**Blink to Get In: Biometric Authentication for Mobile Devices using EEG Signals**

Ekansh Gupta, Mohit Agarwal, Raghupathy Sivakumar  
Georgia Institute of Technology, Atlanta, Georgia  
Email: \{egupta8,magarwal37,siva\}@gatech.edu

**Abstract**—Biometric authentication is becoming popular in a varied range of applications because of its unique specificity for an individual user. In this context, electro-encephalogram (EEG) signals from a user is an interesting candidate for authentication. In this work, we specifically focus on the EEG signal corresponding to the human eye-blink to create an authentication system that could be used to distinguish between multiple users accurately and efficiently while also being burden-less and convenient to the users. We use a dataset of eye-blink related EEG signals, collected from 20 users, to study our solution. Our results show that blink signals can be used for accurately distinguishing between different users and hence can be used for authentication.

**Index Terms**—EEG, blink, authentication, biometric

I. INTRODUCTION

Biometric authentication has fast evolved to be the default authentication mechanism on smartphones and other mobile devices. Apple’s reported statistics show that 89% of its users have a Touch ID enabled smartphone use the Touch ID [1]. There are distinct advantages to using biometrics, including the fact that biometrics are much harder to manipulate and that the burden on the user is very light unlike in password-only authentication where the user is expected to remember long and obfuscated passwords. With users facing an authentication challenge dozens of times in a single day [1], there is a distinct need for an approach that is both lightweight in terms of user burden and strong in terms of security. There are two types of attacks that authentication challenges protect against. The first is against a casual attack, where someone randomly picks up the mobile device and tries to use the device. Current biometric authentication approaches like Touch ID and Face ID are reasonably secure against such casual attacks. However, existing approaches have a bigger vulnerability to targeted attacks. In theory, an attacker can rely on a high-resolution photograph of the user’s fingerprint to compromise Touch ID in a matter of minutes [2], [3]. While Face ID is a much newer biometric authentication mechanism, there already have been successful attempts to compromise certain aspects of it [4]. One of the drawbacks of such morphological biometric solutions is that the biometric template used for the authentication is static and hence any means to get access to that template is sufficient to compromise the authentication process. Thus, there is considerable motivation to continue to develop newer and safer biometric authentication solutions. There are other biometric solutions such as those that rely on the user’s voice [5] where the authentication challenge can be a randomized prompt thus making it difficult to compromise. However, voice biometric solutions have some obvious limitations such as the voice of the user changing because of a cold, etc. [6]. Another class of biometrics is one that relies on physiological data of the user rather than simply morphological data. Extreme examples of physiological data include DNA or saliva composition. While these are more robust in terms of security, they have a high cost of implementation both during initial setup, and for every authentication verification. In this paper, we consider a more accessible physiological data for a user – the user’s electroencephalogram (EEG) data for a specific action – blinking. With EEG growing to be a bonafide input modality in several commercial applications such as healthcare, gaming, and wellness [7], [8], and the consequent wider availability of EEG headsets off-the-shelf, access to a user’s EEG data is easier than it has ever been. At the same time, it is shown that blinks are actions for which the EEG signals are strongly identifiable [9].

Thus, the key question we answer in this paper is the following: Can the user’s EEG signals, captured when the user blinks, be used as accurate and secure biometric authentication data? We answer this question by relying on a dataset of EEG signals collected through controlled experiments [9] with twenty users where the users are prompted to blink, and the corresponding EEG signals are captured through a commercial On-The-Shelf (OTS) EEG headset. We show that a naive approach that relies only on simple features of the blink signal is not accurate enough. We then present a set of systematic strategies to improve the features and show that it is indeed possible to devise an effective authentication solution that relies on a user’s EEG signals captured when the user is blinking.

We use the dataset to evaluate the algorithm, and show that the solution has an accuracy of about 92%. The rest of the paper is organized as follows:

In section II, we briefly cover our motivation behind using EEG and blink signals and provide a mathematical definition of the problem. In section III, we describe the data collection methodology and explain the eye-blink anatomy on EEG. In section IV, we discuss our methodology in detail and evaluate the system along with comparing it against related works. In section V, we summarize the related work in this domain, and finally conclude in section VI.
II. BACKGROUND AND PROBLEM DEFINITION

A. Why biometric and approaches

Biometric authentication is based on the physiological (or biological) or behavioral characteristics of an individual. Biometric approaches are more secure, and are harder to fake over the traditional authentication approaches like passwords or smart-cards. Moreover, the accessibility and ease of use give these methods an extra edge over password-based approaches. Today, facial recognition, voice recognition, fingerprints, and iris tracking are popular authentication technologies, widely used across consumer devices for identification and personal verification. In these methods, a unique template of the user biometric is stored locally on a device (e.g., the mathematical representation of a fingerprint measurement), and is compared against the measurement obtained when someone is attempting to unlock the device. If they are found to be nearly identical, device access is granted to the user.

B. Why EEG and Blink

Despite the promise and ubiquity of currently popular biometric approaches (particularly fingerprint and face recognition), these systems are not foolproof and are shown to be vulnerable. Face recognition based authentication systems can easily be falsified using artificially printed 3D masks [10], [11]. Fingerprint systems are prone to security leaks based on artificial or gummy fingerprints [12]. Given the exposure of these traits (i.e., face pictures, touch prints) to the external world, it is easy to forge and steal the biometric traits of an individual user, e.g., face from social media pictures, and fingerprint from the objects that a user touches. These vulnerabilities motivate the design of a novel biometric authentication system which is unique for the users, and also much harder to clone or fake.

EEG (Electroencephalography) is the measurement of the electrical activity of the brain, captured from the outer surface of the scalp using metal electrodes. Inside the human brain, billions of neurons communicate with each other through electrical impulses, resulting in the residual EEG on the scalp. This neuronal firing pattern captured through EEG is known to be unique [13], and can be used as novel information for biometric-based authentication. Individual differences in human eye-blinking patterns are studied in terms of rate, patterns, frequency, strength, etc [14]–[16]. Eye-blink waveforms on EEG present a very high variability across users [9]. The anecdotal evidence is obtained for the feasibility of developing an eye-blink based authentication system [17]. This modality holds the promise of providing a fast and user-friendly experience to identify and authenticate the users.

C. Discussions on limitations of EEG

One of the major challenges of using EEG as an authentication mechanism is the stability of these signals. The physiological or psychological states can have a significant impact on the EEG. EEG signals in states such as fatigue, feeling angry or upset, may not match the unique EEG template of the user, created while training, and hence would reduce the True Positive Rate (TPR) of the system, essentially restricting the user to access their device. It demands and motivates the research in understanding the variability of EEG under different mental states, and enabling the authentication systems with robustness against such physiological and psychological states. In the current timeline, another limitation with EEG-based authentication is that the system requires the user to wear an EEG wearable headset. Today, biometric sensors are embedded in mobile and computing devices, enabling secure authentication without any external hardware requirements. However, in recent years, these devices have become commercially relevant for day-to-day applications, including education, gaming\textsuperscript{1}, self-regulation\textsuperscript{2} and entertainment\textsuperscript{3}. It is expected that in a decade time-frame, EEG wearables are going to be ubiquitous and will augment the current communication devices.

D. Problem definition and key assumptions

In this work, we consider $N$ users, $u_1, u_2, \cdots, u_N$. Our goal is to develop a system $S$, such that the local copy of $S$ on $i^{th}$ user device, i.e., $S_i$, gives access to only blinks of user $u_i$ and restricts all other users to access the device through their blinks. Hence, for an ideal authentication system design $S$, the below should hold,

$$S_i(u_i) = 1, \forall i \in [1, N]$$

$$S_i(u_j) = 0, \forall i, j \in [1, N], i \neq j$$

In our work, we make the following assumptions while a user is trying to authenticate the system with blinks:

- We assume that the electrode-cap placement for each user is consistent across trials.
- We also assume the consistency of the physiological and psychological state of the user. E.g., the user is not involved in mental-strenuous tasks or is not physically moving her head or facial muscles.

III. DATA COLLECTION METHODOLOGY AND DATASET

To study and characterize the individual differences of blink patterns of users, we have used the EEG-10 dataset collected in our previous work [9]. In EEG-IO, a total of 20 subjects were recruited in the age range of 22 to 30 years old following the approval of the Institutional Review Board (IRB). Subjects were asked to sit in front of a computer screen and wear an electrode-cap (BIOPAC CAP100C was used). We used the electrode gel to establish contact between Fp1, Fp2 electrodes with the scalp. Two additional ear electrodes were used to serve as a reference and noise removal. The electrode cap was further attached to the OpenBCI board [18], sampling the raw EEG signals at 250Hz. The OpenBCI device transmitted the sampled EEG to a computer device over the wireless channel.

Subjects were asked to perform a single eye-blink when presented with a green-cross on the screen. A total of 25 such external stimulations were presented for each subject every 1

\textsuperscript{1}Neurosky Mindwave

\textsuperscript{2}Interaxon Muse headset

\textsuperscript{3}Emotiv EPOC/EPOC+
Figure 1: Eye-Blink EEG profile

3-4s (depending on the subject’s preference). We used the Blink algorithm to extract the eye-blink signatures from the continuous EEG signal [9].

**Eye-Blink profile on EEG:** The act of eye-blinking distorts the electric field around the eyes (due to opposite polarities of the cornea and the retina), and interferes with the EEG signals on the frontal electrodes (mainly Fp1 and Fp2 according to a 10-20 electrode system). This results in a trough-shaped pattern on the EEG captured from the frontal electrodes. The shape of a typical blink signal waveform is shown in Fig. 14. The manifestation of an eye-blink on EEG is highly asymmetric in time. The eye-blinks can be divided into closing and opening of eyes. During a blink, eye-lid closure is generally of short duration resulting in an increased slope in the EEG signal. The later part, i.e., eye-lid opening, is comparatively slower and decelerates before recovering to the full amplitude. The blink slope is directly correlated with the velocity of eye-lid movements while closing or opening the eyes. The strength of the eye-blink is reflected as the amplitude of the signal. Blink duration is the total time taken by the human during the blinking process.

**IV. BLINK BASED AUTHENTICATION**

In this section, we explore the variations of the physiological behavior of the users while the user is performing an eye-blink and its manifestation on the blink profile on EEG. We explore such variations with the goal of handcrafting features that can help us distinguish between users based on their eye-blink patterns.

**A. Pre-processing**

We relied on the Blink algorithm in [9] to extract the eye-blink signatures from the continuous EEG dataset. [19] used a threshold-based method to detect peaks for eye-blink detection, posing vulnerability to other non-blink artifacts, e.g., eye or head movements, etc. In [9], authors have proposed a robust eye-blink detection algorithm that learns the blink template in an unsupervised manner.

We removed the high-frequency components from the EEG blink data by passing it through a low-pass filter. All the frequency components above 10 Hz were discarded. We also manually reviewed the blink patterns of users and removed the data for four users from any further evaluations. These users had very noisy and distorted blink waveforms (as shown in Fig. 2), which could be due to a lot of movement during the experiment, or improper placement of the electrode cap.

**B. Naive features**

[19] performed blink-based authentication with an accuracy of 97.3% on their collected dataset. However, we could not access the dataset collected by the authors. Hence, we used the features described in [19] on our dataset and used them to classify the 16 users based on their eye-blink signals. Specifically, we consider the mean, variance, slope, energy, area, amplitude and duration of the signal.

We implemented the problem in the form of a multi-class classification. A Support-Vector Machine (SVM) with a Radial Basis Function (RBF) kernel was used as the classifier to train this multi-class data. Each of the 16 users contributed about 50 blinks (taken from 2 channels for approximately 25 blinks each) taking the total number of blinks to 800. 80% of each user’s blinks were used for training the multi-class classifier and the remaining 20% of each user’s blinks were used for testing. The training data was then passed through a Principal Component Analysis (PCA) block to extract a combined set of features that would retain 90% of the total variance of all the features. This classification (averaged over 5 different trials) performed with a mean True Positive Rate (TPR) of 53% with the minimum average accuracy for a user going as low as 27%. This can be explained by the distribution of the blink data with respect to these features. Fig. 3 (a),(b),(c) shows that the extracted features are non-separable for users when we consider (a) mean, (b) variance, and (c) peak, duration and area, as the features. Here, each colored cluster of points represents blink signals from a specific user.

We also extended the blink signal into its left and right neighborhoods to gauge if user-separability can be achieved with features extracted from the neighboring signals. We considered Fourier transform and energy bins as the additional features, but the mean TPR remained at 53%. Fig. 3 (d) presents the Fourier bins (summed for three intervals). It can easily be seen that Fourier features also do not present separability across users.

**C. Features capturing finer variations in a blink**

The inferior performance of the features discussed above is due to the simplifying assumptions made on the blink patterns. The features discussed in the above sections, consider eye-blink as an atomic process and computes features based on
central tendency measures, or summary statistics. This fails to capture some of the user-specific variations that happen only during a brief duration. For e.g., during the onset of a blink, the waveform could dip quickly and then slowly reach the minimum or vice versa. This could also be true for a brief duration during the offset. 2 blinks may have the same duration yet one might be relatively flatter or sharper than the other. [20] studied human blinks data and established the fine granularity of temporal and spatial characteristics of human blinks.

Motivated from this, we attempt to capture these features enabling the user separation. Since these details deal with the distribution of points within limited and specific subsets of the entire blink signal, they can be visualized and described using histograms. For e.g., a flatter blink signal will have more data points with values closer to the peak than a sharper blink. Similarly, a slow rising and a slow dropping blink can be differentiated based on a histogram calculated using their slope (single derivative). A blink signal which shoots up fast from the minimum and then rises slowly will have the same average slope as the blink signal which starts slowly from the minimum but becomes steep going forward. However, as shown in Fig. 4, there will be a significant difference between the histogram profiles of the two blinks, showing a more pronounced left extreme in the value histogram for the blink that starts slowly from the minimum.

Based on these intuitions, we incorporated the value-histogram (histogram calculated on the blink signal values) and slope-histogram (histogram calculated on the single derivative of the blink signal) and used their bins (range of values for which the frequency of values is calculated) as additional feature vectors. We see that this increases our True Positive Rate (TPR) of single blink detection to 71%. A more detailed comparison with [19] is presented in Fig. 5 with respect to True Positive Rate and False Positive Rate per user. As seen in the figure, the users show significantly better TPR and FPR (false positive rate) compared to the algorithm in [19]. Our average FPR was 2.03% as opposed to 3.2% of [19] while their average TPR was 52.4% compared to our average TPR of 71%.

D. Multiplicity

The slope and value histograms boosted our TPR to 71%. However, to achieve a frustration-free and usable system, the TPR should at least be 90%. Redundancy can help increase the reliability of a system. To reduce the prediction error, we rely on the multiplicity of blinks, i.e. we would bundle $k$—blinks, and the user would have to blink $k$—times to access the system. Based on the number of blinks (i.e., $k$) we would bundle, each test sample would comprise $k$—blinks. Each blink in this set would be separately used for evaluating the probability vector of the blink belonging to a particular user and then a summation of these probability vectors would be used to decide whether the whole set (of $k$— blinks) belongs to a particular user. There have been similar attempts to use multiplicity to increase the accuracy of blink detection. [19] generates a test sample after averaging 25 blinks from a user. While it fetches a TPR of about 96%, not only is it burdensome and extremely undesirable for a user to blink 25 times to gain access into a system, we show here that our approach can beat this score in as low as 5 blinks.

However, there is a trade-off between the TPR and user convenience, when we increase the required number of blinks by adding redundancy in the system. We ran experiments for multiple values of $k$ and show in Fig. 6 that TPR increases rapidly for a multiplicity value of 2 and 3 after which it starts to taper off and converge. In an independent study performed in [21] to calculate user comfort score for multiple blinks, 3-blinks were rated as comfortable by the users. The confusion matrix for this case is shown in Fig. 7 and the user-specific values
for true positive and false positive percentages are shown in Fig. 8. With our proposed algorithm, the system achieved an aggregate TPR of 92% with 3-blinks, with an aggregate per-user False Positive Rate (FPR) of 0.7%. Hence, we conclude that 3-blinks based authentication is comfortable while being reliable and relatively convenient for a user to manage. With an aggregate accuracy of 92% with 3 blinks, we think this is a sweet spot that can be used.

E. A purely local approximation

In the previous subsection, we’ve achieved the system TPR of 92%. However, the implementation of the proposed approach requires the training data of all users to be stored in a cloud.

This approach explicitly demands that the system needs to be re-trained whenever a new user is added, which is computationally expensive and not scalable from a practical standpoint.

Another interesting and competing approach would be to have a system authenticating a user using only the user’s data. Through this approach, a local copy of the trained classifier weights (trained solely on corresponding user blinks), can be stored locally on the user device to allow the authentication. The local approximation system would be desirable due to its massively reduced computational costs and will also ensure user-privacy. However, as the classification algorithm (i.e. one-class SVM in this case) is not exposed to the blinks of other users, it is less accurate.

We tested our algorithm on a one-class SVM model to see its performance. We tested this using two modes. In the first approach, we train a classifier per-user using 80% of the blink signals for that user and tested it one blink at a time. In the second approach, we bundled multiple blinks during the testing phase and combined their individual predictions using a hard voting mechanism. Each blink within the sample test case set was classified using the one-class SVM and a decision on the whole set was taken based on the class (whether it is an outlier or not) the majority of the blinks were classified into. We used 3, 5 and 7 blinks to test the aforementioned bundling approach and the results are presented in Fig. 9.

The TPR for the one blink one-class case was 60% with an average per-user false positive rate (this estimates the probability of an average user to be able to break into a legitimate user’s system) of about 5.79%. While the true positive accuracy in the cases for 3 and 5 blinks is roughly 73% which improves to about 80% for 7 blinks, the main advantage is seen as the average per-user false positive rate goes down to 4.04% in the 3 blink system, 2.7% in the 5-blink system and to 2.2% in the
One-class with 3-blinks

relied on EOG recordings (placing electrodes around the eye
mean false rejected rate (FRR) of 3.87% on 40 subjects. [24]
99.4% (1.5% improvement over single-level system). [23] com-
combined ERPs obtained through Rapid Serial Visual Presentation
(VEPs - visually evoked potentials) to boost the accuracy to
upon this work and combines eye-blinks based authentication
are the signals with maximum peaks, and detects eye-blinks
on each subject, with 8-12 natural eye blinks in each trial
5
using eye-blink waveforms collected through the Neurosky
make a decision.

This shows that an unsupervised approach for blink-based
authentication can be realized by combining a set of blinks to
make a decision.

V. RELATED WORK

[19] proposes a novel biometric authentication system
using eye-blink waveforms collected through the Neurosky
Mindwave EEG headset on Fp1 electrode. For 25 subjects,
[19] achieved the identification accuracy as 97.3% and error
rate as 3.7%. Here, data collection is performed with 6-8 trials
on each subject, with 8-12 natural eye blinks in each trial
(20-second duration for each trial). It assumes that eye blinks
are the signals with maximum peaks, and detects eye-blinks
based on that, and hence, prone to any other EMG based
artifacts. It also averages 25 user blinks to generate a test
sample which would be very burdensome to a user. [22] builds
upon this work and combines eye-blinks based authentication
with EEG signals during relaxation and visual stimulation
(VEPs - visually evoked potentials) to boost the accuracy to
99.4% (1.5% improvement over single-level system). [23] com-
bined ERPs obtained through Rapid Serial Visual Presentation
(RSVI) with eye-blinks to increase accuracy from 92.4% to
97.6% with a mean false accepted rate (FAR) of 3.90% and a
mean false rejected rate (FRR) of 3.87% on 40 subjects. [24]
relied on EOG recordings (placing electrodes around the eye
corners) and eye-movements (saccades) to authenticate users.
The system was tested on 40 users with accuracy ranging from
96% to 100% across users. Similarly, [25] also uses the EOG
based saccades detection to achieve 90% to 100% accuracy
across 30 subjects. In both works, EOG signals were recorded
from users while following a moving target with their eyes
producing rapid vertical or horizontal eye movements known
as saccades.

VI. CONCLUSIONS

In this paper, we have shown that an efficient and accurate
blink-based authentication method can be developed using
features that capture granular differences in user blinks, as
opposed to the central tendency measures or summary statistics.
We show that such a system can either be a cloud-based
infrastructure that uses the data of multiple users or it could
also operate in an unsupervised manner while only using the
concerned user’s data. Our work performs on a multi-class
classification while combining 3 blinks with a TPR of 92% and
an average per-user FPR of 0.7%. The performance for
the unsupervised classification yields a TPR of 80% and an
average per-user FPR of 2.2%. We plan to extend the future
work in two main directions - (i) consider a more diverse set
of features to improve the TPR while reducing the FPR of the
system (ii) thorough testing of the system for a broader set of
users, with multiple trials, and across different environmental
conditions and mental states.

REFERENCES

id/
biometric-technology.
achilles-heel/147109/.
[6] https://www.sans.org/reading-room/whitespapers/authentication/exploration-
voice-biometrics-1436.
[7] Marijn van Vliet, Arne Robben, Nikolay Chumerin, Nikolay V
Manyakov, Adrien Combaz, and Marc M Van Hulle. Designing a
brain-computer interface controlled video-game using consumer grade
eeg hardware. In 2012 ISNNIP Biosignals and Biobotics Conference:
Biosignals and Robotics for Better and Safer Living (BRC), pages 1–6.
in hardware and software for brain–computer interfaces (bci).
unsupervised algorithm for eye-blink detection in eeg signals. In 2019 57th
Annual Allerton Conference on Communication, Control, and
[10] Javier Galbally and Riccardo Satta. Three-dimensional and two-and-
a-half-dimensional face recognition spoofing using three-dimensional
3d masks. IEEE transactions on information forensics and security,
Hoshino. Impact of artificial “gummy” fingers on fingerprint systems.
In Optical Security and Counterfeit Deterrence Techniques IV, volume
4677, pages 275–289. International Society for Optics and Photonics,
2002.

7-blink system. This progression of the FPR getting lower can
be observed in Fig. 9 where, as we increase the number of
blinks in the test set, we see that the non-diagonal elements of
the matrix (which correspond to misclassifications) reducing in
number and also getting darker (indicating a decrease in their
percentage occurrence).

Figure 9: Bundling blinks in one class classifier

(a) One-class with 1-blink
(b) One-class with 3-blinks
(c) One-class with 5-blinks
(d) One-class with 7-blinks

We use Fp1 and Fp2 for our data collection.


