokki, Giza 12622, Egypt
 ² Electrical Communication and Electronics Systems Department, School of Engineering, October University for Modern Sciences and Arts (MSA), Al Wahat Road, Giza, Egypt

Abstract

Detection of eye movement types whether the movement of the eye itself or blinking has attracted a lot of recent research. In this paper, one method to detect the type of wink or blink produced by the eye is scrutinized and another method is proposed. We discuss what statistical analysis can teach us about detection of eye movement and propose a method based on long short-term memory (LSTM) networks to detect those types. The statistical analysis is composed of two main steps, namely calculation of the first derivative followed by a digitization step. According to the values of the digitized curve and the duration of the signal, the type of the signal is detected. The success rate reached 86.6% in detection of the movement of the eye when those volunteers are not trained on using our system. However, when they are trained, the detection success rate reached 93.3%. The statistical analysis succeeds in achieving detection of all types of eye movement except one type which is the non-intentional blinking. Although rate of success achieved is high, but as the number of people using this system increases, the error in detection increases that is because it is fixed and not adaptive to changes. However; we learnt from statistical analysis that the first derivative is a very important feature to classify the type of an EOG signal. Next, we propose using the LSTM network to classify EOG signals. The effect of using the first derivative as a feature for identifying the type of EOG signals is discussed. The LSTM algorithm succeeds in detecting the type of EOG signals with a percentage equal to 92% for all types of eye movement.

Keywords: Electrooculography signal, Human computer interface, Signal duration, Eye muscle motion, Feature extraction

Introduction

Electro-oculogram (EOG) signals are the electrical signals produced by the muscles of the eye. Due to some illnesses, some humans can only interact with the environment around them through their eyes [1]. Electric signals generated from their eye movement need to be recorded using special sensors. The movement of the eye becomes the only possible way for them to describe their needs and requirements [1]. Classification methods are used to classify the signals measured from the movements of the eyes.

In research, EOG are collected from people to detect mainly three types of human states. The first state is to detect the driver Consciousness level. Second is to detect eye movement types whether the motion of the eye itself or blinking detection. The third is



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicate otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http:// creativecommons.org/licenses/by/4.0/.

to detect the activities which humans are performing from the signals of the eye. In the following paragraphs, we introduce the different research papers which are recently published and are discussing these three types of human states.

In this paragraph, three papers are illustrated which discuss how to detect the state of the brain for drivers. Being absent minded or sleepy during driving is a problem that can cause car crashes [2]. In an attempt to solve this problem, Zhang et al. [2] proposes a system, which is dependent on electric signals coming out of the eyes and brain to detect the status of a driver. They use a deep long short-term memory (LSTM) network to build their model. The paper states that the compared system outperforms other proposed systems [2]. In another paper, Electro-encephalogram (EEG) and EOG signal are used to detect the attention of a driver to driving. Wu et al. [3] proposed a double-layered neural network with sub-network nodes (DNNSNs) to design their model. Their experimental results show that the proposed model outperforms previous models due to the information obtained from EOG and EEG which complement each other to produce better performance [3]. Song et al. [4] investigates the relationship between the electric signals produced from the eye and the physical and mental state of the body. A relation is found when the average, variance, standard deviation and variation coefficient were used to describe the duration of the eye blink [4]. Accordingly, the status of the human body whether he/she has a high load of work, mentally tired, physically tired or not is defined to high degree of accuracy [4].

In the next lines, few papers which discuss how to detect the motion of the eye or blinking for humans are illustrated. Aungsakun et al. [5] discuss the use of eye movement as a method for Human Computer Interface. The eye applies eight types of movements that are discussed. Usakli et al. [6] state that one of the main problems to this application is the removal of noise. The authors propose a method using first derivative, amplitude threshold and area under the curve to differentiate between the eight types of movement. The method is applied on three persons and it reached a success percentage of almost 100% [5]. Rajesh et al. [7] propose a system based on EOG signals to drive motors to be used by amputee. A new method is proposed to detect blinking so that they can be used to drive motors. Initially the method faces a problem that involuntary blinking is not differentiated from voluntarily blinking so both are used to drive the motors [7]. An average technique is used to remove this drawback which succeeds in raising the accuracy of the whole system to be 90.91% [7]. But in reference [8] and [9], a new method is proposed which depends on using images to detect whether an eye is closed or not. The success rate has reached an accuracy of 99.94% when using this method [8]. Although the success rate is very high, no state of the eye is detected other than being closed or open. In reference [10], detection of eye blinking (making a blink with the two eyes at the same time), left eye wink and right eye wink is discussed. The authors used the ZJU dataset and the accuracy of detection is 91.2% [10]. The authors examine the effect of the distance from the camera on the success rate of detection. The optimal distance is found to be 0.5 m at which the highest accuracy is achieved [10].

In the next paper, the detection of the human activity during the day is discussed. Ishimaru et al. [11] discuss the use of a new type of commercially available glasses to detect EOG signals. The proposed method is used to detect four types of human activities from the EOG signals namely reading, writing, eating and talking. In one experiment a time frame of 6 s and in another experiment a time frame of one minute is used to collect data for eye activity [11]. The success rate for identification of the activity is 70% for the first time frame and 100% for the later time frame [11].

In the following lines, the efforts done to solve a classification problem dealing with the EOG signals as the ones mentioned above are illustrated. The removal of the noise is discussed as it is the major drawback that can hinder the use of EOG signal. Lin et al. [12] discuss a new classification technique for eye movement detection. The slope of the EOG signal is passed through a limiter that limits values below a certain one. The system works successively in eye movement detection. However, when using the magnitude of the signals the detection of blinking become difficult that is why the author introduces a correction method to avoid this problem [12]. Abdel-Samei et al. [13] discuss the use of EOG signals as a way to facilitate human computer interaction. A dataset is collected from twenty-seven persons, fourteen of them are males and thirteen are females [13]. EOG sensors are used to detect horizontal and vertical eye movements. Signal processing is used to detect EOG signals and remove noise using band pass filter. The Bo-Hjorth parameter is used to extract features from the collected dataset [13]. For the classification problem, five algorithms are compared among which are K- Nearest Neighbor (KNN) and Support Vector Machine (SVM). The SVM and Cosine KNN are found to be the best classifiers to detect the horizontal and vertical classes of signals [13]. Bodrina et al. [14] discuss the removal of the abnormalities in a detected signal when the EEG and EOG signals are collected. The signals caused by eye movement and blinking introduce noise to EEG and EOG signals that should be removed. The authors use mean standard deviations of the signals read per each electrode to define the places of noise [14]. A signal with noise removal is reconstructed from the original signal. A dataset of readings from 600 persons is collected and the applied method achieves high level of success [14].

Finally, some of the applications in which EOG signal can be used are illustrated. The concept of using EEG and EOG signals to enable users to perform some functions is discussed in many papers such as [6, 15] and [16]. Graphical computer software is developed which presents to the user a number of functions to choose from using eye movements [6, 17]. Sensors with one electrode are used to collect measurements for different type of signals and showed a promising success. Applied methods are used on several users and the results are generally positive which promote further investigation in this field. Barbara et al. [18] discuss the detection of ocular angles. A new method is proposed that takes the distance between the electrodes used for detection and the centers of the left and right retinas into consideration. The method succeeds in detecting the ocular centers with an almost 20% relative error [18]. The same proposed method is used to detect gaze angles. The system succeeds in detecting gaze angles with almost the same relative errors [18].

Electric signals recorded from the brain (EEG) are used to monitor the sleeping modes of humans. In one scenario, the EEG signals can be used alone, while in another the EEG signals can be complemented with the EOG signals to monitor the sleep of some volunteers. Complementing EEG with EOG can enhance the accuracy of any system that aims to interface humans with computers [19]. The EEG–EOG signals are the main focus of some literature review to find out the methods applied on them to control computers [20]. The pre-processing techniques, feature extraction techniques and success evaluation methods have been compiled to represent them to those interested in the topic. In this research work, we are only interested in the EOG signals but further investigation into complementing our work with the benefits of the EEG signals is possible.

In this research work, the previous work of others who investigate how to detect blinking and winking is continued. An EOG sensor is used to record measurements of eye signals. A method is proposed that depends on statistical analysis of the measured signals when a blink or wink is produced to detect the type of eye signal. Finally, proposed features for an LSTM algorithm are illustrated to detect the type of eye movement that are robust and dynamic.

The rest of this paper is divided into five sections. The methodology obeyed in this research and the methods used to apply it are illustrated in section two. The dataset that we use in our calculations is described, and the way used to obtain it is laid out in section three. In section four, the LSTM algorithm is explained. The statistical analysis that is on our dataset is illustrated in section five and the different steps of processing it. In section six, the application of the LSTM algorithm is shown. Finally, the conclusion is drawn and results are compared with the accuracies of others researchers in section seven.

Methods and methodology

In this work, our first aim is to understand the behavior of the blinks and winks of the eye using statistical analysis. When examining the EOG signals, we find that blinks and winks are described into peaks and troughs. The slops of the peaks and troughs are used to define five types of signals. Double winks, single winks, intentional blinks, intentional long blinks and non-intentional blinks are all different types of signals which can be produced by the eye. When the first derivative is calculated, the slopes of peaks and troughs can be defined which can be used to define each type of signal. The maximum value of a positive slope or the absolute of the minimum value of a negative slope are seen to be always greater than 0.025 V/s. So, this value is put as a threshold to digitize the curve of the first derivative. The digitized curve is fed to a computer program to classify the type of the signal.

Statistical methods are used to detect the type of EOG signal generated by the user. A mathematical method is used which depends on the calculation of the first derivative of the EOG signal. The first derivative with respect to time is equal to the slope of the EOG signal [5]. Matlab2010aTM is used to do these calculations. Each type of EOG signal has a specific number of positive slopes and negative slopes to define its shape as is explained later. In addition, the duration of the signal is used to define the type of the signal. Then, the digitization process takes place for the first derivative. For any EOG signal, the values above 0.025 V/s are put equal to positive one. The values below negative 0.025 V/s are put equal to zero. According to the counted number of positive ones and negative ones and negative ones and the duration of the signal, the type of the signal is defined.

Our second aim is to apply the LSTM algorithm on the EOG signals to classify their type. Each EOG signal is fed to the LSTM algorithm accompanied by a group of features. Two types of groups of features are examined. The first group is composed of augmented signals of the original EOG signal. The second group is composed of augmented signals of the first derivative of the original EOG signal. Results are compared using the accuracy and the loss parameters.

Three male participants are invited for the measurement of the EOG signals. They are each informed about the purpose of the study and the procedures. They provided a verbal consent before the measurements because they are the same authors of the manuscript but a written consent can be submitted whenever required. All procedures conform to the Declaration of Helsinki and are approved by the projects committee.

Collection of dataset and tools

In this section, we illustrate the hardware and software which are used to obtain the measurements discussed in this paper. The dataset, which is collected using the hardware, is discussed as well.

Tools

Hardware and software tools are used to obtain our results. In reference [21], a new sensor is proposed to detect the EOG signals to help people to communicate with the world. The new sensor is composed of four electrodes which are placed around the eye to measure the electric signals generated by the eye [21]. The success rate of correct detection reaches 91.25% in some cases [21]. Laport et al. [17] propose a sensor that is composed of one electrode. The accuracy of the proposed sensor is relatively low [17] and is not compared to other sensors which suggests that having one electrode compromise the detection accuracy. In reference [22, 23], the paper proposes using Graphene Electronic Tattoos as electrodes to read eye signals. The electrodes are very accurate and proved their success in collecting signals coming out of the eye [22, 24]. Previous papers propose sensors which in some cases consist of large number of electrodes. In other cases, the accuracy of detection is not up to the required level or is very sophisticated as the one suggested in reference [22, 25].

In this research work, the PSL-iEOG2TM sensor is used to collect our readings. It consists of three electrodes only which make it easy to use by our volunteers. It has a high rate of accuracy in measuring the electric signals generated by the eye. This makes our selected sensor suitable for the aim which we need to study in this paper.

Three volunteers are asked to perform several single winks (SW), double winks (DW), intentional blinks (IB), non-intentional blinks (NB) and intentional long blinks (IBL). The volunteers are trained so that they can generate blinks and winks that can be detected by our computer program with high accuracy. Each volunteer attached three electrodes around his/her eye as shown in Fig. 1. The results are collected using the sensor and a computer (Intel Pentium CPU 3.4 GHz).

The readings are processed using Matlab2010aTM. Once, the processing process includes obtaining the first derivative and the digitization process. In another, the processing process includes obtaining the first derivative and applying the LSTM



Fig. 1 One volunteer is using the EOG sensor to make a DW, SW, ILB, IB or NB. Electrodes are placed around the eye to take the measurement



Fig. 2 An LSTM network with four layers [26]

algorithm on the EOG signals and its features. Microsoft Excel sheetTM is used to plot the graphs shown in this paper.

Collection of dataset

An EOG sensor is used to collect measurements of blinking and winking eye movements for three persons. The EOG sensor produces measurements in volts versus time. The Volt axis is in milli-volts (mV), and the time axis is in seconds (s). For each person, measurements for several DW, SW, ILB, IB and NB are taken. As shown in Fig. 1, one of our volunteers is shown performing one of the types of eye movements. Each volunteer is given four seconds to generate the type of signals he/she prefers with his/her eyes. For each person five files are collected where each file contains several measurements of a single type of blink or wink. For the LSTM algorithm, the dataset is created using original samples and augmented samples as explained later.

Theory of LSTM algorithm

We have four layers in the LSTM network as shown in Fig. 2. The layers are one tanh layer and three sigmoid layers. The output of the tanh layer has a range of values between negative one and positive one, but the output of the sigmoid layer has a range of values between zero and positive one.

The network starts with introducing the state of previous cell C_{t-1} to the LSTM network to decide what to keep and what to forget. This is done by the forget gate layer as shown in Fig. 2. The following equation represents the forget gate layer function [27]:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, X_t] + b_f\right) \tag{1}$$

where W_f and b_f are the weight and constant value, respectively, to improve the performance of the gate. Also, the $[h_{t-1}, X_t]$ is a concatenation operation between h_{t-1} and X_t . The layer looks at the values of the h_{t-1} and X_t to generate a number between zero and positive one using the sigmoid function as seen in Eq. (1). According to the values of f_t some values of C_{t-1} are forgotten and others are remembered.

Next, the LSTM network decides the updated information to store in the new cell state C_t through the input gate. This is done by two layers which are the sigmoid layer and the tanh layer. The sigmoid layer decides what to be replaced and the tanh layer generates the new information through the following equation [27]:

$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \tag{2}$$

where W_i and b_i are the weight and constant value, respectively, to improve the performance of the gate. The following equation is also used [27]:

$$C_t = \tanh(W_c.[h_{t-1}, X_t] + b_c)$$
(3)

where W_c and b_c are the weight and constant value, respectively, to improve the performance of the gate. To update the old state C_{t-1} , the following operation is done to generate the new cell state C_t through the following equation [27]:

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{4}$$

Finally, the fourth layer decides the output of the LSTM network h_t through the following equation [27]:

$$O_t = \sigma(W_o.[h_{t-1}, X_t] + b_o)$$
(5)

where W_o and b_o are the weight and constant value, respectively, to improve the performance of the gate. The following equation is also used [27]:

$$h_t = O_t * \tanh(C_t) \tag{6}$$

The output presented in Eq. 6 is a filtered version of the new cell state C_t . Next, the processing methods for the taken measurements are discussed.

Signal processing using statistical analysis

In this section, we scrutinize the properties of the signals for each type of EOG. The properties of each type are used to the best to detect the type of eye movement. All signals are processed using Matlab2010aTM in three steps as follows. First the measurements are plotted to find out the peaks, troughs and duration of each kind of wink or blink. Second the first derivative is taken for all measurements to allocate slops of signals for each blink or wink. Third a digitization process is done for the first derivative to produce a signal which has only three values negative one, zero and positive one. The five types of signals are to be addressed consequently. We start by the SW type. Next, a plot of the raw measurement of a SW is shown in Fig. 3a.



Fig. 3 a EOG measurement of a single wink (SW) showing one peak and one trough. **b** First derivative (FD) of a single wink (SW) showing three tips of one minimum and two maxima. **c** Digitized first derivative (FD) of a single wink (SW) showing three clearly separated values +1, -1 and zero



Fig. 4 a EOG measurement of a non-intentional blink (NB) showing one peak. **b** First derivative (FD) of a non-intentional blink (NB) showing two tips of one minimum and one maximum. **c** Digitized first derivative (FD) of a non-intentional blink (NB) showing three clearly separated values + 1, -1 and zero

Single wink signal processing

In Fig. 3a, three points are specified as shown. Points 1 and 3 define positive slops while point 2 defines a negative slope. One peak is seen with maxima at t = 0.45s. The peak is located between points 1 and 2. One trough is seen with minima at t = 0.59s. The trough is located between points 2 and 3. We can see the Full Width at Maximum absolute of the Slope (*FWMS*) for the whole SW in Fig. 3a, but its value can only be defined when the slope or derivative of the signal is calculated. Next we calculate the derivative of the signal which is shown in Fig. 3a. The derivative is plotted in Fig. 3b.

The y-axis in Fig. 3b is named as real dV/dt which means that it represents the first derivative of the raw values shown in Fig. 3a without any intermediate processing. From Fig. 3b, three maxima and minima can be seen and are defined by three points as shown. Each maximum defines a positive slope while one minimum defines a negative slope. Two maxima are defined by points 1 and 3. They are located at t = 0.44s and t = 0.62s. One minimum is defined by point 2. It is located at t = 0.52s. We can calculate the *FWMS* for the whole SW by measuring the time difference between points 1 and 3. *FWMS* is equal to 0.18s for the SW. Next, we digitize the first derivative plotted in Fig. 3b. The result is shown in Fig. 3c.

From Fig. 3c, three maxima and minima are seen which are defined by three points as shown. They are the same points defined in Fig. 3b. The difference is that all values are now having one of three possibilities negative one, zero and positive one. This kind of representation helps in defining the points of maximum positive slope and minimum negative slope by the computer. Points of maximum positive slope have a positive one value; points of minimum negative slope have a negative one value and all points in between have a zero value. This means that the computer can detect a SW when it reads two positive ones and one negative one interchanging with each other in a duration of *FWMS* = 0.18s. Next, the raw measurement of a NB is shown in Fig. 4a.

Non-intentional blink signal processing

In Fig. 4a, two points are specified as shown. Point 1 defines positive slop, while point 2 defines negative slope. One peak is seen with maxima at t = 0.53s. The peak is located between points 1 and 2. We can see the *FWMS* for the whole NB in Fig. 4a, but its value can only be defined when the slope or derivative of the signal is calculated. Next we calculate the derivative of the signal which is shown in Fig. 4a. The derivative is plotted in Fig. 4b.

From Fig. 4b, two maximum and minimum are seen which are defined by two points as shown. One maximum defines a positive slope while one minimum defines a negative slope. One maximum is defined by point 1. It is located at t = 0.5s. One minimum is defined by point 2. It is located at t = 0.57s. We can calculate the *FWMS* for the whole NB by measuring the time difference between points 1 and 2. *FWMS* is equal to 0.07s for the NB. Next, the first derivative plotted in Fig. 4b is digitized. The result is shown in Fig. 4c.

From Fig. 4c, two maximum and minimum are seen which are defined by two points as shown. They are the same points defined in Fig. 4b. The difference is that all values are now having one of three possibilities negative one, zero and positive one. This means that the computer can detect a NB when it reads one positive one and one negative one in a duration of *FWMS* = 0.07s. Next, the raw measurement of an IB is shown in Fig. 5a.

Intentional blink signal processing

In Fig. 5a, two points are specified as shown. Point 1 defines positive slop while point 2 defines negative slope. One peak is seen with maxima at t = 0.78s. The peak is located between points 1 and 2. We can see the *FWMS* for the whole IB in Fig. 5a, but its value can only be defined when the slope or derivative of the signal is calculated. Next we calculate the derivative of the signal which is shown in Fig. 5a. The derivative is plotted in Fig. 5b.

From Fig. 5b, two maximum and minimum are seen which are defined by two points as shown. One maximum is defined by point 1. It is located at t = 0.66s. One minimum is defined by point 2. It is located at t = 0.93s. We can calculate the *FWMS* for the whole IB by measuring the time difference between points 1 and 2. *FWMS* is equal to 0.27s for the IB. Next, the first derivative plotted in Fig. 5b is digitized. The result is shown in Fig. 5c.

From Fig. 5c, two maximum and minimum are seen which are defined by two points as shown. They are the same points defined in Fig. 5b. The difference is that all values



Fig. 5 a EOG measurement of an intentional blink (IB) showing one peak. **b** First derivative (FD) of an intentional blink (IB) showing two tips of one minimum and one maximum. **c** Digitized first derivative (FD) of an intentional blink (IB) showing three clearly separated values + 1, -1 and zero



Fig. 6 a EOG measurement of an intentional long blink (ILB) showing one peak. **b** First derivative (FD) of an intentional long blink (ILB) showing two tips of one minimum and one maximum. **c** Digitized first derivative (FD) of an intentional long blink (ILB) showing three clearly separated values +1, -1 and zero



Fig. 7 a EOG measurement of a double wink (DW) showing two peaks and two troughs. **b** First derivative (FD) of a double wink (DW) showing five tips of two minima and three maxima. **c** Digitized first derivative (FD) of a double wink (DW) showing three clearly separated values + 1, -1 and zero

are now having one of three possibilities negative one, zero and positive one. This means that the computer can detect a IB when it reads one positive one and one negative one in a duration of FWMS = 0.27s. Next, the raw measurement of a ILB is shown in Fig. 6a.

Intentional long blink signal processing

In Fig. 6a, two points are specified as shown. Point 1 defines positive slop, while point 2 defines negative slope. One peak is seen with maxima at t = 0.54s. The peak is located between points 1 and 2. We can see the *FWMS* for the whole ILB in Fig. 11, but its value can only be defined when the slope or derivative of the signal is calculated. Next we calculate the derivative of the signal which is shown in Fig. 6a. The derivative is plotted in Fig. 6b.

From Fig. 6b, two maximum and minimum are seen which are defined by two points as shown. One maximum is defined by point 1. It is located at t = 0.32s. One minimum is defined by point 2. It is located at t = 0.76s. We can calculate the *FWMS* for the whole ILB by measuring the time difference between points 1 and 2. *FWMS* is equal to 0.44s for the ILB. Next, the first derivative plotted in Fig. 6b is digitized. The result is shown in Fig. 6c.

From Fig. 6c, two maximum and minimum are seen which are defined by two points as shown. They are the same points defined in Fig. 6b. The difference is that all values are now having one of three possibilities negative one, zero and positive one. This means that the computer can detect a ILB when it reads one positive one and one negative one interchanging in a duration of FWMS = 0.44s. Next, the raw measurement of a DW is shown in Fig. 7a.

Double wink signal processing

In Fig. 7a, five points are specified as shown. Points 1, 3 and 5 define positive slops, while points 2 and 4 define negative slopes. Two peaks are seen with maxima at t = 0.31s and t = 0.48s. The first peak is located between points 1 and 2. The second peak is located between points 3 and 4. Two troughs are seen with minima at t = 0.39s and t = 0.62s. The first trough is located between points 2 and 3. The second trough is located between points 4 and 5. We can see the *FWMS* for the whole DW in Fig. 7a, but its value can only be defined when the slope or derivative of the signal is calculated. Next, the derivative of the signal is calculated and shown in Fig. 7a. The derivative is plotted in Fig. 7b.

From Fig. 7b, five maxima and minima are seen which are defined by five points as shown. Each maximum defines a positive slope while each minimum defines a negative slope. Three maxima are defined by points 1, 3 and 5. They are located at t = 0.3s, t = 0.47s and t = 0.68s. Two minima are defined by points 2 and 4. They are located at t = 0.34s and t = 0.52s. We can calculate the *FWMS* for the whole DW by measuring the time difference between points 1 and 5. *FWMS* is equal to 0.38s for the DW. Next, the first derivative plotted in Fig. 7b is digitized. The result is shown in Fig. 7c.

From Fig. 7c, five maxima and minima are seen which are defined by five points as shown. They are the same points defined in Fig. 7b. The difference is that all values are now having one of three possibilities negative one, zero and positive one. This means that the computer can detect a DW when it reads three positive ones and two negative ones interchanging with each other in a duration of *FWMS* = 0.38s.

Detection accuracy

The statistical features of each type of wink or blink are examined in previous subsection. In this subsection, we illustrate the differences and show how the performance of using these features in eye movement detection behaves. We start by comparing the features of each type of eye signal as shown in Table 1.

As seen in Table 1, the digitized first derivative of the three types of EOG signals, namely NB, IB and ILB, contains one peak of positive one followed by one trough of a negative one. Then, the only way to differentiate between these three types is the time duration (*FWMS*) between the peak and the trough. The average durations (*FWMS*) of all 40 original samples or measurements are listed in Table 1. Each signal type or class is represented by 8 original samples. A clear difference in their values can be seen as the NB, IB and ILB. But, the SW and DW can be clearly differentiated from each other and from the other three types by the number of peaks and troughs. The SW should contain

Signal Type	Number of + ve ones	Number of — ve one	Duration (FWMS) (s)	
NB	1	1	0.07	
IB	1	1	0.27	
ILB	1	1	0.44	
SW	2	1	0.18	
DW	3	2	0.38	

Table 1 The features of each type of EOG signal are tabulated

Signal Type	Accuracy (%)
DW	100
SW	100
ILB	100
IB	60
NB	50

Table 2 Each type of EOG signal is tabulated against the accuracy in its detection

Table 3 The accuracy in detecting random types of EOG signals in case of training our volunteers or not is shown

Training	Accuracy (%)
Without	86.6
With	93.3

two peaks and one trough. The DW should contain three peaks and two troughs. The duration (*FWMS*) of the DW and SW can be considered an extra confirmation on the type of signal.

The three volunteers mentioned in section III are asked to use the PSL-iEOG2TM sensor and record several reading for their eye blinks and winks. Each of the volunteers submits five files containing different type of eye blinks or winks. The statistical analysis is performed on each of the five files. A computer program is run to count the number of positive ones, number of negative ones and the duration of the signal. Then, it defines the type of the signal and accordingly generates a message corresponding to the detected type of signal. The different types of signals and their corresponding messages are defined randomly and stored in a table inside the computer program. The percentage of correct detection of DW, SW and ILB is excellent. The duration of the blink or wink in addition to the number of positive ones and negative ones defined for each type of signal are enough to detect any of them. All three are detected with 100% accuracy as shown in Table 2.

However for the IB and NB, the percentages of correct detection for them are relatively low. For the IB, the accuracy of correct detection is 60% as shown in Table 2. The reason is that the *FWMS* of the IB is smaller than that of the ILB and greater than that of the NB. Also, the signal of any of the three types (IB, NB and ILB) contains one positive one followed by one negative one in the digitized curve. So it is very possible that with different people the IB signal can be detected as either a ILB or a NB. As for the NB, the accuracy of correct detection is 50% as shown in Table 2. The main reason is that the value of the slope of the EOG signal for some people is smaller than the threshold used specifically 0.025 V/s. In some cases, the NB signal is not detected at all for some people. The NB is removed from being used because of its very low percentage of detection.

Three volunteers are asked to use our system to examine its accuracy in case of being trained and not being trained.

For the three volunteers, the system is explained to them and the way it works but they are not trained on how to perform the different types of blinks and winks. The three are left to record their measurements and return four files per each containing their selected type of blink or wink. The statistical analysis described above is run on these files. The accuracy of correct detection for the selected type of blink or wink is 86.6% as shown in Table 3. In another trial, the same three volunteers are trained to perform the different types of blinks and winks. The three are left to record their measurements and return to us four files per each containing their selected type of blink or wink. The accuracy of correct detection for the selected type of blink or wink is 93.3% as shown in Table 3.

We conclude that to detect eye movement temporal properties of eye motion (time domain or x-axis) should be taken in to consideration. Also, spatial properties of eye movement (strength of signal or y-axis) should be taken into consideration. The first derivative of signals coming out of the eye movement is a good representation of changes in eye movement in the time domain. Statistical analysis is excellent in detection of some types of eye movement but not all types. Next, we use the LSTM algorithm to detect types of EOG taking into consideration the importance of the first derivative of the signals of the eye motion. In the next section, the aim is to find out if the training of the LSTM network can enable it to detect the five classes with an acceptable accuracy even for the NB class.

Signal processing using LSTM algorithm

In this section, the LSTM network is used to train a deep neural network to classify time series data. An LSTM network uses the progress of events along the time steps to make predictions about similar progress of events. Kudo et al. [28, 29] uses the Japanese Vowels data set as described to predict the identity of the speaker entered uttering of two Japanese vowels pronounced consecutively. Here, the time series of the eye movement in addition to some features to identify it are used to predict the type of the movement as explained.

Data processing

The LSTM network architecture is defined as follows. A bidirectional LSTM layer with 100 hidden units is specified and output the last element of the sequence. Five classes are defined by including a fully connected layer of size five, followed by a softmax layer and a classification layer. A bidirectional LSTM layer is used since full sequences are available at prediction time [30, 31]. This enables the bidirectional LSTM layer to learn from the full sequences at each time step [32, 33]. A learning rate of 0.001 is used. The values of the parameters used are decided based on trial and error until best accuracies are achieved.

The data are processed to examine the performance when using the LSTM algorithm. For each type of eye movement, we applied data augmentation techniques to obtain few features for each signal. In Fig. 8, each blinking action is processed to obtain a discrete-time series and five other features as shown. One feature is the blinking time series that is normalized to have a maximum value of one. Another feature is the blinking time



Fig. 8 Original SW Blinking action, the normalized version of it and the zero-referenced version of it are shown in solid line. For each type, a reversed version of the signal is shown in dashed line



Fig. 9 The first derivative of the original SW Blinking action is shown in solid line. A reversed version of the signal is shown in dashed line

Table 4 The distribution of the original samples and augmented data is shown

Numbers	Туре	
40	Original samples	
40	Right Shifted (augmented)	
40	Left Shifted (augmented)	
120	Zero Referenced (augmented)	
120	Normalized (augmented)	
360	Reverse (augmented)	
480	First Derivative (augmented)	

series referenced to zero so that it has a minimum value of zero. This is obtained by subtracting the minimum value from the series. The other three features are the reversed order of the original signal, normalized and zero referenced ones.

Other data augmentation techniques are used to obtain extra features for each signal. Each blinking action is processed to obtain a discrete-time series and two other features as shown in Fig. 9. One feature is the first derivative of the blinking time series. Another feature is the reversed order of the first derivative of the blinking time series. The distribution of the original samples and the augmented data (features) are described in the following lines.

Although three volunteers have participated in this study but using data augmentation techniques a reasonable dataset is obtained. Balestriero et al. [34] suggest that augmenting the data has proven to improve the performance of the proposed datasets. As shown in Table 4, the volunteers submit 40 original samples as that shown in Figs. 3a, 4a, 5a, 6a and 7a. The original samples are distributed on the five classes. Using these original samples, 40 right shifted samples are generated which are shifted forward in the time domain. 40 left shifted samples are generated which are shifted backward in the time domain. 120 zero referenced samples are generated as that shown in Fig. 8. 120 normalized samples are generated as that shown in Fig. 8. and 480 first derivative samples are generated as that shown in Fig. 9. The dataset is divided into 80% training samples and 20% testing samples.

Detection accuracy

In this subsection, the accuracy of the classification and the losses done by the LSTM network are discussed. The original blinking action with three other features that are augmented from the original one are used as an input for an LSTM network on first trial. The three features are the normalized, reversed version of the original sample and the reversed version of the normalized sample. On second trial, the original blinking action with two other features that are augmented from the first derivative are used as an input for an LSTM network. The two features are the first derivative sample and the reverse version of the first derivative sample. Both are compared to find out the effect of the first derivative in improving the results of the LSTM network.

As shown in Fig. 10a, the accuracy achieved for the LSTM network through hundred epochs is plotted for the first trial. A best fit shows that accuracy is increasing as the number of iterations increase. The accuracy curve is unstable and keeps fluctuating even as the best fit increases. Training data are not enough to train the network to predict most cases accurately. Learning curve is not stable even as the accuracy increases.

In Fig. 10b, the losses for the LSTM network through hundred epochs are plotted. A best fit shows that loss is decreasing as the number of epochs increase. The loss curve is unstable and keeps fluctuating even as the best fit decreases. Training data is not enough to train the network to predict most cases accurately. Learning curve is not stable even



Fig. 10 The original blinking action, normalized version, zero-referenced version along with their reversed order versions are used. **a** Progress of the accuracy curve as the number of epochs increases for the training dataset is shown. **b** Progress of the loss curve as the number of epochs increases for the training dataset is shown



Fig. 11 The original blinking action and the derivative are used. **a** Progress of the accuracy curve as the number of epochs increases for the training dataset. A best fit shows that accuracy is increasing as the number of epochs increases. **b** Progress of the loss curve as the number of epochs increases for the training dataset. A best fit shows that loss is decreasing as the number of epochs increases

as the losses decrease. It takes 192 s to complete the run. Accuracy of classification for the testing data is 83.33%.

As shown in Fig. 11a, the accuracy achieved for the LSTM network through hundred epochs is plotted for the second trial. A best fit shows that accuracy is increasing as the number of iterations increase. The accuracy curve is stable and doesn't fluctuate much after reaching 40 epochs. Training data are enough to train the network to predict most cases accurately. Learning curve is stable as the accuracy increases.

In Fig. 11b, the losses for the LSTM network through hundred epochs are plotted. A best fit shows that loss is decreasing as the number of epochs increase. The loss curve is more stable and doesn't fluctuate much after reaching 40 epochs. Training data is enough to train the network to predict most cases accurately. Learning curve is stable even as the losses decrease. It takes 160 s to complete the runs. Accuracy of classification for the testing data is 91.67%.

Conclusion

Two methods to detect the type of wink or blink produced by the eye are discussed and proposed as a way for human computer interface. We start by using an ESP-iEOG2 sensor to detect eye blinks and winks. The measured signals are processed using statistical analysis. The analysis is composed of two main steps. First, we take the first derivative of the EOG signal with respect to time. A curve is produced that is composed of peaks and troughs. The peaks represent positive slopes in the raw EOG signal. The troughs represent negative slopes in the raw EOG signal. Then, a digitization process takes place on the curve of the first derivative. Each value above 0.025 V/s is given a value of positive one. Each value below 0.025 V/s is given a value of negative one. Any value in between positive 0.025 V/s and negative 0.025 V/s is given a value of zero. According to the numbers of the positive ones and negative ones in any recorded signal and their duration the type of the signal is detected. Accordingly, the type of the signal is decoded into a message to be generated by the computer program. The NB type of eye signal has a very low percentage of detection and so we advise not to use it. This method is tested on three volunteers. We conclude that training is a very important factor in the success of our system to detect the eye blinks or winks. First derivative is crucial in classifying the type of EOG signal. Next, the LSTM algorithm is used to classify EOG signals. Two types of features are used in the algorithm. One type is the original EOG signal, normalized version of it, zero references version and the reversed order of those versions. Another type is the original signal, first derivative of it and the reversed order of those versions. The formal achieved accuracy of classification for the testing data equal to 83.33% while the latter achieved accuracy of classification for the testing data equal to 91.67%. Testing takes 192 s for the formal while it takes 160 s for the latter. When comparing those results to the recent published ones. Bennett et al. proposed a CNN-LSTM network to be used to detect type of eye movement using images and achieved an accuracy of 83.5% [35, 36]. Reves et al. proposed a method based on brain wave signals to classify type of eye movement and achieved an accuracy of 92% [37, 38]. The proposed features have achieved accuracy better in the former case and almost equal in the later one. The former study has used images to detect eye movement and the later study has used brain waves to make the detection so more rigorous experimental work is required. The dataset must be unified to make the comparison as accurate as possible which is one of the problems in the machine learning field.

The advantage in the use of the LSTM algorithm is accompanying the time series signal with extra features that can improve detection accuracy. The best feature is the inclusion of the first derivative of the original signal which has proven to be effective in improving detection process as shown here and in reference [5].

Acknowledgements

Not applicable.

Author contributions

The author Ahmed M. D. E. Hassanein has done all the methodology steps, interpretation of the data, calculations, analysis, conclusions and writing. The two authors Ahmed G. M. A. Mohamed and Mohamed A. H. M. Abdullah have built the hardware.

Funding

This research was fully funded by the authors.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

The authors declare they have no competing interests.

Received: 30 May 2023 Accepted: 9 September 2023 Published online: 19 September 2023

References

- Reda R, Tantawi M, Shedeed H, Tolba MF (2020) Analyzing electrooculography (eog) for eye movement detection. In: The international conference on advanced machine learning technologies and applications (AMLTA2019) 4, pp 179–189. Springer. https://doi.org/10.1007/978-3-030-14118-9_18
- Zhang G, Etemad A (2021) Capsule attention for multimodal EEG-EOG representation learning with application to driver vigilance estimation. IEEE Trans Neural Syst Rehabil Eng 29(2021):1138–1149. https://doi.org/10.48550/arXiv. 1912.07812
- Wu W, Wu QMJ, Sun W, Yang Y, Yuan X, Zheng W.-L, Lu B-L (2018) A regression method with subnetwork neurons for vigilance estimation using EOG and EEG. IEEE Trans Cognit Dev Syst. https://doi.org/10.1109/TCDS.2018.2889223
- 4. Song J, Wang R, Zhang G, Xiong C, Zhang L, Sun C (2015) Electrooculogram signals analysis for process control operator based on fuzzy c-means. Int J Adv Comput Sci Appl 6(9). https://doi.org/10.14569/IJACSA.2015.060918
- Aungsakun S, Phinyomark A, Phukpattaranont P, Limsakul C (2012) Development of Robust EOG-based humancomputer interface controlled by eight-directional eye movements. Int J Phys Sci 7:2196–2208. https://doi.org/10. 5897/JJPS11.1486

- Usakli AB, Gurkan S, Aloise F, Vecchiato G, Babiloni F (2010) On the use of electrooculogram for efficient human computer interfaces. In: Computational intelligence and neuroscience, vol. 2010, Article ID 135629, 5 pages. https:// doi.org/10.1155/2010/135629
- Anandan NR (2012) Electrooculogram (EOG) signal classification using moving average technique and its application to drive direct current motors. Recent Adv Electr Electron Eng 11(2). https://doi.org/10.2174/235209651066617 0926161127
- Kowalczyk P, Sawicki D (2019) Blink and wink detection as a control tool in multimodal interaction. Multimedia Tools Appl 78. https://doi.org/10.1007/s11042-018-6554-8
- Missimer E, Betke M (2010) Blink and wink detection for mouse pointer control. In: Proceedings of the 3rd international conference on pervasive technologies related to assistive environments, pp 1–8. https://doi.org/10.1145/ 1839294.1839322
- Singh H, Singh J (2018) Real-time eye blink and wink detection for object selection in HCI systems. J Multimodal User Interfaces 12. https://doi.org/10.1007/s12193-018-0261-7
- Ishimaru S, Kunze K, Tanaka K, Uema Y, Kise K, Inami M (2014) Smarter eyewear—using commercial EOG glasses for activity recognition. In: UbiComp 2014—Adjunct Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 153–156. https://doi.org/10.1145/2638728.2638795
- Lin C-T, King J-T, Bharadwaj P, Chen C-H, Gupta A, Prasad M (2019) EOG-based eye movement classification and application on HCI baseball game. IEEE Access, 1–1. https://doi.org/10.1109/ACCESS.2019.2927755
- 13. Abdelsamei A, Ali A, Abd El-Samie, F, Brisha A (2021) Efficient classification of horizontal and vertical EOG signals for human computer interaction. https://doi.org/10.21203/rs.3.rs-471385/v1
- 14. Bodrina N, Sidorov K (2021) The algorithm of automatic localization of EOG artifacts in a multichannel EEG signal1. In: CEUR workshop proceedings. 2021. CEUR-WS.org/vol-2834/shortpaper009.pdf
- Barea R, Boquete L, Ortega S, Guillén M, Rodriguez-Ascariz J (2012) EOG-based eye movements codification for human computer interaction. Expert Syst Appl 39:2677–2683. https://doi.org/10.1016/j.eswa.2011.08.123
- López A, Martin FJ, Yangüela D, Alvarez Peña C, Postolache O (2017) Development of a computer writing system based on EOG. Sensors 17:1505. https://doi.org/10.3390/s17071505
- Laport F, Iglesia D, Dapena A, Castro PM, Vazquez-Araujo FJ (2021) Proposals and comparisons from one-sensor EEG and EOG human–machine interfaces. Sensors (Basel, Switzerland), 21(6):2220. https://doi.org/10.3390/s21062220
- Barbara N, Camileeri T, Camilleri K (2020) EOG-based ocular and gaze angle estimation using an extended Kalman filter. In: 1–5, ACM symposium on eye tracking research and applications; Association for Computing Machinery, 2020, New York, NY, USA. https://doi.org/10.1145/3379156.3391357
- Da Silva Souto CF, Pätzold W, Wolf KI, Paul M, Matthiesen I, Bleichner MG, Debener S (2021) Flex-printed ear-EEG sensors for adequate sleep staging at home. Front Digital Health 3:688122. https://doi.org/10.3389/fdgth.2021.688122.
- 20. Belkhiria C, Peysakhovich V (2020) Electro-encephalography and electro-oculography in aeronautics: a review over the last decade (2010–2020). Front Neuroergonomics 1: 606719. https://doi.org/10.3389/fnrgo.2020.606719.
- 21. Heo J, Yoon H, Park KS (2017) A novel wearable forehead EOG measurement system for human computer interfaces. Sensors 17(7): 1485. https://doi.org/10.3390/s17071485
- 22. Ameri S, Kim M, Kuang I, Perera W, Alshiekh M, Jeong H, Topcu U, Akinwande D, Lu N (2018) Imperceptible electrooculography sensor system for human–robot Interface. npj 2D Mater Appl 2. https://doi.org/10.1038/ s41699-018-0064-4
- Lopez A, Fernandez D, Martin FJ, Valledor M, Postolache O (2016) EOG signal processing module for medical assistive systems. 1–5. https://doi.org/10.1109/MeMeA.2016.7533704
- 24. Merino Monge M, Rivera O, Gonzalez G, Maria I, Cantero A, Zubiete E (2010) A method of EOG signal processing to detect the direction of eye movements. In: Proceedings—1st international conference on sensor device technologies and applications, SENSORDEVICES 2010. 100, 100–105. https://doi.org/10.1109/SENSORDEVICES.2010.25
- Dasgupta A, Routray A (2021) A new multi-resolution analysis method for electrooculography signals. In: IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society. 2021. https://doi.org/10.1109/TNSRE.2021.3117954
- 26. http://colah.github.io/posts/2015-08-Understanding-LSTMs/, visited June 2023
- 27. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9:1735–1780. https://doi.org/10. 1162/neco.1997.9.8.1735
- Kudo M, Toyama J, Shimbo M (1999) Multidimensional curve classification using passing-through regions. Pattern Recognit Lett 20(11–13): 1103–1111. ISSN:0167-8655, 1999. https://doi.org/10.1016/S0167-8655(99)00077-X
- Belkhiria C, Boudir A, Hurter C, Peysakhovich V (2022) EOG-based human-computer interface: 2000–2020 review. Sensors (Basel) 22(13):4914. https://doi.org/10.3390/s22134914
- Bhatnagar S, Gupta B (2022) Acquisition, processing and applications of EOG signals. In: 2022 8th International conference on signal processing and communication (ICSC), Noida, India, pp 266–270, https://doi.org/10.1109/ICSC5 6524.2022.10009179
- https://www.mathworks.com/help/deeplearning/ug/classify-sequence-data-using-lstm-networks.html, visited June 2023
- 32. Zhu W, Tao T, Yan H, Yan J, Wang J, Li S, 信昆仑, K-L (2023) An optimized long short-term memory (LSTM)-based approach applied to early warning and forecasting of ponding in the urban drainage system. Hydrol Earth Syst Sci 27:2035–2050. https://doi.org/10.5194/hess-27-2035-2023
- Hernandez Pérez S, Pérez Reynoso F, Gonzalez-Gutierrez C, Cosío León M, Ortega-Palacios R (2023) EOG signal classification with wavelet and supervised learning algorithms KNN, SVM and DT. Sensors 23:4553. https://doi.org/10. 3390/s23094553
- Balestriero R, Misra I, LeCun Y (2022) A data-augmentation is worth a thousand samples: exact quantification from analytical augmented sample moments. https://doi.org/10.48550/arXiv.2202.08325
- 35. Bennett R, Joshi SH (2021) A CNN and LSTM network for eye-blink classification from mri scanner monitoring videos. In: Annual international conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in

Medicine and Biology Society. Annual International Conference, pp 3463–3466. https://doi.org/10.1109/EMBC4 6164.2021.9629937

- López A, Villar J, Fernández M, Ferrero F (2023) Comparison of classification techniques for the control of EOG-based HCIs. Biomed Signal Process Control 80:104263. https://doi.org/10.1016/j.bspc.2022.104263
- Reyes A, Camacho C, Mateus A, Calderon J (2021) LSTM based brain-machine interface tool for text generation through eyes blinking detection, pp 1–6. https://doi.org/10.1109/CCNC49032.2021.9369597
- Sho'ouri N (2022) Detection of ADHD From EOG signals using approximate entropy and petrosain's fractal dimension. J Med Signals Sens 12(3):254–262. https://doi.org/10.4103/jmss.jmss_119_21

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ► Convenient online submission
- ► Rigorous peer review
- Open access: articles freely available online
- ► High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com