

PAPER • OPEN ACCESS

A hybrid technique for intelligent bank security system based on blink gesture recognition

To cite this article: Raed Awadh Bakunah and Saeed Mohammed Baneamoon 2021 *J. Phys.: Conf. Ser.* **1962** 012001

View the [article online](#) for updates and enhancements.



The banner features a colorful striped border at the top. On the left, the ECS logo is displayed in a green circle. To its right, the text reads: "240th ECS Meeting", "Oct 10-14, 2021, Orlando, Florida", "Register early and save up to 20% on registration costs", "Early registration deadline Sep 13", and "REGISTER NOW" in orange. On the right side of the banner is a photograph of a diverse group of people in a professional setting, with a man in a white shirt and tie clapping and smiling.

ECS **240th ECS Meeting**
Oct 10-14, 2021, Orlando, Florida
**Register early and save
up to 20% on registration costs**
Early registration deadline Sep 13
REGISTER NOW

A hybrid technique for intelligent bank security system based on blink gesture recognition

Raed Awadh Bakunah¹ and Saeed Mohammed Baneamoon^{2,3}

¹ Department of Information Technology, Faculty of Engineering and Information Technology, Al-Rayan University, Yemen

² Department of Computer Engineering, College of Engineering & Petroleum, Hadhramout University, Yemen

³ Department of Computer Science, College of Computers & Information Technology, Hadhramout University, Yemen

¹raidbakunah@gmail.com

^{2,3}sbaneamoon@hu.edu.ye, saeed_banaimoon@yahoo.com

Abstract. Enhancement of bank security system plays an important role at the present time due to increasing attempts to expose it to many infringements by criminals. This paper proposes a hybrid technique of amplitude and duration to recognize its blink gestures to increase the accuracy of the human security system's interaction for intelligent bank security, as this hybrid technique implemented analyzing of face, facial landmark detection, and the duration of the eye aspect ratio (EAR) for blink processes. The proposed hybrid technique is evaluated by experimenting a number of blink gesture with two eyes obtained from web camera and the results shows better and more effective analysis and provide high security bank system.

Keywords: Security system, Blink gesture recognition, blinking and Facial Landmarks.

1. Introduction

Nowadays, most of banks have a number of branches which have expensive jewelry, important documents and cash. Therefore, banking system must has a top level of security to avoid attempts of criminals, where increased numbers of robberies and crimes of banks during working hours when the employees were subjected to dangerous. These crimes sometimes happen under video surveillance systems, which are basically used to provide evidence after crimes, and it cannot do real-time early warning. Thus, there is an urgent need to develop technique for rapid interactive the employee with alarm system to help inform nearest police station to thwart the crime in a timely manner. [1].

By the rapid development of computer vision, image processing and video intelligent technologies, it's now possible to integrate human gestures recognition techniques with video surveillance system to achieve intelligent security system.

Therefore, this paper proposes a human-security interactive system for intelligent bank security, which based on hybrid technique that integrated amplitude and duration of blink gesture recognition of the employee.



2. Related work

Developing of bank security systems become the major concern to the entire world due to increasing crimes of banks. However, several technologies and techniques have been used in developing of security banks' systems. These systems can be categorized into two types; there are employee without help gesture and employees with help gesture.

In employee without help gesture, some systems based only on electronic and embedded technologies as in [2] that detect the Automatic Teller Machine (ATM) theft from robberies using Advanced RISC Machines processor (ARM) controller based on embedded system to process real time data collected by PIR sensor and in [3] where introduced a design of simple security system, the system consists of antenna working at IR frequency band, programmable system on chip and a Silicon Controlled Rectifier (SCR), other systems based on computer vision and image processing technologies that are used deferent methods and techniques as in [4] that presented a real time a multi-camera platform for semantic video interpretation for monitoring, and in [5], [6] and [7] which are designed intelligent monitoring control systems based on video image processing. However, Radio Frequency identification (RFID) system used to provide access of the locker room area to only authorize people as in [8], In addition, there are some systems that used pattern recognition and face recognition as in [9] and [10].

On the other hand, in employee with help gesture, which are depending on gesture of employee to detect robberies as the systems in [11] and [12] that are based on hand gesture technique, this technique is ineffective in state when the hands are tied. Moreover, the system in [13] is very close to our work, it is based on blink gesture recognition, but it use the amplitude parameter of blink which is a little default for some parsons.

However, the duration blink gesture can be integrated with amplitude blink gesture in one system in order to obtaining robust and safety security system with good performance.

3. Background

3.1. Gesture

Gesture is defined as expressive movements of body parts, which have particular messages, to be communicated precisely between a sender and a receiver. The possible location for origin of human gestures or body gestures can be from any part of our body, but the most of -gesture-based studies that are carried out involve the gestures that have originated from the hand [14] [15], face [16], head with eye gaze [17] [18], eyelids [13] [19] and tongue [20], because most normal people transfer the messages through this locations. A gesture is scientifically categorized into two categories: dynamic and static. In order to understand a complete message, the gestures must be recognized through the process of interpreting static and dynamic gestures over a period of time as this process takes place from a specific set of input data. The main entire point of the gesture recognition is to teach a machine to learn different gestures made by humans, and to interpret the same, which means to make the sensible conversations between a human and machines possible.

Presently, the rapid development of image processing, computer vision and video intelligent technologies give a possible to achieve intelligent security systems based on recognize gestures by vision techniques and integrate human gestures recognition techniques with video surveillance systems.

3.2. Eyelids movements

Blinking is defined as a rapid closing and reopening the eyelids of a human eye. Eyes blink differ from person to another in terms of closing and opening speed, in addition to the degree of squeezing them. The main parameters of the blink are duration, rate, and amplitude. Blink duration is the period of time when the eyes were in a closed state, the natural duration lasts approximately 100-400ms, it depends on people's mood. Blink rate is the number of blinks per given timeframe. The natural blink rate

during rest was 17 blinks/min, it increases to 26 during conversation, and it decreases to 4.5 while reading. Finally the blink amplitude which describes the degree of openness of eye, it differs between people and it also depends on mood and head pose [21]. These parameters are shown in 'figure 1'.

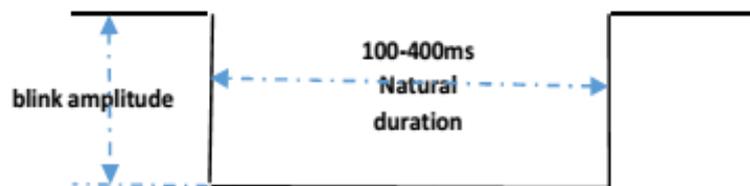


Figure 1. Parameters of natural blink

However, the eyelids motion can be classified to involuntary blink, as natural or spontaneous blinking which, is done without any obvious external stimulus, and the voluntary blink, which implements under the control of the individual.

3.3. Vision-based blink recognition

The states of human eyes contain a lot of useful information, and the research on recognition of unit eye has become a hotspot in human activity recognition. The movements of eye and eyelids are the main method of collecting the visual information, and play an important role to detect the human states, moods and intend. Eyelids recognition not only have research significance in the field of detection for blinking, drossy and fatigue as in [21], [22] and [23] but also have extensive application prospect and profound significance in human-computer interaction (HCI) as computer commands as in [19] and also security systems as in [13].

Many of different methods have been developed for vision-based blink detection that they estimate eye state as either open or closed, or track eye closure events so the output can be either recognition of open or closed eye, or detection of blink. These techniques use only a camera. The characteristics of individual images are used to recognize eye status while a number of sequential frames from video are used to detect blink. They usually have three stages: face detection, eye region detection and a blink detection or recognition. The Viola-Jones algorithm is mostly used for face detection.

The single image processing methods are based on skin color segmentation, edge detection and a parametric model fitting to find the eyelids. In computing of blinks, a combination of eye localization, thresholds for finding the white of the eye and determining whether the white area has been disappearing for a period of time usually included in traditional methods of image processing as in [23].

On the other hand, there are methods that based on matching templates, where the template with open and closed images are identified and the standard eye area correlation coefficient is computed for each image. These methods are used an eye-image template that is compared to part of the input image. The corresponding part of the input image is determined by correlating the template with the current image frame as in [19].

Moreover, some methods use facial landmarks technique to detect the eyelid contours, it relies on extracting the vertical distance between eyelids as in [24] or estimating the eye states by calculating the eye aspect ratio (EAR) for every frame from six landmarks points of each eyelid as in [22].

In this paper we focus on vision-based blink gesture recognition which it makes by eyelids movements.

4. Proposed method

The proposed method based on the hybrid technique that integrate amplitude and duration of blink gesture recognition that analysing facial landmark, EAR and duration of blink. This hybrid technique

is implementing by using python with OpenCV and dlib libraries. The block diagram of complete proposed system is shown in 'figure 2'.



Figure 2. The block diagram of complete system

This methodology includes face and facial landmarks detection and classification for duration of blink as the following:

(1) The facial landmarks detection is the process of localizing the key facial points on a face, including the eyes, eyebrows, nose, mouth, and jawline as in 'figure 3'.

Automatic recognition of the location of landmark points of the face on facial images or videos is the objective of facial landmark detection algorithms [25].



Figure 3. Facial landmark points

(2) Detecting the aspect ratio of the eyes that suggested by Sokopova and Czech is based on the ratio of distances between facial landmarks to the eyes [22]. As the result, a single value is given, which links the distances between the vertical eye landmark points (p_2 , p_3 , p_6 , p_5) to the distances between the horizontal feature points (p_1 , p_4) and it is computed by equation (1).

$$EAR = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2||p_1 - p_4||} \quad (1)$$

Where p_1 to p_6 are the 2D landmark locations.

The ear is mostly constant when eye is open and getting close to zero while closing an eye as shown in 'figure 4'.

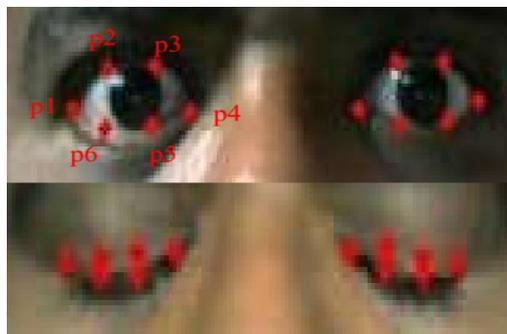


Figure 4. Open and closed eye and its landmark points

(3) This paper proposed voluntary gestures based on duration parameter of the blink for the proposed security interaction system as shown in 'figure 5'.

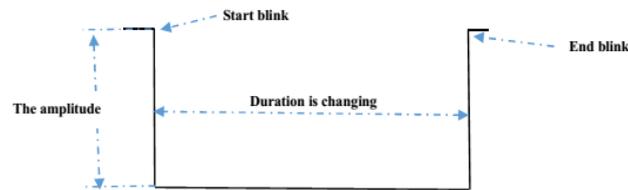


Figure 5. The proposed blink gesture

The duration of the blink are changing differently from normal, it had specific and sufficient duration thresholds values, in order to avoid any trigger of a system by the spontaneous eye blink of human.

5. Experimental proposed method

The first section is localize the face by apply any face detector algorithm in a given frame of a video stream as shown in ‘figure 6’.

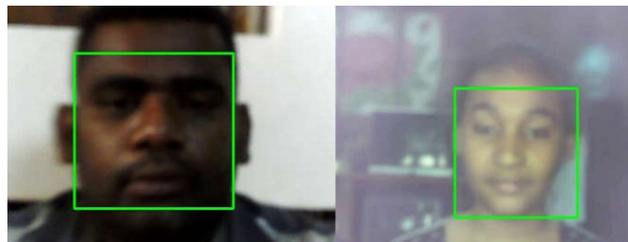


Figure 6. Localized face

Then localize the eyes landmarks as shown in ‘figure 7’. by using 68 key facial landmark detector which is an implementation of one millisecond ace alignment with an ensemble of regression trees in [26].

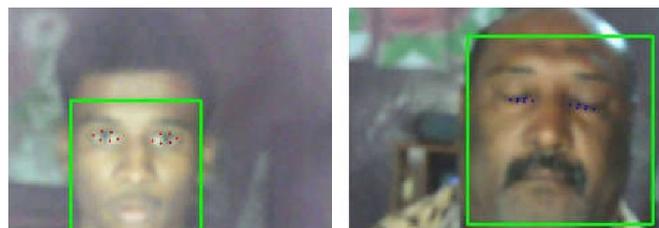


Figure 7. Eyes landmark points

After that Soukupova and Cech’s method is used to measure EAR for each eye, which is ratio of distances between the six facial landmark points. The ear value for each consecutive frame of both eyes is calculated and then thresholds for EAR are set in the code.

The second section is to determine when the blink is considered to be a gesture, to do that we first set two threshold values of ear that must be set in code, higher one for normal open state (NS) and lower one for closed eye (LT). After that we measure the real detected ear value and compare it with LT and NS threshold values, if the comparison is true, then we determine the duration of the detected blink and compare it with predefined Time thresholds (TFseq).

The duration is the number of sequence frames when eyes are in closed state or when the value of ear = < LT threshold, so TFseq predefined threshold is the number of sequence frames must be set in the code.

The duration blink gesture is the change in ear from NS value to (ear < LT), and stay in it predefined pried of time and rise again to NS as shown in ‘figure 8’.

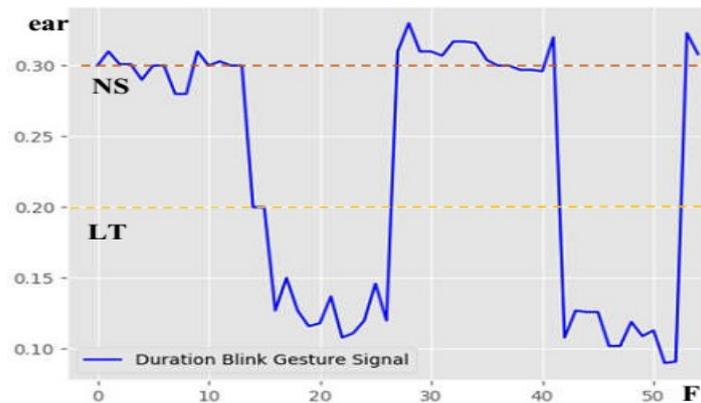


Figure 8. Duration blink gesture signal

When the comparison is correct, the alarm is activated in the system and SMS warning messages are sent. This section is implemented by Python.

6. Experimental result

The hybrid proposed system for blink gesture detection was tested using Intel Core 2 Duo CPU at 2.13 GHz processor 4G RAM on the 30 fps sequences frames from the USB Logitech 270 Pro webcam and Jewa web cameras. The size of the input images sequences were equal to 480, 640 pixels. Testing of the system took place in a room illuminated by 2 fluorescent lamps and daylight from tow windows.

The initial training to perform the predefined gesture (normal blinks and gesture blinks, alternately) is recorded videos for four participants. The web camera was fixed at distance 70 cm away from the face. Finally the experiments in Python, OpenCV, and dlib in videos are implemented and results are shwon in sections: 6.1, 6.2, 6.3 and 6.4.

6.1. Participant vs. EAR experiment results

The ear values for 4 different people were recorded and the ear values for completely closed and normal open state (NS) were measured.

For accuracy the ear was calculated for several times and the mean value was recorded. table 1 shows the mean values. Using the values a graph was plotted as shown in 'figure 9'and the value of Common LT was computed as in equations (2).

$$LT = \frac{\text{Big ear value in closed state} + \text{small ear in normal open state}}{2} \quad (2)$$

Table 1: The mean values of ear for complete closed and normal open of eyes

participants	EAR average for eyes closed	NS
Participant 1	0.104	0.345
Participant 2	0.092	0.323
Participant 3	0.072	0.313
Participant 4	0.103	0.294

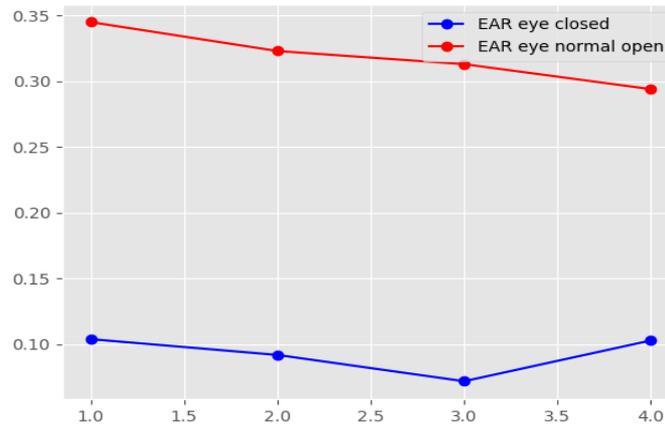


Figure 9. EAR of 4 participants for eyes in open and closed states

6.2. Gestures vs. conditions experiment results

Experiments were done with one participant at low light and wear glasses. Duration Blink Gesture stay detected in low lights as shown in ‘figure 10’ and in wearing glasses as in ‘figure 11’.

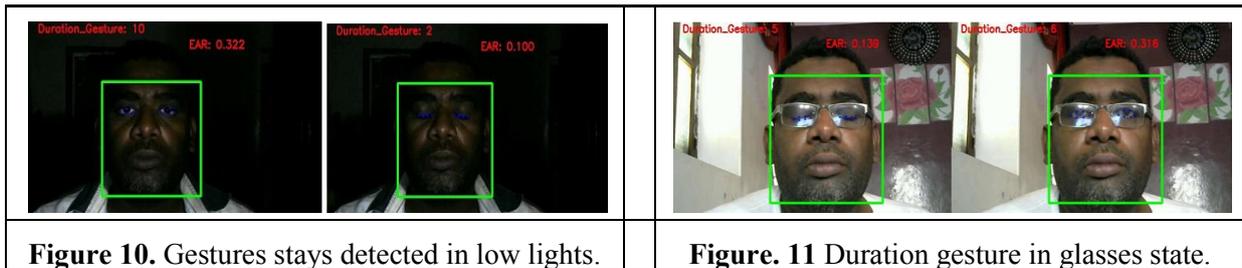


Figure 10. Gestures stays detected in low lights.

Figure. 11 Duration gesture in glasses state.

6.3. Real time capabilities results

The system was able to process the video and recognize gestures in real-time as can be seen from table 2 and table 3 and by using values in table 3 a graph was plotted as shown in ‘figure 12’.

Table 2: Frame rate and processing time for recognize gestures in real-time.

	Baseline	Without a face	Recognition
Mean frame rate	27.3 fps	5.9 fps	5.3 fps
Processing time	0 – 37.2ms	110-142 ms	125–200 ms

Table 3: Gesture processing time for number of frames.

Number of frames of duration gesture	Processing time of duration gesture in seconds
6	1.257s
7	1.467s
8	1.668s
9	1.879s
15	3.137s
16	3.353s
17	3.553s
18	3.718s
20	4.186s

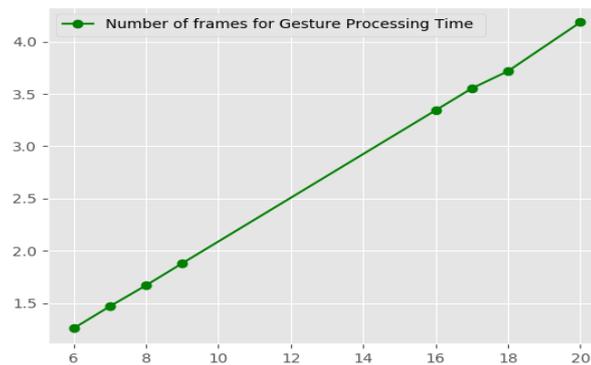


Figure 12. Number of frames for gesture processing time

6.4. Recognition performance experiment results

Duration blink gesture detected with 30fps camera speed. Two types of errors have been identified: false detection where non-existent gesture is detected. The other type of errors is the missing gesture, where existent undetected gesture. The distribution of potential decisions for the gesture detector output is presented as shown in table 4.

Table 4: Possible decisions of gesture output.

Possible decision		Actual state	
		Gesture present	Gesture not present
Result of detection	Gesture detected	TP	FP
	No gesture detected	FN	TN

Correctly detected gestures are indicate as True Positive (TP), false detections are indicate as False Positives (FP) and missed gestures are indicate as False Negatives (FN). Based on these parameters, an accurate measurement of system performance was computed as shown in equation (3) [24].

$$Accuracy = \frac{TP}{TP + FP + TN} \tag{3}$$

System performance and tests are performed with low and good resolution for cameras and other conditions. The results are shown in tables 5 and 6.

Table 5: Accuracy of the duration blink gesture recognition system averaged over all four participants.

Videos	GT	DDG	FN	FP	Accuracy
Participant1	19	17	2		89.47%
Participant2	6	6			100%
Participant3	16	15	1		93.75%
Participant4	10	9	1	1	82%
Averaged Accuracy					91.3%

Table 6: The distribution of the system performance and accuracy for the duration gestures, in states: glasses, without glasses and low lights.

States of video	Duration gesture	Total (GT)	Detected gesture	FN	FP	Accuracy
Middle	20		20			100%
Glasses	20		19	1		95%
Low light	20		20			100%
Averaged Accuracy						98.3%

When GT is Total Duration Gesture (Grown True Gesture) and DAG is detected duration gesture.

Table 7 shows a comparison of proposed system with Gupta, Kumar and Bakunah systems.

Table 7: The comparison of proposed system with Gupta, Kumar and Bakunah works.

Works	Used technique	Observe by suspicious & accomplices	Difficulty level by employee	Acc
Gupta [1]	Hand gesture recognition	easily	easily	95.7%
Kumar [12]	Hand gesture recognition	easily	easily	94.5%
Bakunah [13]	Amplitude blink gesture recognition	difficult	Little difficulty	96.5%
Proposed system	Duration blink gesture recognition	difficult	easily	98.3%

7. Conclusion

In this paper interactive security system based on the hybrid technique of amplitude and duration blink gesture recognition is proposed.

The duration blink gesture is easier to perform by the employee compared with amplitude blink gesture in [13] which is a little difficult. However, the duration blink gesture can be integrated with amplitude blink gesture in one system in order to obtaining robust and safety security system with good performance.

The real time implementation of hybrid blink gesture recognition security system has an accuracy of more than 98% in state of good conditions (as in regularly luminous intensities and middle face orientation with and without glasses) which, is better as compared with the existing results of related works in [11], [12] and [13].

References

- [1] Royanews, [Online]. Available: <https://royanews.tv/news/171426> , 27/12/2018, Updated: 23/7/2020
- [2] Archana K, Reddy P B and Govardhan A 2017 To enhance the security for ATM with the help of sensor and controllers, *International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, Chennai, pp. 1012-1015.
- [3] Elshafiey T M F 2011 Design and implementation of a museum and banks security system using antenna as IR proximity sensor and PSoC technology, *IEEE Symposium on Wireless Technology and Applications (ISWTA)*, Langkawi, pp. 156-161.
- [4] Georis B, Maziere M, Bremond F and Thonnat M 2004 A video interpretation platform applied to bank agency monitoring, *IEEE Intelligent Distributed Surveillance Systems*, London, UK, pp. 46-50.
- [5] Wu Y 2009 Research on bank intelligent video image processing and monitoring control system based on OpenCV, *3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication*, Hong Kong, pp. 211-214.
- [6] Khera N and Verma A 2014 Development of an intelligent system for bank security, *5th International Conference - Confluence The Next Generation Information Technology Summit (Confluence)*, Noida, pp. 319-322.
- [7] Yi H, Liu J and Wang X 2011 Automatic alarm system for self-service bank based on image comparison, *2011 International Symposium on Computer Science and Society*, Kota Kinabalu, pp. 48-50.
- [8] Verma A 2013 A multi-layer bank security system, *International Conference on Green Computing, Communication and Conservation of Energy (ICGCE)*, Chennai, pp. 914-917.

- [9] Joshua B, Grace J and George A M 2016 Open CV pattern based smart bank security sSystem with theft & identification using Android, *2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, Bangalore, pp. 254-257.
- [10] Kumari P 2019 High protection bank locker security system using live image and voice authentication, *International Journal of Scientific Research and Review / UGC Journal No.:* 64650, **07 (03)**, pp. 2616-2619.
- [11] Gupta A, Kumar Y and Malhotra S 2015 Banking security system using hand gesture recognition, *International Conference on Recent Developments in Control, Automation and Power Engineering (RDCAPE)*, Noida, pp. 243-246.
- [12] Kumar S, Yadav G, Singh H P, Malhotra S and Gupta A 2018 SSVM classifier and hand gesture based sign language recognition, *International Conference on Intelligent Circuits and Systems (ICICS)*, Phagwara, pp. 456-460.
- [13] Bakunah R A, Baneamoon S M 2020 An intelligent security system for a bank based on amplitude of blink gesture recognition, *Journal of Emerging Technologies and Innovative Research (JETIR)*, **7(9)**, pp.504-513.
- [14] Shanthakumar V A, Peng C, Hansberger J, Cao L, Meacham S C and Blakely V R 2020 Design and evaluation of a hand gesture recognition approach for real-time interactions, *Multimed Tools Appl* **79**, pp.17707–17730.
- [15] Hussain S, Saxena R, Han X, Khan J A and Shin H 2017 Hand gesture recognition using deep learning, *International SoC Design Conference (ISOCC)*, Seoul, pp. 48-49.
- [16] Pradhan C, Banerjee D, Nandy N and Biswas U 2019 Generating digital signature using facial landmark detection, *International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, pp. 0180-0184.
- [17] Wang K, Zhao R and Ji Q 2018 Human computer interaction with head pose, eye gaze and body gestures, *13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, Xi'an, pp. 789-789.
- [18] Vaitukaitis V and Bulling A 2012, Eye gesture recognition on portable devices, *UbiComp '12: Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pp. 711-714.
- [19] Lak A K 2011 Eye-blink detection system for human–computer interaction, *Universal Access in the Information Society*, Springer **11(4)**, pp 1-11.
- [20] Gupta A, Chitranshi G, Sharma V and Singh H 2017 Controlling devices using facial expression identification and recognition for assisting physically disabled, *7th International Conference on Cloud Computing, Data Science & Engineering - Confluence*, Noida, pp 574-576.
- [21] Soukupov' a T 2016 Eye-blink detection using facial landmarks, *Center for Machine Perception, Technical University Technick'a 2, 16627 Prague 6, Czech Republic, Prague*.
- [22] Soukupov T and Cech J 2016 Real-time eye blink detection using facial landmarks, *21st Computer Vision Winter Workshop Luka' Cehovin, Rok Mandeljc, Vitomir' Struc* (eds.)Rimske Toplice, Slovenia.
- [23] Haq Z A and Hasan Z 2016 Eye-blink rate detection for fatigue determination, *1st India International Conference on Information Processing (IICIP)*, pp 1-5.
- [24] Al-gawwam S 2018 Robust eye blink detection based on eye landmarks and savitzky–golay filtering. *Information*, **9(4)**, p 93.
- [25] Wu Y 2018 Facial landmark detection: A literature survey, *IJCV*, Springer, pp 1-28.
- [26] Kazemi 2014 One millisecond fFace alignment with an ensemble of regression trees, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 1867-1874.