

# Multimodal Human Eye Blink Recognition Using Z-score Based Thresholding and Weighted Features

Puneet Singh Lamba<sup>1,2</sup>, Deepali Virmani<sup>3</sup>, Manu S. Pillai<sup>2</sup>, Gopal Chaudhary<sup>2\*</sup>

<sup>1</sup> University School of Information, Communication and Technology, GGSIPU, Sector 16 C, Dwarka, Delhi (India)

<sup>2</sup> Bharati Vidyapeeth's College of Engineering, Paschim Vihar, New Delhi (India)

<sup>3</sup> Vivekananda Institute of Professional Studies-Technical Campus, New Delhi (India)

Received 15 March 2021 | Accepted 27 August 2021 | Published 8 November 2021



## ABSTRACT

A novel real-time multimodal eye blink detection method using an amalgam of five unique weighted features extracted from the circle boundary formed from the eye landmarks is proposed. The five features, namely (Vertical Head Positioning, Orientation Factor, Proportional Ratio, Area of Intersection, and Upper Eyelid Radius), provide imperative gen (z score threshold) accurately predicting the eye status and thus the blinking status. An accurate and precise algorithm employing the five weighted features is proposed to predict eye status (open/close). One state-of-the-art dataset ZJU (eye-blink), is used to measure the performance of the method. Precision, recall, F1-score, and ROC curve measure the proposed method performance qualitatively and quantitatively. Increased accuracy (of around 97.2%) and precision (97.4%) are obtained compared to other existing unimodal approaches. The efficiency of the proposed method is shown to outperform the state-of-the-art methods.

## KEYWORDS

Eye Blink, Multimodal, Z score Threshold, Weighted Features.

DOI: 10.9781/ijimai.2021.11.002

## I. INTRODUCTION

**E**YE blinking is partly unintended closing and reopening of the eyelid. Muscles that help in closing/opening the eye are Orbicularis oculi and levator palpebrae superioris [1]. Eye blinking assists in cleaning and moistening the eye cornea. Blinking can be categorized into three main classes: spontaneous blinking, reflex blinking, and voluntary blinking. The first two lie under the category of involuntary blinking. On the contrary, voluntary blinking can be invoked intentionally within the control of a subject. The work emphasizes detecting the eye blinks, specifically from voluntary blinking, which can be encrypted to alarm emergency [2]-[3].

Due to its importance in many applications (driver drowsiness, human-computer interaction, micro-expression detection), an exponential rise is seen in the research field related to eye blinking detection. Many innovative, unconventional, and robust methods have been stated in the literature [4]-[11]. In the current scenario, multiple robust real-time facial landmark detectors [12]-[15] that arrest virtually all the distinctive features on a human face are accessible. Several approaches have been proposed to routinely detect blinks [9],[12] from the eyes (from both still [16] and video sequences [17]-[19]). But the disadvantage is that they impose rigid requirements on the setup, in the sense of a relative head alignment, resolution, illumination, etc. This work puts forward an effective method to classify eye-blinks [20]-[22] to overcome the stated complications, extracting five distinct weighted features (explained in detail in section III). In this

work, using the unique circles formed from the eyelids landmarks, five novel features (depicting imperative information) have been extracted, which are further assigned weights, thus resulting in an accurate eye blink detection system. The motivation and research question behind the work is depicted in Fig. 1.

### A. Key Contribution of the Paper

A novel multimodal approach for eye blink detection using low-level eyelid feature extraction and z-score thresholding is proposed.

- Accurate face detection and eye localization method is used (MTCNN – Multitask Cascaded Neural Network).
- Unique circles are formed from the eyelids landmarks.
- The eye blinks are detected using an amalgam of 5 weighted features (Vertical Head Positioning, Orientation Factor, Proportional Ratio, Area of Intersection, and Upper Eyelid Radius) depicting imperative gen (z score threshold), extracted from the circles uniquely formed from the eyelids landmarks.
- A fusion technique is described to fuse the extracted five features into a single binary signal.
- Finally, a Z-score-based thresholding algorithm is proposed to extract peaks from the signal where peaks correspond to eye blinks.

The rest of the work is structured as: Section I familiarizes the blink concept and the essential contributions. Section II reviews the related literature. Feature Extraction of the multimodal eye blink system is elaborated in section III. The methodology used, database, performance parameters, investigational setup, and blink prediction are described in section IV. Experimental results and conclusions are analyzed in sections V and VI.

\* Corresponding author.

E-mail address: gopal.chaudhary88@gmail.com

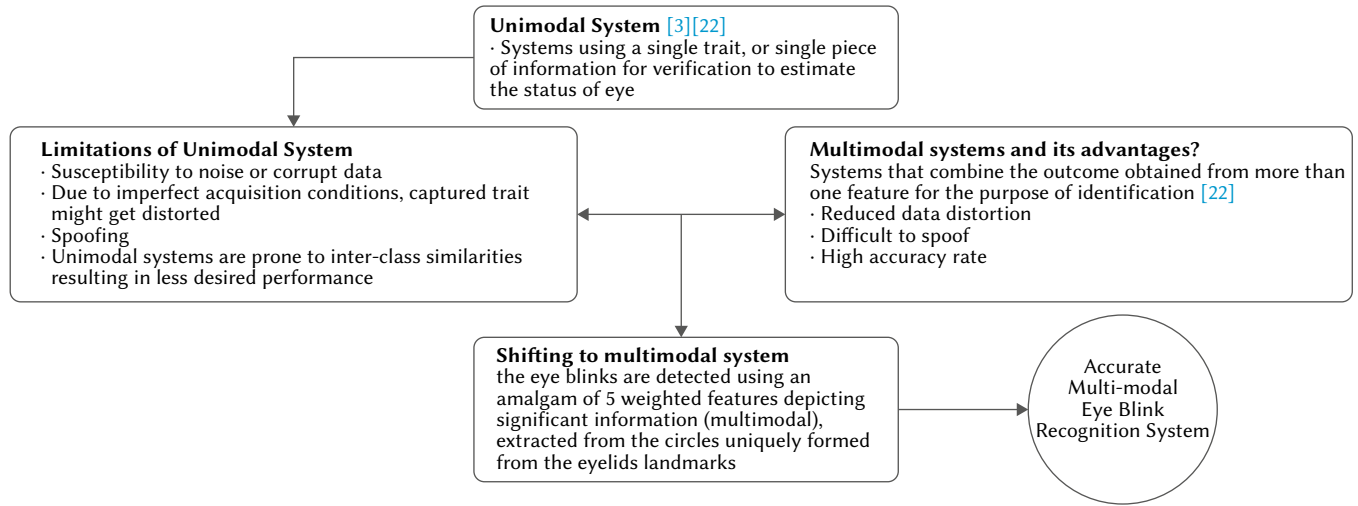


Fig.1. Motivation and research question.

TABLE I. RELATED WORK WITH PERFORMANCE ATTRIBUTES

S. no.	Author and Year	Year	Dataset Used	Performance	Method employed
Z1	Michael Chau and Margrit Betke. 2005 [39]	2005	Eight test subjects	Accuracy: 95.3	Templates (open/close eye) are used in determining the state of an eye.
Z2	Diego Torricelli et al. 2009 [40]	2009	ZJU eye blink dataset	Accuracy: 95.7	Eye-Blinking is sensed without any constraints on the head pose or target position concerning the camera.
Z3	W.O. Lee et al 2010 [8]	2010	ZJU eye blink dataset	Precision: 94.4	Multiple features are extracted to sense the state of the eye. 1) the fraction of elevation breadth of the eye 2) the aggregate alteration of the no. of black pixels around the eye area using an adaptive threshold in consecutive image frames.
Z4	Kohei Arai and Ronny Mardiyanto. 2010 [21]	2010	Individual users used	N.A	Gabor filters are applied to detect the eye arcs, and eye-blink is sensed by figuring the arcs' distance.
Z5	Tomas Drutarovsky and Andrej Fogelton. 2014 [5]	2014	ZJU eye blink dataset	Accuracy: 93.45	Vertical motions in the eye are examined to sense the blinks.
Z6	Fengyi Song et al. 2014 [6]	2014	ZJU eye blink dataset	Accuracy: 96.8	The strength of multiple features collectively characterizes information (eye patches).
Z7	Tereza Soukupova and Jan Cech. 2016 [1],[7]	2016	ZJU eye blink dataset	N.A	Eye Aspect Ratio (EAR) is calculated to determine the eye status (open/closed) in a video sequence
Z8	Federico M. Sukno et al. 2016 [20]	2016	AV@CAR	N.A	The blink detection procedure outputs freq, durations, and some fraction metrics.
Z9	P. Singh and D. Virmani. 2018 [3]	2018	ZJU eye blink dataset	Accuracy: 97	The blinks are perceived using a 13-dim sequential window calculated using eye facet correlation (Ratio of diagonal distance and width of eye landmarks), which outclasses the threshold method and [7].
Z10	P. Singh and D. Virmani. 2020 [22]	2020	ZJU eye blink dataset	Accuracy: 99.02 Precision: 99.65	A precise multimodal eye-blink recognition method using feature-level fusion (MmERMFLF) to detect the eye status.
Z11	Al-gawwam, S., & Benaissa, M. [37]	2018	ZJU eye blink dataset, eyeblink and talking face	Precision: 100%, 96.65% and 98.38%	The method approximates the facial-landmark positions and excerpts the vertical distance between eyelids. A Savitzky-Golay (SG) filter is used to level the signal while keeping the peak info to sense eye-blinks.

## II. RELATED STUDY

In recent times a lot of research has been conducted in different biometric modalities wherein different modalities have proven their prevalence. Recognition systems (RS) (gait, voice, and face) are efficient but lack one or the other aspects. In Gait RS data procurement is ineffectual. Voice RS is an effectual process; but info procurement is a cumbrous errand. Face RS is burdened by the inevitable dispute of aging. But due to their stability and efficiency, eye blink recognition

systems does not suffer from the limitations mentioned above. Table I reveals prior work based on functionality engaged in detecting blinks. Z1–Z11 uses the ZJU dataset to detect the blinks accurately. A solitary feature engages in computing an eye status in [Z1, Z2, Z4, Z5, Z7–Z9, Z11]. While multiple features are used in [Z3, Z6, and Z10] to sense blinks, subsequently creating a multimodal environment. This work emphasizes feature-level fusion (FLF) of weighted features extricated to sense the eye blinks accurately.

In recent times eye-blink has been detected using two primary techniques: electrooculography (EOG) [23]-[24] and videooculography (VOG) [25]-[34]. Ag/AgCl electrodes are used in the former technique and equipped around the eye area to detect eye movements and blinks [23], [35], [36], [38]-[43]. Any movement (left, right, up, down) can be perceived. However, the variability of electrooculography rests on factors that are hard to control [24]. The electrode placement also causes troublesomeness for the user, and the method is relatively costly [28]. In the latter technique, the subject is in sight of the line of the camera installed, and the frames are processed one by one, resulting in the eye's final status (open/close) using an algorithm [3], [22]. All significant research in this area nowadays uses the latter method (VOG). Several VOG-based blink detector techniques are available but lack in one or the other area. This paper proposes a novel multimodal approach for eye blink detection using low-level eyelid feature extraction and z-score thresholding. Firstly, accurate face detection and eye localization method are used (MTCNN). Multitask Cascaded Neural Network or MTCNN is a convolutional neural network architecture jointly trained for facial landmark detection and alignment, achieving superior accuracy over the state-of-the-art techniques on the challenging Fddb and WIDER FACE benchmark for face detection, and AFLW benchmark for face alignment, while keeping real time performance [44]-[46].

Further, unique circles are formed from the eyelid's landmarks, and eye blinks are detected using an amalgam of 5 weighted features (Vertical Head Positioning, Orientation Factor, Proportional Ratio, Area of Intersection, and Upper Eyelid Radius) depicting imperative gen (z score threshold). A unique fusion technique is described to fuse the extracted five features into a single binary signal. Finally, a Z-score-based thresholding algorithm is proposed to extract peaks from the signal where peaks correspond to eye blinks. No work has been done on sensing blinks with low-level eyelid feature extraction and z-score thresholding as per our familiarity. Neither multiple features (multimodal system) have been employed to detect the eye status (open/close) from video sequences.

### III. FEATURE EXTRACTION USING MULTIMODAL EYE BLINK DETECTION SYSTEM

For sensing the eye blinks, feature extraction remains the most prominent step. This section focuses on detecting the face and eyes using landmarks and feature extraction. Face and eye detection using landmarks are explained in sec. A and the proposed feature extracted are explained in detail and relegated in sec. B.

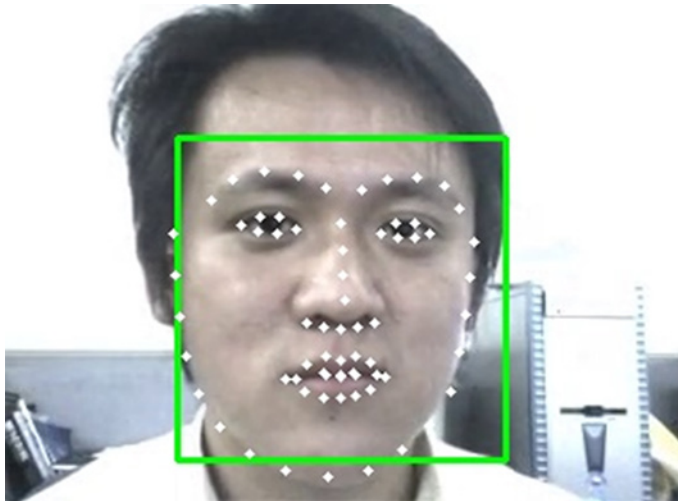


Fig. 2. Face landmarks extracted from the input video stream (ZJU dataset).

#### A. Face Detection and Eye Landmark Prediction

Traditional systems and approaches use techniques that directly localize eyes in an image or prefer detecting face regions first and then localizing eyes within the detected face region. The proposed approach in this work increases true positives for blinking detection and reduces false positives. An accurate eye blink detection system begins from accurate face detection and eye localization, for which we chose the widely used face detection deep learning model MTCNN. Once the face is localized from the input image, we extract eyelid landmarks from the face region detected using dlib, as shown in Fig. 2.

#### B. Feature Extraction

For extracting features, the landmarks around the eyelids are used. The 12 landmarks detected as the eyelids (6 for each eye) are separated into upper eyelid landmarks and lower eyelid landmarks. The four edge landmarks of the eyes belong to both groups. We first calculate the line's midpoint joining the two points lying on the eyelids in each group. Then, along with the edge points and the midpoint just calculated, a unique circle that passes through these 3 points is created. Fig. 3 shows the entire procedure where white dots are the original landmarks and the line's midpoint is shown in red. Circle formation for both open and closed eyes are depicted in Fig. 4.

##### 1. Feature Extracted

Five sets of features are extracted from the four circles generated:

###### a) Vertical Head Positioning (VHP)

For approximating the positioning of the head/eyes, we calculate the distance between the center of the upper and lower eyelid circles for both the eyes.

The mean of these two distances gives VHP:

$$VHP = \frac{l_1 + l_2}{2} \quad (1)$$

where  $l_1$  and  $l_2$  are the distances between centers of the left and right eye circles as shown in Fig. 5.

###### b) Orientation Factor (OF)

For calculating the head/eyes orientation, we calculate the distance between the center of the upper eyelid circle and lower eyelid circle of the left eye and upper eyelid circle and lower eyelid circle of the right eye, respectively.

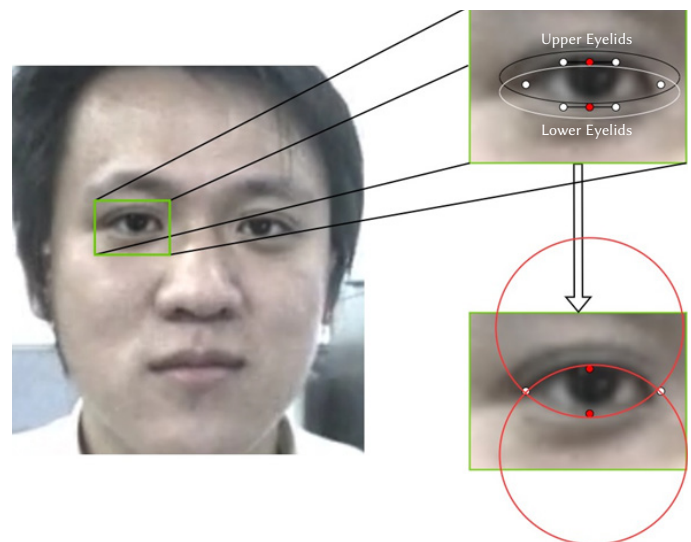


Fig. 3. Upper and lower eyelid landmarks and the circle generated using the points.

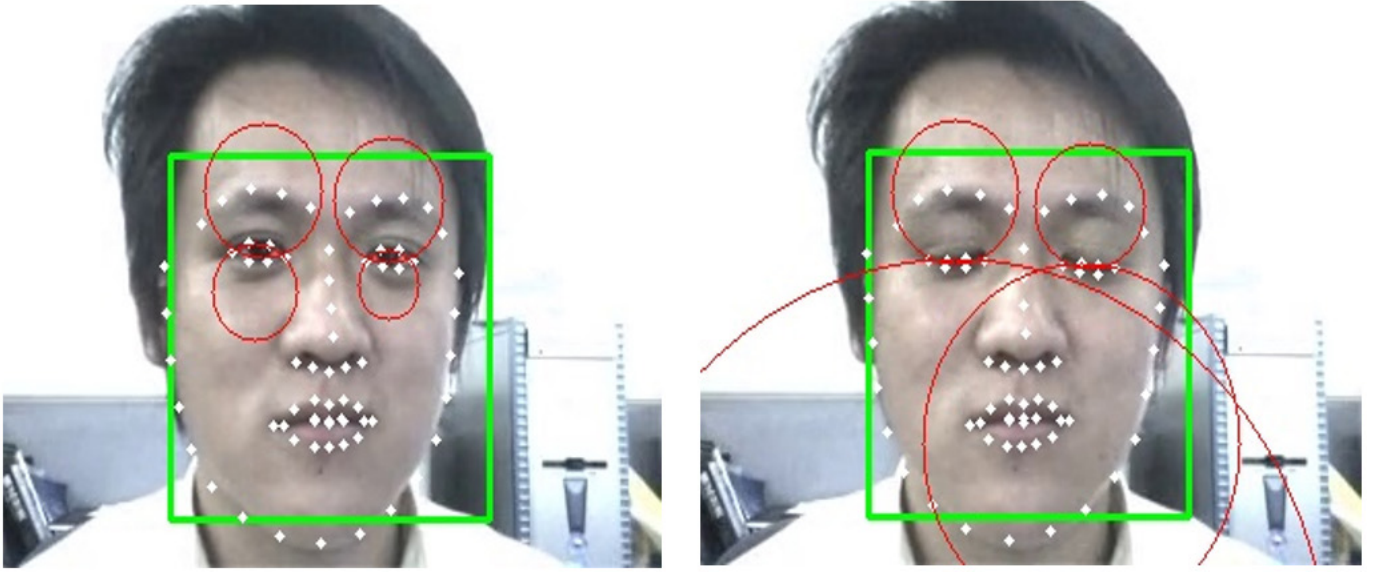


Fig. 4. Circles generated using both the eyelid landmarks for open (left) & closed (right) eyes.

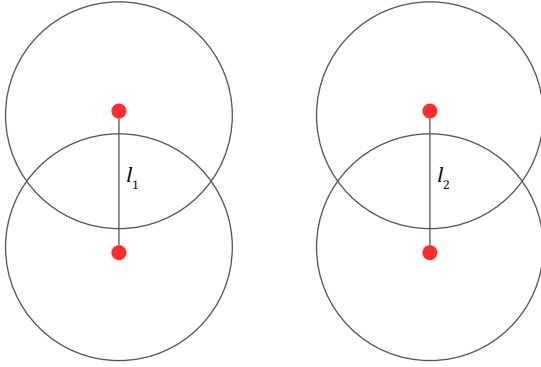


Fig. 5. Distance between the centers of the left and right eye.

The ratio of these two distances gives  $OF$ :

$$OF = \frac{l_1}{l_2} \quad (2)$$

where  $l_1$  and  $l_2$  are distances between the centers of the circles as shown in Fig. 6.

#### c) Proportional Ratio (PR)

To calculate the circle size proportional to the triangle perimeter formed by the eyelid's landmarks, we calculate the upper eyelid circle radius to both eyes' lower radius eyelid circle.

The mean of these ratios gives  $PR$  as follows:

$$PR = \frac{r_1 + r_2}{2} \quad (3)$$

$$r_1 = \frac{lr_{up}}{lr_{down}} \quad (4)$$

$$r_2 = \frac{rr_{up}}{rr_{down}} \quad (5)$$

where  $lr_{up}$ ,  $lr_{down}$ ,  $rr_{up}$  and  $rr_{down}$  are radius of left upper eyelid circle, left lower eyelid circle, right upper eyelid circle and right lower eyelid circle respectively as depicted in Fig. 7.

#### d) Area of Intersection (AOI)

$AOI$  is the area of intersection between the upper eyelid circle and lower eyelid circle of either eye. In our work, we have used the  $AOI$  of the left eye. For experiment purposes, either of the eyes can be used.

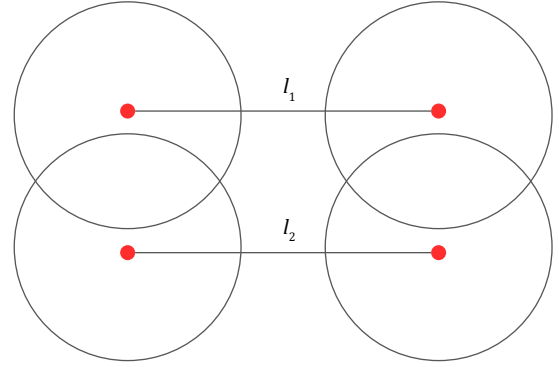


Fig. 6. Distance between the center of the upper eyelid circle and lower eyelid circle of the left eye and upper eyelid circle and lower eyelid circle of the right eye.

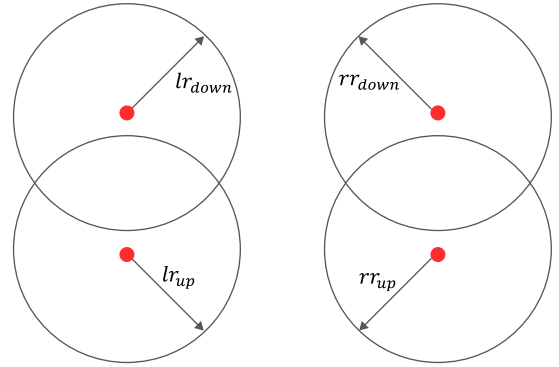


Fig. 7. Radius of left upper eyelid circle, left lower eyelid circle, right upper eyelid circle, and right lower eyelid circle.

We can also take the mean of  $AOI$  of both the eyes. Fig. 8 shows the  $AOI$  of an upper and lower eyelid.

#### e) Upper Eyelid Radius (UER)

$UER$  is the upper eyelid radius of either eye. In our experiments, we have used  $UER$  of the left eye. Either of the eyes can be used for experiments, as in the case of  $AOI$ . We can also take the mean of  $UER$  of both the eyes.



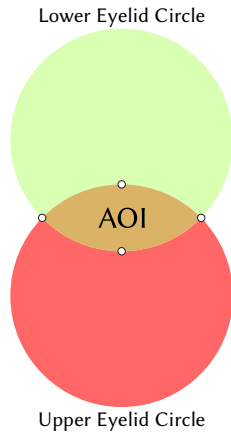


Fig. 8. The AOI of an upper and lower eyelid.

#### IV. RESEARCH METHODOLOGY USED

This section focuses on the methodology that is employed by the method for sensing the blinks. First of all, the dataset employed for the experimental purpose is discussed (section A). Then the performance parameters that judge the efficiency of the method are explained in section B. Experimental setup, blink detection/prediction are explained in detail in subsequent sections.

##### A. Dataset Used

One state-of-the-art dataset ZJU (eyeblink) [18], is used to evaluate the proposed method's performance. The dataset consists of 80 video sequences of 20 users, as depicted in Table II. All the video sequences (4 to 5 sec) are recorded in 30 fps having two to six voluntary blinks. In total, there are 255 blinks in the entire dataset.

TABLE II. ZJU EYE-BLINK DATASET

No. of Users	Four Clips per person			
	No. of Clips	View	Spectacles	Total Blinks
20	1	Front	N	255
	1	Front	Y	
	1	Front	Y	
	1	Front	N	

##### B. Performance Parameters

In this paper, eye blinks are detected using an amalgam of 5 weighted features depicting imperative gen (z score threshold). A unique fusion technique is described to fuse the extracted five features into a single binary signal. Finally, a Z-score-based thresholding algorithm is proposed to extract peaks from the signal where peaks correspond to eye blinks. The performance of the method discussed is evaluated using several parameters as shown below:

1. **Recall (R)** can be defined as the ratio of the correct categorized cases of a class to the total number of class cases.

$$P = \frac{TP}{TP + FN} \quad (6)$$

2. **Specificity (Sp)** It deals with the proportion of actual negatives (true) that are correctly identified as such.

$$Sp = \frac{TN}{FP + TN} \quad (7)$$

3. **Precision (P)** can be defined as the ratio of correct positive predictions to the total number of positive predictions.

$$P = \frac{TP}{TP + FP} \quad (8)$$

4. **Negative Predictive Value (NPV)**: The ratio of accurate negative class predicted to the over-all amount of negative predictions.

$$NPV = \frac{TN}{TN + FN} \quad (9)$$

5. **False Positive Rate (FPR)**: The ratio of false-positive class predictions to the overall negatives.

$$FPR = \frac{FP}{FP + TN} = 1 - Sp \quad (10)$$

6. **False Discovery Rate (FDR)**: The ratio of false-positives predictions to the overall positives.

$$FDR = \frac{FP}{FP + TP} \quad (11)$$

7. **False Negative Rate (FNR)**: The ratio of false-negatives predictions to the aggregate number of positives.

$$FNR = \frac{FN}{FN + TP} \quad (12)$$

8. **Accuracy** can be defined as the ratio of correctly classified instances to the total number of instances.

$$P = \frac{TP + TN}{P + N} \quad (13)$$

9. **F1 Score (F1)** can be defined as weighted avg. of precision and recall

$$P = \frac{2 * TP}{* TP + FP + FN} \quad (14)$$

10. **AUC Score** is used to measure the performance of a model at different threshold settings.

(**Note:** TP: True +, TN: True -, FP: False +, and FN: False -)

##### C. Experimental Setup

We have used ZJU (eye-blink) dataset (explained in sec. A) for our experiments. The five extracted set of features supports blinking with sudden peaks in their respective plots. Fig. 9 shows the plot (five extracted features (explained in sec. III) vs. the number of frames) from one of the dataset's input videos.

It can be inferred from the plots that frames 36 to 43 exhibit a peak in all four plots except Orientation Factor (OF) plot. It remains uniform with a value greater than 1, indicating standard behavior of other features. Similar trends can be seen at frames 85 to 92 and 130 to 140, indicating blinks at those frames in the input video sequence. The case discussed in Fig. 9 is a clear case where all the features show a positive consent as the subject is sitting in the camera's line of sight and not wearing any spectacles. On the contrary, Fig. 10 shows the plot of all five features for another input video sequence of the dataset with disturbance and other perturbations. The subject in the input video sequence is blinking in a dark background. Here, feature OF is not uniform, and in certain intervals, its value drops below 1. This behavior is an indication that other features are not behaving normally and are affected by the perturbation in the input video sequence, except Upper Eyelid Radius (UER), which still produces peaks at four intervals supporting blinks in the video input. With further experiments, we found that when the value of Orientation Factor (OF) drops below 1, Upper Eyelid Radius (UER) is the most dominant one to report a blink with a peak, while the other three features produce sloppy peaks. We used weighted signal averaging or convolution conditioned on OF to implement this behavior in an automated system. The same behavior is seen when the subject is wearing spectacles also.

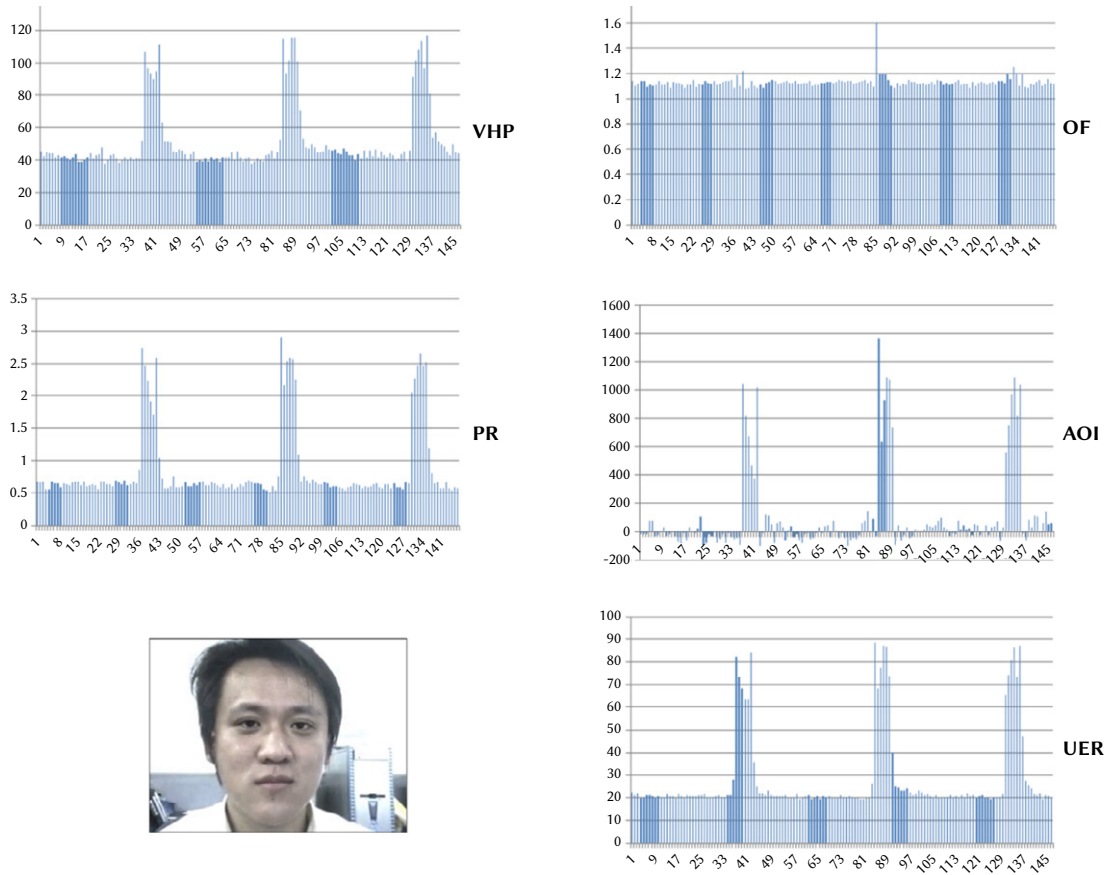


Fig. 9. The plot of all five features extracted vs. frames from an input video sequence along with a frame from the video.

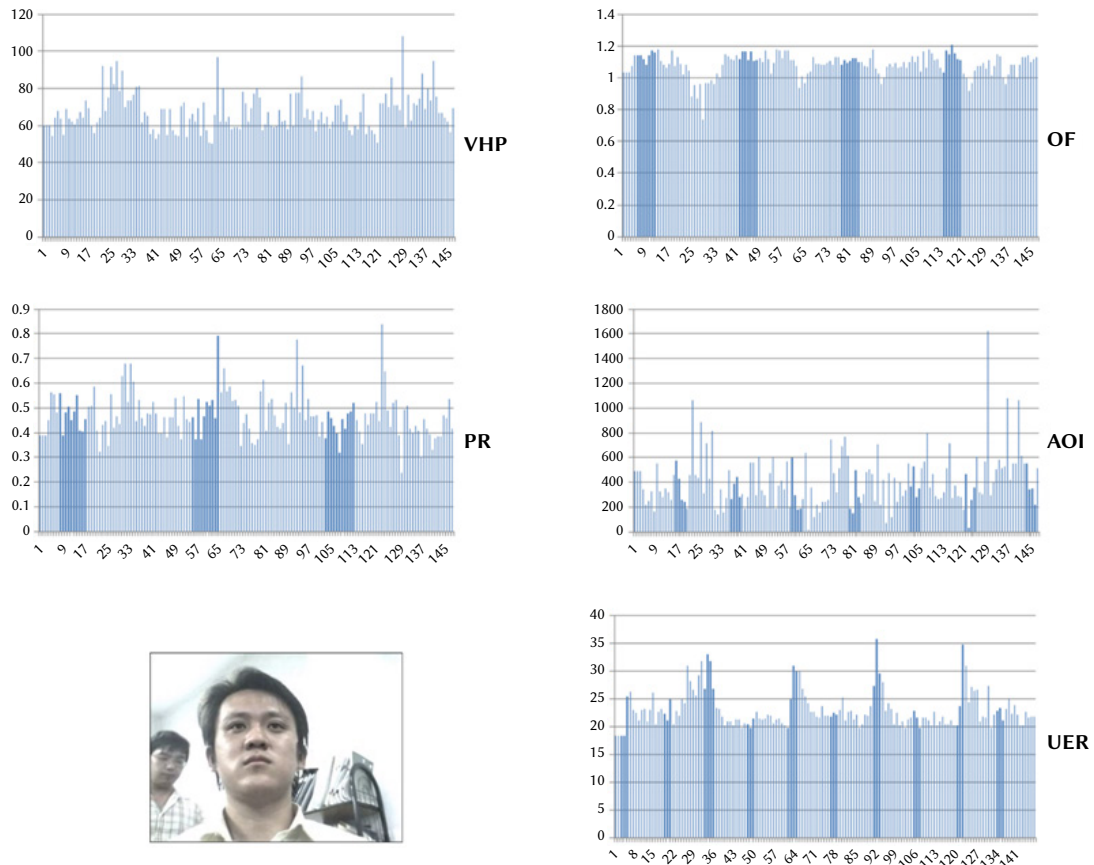


Fig. 10. The plot of all five features for an input video sequence with disturbance and other perturbations along with a frame from the video.

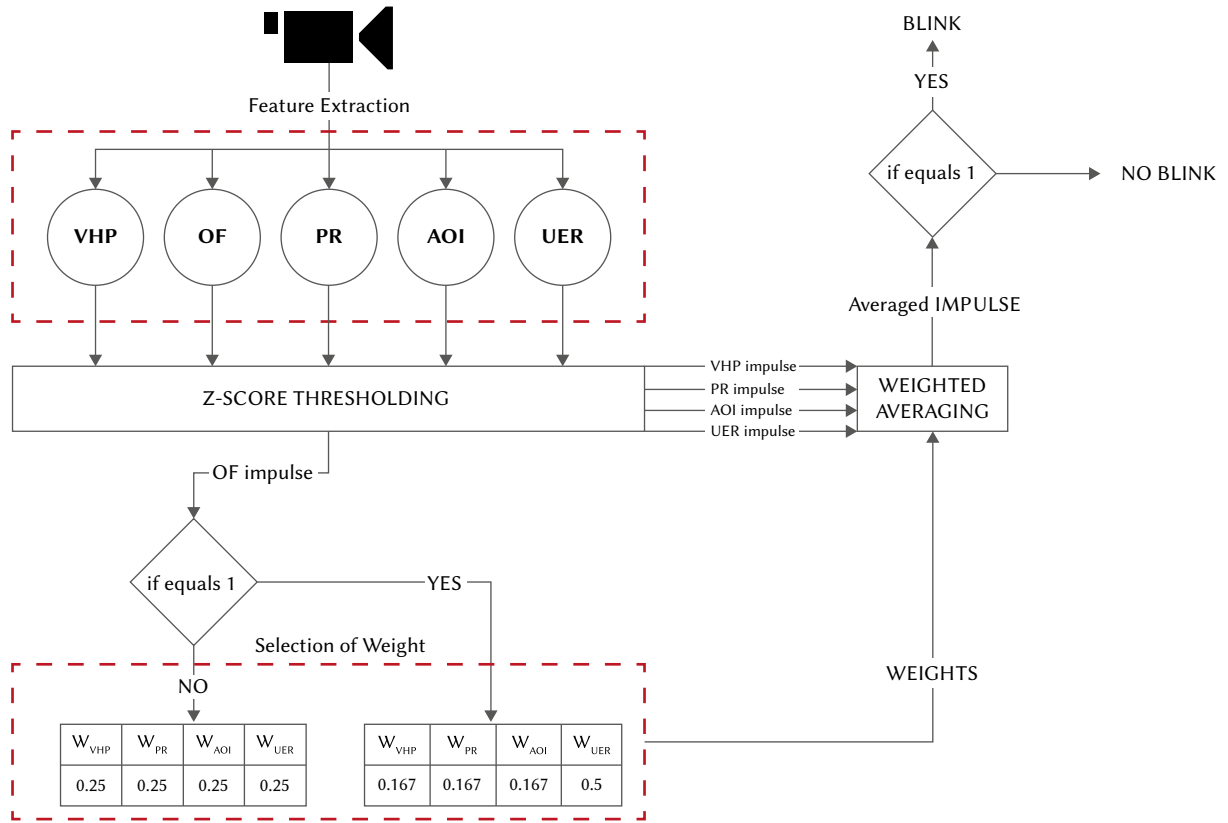


Fig. 11. Flowchart of the multimodal eye-blink detection system.

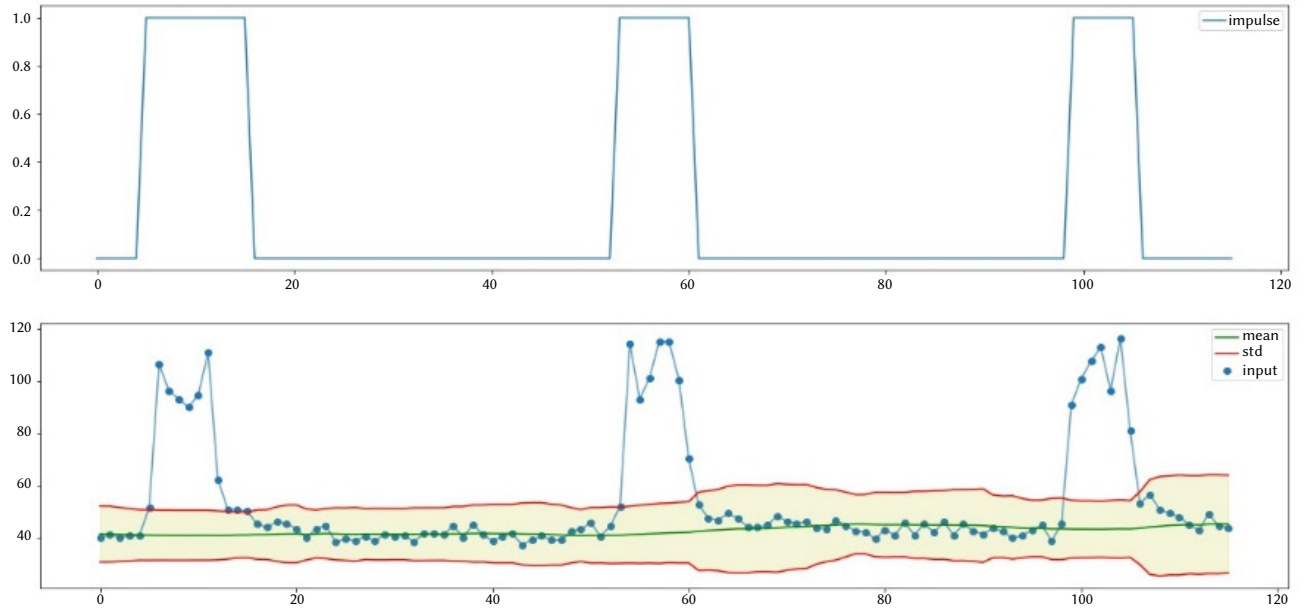


Fig. 12. An arbitrary input signal and its generated unit impulse signal using the method.

#### D. Blink Detection

Automated blink detection is a methodology consisting of peak detection and weighted averaging. We generate a unit impulse signal for peaks in all four features individually and interpolate these signals conditioned on the impulse signal of Orientation Factor (OF). Fig. 11 depicts the entire workflow with a flowchart.

- a) **Peak Detection:** A simple way to detect peaks in a signal is hard thresholding with a predefined value. But in our experiments, we

found that the scale of the extracted features depends vastly on multiple factors, like, camera position, head positioning, landmark detector accuracy, etc. Therefore, direct hard thresholding won't be able to detect peaks for our input in real-time settings. Therefore, we derived a more robust, yet simple z-score based peak detection algorithm inspired by probability distributions properties [41]-[43] for peak detection. Rather than keeping a set threshold value for signaling peaks in a given input signal, the algorithm calculates the moving mean and moving standard deviation of the

input signal. When a data point lies the preset threshold times the standard deviation away from the mean, the algorithm signals a peak. The approach of calculating a moving mean and standard deviation keeps the method's integrity at different scales. The algorithm can be made more robust to outliers and signal length by calculating the moving mean and standard deviation by only looking at the last 'k' values, where k is a whole number. Fig. 12 shows an arbitrary input signal and its generated unit impulse signal using our method.

b) **Weighted averaging of signals:** Fig. 13(a) shows the impulse signals of VHP, PR, AOI, and UER, and Fig. 13(b) shows the impulse

signals of OF generated using hard thresholding at 1 (if the signal is less than 1, an impulse is generated). When there is an impulse in OF, it means that other features except UER will struggle to report a blink with a peak, whereas when OF has no impulse signals, it suggests the normal behavior which peaks in all other four features. Fig. 14(a) and Fig. 14(b) show the same feature impulses for an input video sequence with abnormal behavior. Table III shows the feature weights we used to average the impulse signals.

For  $OF = 0$ , we have averaged over all the four impulse signals, and for  $OF = 1$ , we have given UER a weight of 0.5, and the rest three impulse signals are given equal weightage out of 0.5, i.e.,  $0.5/3 = 0.167$ .

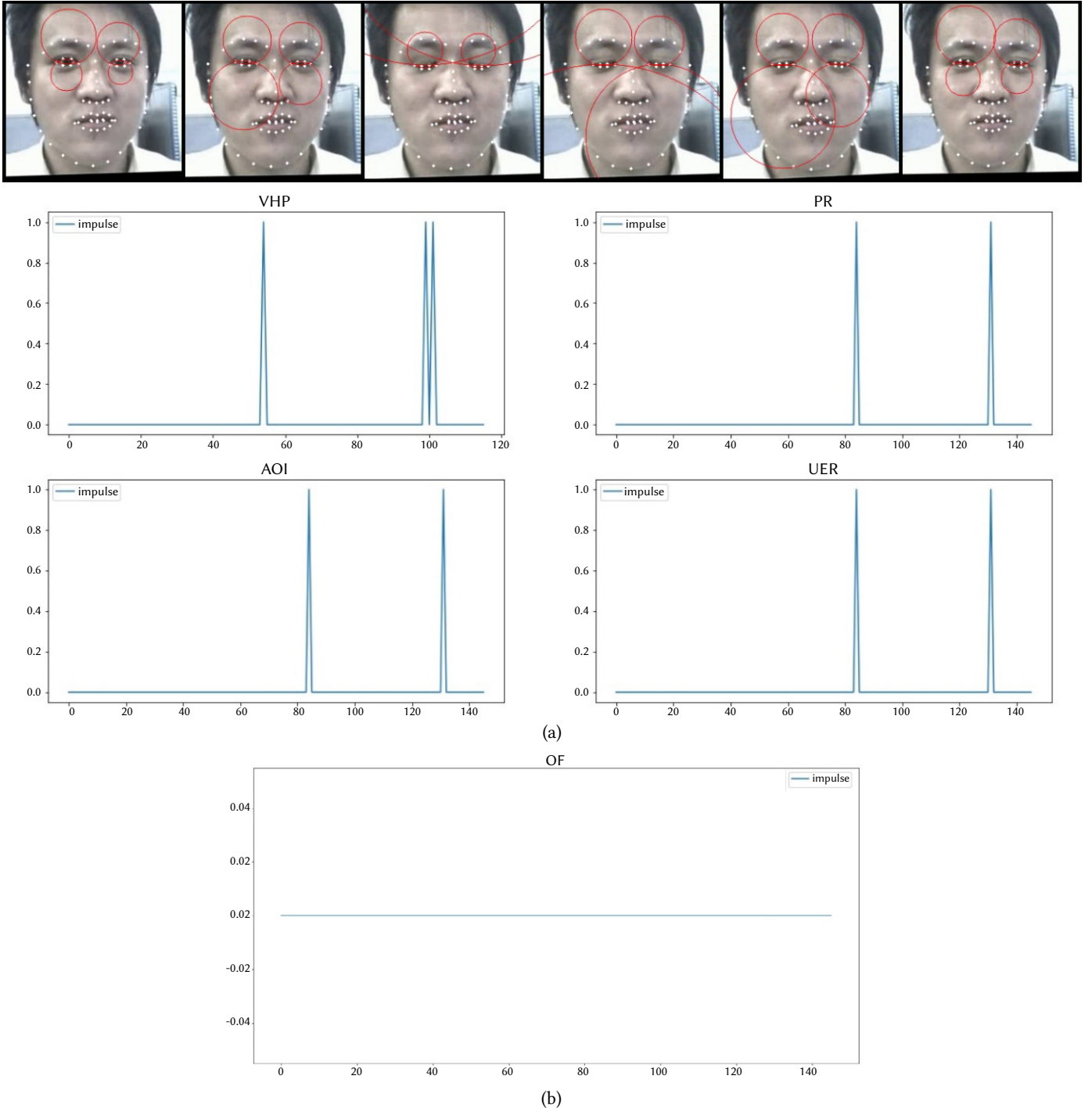


Fig. 13(a). The impulse signals of VHP, PR, AOI and UER. (b). The impulse signals of OF generated using hard thresholding at 1 (if the signal is less than 1, an impulse is generated).



TABLE III. THE FEATURE WEIGHTS WE USED TO AVERAGE THE IMPULSE SIGNALS

$w_i$	if OF = 0				if OF = 1			
	$W_{VHP}$	$W_{PR}$	$W_{AOI}$	$W_{UER}$	$W_{VHP}$	$W_{PR}$	$W_{AOI}$	$W_{UER}$
	0.25	0.25	0.25	0.25	0.167	0.167	0.167	0.5

Equation (15) gives the final output averaged signal that reports a peak for a blink in the input video sequence. The step-by-step process is depicted in algorithm 1.

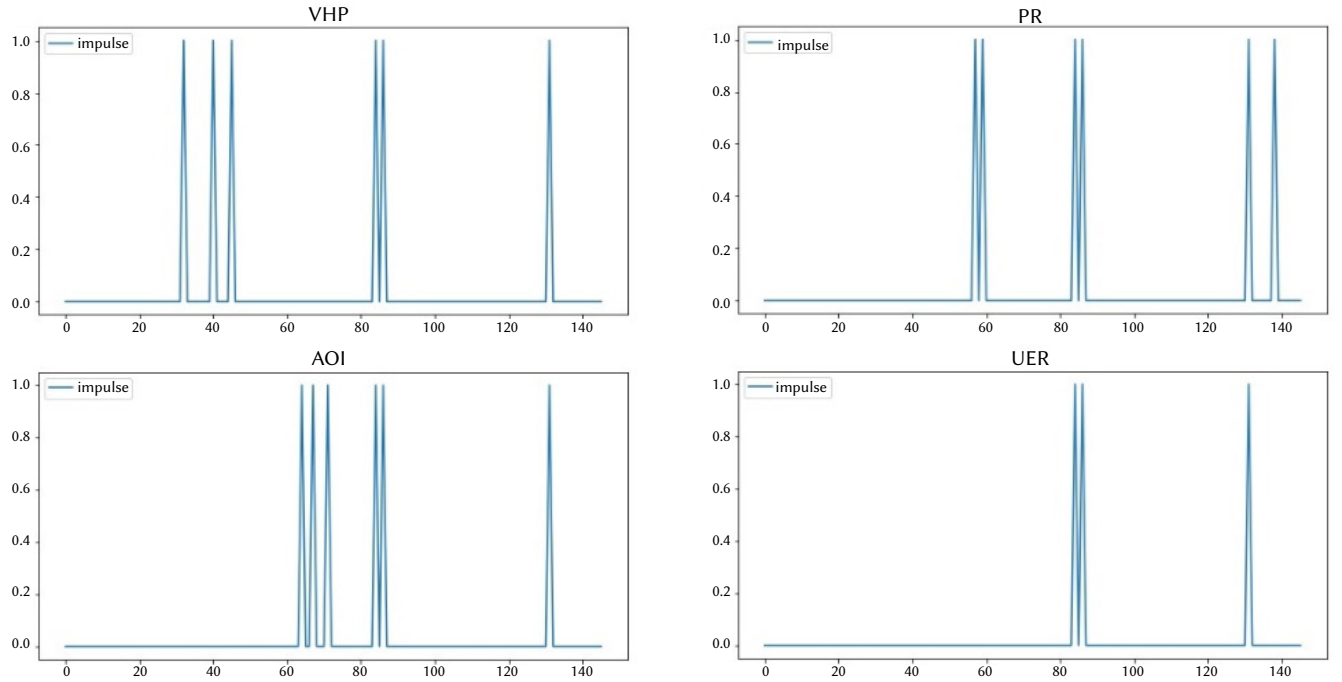
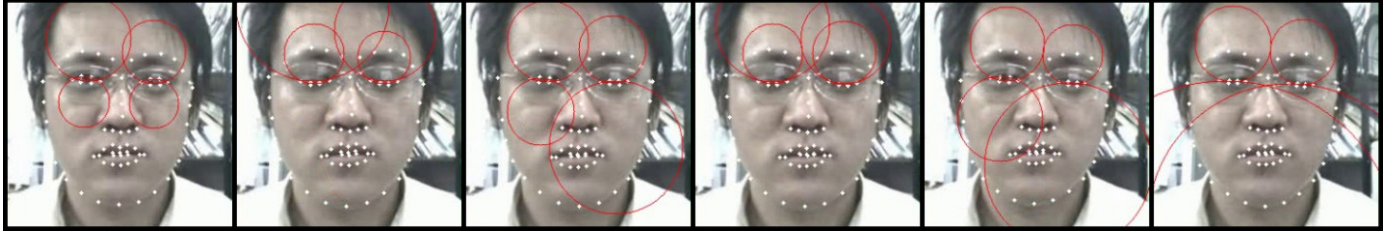
$$y^t = \sum_{i=1}^4 w_i x_i^t \quad (15)$$

Where  $y^t$ : output averaged signal at time  $t$

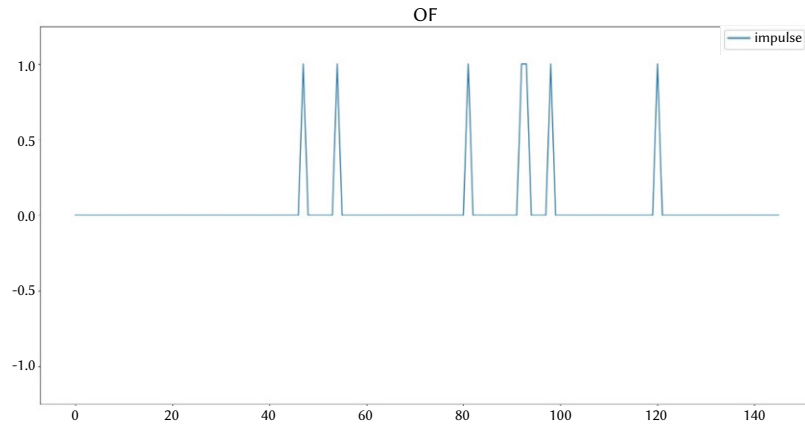
$W_i$ : the weight of feature  $i$  from Table III, which is conditioned on OF feature at time  $t$ .

#### E. Blink Prediction

Once the final signal  $y^t$  is obtained, we hard threshold the signal with a value 0.75 such that when the signal value is greater than or equal to 0.75, a unit impulse is considered, and a blink is predicted. The threshold we used favors predicting a blink only when 75% of features



(a)



(b)

Fig. 14(a). The impulse signals of VHP, PR, AOI and UER. (b). The impulse signals of OF generated using hard thresholding at 1(if the signal is less than 1, an impulse is generated).

report a peak. Future implementations can use various threshold values for the confidence of prediction while incorporating more handcrafted and deep learning extracted features.

**Algorithm 1:** Algorithm for Multimodal Eyeblink Detection system:

1. Capture video stream frames from the input camera.
2. For each frame:
  - a. Localize face from the frame using Multitask Cascaded Neural Network (MTCNN).
  - b. Detect Eye landmarks from the extracted face region of the frame.
3. For each eye:
  - a. Separate the 12 eye landmarks detected into the upper eyelid and lower eyelid landmarks.
  - b. For each group, calculate the midpoint of the line joining the two points lying on the eyelids.
  - c. Then, along with the edge points and the midpoint just calculated, create a unique circle that passes through these 3 points using the algorithm presented in appendix A.
  - d. Using the two circles created, calculate the five sets of features: Vertical Head Positioning, Orientation Factor, Proportion Ratio, Area of Intersection, and Upper Eyelid Radius.
  - e. Perform Z-Score thresholding on all these five feature sets.
  - f. Select the set of weights to be applied to normalize the impulse signals resulting from the z-score thresholding using table III and perform the weighted averaging.
4. Combine the final signal for both the eyes using signal averaging. (Optional)
5. Report a blink if the final signal's magnitude is greater than the threshold (0.75) chosen.

## V. EXPERIMENTAL RESULTS

A multimodal human eye blink detection system is implemented using five weighted features, namely Vertical Head Positioning, Orientation Factor, Proportional Ratio, Area of Intersection, and Upper Eyelid Radius. The five unique features are extracted from a unique circle formed from eye landmarks. The features extracted are used to depict the positioning and orientation of a subject's head/eyes being tested. While observing the features experimented on ZJU dataset, two sets of patterns have resulted, as shown in Fig. 9 and 10. In Fig. 9, a clear pattern is observed wherein all features contributed equally

and showed expected yields. Whereas in Fig. 10, a different pattern is observed where the UER feature is given more weightage than the rest of the features. A clear pattern is observed when the subject is not wearing spectacles in decent light conditions.

On the contrary, when a subject is wearing spectacles or the light conditions are not good, a blurred pattern is observed. A multimodal eyeblink detection system is proposed and implemented to balance out the two patterns, as shown in Fig. 11. For testing the proposed system's performance, a state-of-the-art dataset, ZJU, is used. Table IV shows how the proposed system has performed in terms of performance parameters. A reputable accuracy and precision of 97.2% and 97.4% are achieved. For testing purposes, 60% of the dataset is employed. ROC graph along with an AUC score of 0.972 resulted is shown in Fig. 15.

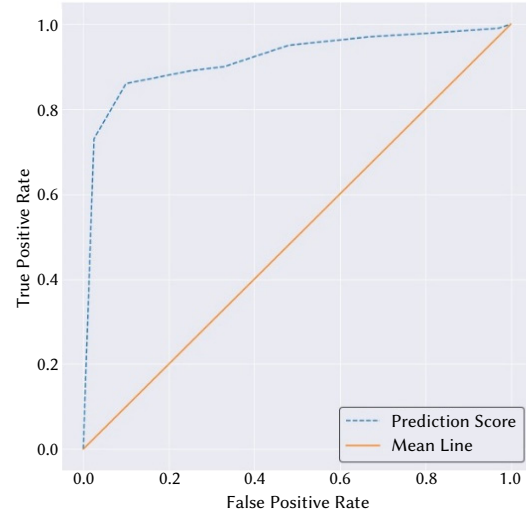


Fig. 15. ROC Curve.

### A. Analysis of the Results

Table V shows a comparative accuracy comparison of the proposed work with existing methodologies. It can be clearly seen that the proposed method has outperformed the existing methodologies by a margin. Increased accuracy and precision of 97.2% and 97.4% is achieved when a multimodal approach with a weighted feature set is used. A remarkable false positive, discovery and a negative rate of 0.025, 0.025, and .003 is perceived in Table IV.

TABLE IV. PERFORMANCE PARAMETERS FOR THE MULTIMODAL EYE-BLINK DETECTION SYSTEM

Recall	Specificity	Precision	Negative Predictive Value	False Positive Rate	False Discovery Rate	False Negative Rate	Accuracy	F1 Score	AUC score
0.97	0.975	0.974	0.970	0.025	0.025	0.03	0.972	0.972	0.972

TABLE V. COMPARATIVE ACCURACY COMPARISON OF THE PROPOSED WORK WITH EXISTING METHODOLOGIES

S. NO.	Author and Work Reference	Performance	Accuracy and precision of the proposed work
1	Michael Chau and Margrit Betke. 2005 [39]	Accuracy: 95.3	Accuracy: 97.2 Precision: 97.4 Recall: 97.0
2	Diego Torricelli et al. 2009 [40]	Accuracy: 95.7	
3	W.O. Lee et al 2010 [8]	Precision: 94.4	
4	Tomas Drutarovsky and Andrej Fogelton. 2014 [5]	Accuracy: 93.45	
5	Fengyi Song et al. 2014 [6]	Accuracy: 96.8	
6	Tereza Soukupova and Jan Cech. 2016 [7]	Recall: 92.9	
7	Federico M. Sukno et al. 2016 [20]	Frame classification acc.:97.1	
8	P. Singh and D. Virmani. 2018 [3]	Accuracy: 97	

## VI. CONCLUSION

A multimodal eye blink recognition system using Z-score-based thresholding and weighted features was presented in this work. The eye blinks were detected using a blend of 5 weighted features (Vertical Head Positioning, Orientation Factor, Proportional Ratio, Area of Intersection, and Upper Eyelid Radius) depicting imperative gen (z score threshold), extracted from the circles uniquely formed from the eyelids landmarks. For testing the performance of the method, ZJU eye-blink dataset was used. While implementing the proposed method with the said dataset, it was observed that when there is an impulse in OF, it means that other features except UER will struggle to report a blink with a peak. In contrast, when OF has no impulse signals, it suggests the expected behavior, which peaks in all four features. The multimodal system's performance was increased to 97.2% (accuracy) with a precision of 97.4%. Other performance parameters also showed a decent routine. As a future scope, more features can be incorporated to increase the performance attributes further.

## APPENDIX

### Procedure to create a unique circle given 3 points:

The algorithm to create a circle given 3 points is as follows:

1. Let  $(x_a, y_a), (x_b, y_b)$  &  $(x_d, y_d)$  be 3 points from which a circle is to be created. Now, consider the general equation of a circle as  $(x - x_c)^2 + (y - y_c)^2 - r^2 = 0$  where  $(x_c, y_c)$  is the center of the circle, and  $r$  is the radius.
2. Given 3 points, plug the values into the equation of the circle.
 
$$(x_a - x_c)^2 + (y_a - y_c)^2 - r^2 = 0 \quad (A1)$$

$$(x_b - x_c)^2 + (y_b - y_c)^2 - r^2 = 0 \quad (A2)$$

$$(x_d - x_c)^2 + (y_d - y_c)^2 - r^2 = 0 \quad (A3)$$
3. Solve the linear equations formed for  $x_c, y_c$  by subtracting equation (A1) from (A2) and (A3).
4. Plug the values for the center  $(x_c, y_c)$  in any of the three quadratic equations and solve for  $r$

### An example:

$$\begin{aligned} x_a, y_a &= 1, 1 \\ x_b, y_b &= 2, 4 \\ x_d, y_d &= 5, 3 \end{aligned}$$

$$\begin{aligned} (1 - x_c)^2 + (1 - y_c)^2 - r^2 &= 0 \\ (2 - x_c)^2 + (4 - y_c)^2 - r^2 &= 0 \\ (5 - x_c)^2 + (3 - y_c)^2 - r^2 &= 0 \end{aligned}$$

$$\begin{aligned} -2x_c - 6(y_c - 3) &= 0 \\ (y_c + 7) - 6x_c &= 0 \end{aligned}$$

$$(x_c, y_c) = (3, 2)$$

$$\begin{aligned} (1 - 3)^2 + (1 - 2)^2 - r^2 &= 0 \\ 5 - r^2 &= 0 \\ r &= \sqrt{5} \end{aligned}$$

## REFERENCES

- [1] T. Soukupova, "Eye Blink Detection using Facial Landmarks," Diploma Thesis, Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University, Prague, 2016.
- [2] Puneet Singh Lamba and Deepali Virmani, "Information Retrieval from Emotions and Eye Blinks with Help of Sensor Nodes," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 4, pp. 2433-2411, 2018.
- [3] Puneet Singh Lamba and Deepali Virmani, "Reckoning number of eye blinks using eye facet correlation for exigency detection," *Journal of Intelligent & Fuzzy Systems*, vol. 35, no. 5, pp. 5279-5286, 2018.
- [4] Stephen Milborrow and Fred Nicolls, "Locating Facial Features with an Extended Active Shape Model," *Computer Vision - ECCV 2008, Lecture Notes in Computer Science*, Springer, 5305 (2008), pp. 504-513.
- [5] T. Drutarovsky and A. Fogelton, "Eye blink detection using variance of motion vectors," In: *Computer Vision - ECCV Workshops*, 2014.
- [6] F. Song, X. Tan, X. Liu and S. Chen, "Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients," *Pattern Recognition, The Journal of the Pattern Recognition Society*, vol. 47, pp. 2825-2838, 2014.
- [7] T. Soukupova and J. Cech, "Real-Time Eye Blink Detection using Facial Landmarks," *21st Computer vision Winter Workshop, Luka Cehovin, RokMandelj, VitomirStruc (eds.) RimskeToplice, Slovenia, February 3-5, 2016*.
- [8] W.H. Lee, E.C. Lee and K.E. Park, "Blink detection robust to various facial poses," *Journal of Neuroscience Methods*, vol. 193, no. 2, pp. 356-372, 2010.
- [9] N. Alsaedi, and D. Wloka, "Real-Time Eyeblink Detector and Eye State Classifier for Virtual Reality (VR) Headsets (Head-Mounted Displays, HMDs)," *Sensors*, vol. 19, no. 5, 2019.
- [10] R. Sanyal and K. Chakrabarty, "Two Stream Deep Convolutional Neural Network for Eye State Recognition and Blink Detection," *2019 3rd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, Kolkata, India, 2019, pp. 1-8.
- [11] C. Zhu and C. Huang, "Adaptive Gabor algorithm for face posture and its application in blink detection," *2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, Chongqing, China, 2019, pp. 1871-1875.
- [12] Muhammad Tayab Khan, Hafeez Anwar, Farman Ullah, et al., "Smart Real-Time Video Surveillance Platform for Drowsiness Detection Based on Eyelid Closure," *Wireless Communications and Mobile Computing*, 2019.
- [13] A. Asthana, S. Zafeoriou, S. Cheng, and M. Pantic, "Incremental face alignment in the wild," in *Conference on Computer Vision and Pattern Recognition*, 2014.
- [14] Gianluca Donato, Marian Stewart Bartlett, Joseph C. Hager, Paul Ekman, and Terrence J. Sejnowski, "Classifying Facial Actions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 10, pp.974-989, 1999.
- [15] Ying-li, Takeo Kanade, and Jeffrey, "Recognizing Action Units for Facial Expression Analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 1-19, 2001.
- [16] Choi, I., Han, S., & Kim, D. (2011, July), "Eye detection and eye blink detection using adaboost learning and grouping," In *2011 Proceedings of 20th International Conference on Computer Communications and Networks (ICCCN)*, 2011, pp. 1-4. IEEE.
- [17] Singh, H., & Singh, J., "Real-time eye blink and wink detection for object selection in HCI systems," *Journal on Multimodal User Interfaces*, vol. 12, no. 1, pp. 55-65, 2018.
- [18] G. Pan, L. Sun, Z. Wu, and S. Lao, "Eyeblink-based anti-spoofing in face recognition from a generic webcam," In *ICCV*, 2007.
- [19] L. Tan, K. Yu, F. Ming, X. Cheng & G. Srivastava, "Secure and Resilient Artificial Intelligence of Things: a HoneyNet Approach for Threat Detection and Situational Awareness," *IEEE Consumer Electronics Magazine*, 2021.
- [20] F. M. Sukno, S.-K. Pavani, C. Butakoff and A.F. Frangi, "Automatic assessment of eye blinking patterns through statistical shape models," In *ICVS*, 2009.
- [21] K. Arai, R. Mardiyanto, "Real Time Blinking Detection Based on Gabor Filter," *International Journal of Human Computer Interaction*, vol. 1, no. 3, pp. 33-45, 2010.
- [22] Lamba, P.S., Virmani, D. & Castillo, O., "Multimodal human eye blink recognition method using feature level fusion for exigency detection," *Soft Comput* vol. 24, pp. 16829-16845, 2020, <https://doi.org/10.1007/s00500-020-04979-5>
- [23] Usakli, A. B., Gurkan, S., Aloise, F., Vecchiato, G., & Babiloni, F., "On the use of electrooculogram for efficient human computer interfaces," *Computational intelligence and neuroscience*, 2010.
- [24] Pander, T., Przybyla, T., & Czabanski, R. (2008, May), "An application of detection function for the eye blinking detection," In *2008 Conference on*



*Human System Interactions*, 2008, pp. 287-291. IEEE.

- [25] Królak, A., & Strumillo, P., "Eyeblink detection system for human-computer interaction," *Universal Access in the Information Society*, vol. 11, no. 4, pp. 409-419, 2012.
- [26] Kraichan, C., & Pumrin, S. (2014, May), "Face and eye tracking for controlling computer functions," In *2014 11th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 2014, pp. 1-6. IEEE.
- [27] MacKenzie, I. S., & Ashtiani, B., "BlinkWrite: efficient text entry using eye blinks," *Universal Access in the Information Society*, vol. 10, no. 1, pp. 69-80, 2011.
- [28] Arai, K., & Mardiyanto, R. (2011, April) "Eye-based HCI with full specification of mouse and keyboard using pupil knowledge in the gaze estimation," In *2011 Eighth International Conference on Information Technology: New Generations*, 2011, pp. 423-428. IEEE.
- [29] Khilari, R. (2010, December), "Iris tracking and blink detection for human-computer interaction using a low resolution webcam," In *Proceedings of the Seventh Indian Conference on Computer Vision, Graphics and Image Processing*, 2010, pp. 456-463.
- [30] Grauman, K., Betke, M., Lombardi, J., Gips, J., & Bradski, G. R., "Communication via eye blinks and eyebrow raises: Video-based human-computer interfaces," *Universal Access in the Information Society*, vol. 2, no. 4, pp. 359-373, 2003.
- [31] Siriluck W, Kamolphiwong S, Kamolphiwong T, "Blink and click," In: *Proceedings of the 1st international convention on rehabilitation engineering and assistive technology: in conjunction with 1st Tan Tock Seng Hospital neurorehabilitation meeting*, 2007, pp. 43-46.
- [32] Pimplaskar, D., Nagmode, M. S., & Borkar, A., "Real time eye blinking detection and tracking using OpenCV," *Technology*, vol. 13, no. 14, pp. 15, 2015.
- [33] Missimer, E., & Betke, M. (2010, June), "Blink and wink detection for mouse pointer control," In *Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments*, 2010, pp. 1-8.
- [34] Su, M., Yeh, C., Lin, S., Wang, P., & Hou, S. (2008, July), "An implementation of an eye-blink-based communication aid for people with severe disabilities," In *2008 International Conference on Audio, Language and Image Processing*, 2008, pp. 351-356. IEEE.
- [35] Venkataramanan, S., Prabhat, P., Choudhury, S. R., Nemade, H. B., & Sahambi, J. S. (2005, January), "Biomedical instrumentation based on electrooculogram (EOG) signal processing and application to a hospital alarm system," In *Proceedings of 2005 International Conference on Intelligent Sensing and Information Processing*, 2005, pp. 535-540. IEEE.
- [36] Kumar, D., & Poole, E. (2002, October), "Classification of EOG for human computer interface," In *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society*[*Engineering in Medicine and Biology*, Vol. 1, pp. 64-67. IEEE.
- [37] Al-gawwam, S., & Benaissa, M., "Robust eye blink detection based on eye landmarks and Savitzky-Golay filtering," *Information*, vol. 9, no. 4, pp. 93, 2018.
- [38] Han, Y. J., Kim, W., & Park, J. S., "Efficient eye-blinking detection on smartphones: A hybrid approach based on deep learning," *Mobile Information Systems*, 2018.
- [39] M. Chau and M. Betke, "Real time eye tracking and blink detection with USB cameras," *Technical Report 2005-12, Boston University Computer Science*, May 2005.
- [40] D. Torricelli, M. Goffredo, S. Conforto and M. Schmid, "An adaptive blink detector to initialize and update a view- based remote eye gaze tracking system in a natural scenario," *Pattern Recogn Lett*, vol. 30, no. 12, pp. 1144-1150, 2009.
- [41] Trobo, I. P., Díaz, V. G., Espada, J. P., Crespo, R. G., & Moreno-Ger, P., "Rapid modeling of human-defined AI behavior patterns in games," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 7, pp. 2683-2692, 2019.
- [42] Khari, M., Kumar, P., Burgos, D., & Crespo, R. G., "Optimized test suites for automated testing using different optimization techniques," *Soft Computing*, vol. 22, no. 24, pp. 8341-8352, 2018.
- [43] García-Díaz, V., Espada, J. P., Crespo, R. G., G-Bustelo, B. C. P., & Lovelle, J. M. C., "An approach to improve the accuracy of probabilistic classifiers for decision support systems in sentiment analysis," *Applied Soft Computing*, vol. 67, pp. 822-833, 2018.
- [44] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y., "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499-1503, 2016.
- [45] Harish, B. S., Maheshan, M. S., & Nagadarshan, N., "A Convolution Neural Network Engine for Sclera Recognition," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 1, pp. 78-83, 2020. <http://doi.org/10.9781/ijimai.2019.03.006>
- [46] Verma, K. K., Singh, B. M., Mandoria, H. L., & Chauhan, P., "Two-Stage Human Activity Recognition Using 2D-ConvNet," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 2, 2020. <http://doi.org/10.9781/ijimai.2020.04.002>



Puneet Singh Lamba

Puneet Singh Lamba received the Master degree (M.tech) in Information Technology from University School of Information And Communication Technology, Guru Gobind Singh Indraprastha University in 2013. He is a PhD candidate in Information Technology at USICT, GGSIPU. His research interests are in the areas of Sensor Networks and Adhoc Networks. He has more than 11 years of

teaching experience and is currently working as Assistant Professor in Bharati Vidyapeeth's College of Engineering New Delhi, India. He has published more than 10 research papers in International journals / National journals / International conferences of repute.



Deepali Virmani

Deepali Virmani has done, B. Tech, Computer Science from MDU, Rohtak in 2001, M. Tech in Information Technology in 2005, from GGSIPU and Ph.D Wireless Sensor Network from Delhi University in 2013. She has a teaching experience of 17 years. She is actively engaged in teaching and research in areas of Computer Science since 2001. She has published more than 73 research papers in

International journals / National journals / International conferences of repute. Her research interests are in the areas of Sensor Networks, Data Mining and Security. She has guided more than 45 B.Tech projects. Presently she is guiding 5 Ph.D scholars registered with reputed universities like GGSIPU, UPTU. She is branch counselor BPIT -IEEE student chapter and BPIT-CSI student branch. She is on the reviewer panel / editorial board of many International Journals. She has organized many professional activities like FDPs, Workshops, expert lectures. She has been the session chair in National / International conferences.



Manu S Pillai

Manu S Pillai is currently working as an Instructor and Product Engineer for Machine Learning and Data Science in Coding Blocks, New Delhi, India. He is also associated with Information Technology Department, Bharati Vidyapeeth's College of Engineering, Guru Gobind Singh Indraprastha University, Delhi, India. His current research interests include Computer Vision, Image Processing & Deep Learning.



Dr. Gopal Chaudhary

Dr. Gopal Chaudhary is currently working as an assistant professor in Bharati Vidyapeeth's College of Engineering, Guru Gobind Singh Indraprastha University, Delhi, India. He holds a Ph.D. in Biometrics at the division of Instrumentation and Control engineering, Netaji Subhas Institute of Technology, University of Delhi, India. He received the B.E. degree in electronics and communication

engineering in 2009 and the M.Tech. degree in Microwave and optical communication from Delhi Technological University (formerly known as Delhi College of Engineering), New Delhi, India, in 2012. He has 50 publications in refereed National/International Journals & Conferences (Elsevier, Springer, Inderscience) in the area of Biometrics and its applications. His current research interests include soft computing, intelligent systems, information fusion and pattern recognition. He has organized many conferences and leading special issues in Taylor and Francis, Springer, IOS Press, Computers Materials & Continua etc.