New Features for Eye-Tracking Systems: Preliminary Results

1st Audi I. Al-Btoush
Faculty of Information Technology
Mutah University
Karak, Jordan

2nd Mohammad A. Abbadi
Faculty of Information Technology
Mutah University
Karak, Jordan

3rd Ahmad B. Hassanat
Faculty of Information Technology
Mutah University
Karak, Jordan

4th Ahmad S. Tarawneh
Department of Algorithm and Their Applications
Eötvös Loránd University
Budapest, Hungary

5th Asad Hasanat
Queen Rania Hospital
Ministry of Health
Petra, Jordan

6th V. B. Surya Prasath
Cincinnati Childrens Hospital Medical Center
Division of Biomedical Informatics
Cincinnati OH 45229, USA

***Abstract*—Due to their large number of applications, eye-tracking systems have gain attention recently. In this work, we propose 4 new features to support the most used feature by these systems, which is the location (x, y). These features are based on the white areas in the four corners of the sclera; the ratio of the whites area (after segmentation) to the corners area is used as a feature coming from each corner. In order to evaluate the new features, we designed a simple eye-tracking system using a simple webcam, where the users faces and eyes are detected, which allows for extracting the traditional and the new features. The system was evaluated using 10 subjects, who looked at 5 objects on the screen. The experimental results using some machine learning algorithms show that the new features are user dependent, and therefore, they cannot be used (in their current format) for a multiuser eye-tracking system. However, the new features might be used to support the traditional features for a better single-user eye-tracking system, where the accuracy results were in the range of 0.90 to 0.98.**

***Index Terms*—Eye tracking, Eye-gaze, Hough transform, Face detection, Eye detection, Machine learning.**

I. Introduction

Eyes are a rich source of information for humans and used for perceiving, watching, reading, learning, playing, etc. The steady stream of information provided by our eyes help us in daily interactions with surrounding objects. Eyes are critical to different biometrics applications by the analyses of eye-movement, also in human-computer interaction (HCI). Analyzing and capturing eye movements can help disabled people to control the computer effectively [1]. However, eye tracking faces considerable challenges, due to human eyes unique perceptive nature and physiology. There are six different muscles devoted to each eye that control the movements, which allows the eye to converge, diverge, elevate, depress, and roll. These muscles are controlled voluntarily and involuntarily to track objects and correct for simultaneous head movements [2]. The essential elements of the human eye are as follows:

- The cornea, which is a transparent coat in front of the eyeball.
- The iris, which is the structure that controls the size of pupil, similar to the aperture in a camera that let light goes inside. Irises are coloured and different from person to person.
- The pupil is the back circle located in the centre of the iris.
- The sclera is the tough outer surface of the eyeball and appears white in the eye image.
- The limbus is the boundary between the sclera and cornea.

A. Related Works

Eye tracking systems are based on some devices to track the movement of the eyes for the purpose of knowing exactly where the person is looking, as well as to know for how long they were looking [3]. Applications of eye tracking include the ability to control the position of the cursor on the screen relative to the users eye position, or navigating the computer display without the need for mouse or keyboard input, which may be beneficial for certain types of users, such as disabled individuals [4]. The importance of eye tracking systems stems from their diverse applications such as applications for disabled persons, Mobile applications, Software Engineering Applications and Marketing Applications. There is a plethora of literature on eye tracking systems, which were built for these applications. Applications for disabled persons include the work of [5], who developed a smart wheelchair that can be controlled by eyes. Vickers and co-workers [6] designed GuitarHero game for persons with physical disabilities by using only eye movements. Other works in this domain include [7]–[10]. Mobile applications includes the work of [11], who presented an eye tracking software that works on Smartphone and Tablets, without the need for additional sensors. Drewes et al [12] discussed a user study to see whether people can interact with applications using eye-gaze tracking on mobile phones. Software Engineering Applications include the work of Yusuf et al [13] used eye-tracking technique that assessed how well a subject comprehends class diagrams and deter-

mines their characteristics. Similar studies were conducted by Sharif and Maletic [14] and Guhneuc [15]. Marketing Applications include the work of Khushaba et al [16], where the eye tracker system was used in this case solely to map the transition between the choice sets and the actual choice of an object. This technology is especially useful when studying the behaviors of the clients on Internet shops. Website designers have the possibility to understand how consumers see and read websites, what people look at, which elements they want, and which are omitted. The eye tracking system also provides answers to questions on whether respondents noticed the elements key object (logo, the "buy" button, etc.) [17].

B. Contribution

Intuitively, most of these systems are based on finding the (x, y) location of the pupil as a feature for tracking the eye. However, such features are affected by the head movement, particularly if the user was gazing on the same thing; this unintentional behavior allows the system to make mistakes. In this paper we propose the use of four more features to decide where the eye is actually looking, these features are based on the white areas in the four corners of the sclera, the ratio of the whites area (after segmentation) to the corners area is used as a feature coming from each corner, thus four new representative features are resulted: R1, R2, R3, and R4. As can be seen from Figure 1, the ratios are changed depending on the pupil location; these ratios should be less affected by the unintentional movements of the head and therefore alleviate the problem.

We organized the rest of the paper as follows. Section II introduces the new features with illustrative examples. Section III provides supporting experimental results with our eye-tracking system. Section IV concludes the paper.

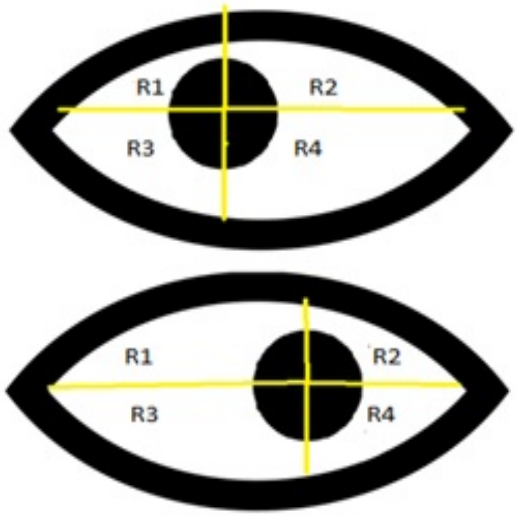

Fig. 1. The new four features, ratios of whites area to corners area.

II. The proposed method

In addition to the (x, y) location of the pupil, we propose the use of the ratios of the number of the white pixels of the four corners of the sclera to the number of the black pixels from each corner. To get these features we need to detect the face, first then the eye, because the face detectable than the eye, since it has more features to be targeted by the detection algorithm [18], [19]. After detecting the face, the eye can be easily detected, and then we can easily get the features that we are after. Figure 2 shows the block diagram of the system.

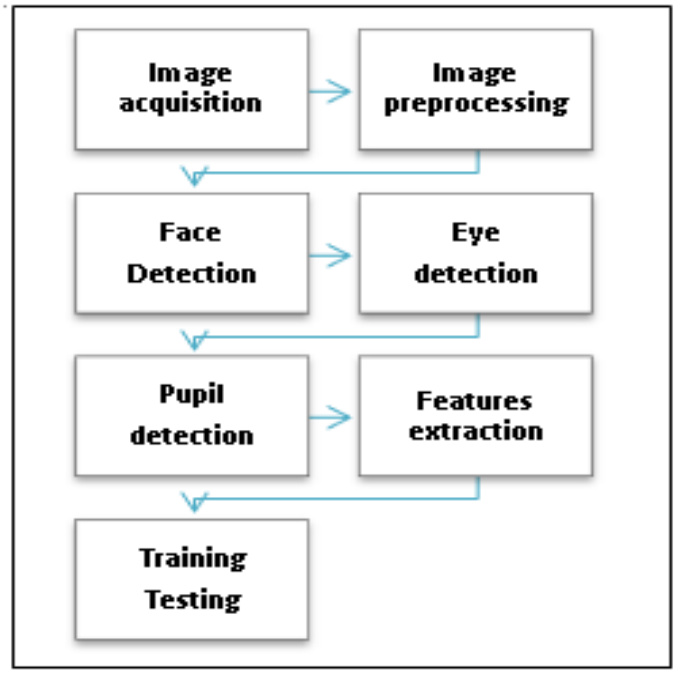

Fig. 2. Block diagram of a typical eye-tracking system.

A. Image acquisition

The images were acquired using a normal webcam, with no extra specifications. Due to the lack of publicly available datasets in the area of eye-tracking, we built an in-house dataset for this work. The dataset consists of the observations of 10 subjects (7-male and 3-female), they all asked to look at 5 objects located at the center and the four corners of the screen from the same distance to the screen (0.5 m), the objects are the letters A, B, C, D and E. the webcam is located at the top of the screen to record the eyes of the subjects, the subjects were asked to look at each character 10 times, resulting in 500 observations, see Figure 3.

B. Image preprocessing

In this stage, any image processing can be used to remove optical noise, and fix illumination; however, we did not use any of these because we recorded our database in office environment, where the light was good enough. In this work, the colour images were converted to gray scale images, because the next stages are not going to benefit from the colour information.

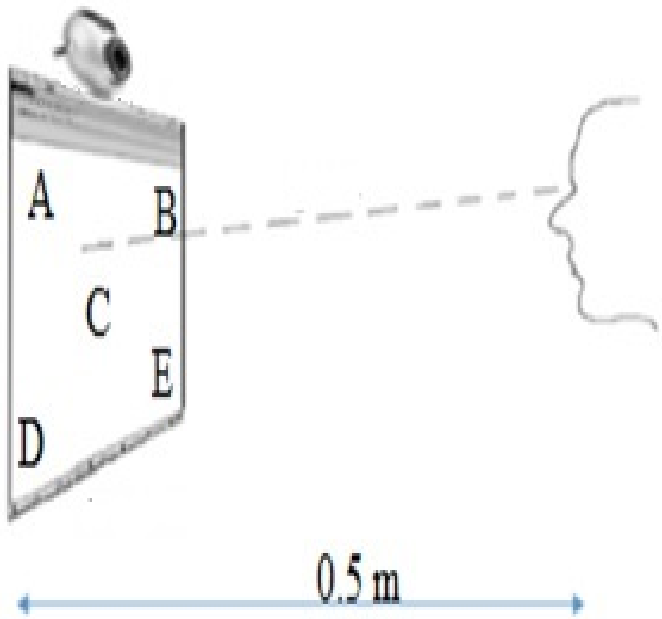

Fig. 3. Database collection.

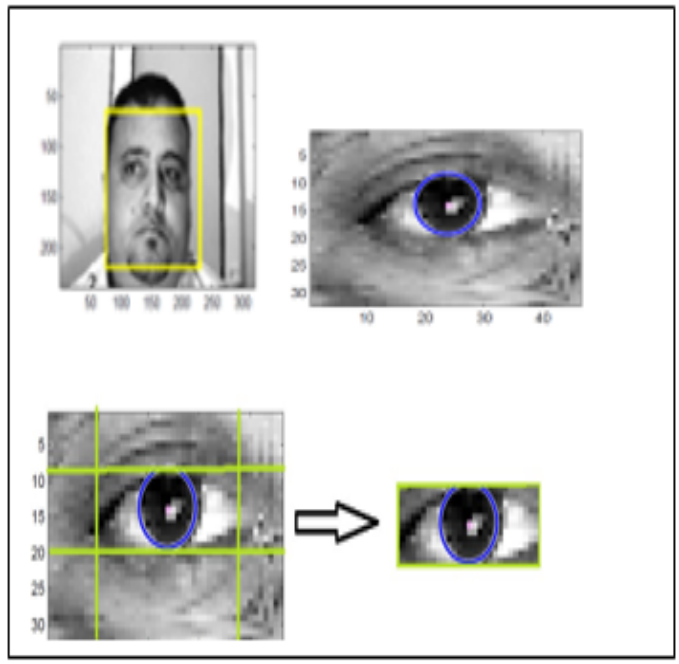

Fig. 4. (Upper-left) face detection, (upper-right) eye detection, (lower-left) pupil detection, (lower-right) region of interest (ROI).

C. Face and Eye Detection

Face detection is used in a variety of applications that are related to human faces in digital images. Here we opt for Viola-Jones [20] object detection framework to detect the face then one of the eyes, we used the Voila-Jones face detection method because it is fast, robust and highly accurate [21], [22]. Normally, there is no difference between left and right eyes, and therefore we get almost the same features, therefore we opt for detecting and working on the left eye.

D. Pupil detection

We next employ the circular Hough transform for detecting the circular boundary of the pupil, detecting the pupil is essential for determining the (x, y) location of the eye and the region of interest (ROI), from which other features will be extracted. Hough transform is used here since it is robust to the presence of noise, occlusion and varying illumination [23]. ROI is cropped by removing all rows located above and under the detected circle, in addition to cropping the width to be 6 times the radius of the detected circle, so as to prevent other unrelated facial parts from appearing in the ROI, see Figure 4.

E. Features extraction

In this work we extracted 6 features, the tradition (x, y) location of the pupil, and the new features R1, R2, R3 and R4, see Figure 5.

Both x and y values are normalized to the range [0, 1], so both are translated based on the image of the detected eye and not the new cropped image (see Figure 5), and then scaled to be in the range [0, 1]; as the x value is divided by the width of the eye's image, and the y value is divided by the height of the eye's image. The new four features depend mainly on counting the white pixels in the cropped image, which are supposed to represent the area of sclera at the four corners; therefore we opt for Otsu thresholding to obtain a binary image, see Figure 6, where the number of the white pixels represents the white part of the eye at a specific corner, and the number of the black pixels represents the other parts of the eye at the same corner. The number of the white pixels can be calculated using,

$$W_i = \sum_{j=1}^{N_i} \begin{cases} 1 & \text{, pixel is white} \\ 0 & \text{, otherwise} \end{cases} \tag{1}$$

and the number of the black pixels can be calculated using,

$$B_i = \sum_{j=1}^{N_i} \begin{cases} 1 & \text{, pixel is black} \\ 0 & \text{, otherwise,} \end{cases} \tag{2}$$

where i is the index of the feature (i.e. the number of the coroner [1,4]), n_i is the number of pixels in the corner i.

Consequently, we can calculate the feature i using,

$$R_i = \frac{W_i}{W_i + B_i}, \tag{3}$$

where R represents the ratio of the white area that appears at the coroner i while gazing, this ratio is calculated for each corner.

III. Experimental Results and Discussion

All the aforementioned methods were implemented with MATLAB code, we execute the code on a laptop with webcam built in, Inter(R) Core (TM) i5-6500T, CPU @ 2.50 GHz, installed memory (RAM) 4.00 GB and 64-bit windows 10

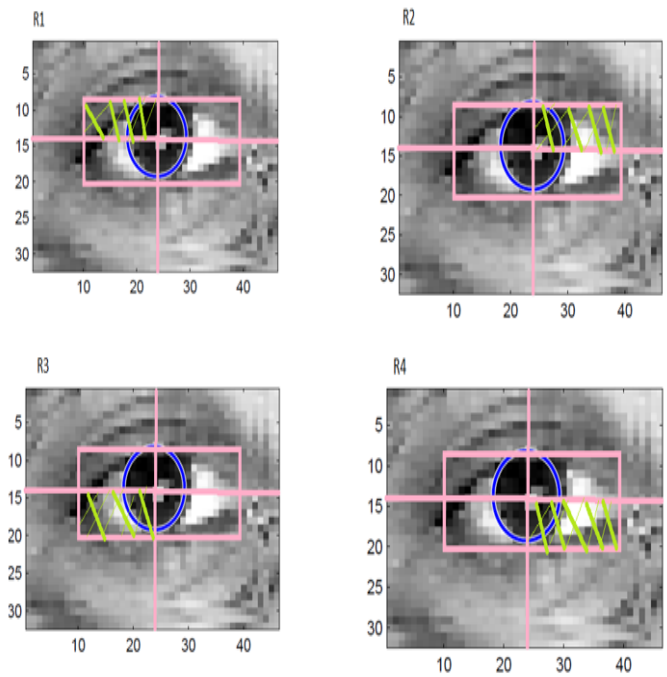

Fig. 5. The new features, the shaded areas represent the computed features; (upper-left) R1, (upper-right) R2, (lower-left) R3, and (lower - right) R4.

operating system. The code is executed for each subject as mentioned in the Data sub-section to get the six features, Table I shows the mean of these six features for each of the 5 objects/classes (A, B, C, D, and E).

TABLE I
THE MEAN OF THE FEATURES FOR EACH CLASS.

Class	X	Y	R1	R2	R3	R4
A	0.487	0.436	0.683	0.223	0.376	0.297
B	0.699	0.416	0.590	0.063	0.526	0.020
C	0.654	0.512	0.841	0.202	0.542	0.020
D	0.532	0.524	0.783	0.241	0.387	0.211
E	0.736	0.550	0.868	0.463	0.365	0.326

The data shown in Table I is the mean of each feature for each class from the 500 examples recorded in the database. To validate the effectiveness of the proposed features we made three experiments to validate the eye-tracking system, the first is designed using the traditional features (x, y) only, the second is designed to using the proposed features (R1, R2, R3 and R4), and the third is designed using all of the six features. The data of each experiment is trained using a number of machine learning algorithms from the WEKA data mining tool [24], namely, multilayer perceptron (MLP), J48 decision tree classifier and Random Forest (RF). Table II shows the accuracy results of the three experiments.

As can be seen from Table II, using the new features alone allows the accuracy results to be decreased dramatically, and when added to support the traditional (x, y) features, the eye-tracking accuracies were not increased or increased/decreased insignificantly. Actually, these results were not expected, intuitively, if the white ratios were correctly calculated the new features would increase the accuracy of the system. However, calculating these ratios correctly is dependent on finding the true scleras pixels, the Otsu segmentation method used for this purpose is not perfect for this task, as there are many false acceptance and false rejection errors associated with this task. This is due to many issues such as the camera quality (webcam in our case), Illumination, optical noise, and most importantly the different appearance of the subjects eyes, which is different from person to another; this is due to different factors such as different lengths of eye lashes, the appearance/disappearance of eye bow, different eye sizes, shadows, etc. Moreover, we noticed that most of the subjects are having redness in their sclera. Figure 6 shows the segmentation results, which is used to find the R ratios.

TABLE II
ACCURACY RESULTS OF THE CLASSIFICATION ALGORITHMS, USING 10-FOLD CROSS VALIDATION, RESULTS ARE IN PERCENTAGE (%).

Features	MLP	J48	RF
X, Y	95.6	96.2	97
R1234	55.4	60.6	65.6
X,Y, R1234	95.2	95.6	97.2

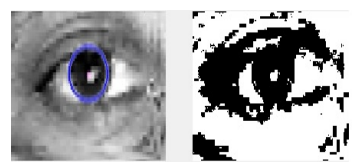

Fig. 6. (left) cropped eye image, (right) segmentation result of Otsu method.

As can be seen from Figure 6, regarding the detection of the white pixels of the sclera, there are many false accepted and rejected pixels, which affects the calculations of the R ratios, and makes them unstable features due to the image processing technique used. The instability of the proposed features is also supported by the data in Table I, for example, when the 10 subjects were looking at A their pupil moves to the upper-left corner, which makes $R1 < R4$, which is not the case, since R1=0.68 and R4=0.30. In order to see if the appearance of the subjects affects the stability of the R features, we trained/tested the data of each subject solely, we used the RF algorithm since it obtained the highest results (see Table II), with 10-fold cross validation on the 50 examples of each subject. Table III shows the eye-tracking accuracy results for each subject.

As can be seen from Table III, the results favor the new features when training/testing subjects solely, this means that the new features are user dependent, and they can be used to support the traditional features, providing that the system works for a single user. However, to train the eye-tracking system to be multiuser; it is better to use the (x, y) features only, because finding these features accurately is less affected by the eye appearance and other image processing problems. Since the new features are based on the appearance of the users eye and sclera, the system might befit from a preprocessing

TABLE III
ACCURACY RESULTS OF THE RF FOR EACH SUBJECT, USING 10-FOLD CROSS VALIDATION, RESULTS IN PERCENTAGE (%).

Subject	(X, Y)	(X,Y, R1234)
1	98	98
2	96	96
3	96	94
4	94	98
5	94	94
6	96	96
7	98	98
8	88	90
9	92	98
10	90	90
Avg	94.4	95.2

step that includes image enhancement, removal of scleras redness and eyelashes, in addition to finding another method that can correctly assess the white area of the sclera in the four corners of the eye.

IV. CONCLUSION

In this paper we propose new features for eye-tracking systems to support the tradition (x,y) features, these features are based on finding the area of the sclera in each corner of the eye. We designed a simple eye-tracking system using a simple webcam, where the users faces and eyes are detected, which allows for extracting the traditional and the new features. The system was evaluated using 10 subjects, who looked at 5 objects on the screen. The experimental results using some machine learning algorithms show that the new features are user dependent, and therefore they cannot be used (in their current format) for a multiuser eye-tracking system. However, the new features might be used to support the traditional features for a better single-user eye-tracking system. The limitation of this study includes the use of a simple method to assess the area of the sclera at the four corners of the eye, which is currently based on the simple Otsu segmentation method. Such a simple method is easily affected by the appearance of the eye, which might be different from user to another, the system might benefit from a preprocessing step that includes image enhancement, removal of scleras redness and eyelashes. Therefore, our future work will focus on finding an alternative robust method that can correctly assess the white area of the sclera in the four corners of the eye, for this purpose, we may use the methods proposed by [25]–[28].

ACKNOWLEDGMENT

AST would like to acknowledge Tempus Public Foundation for sponsoring his Ph.D. study.

REFERENCES

[1] Y.-m. Cheung and Q. Peng, "Eye gaze tracking with a web camera in a desktop environment," *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 4, pp. 419–430, 2015.

[2] R. H. Carpenter, *Movements of the Eyes, 2nd Rev*. Pion Limited, 1988.

[3] R. G. Lupu and F. Ungureanu, "A survey of eye tracking methods and applications," *Buletinul Institutului Politehnic din Iasi, Automatic Control and Computer Science Section*, vol. 3, pp. 72–86, 2013.

[4] M. Betke, J. Gips, and P. Fleming, "The camera mouse: Visual tracking of body features to provide computer access for people with severe disabilities," *IEEE Transactions on neural systems and Rehabilitation Engineering*, vol. 10, no. 1, pp. 1–10, 2002.

[5] A. P. Ghode and R. Yawale, "Eye gaze tracking system," *International Research Journal of Engineering and Technology*, vol. 04, no. 07, 2017.

[6] S. Vickers, H. Istance, and M. Smalley, "Eyeguitar: making rhythm based music video games accessible using only eye movements," in *Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology*. ACM, 2010, pp. 36–39.

[7] X. Zhang, H. Kulkarni, and M. R. Morris, "Smartphone-based gaze gesture communication for people with motor disabilities," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2017, pp. 2878–2889.

[8] P. Ghude, A. Tembe, and S. Patil, "Real-time eye tracking system for people with several disabilities using single web cam," *IJCAT International Journal of Computing and Technology*, vol. 1, no. 2, 2014.

[9] R. Anacan, J. G. Alcayde, R. Antegra, and L. Luna, "Eye-guide (eye-gaze user interface design) messaging for physically-impaired people," *arXiv preprint arXiv:1302.1649*, 2013.

[10] M. N. V. Khope, M. Burange, and A. Gadicha, "Implementation of real time image processing for a human eye computer interaction system," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 4, no. 11, pp. 235–239, 2016.

[11] K. Krafka, A. Khosla, P. Kellnhofer, H. Kannan, S. Bhandarkar, W. Matusik, and A. Torralba, "Eye tracking for everyone," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2176–2184.

[12] H. Drewes, A. De Luca, and A. Schmidt, "Eye-gaze interaction for mobile phones," in *Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology*. ACM, 2007, pp. 364–371.

[13] S. Yusuf, H. Kagdi, and J. I. Maletic, "Assessing the comprehension of uml class diagrams via eye tracking," in *Program Comprehension, 2007. ICPC'07. 15th IEEE International Conference on*. IEEE, 2007, pp. 113–122.

[14] B. Sharif and J. I. Maletic, "An eye tracking study on camelcase and under_score identifier styles," in *2010 IEEE 18th International Conference on Program Comprehension*. IEEE, 2010, pp. 196–205.

[15] Y.-G. Guéhéneuc, "Taupe: towards understanding program comprehension," in *Proceedings of the 2006 conference of the Center for Advanced Studies on Collaborative research*. IBM Corp., 2006, p. 1.

[16] R. N. Khushaba, C. Wise, S. Kodagoda, J. Louviere, B. E. Kahn, and C. Townsend, "Consumer neuroscience: Assessing the brain response to marketing stimuli using electroencephalogram (eeg) and eye tracking," *Expert Systems with Applications*, vol. 40, no. 9, pp. 3803–3812, 2013.

[17] M. Wedel and R. Pieters, "A review of eye-tracking research in marketing," in *Review of marketing research*. Emerald Group Publishing Limited, 2008, pp. 123–147.

[18] A. B. A. Hassanat, "Visual speech recognition," in *Speech and Language Technologies*, I. Ipsic, Ed. Rijeka: IntechOpen, 2011, ch. 14. [Online]. Available: https://doi.org/10.5772/19361

[19] A. Hassanat and S. Jassim, "A special purpose knowledge-based face localization method," in *Mobile Multimedia/Image Processing, Security, and Applications 2008*, vol. 6982. International Society for Optics and Photonics, 2008, p. 69820M.

[20] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1. IEEE, 2001, pp. I–I.

[21] A. Liu, Z. Li, L. Wang, and Y. Zhao, "A practical driver fatigue detection algorithm based on eye state," in *Microelectronics and Electronics (PrimeAsia), 2010 Asia Pacific Conference on Postgraduate Research in*. IEEE, 2010, pp. 235–238.

[22] A. Gupta and R. Tiwari, "Face detection using modified viola jones algorithm," *International Journal of Recent Research in Mathematics Computer Science and Information Technology*, vol. 1, no. 2, pp. 59–66, 2014.

[23] A. Shrivas and P. Tuli, "Analysis of iris images for iris recognition system," in *National Conference on Innovative Paradigms in Engineering & Technology*, 2012.

[24] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *SIGKDD Explorations*, vol. 11, no. 1, pp. 10–18, 2009.

[25] A. B. Hassanat and S. Jassim, "Color-based lip localization method," in *Mobile Multimedia/Image Processing, Security, and Applications 2010*, vol. 7708. International Society for Optics and Photonics, 2010, p. 77080Y.

[26] A. B. Hassanat, M. Alkasassbeh, M. Al-awadi, and A. Esra'a, "Color-based object segmentation method using artificial neural network," *Simulation Modelling Practice and Theory*, no. 64, pp. 3–17, 2016.

[27] A. B. Hassanat, V. B. S. Prasath, M. Al-kasassbeh, A. S. Tarawneh, and A. J. Al-shamailh, "Magnetic energy-based feature extraction for low-quality fingerprint images," *Signal, Image and Video Processing*, vol. 12, no. 8, pp. 1471–1478, 2018.

[28] A. Hassanat, V. B. S. Prasath, M. Abbadi, S. Abu-Qdari, and H. Faris, "An improved genetic algorithm with a new initialization mechanism based on regression techniques," *Information*, vol. 9, no. 7, p. 167, 2018.