2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Cost Effective Real-time System for cognitive computing using Personalized Eye Blink Detection from Camera

Rahul Dasharath Gavas, Somnath Karmakar, Debatri Chatterjee, Ramesh Kumar Ramakrishnan, Arpan Pal
TCS Research, India

***Abstract*— Eye blink is indicative of various mental states. Generally, vision based approaches are used for detecting eye blinks. However, performance of such approaches varies across participants. Standard eye tracker or eye glasses used for detecting blinks, are very costly. Here, we are proposing a personalized vision based eye blink detector system. Proposed approach is ubiquitous and unobtrusive in nature and can be implemented using standard webcams/mobile camera, making it deployable for real world scenarios. Our approach has been validated on a set of data collected from our lab and on an open data set. Results show that in both cases, our system performs well for various conditions like natural/artificial light, with or without spectacles. We achieved a F_{score} of 0.98 for own collected data and 0.91 for open dataset, which outperform state of the art approaches.**

I. Introduction

Eye tracking is widely used to understand various cognitive processes [1]. Various features like gaze tracking [2], pupillometry [3], [4] and blinks [5] have been used in cognitive science and health care applications [6]. Eye blinks are comparatively easier modality among the three and hence, is found to be most widely used one. Eye blink is the rapid closing and reopening of the eyelid. Apart from its biological importance, blinks are correlated with attention [5], stress or fatigue [7], mental workload [8] and are used for building systems that allows the user to control computer applications [9], driver drowsiness detection [10], assessing progression of diseases like Parkinson's disease [11], detecting DeepFakes [12], to avoid dry eyes and computer vision syndrome [7], [13]. For such real world applications, accurate detection of eye blinks in real time is absolutely necessary.

There are active and passive devices that are used to detect blinks. Active devices are special hardware like eye glasses [14], infrared cameras [15]. These devices are often costly and hence not suitable for mass deployment. Moreover, these devices are intrusive in nature and might not be suitable for prolonged use. Passive systems use devices like webcams and mobile cameras. The underlying technologies used for active and passive devices are typically ultrasound [16], electroencephalography [17], infrared light [18], electromyograph (EMG) [19], electro-oculogram (EOG) [20], and video signals [21]. Usage of standard video camera for detection of eye blinks makes the system ubiquitous, unobtrusiveness and cost effective. Several techniques are used in literature to detect blinks from a video. However, the accuracy of detection are significantly affected by head orientation, resolution of images, motion dynamics, usage of spectacles, face illumination and ambient light. The performance of such systems largely depend on accurate identification of eye regions. Due to change in shape and position of eyes, the performance also varies a lot across individuals. Here, we are proposing a personalized camera based blink detection system having significantly higher accuracy across participants. Initial one minute of video of each user is used to derive person specific blink features which are then used for blink detection in subsequent sessions. The main contributions of our work are:

- designing a personalized camera based eye blink detection system
- a system for improved eye blink detection in real world scenarios.
- pave way for cost effective, unobtrusive and ubiquitous approach to detect cognitive/mental states

We have validated our approach on our own collected dataset as well as on a publicly available dataset. Results show that our system performs better compared to state of the art approaches and gives better F_{score} across participants. We have also analyzed the performance of our system for conditions like artificial/natural light, with/without spectacles and natural/simulated blinks. For all these scenarios, our system performs equally well and hence can be used in any real world scenarios for detecting blinks. Moreover, the system does not require any specialized hardware hence is mass deployable.

II. Related work

Vision based blink detection approaches can be classified into *sequential methods*, *appearance based methods* and *facial landmark based* approaches. In *sequential methods*, face and eyes are detected using Viola-Jones algorithm [22] or HAAR classifiers [23]. Post face detection, the motion in the eye region is estimated using sparse tracking [7], [24], or by computing frame to frame intensity differences and using thresholds to identify blinks [25], [26]. *Appearance-based* methods detect blinks from an image using active shape models [27] or by using templates of open/close eyes [28] and fitting of models to detect the eye lids [29]. Accuracy of these approaches are significantly affected by head orientation, image resolution, usage of spectacles, face illumination and ambient light. In *facial landmark based* detection [30], unique locations of face (e.g eye corner) are detected. Initially, researchers found this approach to be challenging. Recently, there have been significant improvements in this area and as a result, more robust facial landmark detectors are available now [31]. However, their performance is not consistent across individuals.

III. METHODOLOGY

A. Proposed eye blink detection system

There are three types of blinks- spontaneous/ natural blinks, reflexive blinks occurring due to external stimulus like air and voluntary blinks resulting from intentional eye closing. We have focused mainly on detection of spontaneous and voluntary eye blinks under various external conditions. Main components of our personalized eye blink detector application are shown in Fig. 1.

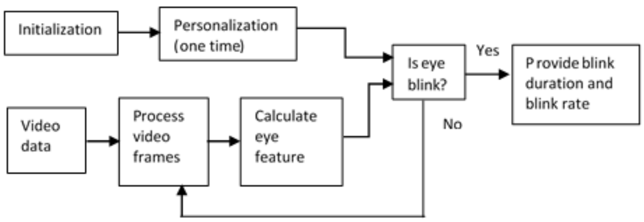

Fig. 1: Main components of the proposed system

1) Initialization: The application can run on any laptop/desktop (windows or Linux) having an integrated or external webcam. On starting, the application enables the webcam and a screen (Fig. 2) appears with the video of user's eye region. It prompts the user to adjust (through a slider) the brightness and contrast levels, so that the eye region is clearly visible. Next, it asks whether the user is using spectacles and if the ambient light is natural/artificial light. This data is used to analyze the system performance under different scenarios.

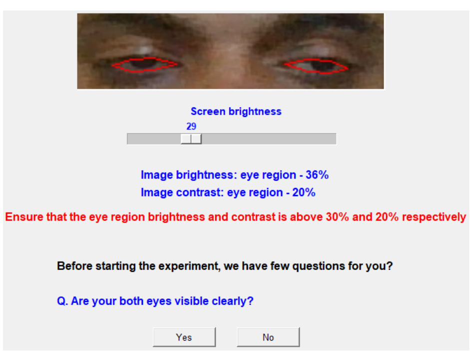

Fig. 2: Initialization screen of the proposed system

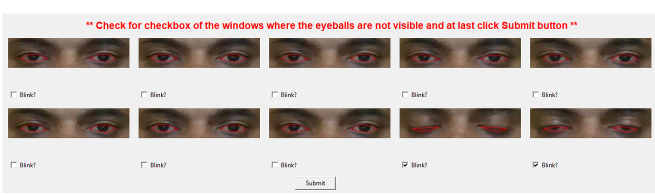

Fig. 3: Annotation screen containing eye blinks

2) Personalization: Personalized blink features are extracted from the eye videos collected. First, eye landmarks are detected for every video frame using the Dlib shape predictor. Dlib provides 68 facial landmarks [31] of which we are using 6 as shown in Fig.4. Next, eye aspect ratio

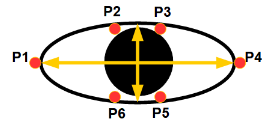

Fig. 4: Eye landmarks using Dlib shape predictor

(EAR) [30] is calculated using (1). EAR is high when the eyes are open and decreases when the eyes are closed. EAR is computed separately for left and right eyes and the average of these two is used.

$$EAR_e = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2||p_1 - p_4||} \tag{1}$$

Proposed approaches of personalization:-The average EAR values are first interpolated by averaging the neighboring values. Continuous EAR time series thus obtained is used for further analysis. After data collection, an annotation screen (Fig. 3) appears consisting of windows of 100 ms of the eye video. The users are instructed to identify windows where more than 50% of the eyeballs are covered and seems to be a blink. We have used three approaches for personalization.

i) Annotation based approach ($A1$) - The blink and no-blink frames are segregated. The average (B_m) and standard deviation (B_s) of EAR values corresponding to blinks and non-blink portions (NB_m) are computed. These values are used for detecting the blinks in test data (yc) in run time. A decision vector$D$ is computed in the test phase (y) as,

$$D_i = \begin{cases} 1, & \text{if } yc_i < (B_m + B_s) \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

$\forall, i = 1, 2, ..., N$ number of EAR values in yc.

ii) Threshold based approach ($A2$) - Factors like ambient lighting conditions, distance from camera and usage of spectacles change the trend of EAR time series making it non-stationary in nature. One such signal is shown in Fig. 5. Under such circumstances, $A1$ might fail considerably. Hence, we performed baseline correction on the EAR signal using the asymmetric least squares smoothing approach [32]. The baseline signal (x_b) is computed and used to get the corrected signal (x_c) as,

$$xc = x - x_b \tag{3}$$

The baseline corrected signal is also shown in Fig. 5. Next, the minimum (xc_{min}) and average (xc_m) of x_c are calculated. Finally, the user specific threshold value (t) is calculated as,

$$t = xc_m - k \times (xc_m - xc_{min}) \tag{4}$$

where constant k is empirically derived as 0.25 on a subset of user data. D for test data (yc) is computed as

$$D_i = \begin{cases} 1, & \text{if } yc_i < t \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

$\forall, i = 1, 2, ..., N$ number of EAR values in the new data.

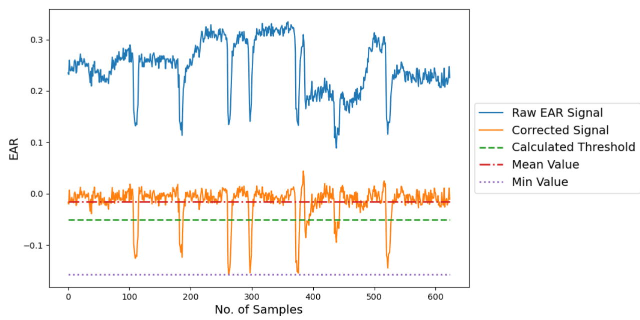

Fig. 5: Threshold computation from EAR series using A2

iii) Signal level based approach ($A3$) - While working with open dataset, we observed that constant k used in $A2$ varies with demography. Hence, we propose a method which is independent of k. We computed the average of blink (xc_{Mb}) and no blink portions (xc_{Mnb}) of the signal x_c. These values are used to calculate D on the test data (y_c) as,

$$D_i = \begin{cases} 1, & \text{if } |xc_{Mb} - yc_i| <= |xc_{Mnb} - yc_i| \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

$\forall, i = 1, 2, ..., N$ number of EAR values in the new data.

For all three approaches, instances of '1's in D corresponding to at least 100 ms are considered as blinks.

B. Data collection for validation of proposed system

We collected 32 sessions of data from various users (mean age: 30 years) of our lab. Participants signed an informed consent form. The experimental protocol was approved by our internal IRB. They downloaded and ran the application on their own computers. User specific thresholds were calculated in run time and only those values were shared with us. We collected data under 4 conditions - i) with spectacles (WS), ii) Without spectacles (WoS) iii) Artificial light (AL) where data was collected under florescent lights and iv) Natural light (NL) where data was collected under daylight. Each session comprises of three steps as explained below:

Step 1 - Data for deriving personalized thresholds - a white screen appears with a black "+" at the center and 5 beep sounds (2500 Hz of 500ms duration) occurred. Users had to blink as soon as they hear the beep and the eye video was captured. Next, the annotation screen appeared and users annotated the frames containing blinks. User specific threshold values were derived from the video and stored.

Step 2 - Simulated blink - Here we followed same protocol as *step 1*. In total, 6 beeps with variable inter-beep gaps were played. No annotation screen was shown in this step. The data collected was used to analyze the system performance for detection of simulated blinks.

Step 3 - Natural blink - Stimulated blinks have somewhat known patterns in terms of occurrence, duration and intensity. Thus, we analyzed the performance of our system for natural blinks. Here, the users were free to blink naturally and then enter the number of blinks in the application as the ground truth. Three such trials were conducted one after another.

C. Test phase - real time detection of eye blinks

The videos collected in *step 2* and *step 3*were processed frame by frame to construct EAR time series and perform baseline correction. These EAR values were compared with user specific thresholds and a decision (blink or no-blink) was taken. For each detected blink, corresponding blink amplitude and duration were also reported.

TABLE I: Average F_{score} (SD) of proposed system and comparison with SOA approaches

Methods	**With spec**	**W/o spec**	**Artif. light**	**Nat. light**	**Simu. blink**	**Nat. blink**
Annotation based (A1)	0.81 (0.26)	0.89 (0.19)	0.85 (0.22)	0.87 (0.22)	0.86 (0.2)	0.85 (0.26)
Threshold based (A2)	**0.97 (0.07)**	**0.98 (0.03)**	**0.98 (0.04)**	**0.98 (0.05)**	**0.94 (0.18)**	**0.97 (0.1)**
Signal level based (A3)	**0.98 (0.05)**	**0.99 (0.03)**	**0.98 (0.04)**	**0.98 (0.04)**	**0.95 (0.18)**	**0.98 (0.11)**
SOA1 [33]	0.88 (0.15)	0.9 (0.15)	0.91 (0.14)	0.87 (0.17)	0.83 (0.24)	0.9 (0.15)
SOA2 [34]	0.48 (0.37)	0.37 (0.38)	0.41 (0.4)	0.40 (0.37)	0.28 (0.28)	0.45 (0.4)
SOA3 [23]	0.69 (0.36)	0.7 (0.33)	0.74 (0.34)	0.67 (0.34)	0.71 (0.34)	0.68 (0.35)

IV. Results

System performance for two types of light (NL and AL), two types of blinks (simulated and natural) and usage of spectacles (WS and WoS) were analyzed using F_{score} which is the harmonic mean of precision and recall. Average F_{scores} across participants for various scenarios are presented in Table I. It is observed that, approaches A2 and A3 perform similarly giving maximum F_{score} across all scenarios. We also compared our approaches with few state of the art approaches. Huda et al. [33] (SOA1) and Mallikarjuna [34] (SOA2) have used fixed threshold values of 0.24 and 0.3 respectively whereas A. Mohammed used HAAR classifier [23] (SOA3) based approach. We applied these SOA approaches on our dataset and the average F_{scores} across participants are presented in Table I. The F_{scores} for various scenarios across participants using our approaches and SOA approaches are shown in Fig. 6. It is seen that

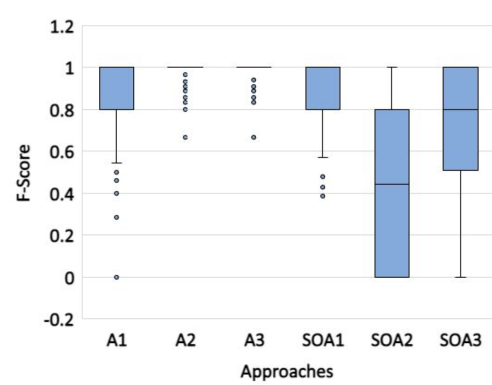

Fig. 6: System performance using proposed and SOA approaches

A1 is performing similar to SOA1, however, A2 and A3 outperforms all SOA approaches. Thus, we conclude that our approach is performing significantly better than the SOA.

TABLE II: System performance (F_{score}) for Eyeblink8 dataset

Approach	A1	A2	A3	SOA1	SOA2	SOA3
Average	0.55	**0.91**	0.86	0.62	0.25	0.21
SD	0.26	**0.06**	0.11	0.25	0.27	0.26

Finally, we analyzed the performance on publicly available Eyeblink8 [35] dataset. SOA1, SOA2 and SOA3 have also used this dataset. The dataset contains blink data from 8 participants and annotations as ground truth. Results obtained using our approaches are given in Table II. The data set has varying sizes of data for different users ranging from 2.5 minutes to 9 minutes and three participants have recordings for less than 3 minutes (minimum is 2 min 40 sec.). To bring uniformity across participants, we selected a frame size of duration 2 min 30 seconds. The first 60 seconds of data was use to personalize the system. It is observed, that both A2 and A3 outperform SOA approaches for this dataset also, however, since $A3$ relies on self-annotation, it did not perform well compared $A2$. Thus,$A2$ is more robust compared to $A3$.

V. Conclusions and Future Scope

Camera based approaches are used by researchers to detect eye blinks. However, accuracy of such camera based systems varies a lot across participants and external conditions. In the present work, we have proposed a camera based personalized system for detection of eye blinks in real time. Three approaches have been proposed for personalization which were validated on a set of video data collected under various conditions like with/without spectacles, artificial/natural light, and natural/simulated blinks. We also analyzed the performance of the proposed system on Eyeblink8 dataset. Results show that proposed signal level based approach and threshold based approach are good and outperform state of the art approaches. However, Threshold based approach is the most robust one, as it does not require any manual annotation for personalization. The proposed system can be implemented in any computers having an internal or external webcam and hence is easy to deploy. In future, we would like to use the proposed system in various applications like assessment of human cognition, behavior and so on.

References

[1] Gavas et al, “Effect of cognitive load on a random sequence generation task,” in *IEEE (TENSYMP)*. IEEE, 2019, pp. 243–248.

[2] Gavas et al, “Enhancing the usability of low-cost eye trackers for rehabilitation applications,” *PloS one*, vol. 13, no. 6, 2018.

[3] Gavas et al, “Cognitive load and metacognitive confidence extraction from pupillary response,” *Cognitive Systems Research*, vol. 52, pp. 325–334, 2018.

[4] Gavas et al, “Estimation of cognitive load based on the pupil size dilation,” in *IEEE SMC*. IEEE, 2017, pp. 1499–1504.

[5] Gavas et al, “Blink rate variability: a marker of sustained attention during a visual task,” in *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2020, pp. 450–455.

[6] Chatterjee et al, “Evaluating age-related variations of gaze behavior for a novel digitized-digit symbol substitution test,” *Journal of Eye Movement Research*, vol. 12, no. 1, 2019.

[7] Divjak et al, “Eye blink based fatigue detection for prevention of computer vision syndrome.,” in *MVA*, 2009.

[8] Paprocki and Lenskiy, “What does eye-blink rate variability dynamics tell us about cognitive performance?,” *Frontiers in human neuroscience*, vol. 11, pp. 620, 2017.

[9] Khasnobish et al, “Eyeassist: A communication aid through gaze tracking for patients with neuro-motor disabilities,” in *2017 IEEE PerCom*. IEEE, 2017, pp. 382–387.

[10] Sahayadhas et al, “Detecting driver drowsiness based on sensors: a review,” *Sensors*, vol. 12, no. 12, pp. 16937–16953, 2012.

[11] Garcia-Agundez et al, “Eye aspect ratio based blink rate detection and its potential use for parkinson’s disease,” 2018.

[12] Jung et al, “Deepvision: Deepfakes detection using human eye blinking pattern,” *IEEE Access*, vol. 8, pp. 83144–83154, 2020.

[13] Yan et al, “Computer vision syndrome: A widely spreading but largely unknown epidemic among computer users,” *Computers in human behavior*, vol. 24, no. 5, pp. 2026–2042, 2008.

[14] Bulling et al, “Wearable eog goggles: eye-based interaction in everyday environments,” in *CHI’09 Human Factors in Computing Systems*, pp. 3259–3264. 2009.

[15] Prabaswari et al, “Eye tracker evaluation on google classroom using use questionnaire,” in *Journal of Physics: Conference Series*. IOP Publishing, 2021, vol. 1764, p. 012181.

[16] Rakshit and Sinharay, “Affordable design realization for ultrasound based non-contact eye blink event detection,” in *IEEE SENSORS*. IEEE, 2018, pp. 1–4.

[17] Sinha et al, “Artifact removal from eeg signals recorded using low resolution emotiv device,” in *IEEE SMC*. IEEE, 2015, pp. 1445–1451.

[18] Ren et al, “Comparison of the use of blink rate and blink rate variability for mental state recognition,” *IEEE Tran. on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 5, pp. 867–875, 2019.

[19] Frigerio et al, “Surface electromyographic mapping of the orbicularis oculi muscle for real-time blink detection,” *JAMA facial plastic surgery*, vol. 16, no. 5, pp. 335–342, 2014.

[20] Kumar and Sharma, “Electrooculogram-based virtual reality game control using blink detection and gaze calibration,” in *ICACCI*. IEEE, 2016, pp. 2358–2362.

[21] Dyana Muller, “A review of video-based and machine learning approaches to human eye blink detection in video,” 2019.

[22] Yi-Qing Wang, “An analysis of the viola-jones face detection algorithm,” *Image Processing On Line*, vol. 4, pp. 128–148, 2014.

[23] Assit Prof Aree A Mohammed, “Efficient eye blink detection method for disabled-helping domain,” *Eye*, vol. 10, no. P1, pp. P2, 2014.

[24] Drutarovsky and Fogelton, “Eye blink detection using variance of motion vectors,” in *European Conference on Computer Vision*. Springer, 2014, pp. 436–448.

[25] Torricelli et al, “An adaptive blink detector to initialize and update a view-basedremote eye gaze tracking system in a natural scenario,” *Pattern Recognition Letters*, vol. 30, no. 12, pp. 1144–1150, 2009.

[26] Lee et al, “Blink detection robust to various facial poses,” *Journal of neuroscience methods*, vol. 193, no. 2, pp. 356–372, 2010.

[27] Sukno et al, “Automatic assessment of eye blinking patterns through statistical shape models,” in *International Conference on Computer Vision Systems*. Springer, 2009, pp. 33–42.

[28] Chau and Betke, “Real time eye tracking and blink detection with usb cameras,” Tech. Rep., Boston University Computer Science Department, 2005.

[29] Yang et al, “Robust eyelid tracking for fatigue detection,” in *International Conference on Image Processing*. IEEE, 2012.

[30] Soukupova and Cech, “Eye blink detection using facial landmarks,” in *21st computer vision winter workshop, Slovenia*, 2016.

[31] Davis E King, “Dlib-ml: A machine learning toolkit,” *The Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.

[32] Eilers and Boelens, “Baseline correction with asymmetric least squares smoothing,” *Leiden University Medical Centre Report*, vol. 1, 2005.

[33] Huda, Tolle, and Fitri, “Mobile-based driver sleepiness detection using facial landmarks and analysis of ear values,” 2020.

[34] Mallikarjuna Rao Nishanth, “Liveness detection based on human eye blinking for photo attacks,” *IJEAT*, vol. 9, 2019.

[35] Tomas and Andrej, “Eye blink detection using variance of motion vectors,” in *European Conference on Computer Vision*. Springer, 2014, pp. 436–448.