2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)
Allerton Park and Retreat Center
Monticello, IL, USA, September 24-27, 2019

Blink: A Fully Automated Unsupervised Algorithm for Eye-Blink Detection in EEG Signals

Mohit Agarwal
Electrical and Computer Engineering
Georgia Institute of Technology
magarwal37@gatech.edu

Raghupathy Sivakumar
Electrical and Computer Engineering
Georgia Institute of Technology
siva@gatech.edu

Abstract— **Eye-blinks are known to substantially contaminate EEG signals, and thereby severely impact the decoding of EEG signals in various medical and scientific applications. In this work, we consider the problem of eye-blink detection that can then be employed to reliably remove eye-blinks from EEG signals. We propose a fully automated and unsupervised eye-blink detection algorithm, *Blink* that self-learns user-specific brainwave profiles for eye-blinks. Hence, *Blink* does away with any user training or manual inspection requirements. *Blink* functions on a single channel EEG, and is capable of estimating the start and end timestamps of eye-blinks in a precise manner. We collect four different eye-blink datasets and annotate 2300+ eye-blinks to evaluate the robustness performance of *Blink* across headsets (OpenBCI and Muse), eye-blink types (voluntary and involuntary), and various user activities (watching a video, reading an article, and attending to an external stimulation). The *Blink* algorithm performs consistently with an accuracy of over 98% for all the tasks with an average precision of 0.934. The source code and annotated datasets are released publicly for reproducibility and further research. To the best of our knowledge, this is the first ever annotated eye-blink EEG dataset released in the public domain.**

I. Introduction

Electroencephalography (EEG) signals captured from the scalp of the brain in the form of non-stationary electric potentials provides a window into the neural activity in the brain. Until two decades ago, the applications for EEG were limited to clinical and medical diagnostics, including epilepsy, Alzheimer's, coma, brain disorder, etc. More recently, significant advances in wearable hardware and sensing technologies have enabled the recording of high-quality EEG data using off-the-shelf headsets. EEG measures are typically recorded and correlated with cognitive and physiological processes to gain deeper insights into the functionality of such processes. The Human Connectome Project (HCP)[1] is one such initiative to map the human brain on a neuronal level through various neural measures, including EEG. Meanwhile, the widespread availability of wearable EEG headsets is also gaining the attention of consumers interested in the analysis of their mental health, mindfulness, meditation and sleep statistics.

However, EEG signals are quite vulnerable to distortion caused by other interfering electrical fields. Specifically, eye-blinks produce a very strong interfering electric field (as the retina and cornea form an electric dipole [1], [2]) severely impacting the signal-to-noise ratio (SNR) of recorded EEG measurements. The presence of eye-blink artefacts in the EEG signal leads to confused or possibly false EEG interpretations. Hence, the detection and removal of eye-blink components can be significantly useful in any EEG analysis. Several algorithms have been proposed in the literature to identify eye-blinks, but they are characterized by one or more of the following limiting requirements - (i) a partly manual inspection for thresholds or template selection, (ii) a user training phase, (iii) a high number of EEG channels, and (iv) Electrooculography (EOG) data requiring additional electrodes above and below the eyes.

In this context, we first show that the brainwaves generated when a user eye-blinks are detectable with a high degree of robustness. We then propose a fully automatic and unsupervised (i.e. *without requiring any training from the user*) eye-blink detection algorithm, *Blink*, to identify accurate timestamps of eye-blinks in the EEG data. The precise timestamping of eye-blinks in the EEG data maximizes the availability of clean EEG for analysis, and can provide insights into eye-blink duration and eye-blink interval. *Blink* relies on the natural frequency of occurrence of eye-blinks to self-learn brainwave profiles for each specific user's eye-blinks, and hence does away with any user training requirements. *Blink*'s design requires only a single EEG channel to operate.

Through extensive user experiments, we show that *Blink* can detect eye-blinks robustly across different EEG headsets and various user activities. We use two different commercially available BCI platforms—Muse and OpenBCI—to show the generalizability of *Blink* over EEG headsets. We use controlled and uncontrolled user studies to evaluate the performance of *Blink* over involuntary and voluntary eye-blinks, respectively. Overall, we collected four different user EEG datasets (Table I) with real users containing more than 2300 eye-blink waveforms. We show that *Blink* detects eye-blinks with an accuracy of over 98% for all four datasets along with a high degree of precision.

We have publicly released our collected datasets, and code[2] for the *Blink* algorithm so that the results presented in this paper can be reproduced[3]. To the best of our knowledge, this is the first-ever annotated eye-blink EEG dataset released in the public domain. We later discuss a methodology for

[1]http://www.humanconnectomeproject.org/

[2]Datasets and code are available at `https://github.com/meagmohit/BLINK`.

[3]User data is anonymized to ensure the privacy

Dataset	Device	Type	Users	Total	Activity
EEG-IO	OpenBCI	Involuntary	20	500	external stimulation
EEG-IM	Muse	Involuntary	20	500	external stimulation
EEG-VV	OpenBCI	Voluntary	12	750	watching video
EEG-VR	OpenBCI	Voluntary	12	600	reading article

TABLE I: EEG datasets collected for *Blink* evaluation

using *Blink* as-is in an online fashion to enable real-time eye-blink detection. This can widen the applicability of *Blink* in the domains of Brain-Computer Interface (BCI) based communication and control, and real-time EEG data processing.

The rest of the paper is organized as follows; in Section 2 we review the related work at the intersection of EEG and eye-blinks. In Section 3, we describe the manifestations of eye-blinks on EEG data, and discuss the complexity of the problem in hand. Section 4 presents the *Blink* algorithm, followed by its evaluation in Section 5. Finally, we discuss the limitations of *Blink* and future work in Section 6 and conclude the paper.

II. Related Work

Several related works lie at the intersection of EEG and eye-blinks, which can be broadly classified into two categories: (i) removing eye-blink artifacts from the EEG signal, and (ii) detecting the time instants of eye-blinks in EEG. From a technical perspective, both categories are quite different from each other. The former removes the eye-blink components from EEG resulting in pure cerebral data, however, is unable to locate the time instants of eye-blinks. The latter locates the time instants but is incapable of removing the distortion without losing the cerebral data within the eye-blink duration. Several hybrid approaches have been proposed in the literature first to identify the eye-blinks and removing the related component to clean the signal [3], [4].

A. Eye-Blink component removal methods

Multiple strategies are proposed in the literature to purify the EEG waveform using Blind Source Separation (BSS) based methods. These methods [5], [6], [7], [8], [9] vivisect EEG waveform into additive subcomponents using BSS algorithms like Independent Component Analysis (ICA) and remove the non-cerebral (mostly eye-blink) component from the EEG using template matching. The templates are created with labeled eye-blink examples which are proved to be consistent across users. These methods perform very well but require maintaining a large database of templates, and sampling from a large number of electrodes to find the multiple subcomponents. [10] is one such semi-automatic process requiring the manual labeling and selection of a template. Some of these works even require putting extra electrodes over and above the eye, also known as Electrooculography (EOG) [7]. EyeCatch [11] uses a similar strategy to detect eye-blinks specifically. It analyses and compares the IC scalp maps with the half-million scalp maps present in their database. ICA based approaches are advantageous in circumventing the limitations of conventional artifact detection methods, however, the can be only used in dense EEG systems due to their strict requirements of a high number of EEG channels.

B. Eye-blink identification methods

A very trivial approach to detect eye-blink timestamps is to continuously monitor the EEG signal and detect eye-blink if the amplitude crosses a preset threshold value. Improved approaches in the literature extract relevant features to apply a threshold. In [12], various statistic based features were calculated for data artifacts in five aspects of the EEG data: channels, epochs, ICs, single-channel single-epochs, and aggregated data (i.e., across subjects). A threshold of ± 3 was used for the Z-score for each feature to detect the eye-blink artifact. [12] was shown to perform with a score of 94.47 and 98.96 for sensitivity and specificity respectively on simulated data over 128-channels. The performance of [12] drops significantly with a reduced number of electrodes (i.e., 32). [13] employs the use of extreme statistics and used p-value as the threshold parameter to detect the eye-blink artifacts on 29-channel EEG data. An automatic threshold of $\mu + 2\sigma$ is used along with channel correlation (in Fp1 and Fp2) electrodes in [14]. [15] proposed the use of multi-window summation of derivatives approach and compared against the correlation, Dynamic Time Warping (DTW) and Root Mean Square Error (RMSE) based approaches. The Similar threshold-based approach was used in [16] along with DTW. [17] applies threshold-based peak detection technique for activating the home lighting system. Such threshold-based techniques were also used in [18], [13] over the frequency spectrum. Power Spectrum Density (PSD) of a moving window was compared to a threshold to detect eye-blink artifacts. The performance of such methods suffer due to a high variance in eye-blink duration, and the peak not falling in the middle of the window. Threshold-based approaches are highly sensitive to the chosen features and preset threshold, which could vary highly across devices and subjects.

Fingerprint or template matching based methods are widely used in the field of pattern recognition. In these approaches, an eye-blink template (or fingerprint) is first obtained and then matched with the continuous EEG data using a moving window. If the similarity measure crosses a preset threshold value, an eye-blink signal is detected in the particular window. These methods are highly sensitive to the chosen template and the similarity metric. [19] applied Dynamic Positional Warping (DPW), a variant of DTW and demonstrated the accuracy improvements over DTW [20], RMSE and correlation [21] as the similarity metric. The templates are typically chosen either through manual inspection or generated with an algorithm. [19] selected five templates from the ground truth dataset, and hence is not fully unsupervised.

Supervised learning based methods design a specific kind of neural network architecture (or deep architecture) for learning the distinctive and similar patterns based on the training data [3], [22], [23]. [23] uses Support Vector Machines (SVMs) for identification of eye-blink artifacts with a moving window of 450ms. [24] uses segmentation of a 1-second window, and applies RBF network on three extracted features achieving an accuracy of 75.3%. Such techniques demand user training and are heavy in computation (for training) and memory (weight storage).

Other algorithms that work on purely statistical techniques do not estimate the eye-blink positions but rather count them [25], [26] or are highly sensitive to the input parameters. [26] does not specifically detect eye-blinks but any spiked artifacts. This can result in high false positives as a result of the eye and head movements. Sensitivity to the input parameters defeats the universality point. [4] proposed a complicated approach of combining high-speed eye tracker to timestamp eye-blinks and further removed artifacts caused by eye-blinks and movements. [27] proposed a novel combination of ICA with mutual information and wavelet analysis to achieve 97.8% accuracy using 6 EEG and 2 EOG electrodes. [28] detects eye-blink artifacts with 90% specificity and 65% sensitivity using extended Kalman filter. [29] performs DTW score clustering during wearable EEG-based cognitive workload assessment tests to achieve an accuracy of 96.42%. Despite the attractive performance rates, the proposed method is not suitable due to the requirements of user training and 7- EEG channels. [30] relies purely on statistical techniques but requires EEG signal for an extended period (offline), to extract eye-blink profile.

Regression-based methods require measuring EOG electrodes to correctly estimate the regression coefficients [31], [32], [33]. This again puts additional hardware requirements on the available EEG architectures in the market and are not suitable for our case, hence, we skip the discussion of such approaches.

Thus, there does not exist any eye-blink detection algorithm (through EEG) that fits the requirements of universality, no supervised-training, no manual involvement, small form-factor, and near-perfect detection accuracy. In this context, we later present in the paper, a novel solution and compare against a specific related work, BLINKER [30].

III. Eye-Blink Characteristics and Detection Challenges

A. Blink waveform characteristics

A typical eye-blink waveform on the frontal EEG is visually similar to a trough waveform in the voltage-time domain. Fig. 1 shows a snapshot of such waveform at frontal electrode position (Fp1 in this case, according to the 10-20 electrode system) referenced to the earlobe electrodes (x-axis: time-domain, y-axis: voltage-domain). The eye-blink waveform can be characterized by its (i) waveform pattern, (ii) eye-blink amplitude, and (iii) eye-blink duration. An eye-blink waveform pattern is defined as the voltage variation with time during a natural or forced eye-blink. The depth of the trough in the waveform pattern is known as the eye-blink amplitude. Eye-blink duration is simply the time taken by the user to perform the eye-blink.

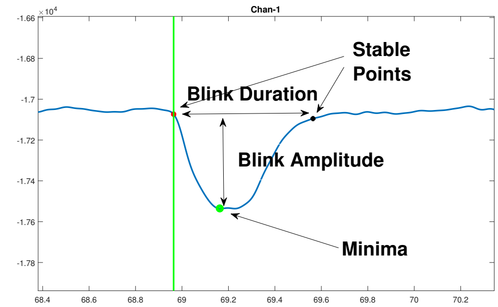


Fig. 1: A typical eye-blink waveform

B. Detection Challenges

Detecting eye-blinks is ostensibly easy as eye-blink waveforms are visually prolific in features (as in Fig.1). The normalized eye-blink waveform pattern (in time and voltage domains, i.e., single-unit time duration and single-unit voltage deviation) is consistent across multiple eye-blinks of a single user, and also across different users. We can see this from Fig. 2, that the similarity (correlation) of eye-blink templates without considering amplitude deviation in correlation metric is similar for intra-subject eye-blinks (multiple eye-blinks of a single user) and inter-subject (eye-blinks across users). In reality, state-of-the-art technologies present EEG waves inter-weaved with high-power noise (including inherent signal noise and measurement sensor noise). The variability across user-specific eye-blink waveforms are so high across users (considering the amplitude deviation for eye-blink waveforms) that if compared on the same scale, what looks like an eye-blink waveform for one user is simply noisy perturbations for another user. The high variability is not just limited to across users, but also is exhibited across different eye-blink waveforms of a specific user (Fig. 2 shows that when amplitude deviation is considered in the correlation metric, the correlation drops significantly in the case of eye-blinks across users (inter-subject))[4]. This high variability among the eye-blink patterns poses the first challenge of *designing a single universal algorithm that can account for the user and state variability, without an explicit requirement of fine-tuning algorithmic parameters.*

One might argue for the deployment of supervised training based approaches (e.g., neural networks, deep learning) to tackle the user-variability and noise issues like in the image or speech recognition problems. However, such a solution strategy is undesirable for wearable BCIs, where *user comfort* is an important consideration. Supervised training based approaches require users to go through an extensive training phase that directly impacts the usability and hence, the consumer adoption of such devices. The second challenge, thus, is to *devise solutions that eliminate the user-training phase (essentially eliminating all supervised training based approaches*).

The above challenges when coupled with the *small form-factor constraints* (usage of fewer channels), and *high accu-*

[4]For this result, we used *EEG-VR* dataset (Table I)

racy requirements with low false positives (high precision - robust detection to avoid user frustration), considerably elevates the complexity of this problem. In summary, the key challenges in developing an eye-blink detection algorithm are the following: (i) Universality, (ii) No supervised training, (iii) Small form-factor and (iv) Accurate performance.

IV. THE *Blink* DETECTION ALGORITHM

We propose an algorithm *Blink* that is capable of robust eye-blink detection without requiring any training from the user. *Blink* is presented in Algorithm 1 along with subroutine 1.

A. Assumptions

Blink operates on two assumptions

1) Consistency of eye-blink patterns: It assumes that the eye-blink patterns are consistent for a single user for a short period (i.e., during data recording). However, no such assumption is made for different users (or different recordings) and hence allows for user and session variability. To validate this assumption, we utilize the *EEG-IO* dataset (Table I), which provides us with the timestamps of true eye-blinks. For the user EEG data with given eye-blink waveforms, we extract a template eye-blink signal (or fingerprint) based on the given eye-blinks, and compute the correlation of template with (a) noise waveforms (but similar to trough pattern) shown as *crosses* and (b) the given eye-blink waveforms shown as *circles* in Fig. 3. Based on the correlation threshold comparison[5], if the waveforms are classified as eye-blink or noise using a threshold, we mark the corresponding incorrectly classified waveforms using red ink. The template extraction and correlation is done for users separately (total 10 subjects are shown in Fig. 3, best-5 and worst-5 are shown) in Fig. 3(i) and Fig. 3(iii), and finally for all the subjects together i.e., one template eye-blink waveform for all users (global fingerprint) in Fig. 3(ii). When subjects are treated separately, eye-blink waveforms can be assumed consistent i.e., a single template can represent all the eye-blink waveforms robustly and hence can distinguish from the noisy trough patterns. However, this is not true for multiple users due to the high overlap between eye-blink and noise correlation with the template, as in Fig. 3(ii). Similarly, if amplitude deviation is not considered, the overlap between noise and eye-blink waveforms is significantly high, adversely affecting the detection performance (Fig. 3(iii)). This establishes the consistency in eye-blink patterns for a particular user and can be leveraged to detect eye-blinks from the raw EEG feed efficiently.

2) No other repetitive waveforms: There are no other repetitive waveforms in the input signal that present the same characteristics as an eye-blink waveform. This is a valid assumption, as frontal electrodes are mostly corrupted by eye-blinks, eye movements, facial muscles, and head movements. The pattern of other waveforms is either non-repetitive and random or dissimilar to the eye-blink waveform (trough-shaped).

[5]A threshold was selected to minimize the number of incorrect classifications. For each waveform, its correlation was compared with the threshold to label as *eye-blink waveform* or *noise waveform*

Algorithm 1: *Blink*[6]: an eye-blink detection algorithm based on feature detection and cluster-analysis

Input : E: EEG raw data, f_s: Sampling frequency
Output : $[t_{start}]$: start time of all eye-blinks, $[t_{end}]$: end time of all eye-blinks

1 Preprocess: lowpass filter E
2 $[t_{peaks}] \leftarrow peak\_detect(E, delta = 0)$
3 $[t_{start}], [t_{min}], [t_{end}] \leftarrow identify\_stable\_points(E, delta = 0, [t_{peaks}])$
4 **for** $i = 1, 2, \cdots, size([t_{min}])$ **do**
5   **for** $j = i + 1, i + 2, \cdots, size([t_{min}])$ **do**
6         $sig_a \leftarrow E[t^{(i)}_{start} : t^{(i)}_{min} : t^{(i)}_{end}]$
7         $sig_b \leftarrow E[t^{(j)}_{start} : t^{(j)}_{min} : t^{(j)}_{end}]$
8         $corr_{mat}[i, j] \leftarrow correlate(E, sig_a, sig_b)$
9         $power_{mat}[i, j] \leftarrow max(\frac{std(sig_a)}{std(sig_b)}, \frac{std(sig_b)}{std(sig_a)})$
10   **end**
11 **end**
12 $[index_{blinks}] \leftarrow high\_corr\_comp([[corr_{mat}]], [[power_{mat}]])$
13 $stable_{th}, delta \leftarrow blink\_typify([t_{start}], [t_{min}], [t_{end}], [index_{blinks}])$
14 $[t_{peaks}] \leftarrow peak\_detect(E, delta)$
15 $[t_{start}], [t_{min}], [t_{end}] \leftarrow stable\_points(E, stable_{th}, t_{peaks})$
16 Repeat *steps 5 to 15*
17 **return** $[t_{start}], [t_{end}]$

B. Blink *Algorithm*

Some properties of *Blink* algorithm are, (i) *Blink* relies on the natural frequency of occurrence of eye-blinks to self-learn brainwave profiles for each specific user's eye-blinks, and hence does away with any user training requirements (it performs unsupervised learning); (ii) *Blink* requires raw EEG data as input and returns the start and end positions of the eye-blinks in the EEG data. Thus, *Blink* can easily provide insights into the eye-blink duration, and eye-blink interval; (iii) *Blink* design requires only single-channel data. However, in the case of multiple channels the results can be combined (in an OR fashion) to achieve more accurate results;

Algorithm Explanation: The pre-processing step (*line 1*) is to apply a low-pass filter to suppress high-frequency noise and smoothening the signal. The first step of the algorithm is to find local minimas and stable points (Fig. 1). Subroutine 1 ($peak\_detect$) finds the local minimum points in the signal separated at least by $2w$ units in the time-domain (*line 2*). With each minimum point found, the algorithm searches for nearby stable points (*line 3*), where the signal fully recovers from the eye-blink trough (as shown in Fig. 1). This is performed in function ($stable\_points$) where the vicinity of

[6][] and [[]] represents 1-D and 2-D array respectively in the algorithm, std represents the standard deviation

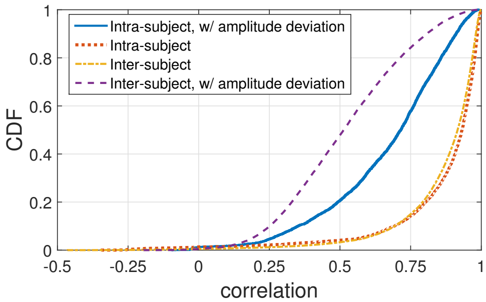


Fig. 2: Correlation of eye-blink waveforms for (i) multiple eye-blinks of a single user, (ii) eye-blinks across users, each with and without considering the amplitude deviation

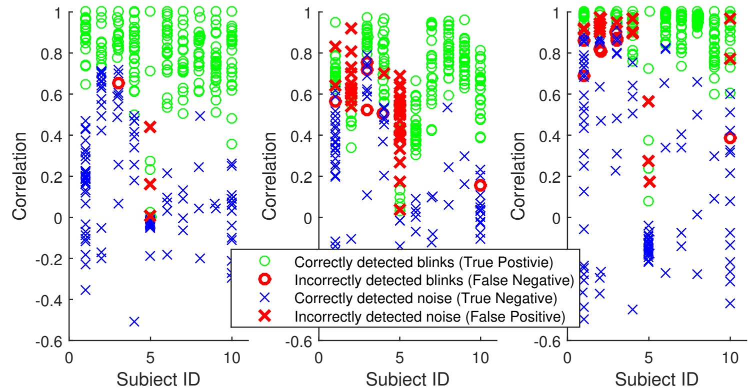


Fig. 3: Correlation with template eye-blink waveform for given eye-blinks and trough-shaped noise, (i) template is constructed independently with amplitude deviation (intra-subject with amplitude deviation) , (ii) template is constructed together for all subjects with amplitude deviation (inter-subject with amplitude deviation), (iii) template is constructed independently without amplitude deviation in correlation (intra-subject)

Subroutine 1: Subroutine *peak_detect* for *Blink* algorithm

Input : E: EEG raw data, $delta$: threshold for peak detection
Parameters: w : size of the moving window

Initialize $[t_{min}]$ with all local minimas in E
if $delta$ *is* 0 **then**
    return subset of $[t_{min}]$ such that consecutive elements are separated by w units in time-domain
else
    return subset of $[t_{min}]$ such that consecutive elements are separated by $delta$ units in voltage-domain
end

---

each local minima is scanned to estimate the noise power (or $stable_{th}$), which in turn is used to compute aforementioned stable points such that the signal power from minima to a stable point crosses *stable_th*, but is limited after stable points. If, for any particular minima two stable points are not found (one on the left, and other on the right), such local minimum points are discarded for further eye-blink investigation, and a set of stable points are returned for every other local minimum.

At this point (*line 3*), the algorithm has a set of trough patterns (each pattern consists of one local minimum and two stable points), which are further interpolated (as time length is different for each pattern) and linearly correlated on a one-to-one basis (*line 4-11*) to compute the cross similarity matrix in the time-domain (eye-blink shape) and the voltage-domain (eye-blink amplitude).

Further, highly correlated components of such patterns is computed (*line 12*, $high\_corr\_comp$) based on the time-domain similarity and a correlation threshold (which is kept low for robust detection) to find the matching repetitive patterns. The repetitive patterns might look similar (in the time-domain) but could correspond to eye-blink waveform (high amplitude) or simply noise (less amplitude), which is further separated into two different clusters, and the high trough amplitude cluster is returned as potential eye-blinks. To make the algorithm more robust, resultant eye-blink patterns are profiled (smartly characterized) to have a better estimate of the noise power and the eye-blink amplitude (*line 17* $blink\_typify$). Finally, a second pass is done to recover any missed eye-blink patterns (*line 14-16*), with the additional information of eye-blink SNR (signal-to-noise ratio) and user eye-blink profile. Thus, in the end, *Blink* algorithm robustly detects all eye-blink patterns along with their start and end times.

Subroutine *peak_det* detects the minimas in the signal data separated at least by $2w$ units. The subroutine, if provided with a non-zero *delta* threshold, identifies the minimas which have at least of delta-amplitude difference with immediate maximas.

A careful inspection of the algorithm reveals that the parameters of the *Blink* algorithm (and corresponding subroutines) are filter orders, different moving window sizes (time-domain), and correlation thresholds, which are not required to be tuned to different users, and thus allowing for the user-agnostic universality of the algorithm.

V. Evaluation

In this section, we first explain the user experiments conducted along with the correspondingly collected EEG data. We then evaluate the *Blink* algorithm to validate its near-zero detection error with low false positives.

A. Experimental Protocol and EEG Dataset Description

We have conducted four different user experiments to evaluate the robustness of the *Blink* algorithm under a variety

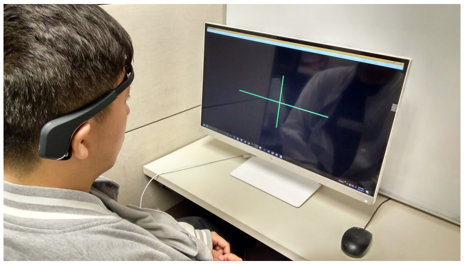


Fig. 4: User-Evaluation Setup

of EEG headsets and tasks. All the research protocols for the user data collection were reviewed and approved by the Institutional Review Board of Georgia Institute of Technology. The subjects for the study were recruited from mixed demographics with an age range between 22 to 30 years old and were either full-time students or full-time employees. Upon arrival, the experimental protocol was explained to the subjects, and the subjects were provided with consent forms and a demographic questionnaire. They were compensated with Amazon gift cards (10 USD value) for their successful participation in the study. The experimental paradigms and the collected EEG datasets are explained below:

A. Guided single eye-blink experiments: We collected raw EEG traces from 20 subjects in a guided (i.e., software instructed) environment where subjects were asked to perform a single eye-blink when instructed. Subjects were asked to sit comfortably in front of a computer screen and wear a BIOPAC 100C electrode cap [34]. Electrode gel was used to ensure the surface contact between the Fp1 and Fp2 (as per the 10-20 electrode system) electrodes on the scalp and forehead. Two silver ear-clip electrodes were additionally placed on the left and right earlobes to serve as a reference and to aid in the noise cancellation. The electrode cap was attached with the OpenBCI platform, which sampled the raw EEG at 250Hz. The digital signals were shipped to a desktop machine over the wireless channel. We used OpenViBE software (developed by Inria [35]) to present the on-screen stimulations and collect the user EEG data with synchronized timestamps. We also recorded a video of the subjects performing the experiments. The subjects were asked to perform a single eye-blink ONLY if a *green plus* appears on the screen (fig. 4). One experimental session presented 25 such external stimulations to perform eye-blinks every 3-4s depending on the subject's preference, resulting in the experiments lasting for 75 to 100 seconds per user. We repeated the same experimental protocol with Muse headset [36]. Muse headset is a dry-electrode headset and does not require a sticky gel to maintain the scalp contact. The Muse electrodes were moistened with water before the headset was worn by the user. We used the Muse Monitor application [37] on an Android platform to collect the user EEG data, however, the stimulations on a computer screen were still provided using the OpenViBE platform[7]. For both of the experiments, the video feed was manually reviewed, and true labels of the eye-blinks were marked for providing the *ground truth*[8]. These datasets collected from OpenBCI and Muse headsets were termed as *EEG-IO* and *EEG-IM* (Table I), and were used to evaluate the performance of *Blink* on involuntary eye-blinks and different EEG headsets.

B. Unguided eye-blink experiments: We also conducted uncontrolled user experiments with 12 subjects for the OpenBCI device where subjects were asked to (i) watch a video, and (ii) read an article, each for 5 minutes. These datasets were termed as *EEG-VV* and *EEG-VR* (Table I). In unguided experiments, no external stimulations were provided. Other experimental and annotation methodologies were similar to the previous experiment. As the manual annotation process was time demanding, we annotated only first 200 seconds of the unguided data, to use it for evaluating *Blink* on voluntary eye-blinks and different user activities.

For all the collected datasets, ground truth, i.e., annotation was performed before evaluating the *Blink* algorithm to ensure the unbiased evaluation.

B. Blink *Algorithm Performance*

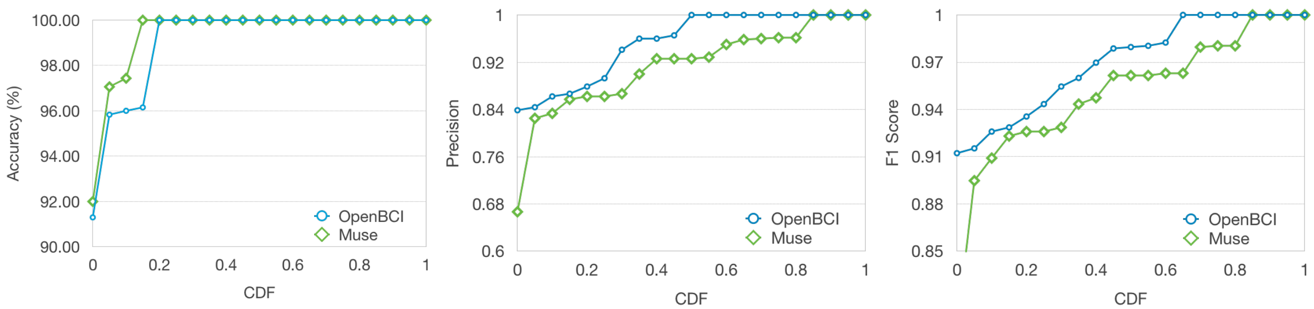
We evaluate the performance of *Blink* algorithm using three different metrics. *Accuracy* measures the percentage of correctly detected eye-blinks out of total given eye-blinks (true positives). *Precision* refers to the number of correctly detected eye-blinks out of the total detected eye-blinks. *F1 score* represents the harmonic mean of precision and recall. An ideal detection algorithm would perform with 100% accuracy, with precision and F1 score of 1 and 1 respectively

The collected EEG datasets were analyzed offline by implementing *Blink* algorithm (algorithm 1) in Python. We analyzed the results for two frontal channels (Fp1 and Fp2) whose results were combined in an OR fashion. We used a 4th order Butterworth low pass filter (*algorithm 1: line 1*) with a frequency of 10 Hz. The correlation threshold for computing highly correlated components ($high\_corr\_comp$, algorithm 1: line 12), was kept to 0.2 (low value), as to allow more potential eye-blinks for robust profiling.

1) Involuntary Eye-Blinks: We compute and present the detection performance of the *Blink* algorithm on involuntary eye-blinks (i.e., *EEG-IO* and *EEG-IM* dataset from Table I) in fig. 5 in the form of cumulative distribution for both platforms. The mean algorithm accuracy for all 20 subjects is near perfect (98.96% for OpenBCI, and 99.2% for Muse). The mean accuracy of (top-5, worst-5) subjects is (100%, 96.00%) for OpenBCI traces, and (100%, 97.2%) for Muse traces. The top-5 and worst-5 accuracies do not differ much, which validates the universality of the algorithm. Mean precision is above 0.9 for both the devices (0.951 for OpenBCI, 0.913 for Muse). Similar (top-5, worst-5) precision scores are (1.0, 0.858) for OpenBCI and (0.993, 0.801) for Muse. F1 score assigns a weighted score of accuracy and false positives. We received an average F1 score of 0.968 and 0.944 for OpenBCI and Muse, respectively, which confirms

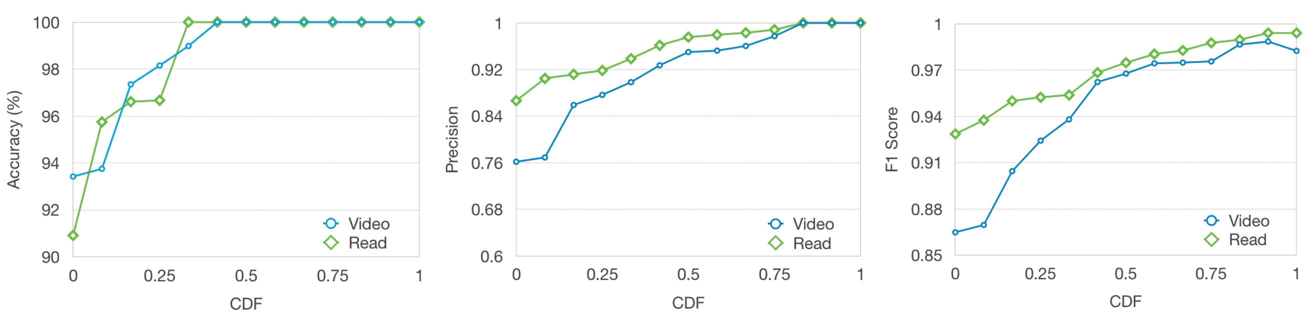
[7]OpenViBE software does not provide the drivers of Muse headset for collecting EEG directly from the headset and hence, we used the Muse Monitor application.

[8]We performed the manual labeling as we found from the video feed that subjects blinked their eyes even when the green plus was not shown on the screen



(a) Accuracy (b) Precision (c) F1 Score

Fig. 5: Detection performance results of *Blink* algorithm on involuntary eye-blinks



(a) Accuracy (b) Precision (c) F1 Score

Fig. 6: Detection performance results of *Blink* algorithm on voluntary eye-blinks

Dataset	Accuracy	Precision	F1 Score
EEG-IO	98.96 (± 2.32) %	0.950 (± 0.062)	0.968 (± 0.031)
EEG-IM	99.2 (± 1.92) %	0.913 (± 0.079)	0.944 (± 0.046)
EEG-VV	98.47 (± 2.44) %	0.922 (± 0.083)	0.950 (± 0.046)
EEG-VR	98.32 (± 2.86) %	0.952 (± 0.043)	0.967 (± 0.022)

TABLE II: A summary of *Blink* performance over collected datasets

the robustness of the algorithm. Moreover, the results for Muse and OpenBCI do not differ much, which validates the extensibility of the algorithm across other BCI platforms.

2) Voluntary Eye-Blinks: EEG-VV and *EEG-VR* datasets (Table I) were used to evaluate the performance of *Blink* algorithm on natural eye-blink patterns when users were watching a video or reading an article. Fig. 6 presents the performance of *Blink* to detect involuntary eye-blinks in the form of cumulative distribution for both user activities. Averaged over 12 subjects, we achieved an accuracy of 98.4% and 98.3% for video and read activities respectively. The corresponding average precision measures and F1 scores are (0.92, 0.94) for video, and (0.95, 0.96) for reading activity. The consistent performance of *Blink* on natural eye-blinks over different activities show the robust performance and applicability of *Blink* in practical uses.

A summary of the *Blink* performance is presented in Table II over the collected datasets.

C. Comparison of Blink *performance with related work*

1) Comparison with BLINKER[30]: For comparing the algorithm performance with BLINKER[30], we look at the mean of accuracy, false positive rate and F1 score for 7 subjects in *EEG-IO* dataset. BLINKER requires long EEG traces, and runs successfully only on the dataset from 7 subjects, hence we use 7 subjects out of 20 for result comparison in Table III). We can see the significant difference in eye-blink detection performance of *Blink* and BLINKER (Table III). *Blink* performs perfectly (100% mean accuracy, 0.952 precision), but BLINKER[30] performs 44.05% accurate with the precision of 0.558.

	Accuracy	Precision	F1 Score
Blink	100%	0.952	0.97
BLINKER[30]	44.05%	0.558	0.69

TABLE III: Performance comparison with BLINKER[30]

2) Comparison with the basic threshold approach: While we know that threshold-based comparison approaches are highly ineffective, a curious reader might be interested in the merits of the proposed algorithm. Hence, for the completeness, we implemented a naive statistical algorithm to detect eye-blinks (used frequently in EEG community [38], [39]) by comparing the signal variance (or standard deviation) with a threshold. For *EEG-IO* dataset of 20 subjects, the best threshold value was learned (which results in the highest F1 score), and the corresponding accuracy obtained was 6.83%, precision being 0.441 with an F1 score of 0.66.

3) Comparison with learning approaches: Having previously established the inadequacy of learning approaches to detect eye-blinks for our solution (requirement of user training), we compare the *Blink* performance with learning

	Accuracy	Precision	F1 Score
Blink	98.15%	0.951	0.96
SVM	46.49%	0.559	0.69
k-NN	67.82%	0.664	0.75
basic threshold	6.83%	0.441	0.66

TABLE IV: Performance comparison with learning approaches

Algorithm	Performance	Limitations
[19]	87.13%	requires training phase
[15]	97.0% TPR (for 10% FPR)	not fully automatic
[27]	99.9% sensitivity, 94.7% specificity	uses 6 EEG, 2 EOG
Blink	$\geq$ 98% accuracy	

TABLE V: Reported performance and limitations of the related work

approaches, namely (i) SVM [40], and (ii) k-NN (k-Nearest Neighbors) [41] to establish a baseline. For this comparison, we use a moving window of $0.5f_s$ length with a stride of $0.1f_s$ to bucket the features as *eye-blinks* and *no-blinks* based on the given labels. We split the *EEG-IO* dataset in 80:20 ratio for training and testing. For SVMs, the linear kernels were used, and the number of nearest neighbors was set to 5 for k-NNs. For SVMs, we receive an accuracy of 46.49%, precision of 0.559 and f1 score of 0.69. Similarly, for k-NNs, the obtained metrics are 67.82%, 0.664, and 0.75, respectively.

4) Reported performance comparison with the related work: After attempting to run codes released with previous works [30], [15], we concluded that every proposed algorithm is followed by process of optimizing the algorithm parameters on their collected dataset. Hence in Table V, we present the reported performance metrics of the selected related works (optimized on their collected dataset) along with their limitations and compare against the *Blink* performance. We can see that although [15] and [27] report comparable accuracies, they have limitations of not being fully automatic, or requiring multiple EEG and EOG electrodes respectively.

VI. Discussion

A. Towards an online algorithm

Blink algorithm is designed and presented as an offline algorithm. However, the *Blink* algorithm can be used as-is in an online fashion. Real-time eye-blink detection widens the applicability of such an approach in the domain of BCI based communication, control, neurogaming, etc., and real-time EEG data processing. By design, the *Blink* algorithm assumes the presence of a few (3+) similar eye-blinks in the EEG signal. Leveraging this fact, the proposed approach can be used as-is in an online manner by applying *Blink* algorithm on a moving window with sufficient length ($\geq$ 30 seconds)[9].

We intend to extend this work by exploring the feasibility of optimizing this algorithm to operate in an online, real-time manner without using the moving window approach with repetitive computations. One of the direction to extend this approach is to dynamically build the correlation matrix and improve cluster formation upon peak detection during the continuous real-time monitoring of the EEG data. It is a challenging task to allow *Blink* algorithm to detect eye-blinks without compromising the performance instantly, and will be studied in the future work.

[9]The average human eye-blink rate is 17 blinks/min in the rest condition [42]

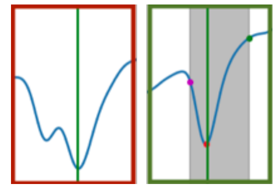


Fig. 7: Failure cases of *Blink* algorithm: (Left) abrupt eye-blink pattern not detected by *Blink* algorithm, (Right) the regular eye-blink pattern exhibited by the user

B. Limitations of the Blink algorithm

Despite the attractive performance score of *Blink* algorithm, *Blink* still fails to detect ~50 eye-blink samples out of 2300 eye-blinks. We analyzed the undetected eye-blinks and concluded that failure cases, although being quite low (<2%) are mostly caused by the invalidity of the assumption of consistent eye-blink patterns within a subject. Occasionally, an irregular eye-blink pattern was observed in the user data, which is quite dissimilar to the regular eye-blink pattern exhibited by the user. Fig.7 shows the cleaned irregular eye-blink pattern side by side to the regular eye-blink pattern. With the datasets collected in this work (Table I), we plan to statistically evaluate the assumption of consistency in eye-blink patterns and improve the *Blink* algorithm to consider such cases.

VII. Conclusions

In this work, we study the problem of eye-blink detection in EEG signals. In our literature review, we find that regardless the abundance of research in this area, the applicability of the proposed algorithms is limited due to one or more requirements of multiple EEG channels, EOG channels, user-training phase and manual inspection for the robust detection. In this context, we propose a fully automated unsupervised algorithm, *Blink*, to detect eye-blinks in the EEG data. Our approach self-learns brainwave profiles for each specific user's eye-blinks, and hence does away with any user training or manual inspection requirements. *Blink* capable of functioning on a single channel EEG accurately, estimates the start and end timestamps of eye-blinks very precisely. We collected four different EEG datasets to evaluate the robustness of algorithm across various EEG headsets, user activities, and eye-blink types, and show that *Blink* performs with an accuracy of over 98% in all cases along with an average precision of 0.934.

VIII. Acknowledgements

This work was supported in part by the National Science Foundation under grants 1701115, 1813242, and 1837369, and the Wayne J. Holman Endowed Chair.

References

[1] P. Berg and M. Scherg, "Dipole models of eye movements and blinks," *Electroencephalography and clinical neurophysiology*, vol. 79, no. 1, pp. 36–44, 1991.

[2] P. Berg, "The residual after correcting event-related potentials for blink artifacts," *Psychophysiology*, vol. 23, no. 3, pp. 354–364, 1986.

[3] L. Shoker, S. Sanei, and J. Chambers, "Artifact removal from electroencephalograms using a hybrid bss-svm algorithm," *IEEE Signal Processing Letters*, vol. 12, no. 10, pp. 721–724, 2005.

[4] B. Noureddin, P. D. Lawrence, and G. E. Birch, "Online removal of eye movement and blink eeg artifacts using a high-speed eye tracker," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 8, pp. 2103–2110, 2011.

[5] Y. Li, Z. Ma, W. Lu, and Y. Li, "Automatic removal of the eye blink artifact from eeg using an ica-based template matching approach," *Physiological measurement*, vol. 27, no. 4, p. 425, 2006.

[6] S. Hoffmann and M. Falkenstein, "The correction of eye blink artefacts in the eeg: a comparison of two prominent methods," *PLoS One*, vol. 3, no. 8, p. e3004, 2008.

[7] C. A. Joyce, I. F. Gorodnitsky, and M. Kutas, "Automatic removal of eye movement and blink artifacts from eeg data using blind component separation," *Psychophysiology*, vol. 41, no. 2, pp. 313–325, 2004.

[8] D.-h. Zhu, J.-j. Tong, and Y.-q. Chen, "An ica-based method for automatic eye blink artifact correction in multi-channel eeg," in *Information Technology and Applications in Biomedicine, 2008. ITAB 2008. International Conference on*. IEEE, 2008, pp. 338–341.

[9] T. Jung, C. Humphries, T.-W. Lee, S. Makeig, and M. McKeown, "Extended lca removes artifacts from electroencephalographic recordings," *Adv. Neural Inform. Process. Syst. 10*, vol. 10, pp. 894–900, 1998.

[10] F. C. Viola, J. Thorne, B. Edmonds, T. Schneider, T. Eichele, and S. Debener, "Semi-automatic identification of independent components representing eeg artifact," *Clinical Neurophysiology*, vol. 120, no. 5, pp. 868–877, 2009.

[11] N. Bigdely-Shamlo, K. Kreutz-Delgado, C. Kothe, and S. Makeig, "Eyecatch: Data-mining over half a million eeg independent components to construct a fully-automated eye-component detector," in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. IEEE, 2013, pp. 5845–5848.

[12] H. Nolan, R. Whelan, and R. B. Reilly, "Faster: fully automated statistical thresholding for eeg artifact rejection," *Journal of neuroscience methods*, vol. 192, no. 1, pp. 152–162, 2010.

[13] A. Klein and W. Skrandies, "A reliable statistical method to detect eyeblink-artefacts from electroencephalogram data only," *Brain topography*, vol. 26, no. 4, pp. 558–568, 2013.

[14] A. Egambaram, N. Badruddin, V. S. Asirvadam, E. Fauvet, C. Stolz, and T. Begum, "Unsupervised eye blink artifact identification in electroencephalogram," in *TENCON 2018-2018 IEEE Region 10 Conference*. IEEE, 2018, pp. 2148–2152.

[15] W.-D. Chang, H.-S. Cha, K. Kim, and C.-H. Im, "Detection of eye blink artifacts from single prefrontal channel electroencephalogram," *Computer methods and programs in biomedicine*, vol. 124, pp. 19–30, 2016.

[16] S. Maruthachalam, M. Kumar, and H. Murthy, "Time warping solutions for classifying artifacts in eeg," 07 2019.

[17] M. bin Abd Rani *et al.*, "Detection of eye blinks from eeg signals for home lighting system activation," in *2009 6th International Symposium on Mechatronics and its Applications*, 2009.

[18] A. Delorme, T. Sejnowski, and S. Makeig, "Enhanced detection of artifacts in eeg data using higher-order statistics and independent component analysis," *Neuroimage*, vol. 34, no. 4, pp. 1443–1449, 2007.

[19] W.-D. Chang and C.-H. Im, "Enhanced template matching using dynamic positional warping for identification of specific patterns in electroencephalogram," *Journal of Applied Mathematics*, vol. 2014, 2014.

[20] A. Aarabi, K. Kazemi, R. Grebe, H. A. Moghaddam, and F. Wallois, "Detection of eeg transients in neonates and older children using a system based on dynamic time-warping template matching and spatial dipole clustering," *Neuroimage*, vol. 48, no. 1, pp. 50–62, 2009.

[21] S. Kim and J. McNames, "Automatic spike detection based on adaptive template matching for extracellular neural recordings," *Journal of neuroscience methods*, vol. 165, no. 2, pp. 165–174, 2007.

[22] M. Nakanishi, Y. Mitsukura, Y. Wang, Y.-T. Wang, and T.-P. Jung, "Online voluntary eye blink detection using electrooculogram," in *International Symposium on Nonlinear Theory and Its Applications, Palma, Mallorca, Spain*, 2012.

[23] R. Ghosh, N. Sinha, and S. K. Biswas, "Automated eye blink artefact removal from eeg using support vector machine and autoencoder," *IET Signal Processing*, vol. 13, no. 2, pp. 141–148, 2018.

[24] S. Rihana, P. Damien, and T. Moujaess, "Eeg-eye blink detection system for brain computer interface," in *Converging Clinical and Engineering Research on Neurorehabilitation*. Springer, 2013, pp. 603–608.

[25] R. Paprocki, T. Gebrehiwot, M. Gradinscak, and A. Lenskiy, "Extracting blink rate variability from eeg signals," *arXiv preprint arXiv:1603.03031*, 2016.

[26] Z. Tiganj, M. Mboup, C. Pouzat, and L. Belkoura, "An algebraic method for eye blink artifacts detection in single channel eeg recordings," in *17th International Conference on Biomagnetism Advances in Biomagnetism–Biomag2010*. Springer, 2010, pp. 175–178.

[27] H. Ghandeharion and A. Erfanian, "A fully automatic ocular artifact suppression from eeg data using higher order statistics: Improved performance by wavelet analysis," *Medical engineering & physics*, vol. 32, no. 7, pp. 720–729, 2010.

[28] M. Rohál'ová, P. Sykacek, M. Koskaand, and G. Dorffner, "Detection of the eeg artifacts by the means of the (extended) kalman filter," *Meas. Sci. Rev*, vol. 1, no. 1, pp. 59–62, 2001.

[29] O. Dehzangi, A. Melville, and M. Taherisadr, "Automatic eeg blink detection using dynamic time warping score clustering," in *Advances in Body Area Networks I*. Springer, 2019, pp. 49–60.

[30] K. Kleifges, N. Bigdely-Shamlo, S. E. Kerick, and K. A. Robbins, "Blinker: Automated extraction of ocular indices from eeg enabling large-scale analysis," *Frontiers in neuroscience*, vol. 11, 2017.

[31] G. Gratton, M. G. Coles, and E. Donchin, "A new method for off-line removal of ocular artifact," *Electroencephalography and clinical neurophysiology*, vol. 55, no. 4, pp. 468–484, 1983.

[32] R. J. Croft and R. J. Barry, "Eog correction: Which regression should we use?" *Psychophysiology*, vol. 37, no. 1, pp. 123–125, 2000.

[33] ——, "Removal of ocular artifact from the eeg: a review," *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 30, no. 1, pp. 5–19, 2000.

[34] "Biopac cap-100c," https://www.biopac.com/product/eeg-caps-for-cap100c/, 2019.

[35] Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, O. Bertrand, and A. Lécuyer, "Openvibe: An open-source software platform to design, test, and use brain–computer interfaces in real and virtual environments," *Presence: teleoperators and virtual environments*, vol. 19, no. 1, pp. 35–53, 2010.

[36] "Interaxon muse headband," http://www.choosemuse.com, 2019.

[37] "Muse monitor application," https://musemonitor.com/, 2019.

[38] (2019) Automated eye blink detection online. [Online]. Available: http://openbci.com/community/automated-eye-blink-detection-online-2/

[39] S. Hoffmann and M. Falkenstein, "The correction of eye blink artefacts in the eeg: a comparison of two prominent methods," *PLoS One*, vol. 3, no. 8, p. e3004, 2008.

[40] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[41] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.

[42] A. R. Bentivoglio, S. B. Bressman, E. Cassetta, D. Carretta, P. Tonali, and A. Albanese, "Analysis of blink rate patterns in normal subjects," *Movement Disorders*, vol. 12, no. 6, pp. 1028–1034, 1997.