iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XI **Month of publication:** November 2024

DOI: https://doi.org/10.22214/ijraset.2024.49545

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XI Nov 2024- Available at www.ijraset.com

Eyeball Cursor Movement using Deep Learning

Telukuntla Sharmili[1], Vallepu Srinivas[2], Vangapally Umesh[3], Dr. G. Surya Narayana[4]
Department of CSE, Vardhaman College of Engineering, Hyderabad,India

Abstract: Since its inception, computers have developed to the point that we may use them wirelessly, remotely, and in a variety of other ways. Nevertheless, these innovations are only for those who are physically fit and without physical disabilities. Other methods are needed to make computers more usable and accessible for those with physical disabilities. We have created a system that uses eyeball movement instead of the traditional manner of pointer movement with a mouse in order to create an environment that is fair to everyone who is capable of it. This will make it easier for those who suffer from physical conditions to use computers, advancing knowledge and technology. In this system, the laptop camera is used to take the pictures, which are then used to recognise the eyeball positions and movements using face indexing, as well as to conduct the actions that a real mouse could carry out using OpenCV. Because there are no pre-installations needed, this project is economical. Those who are paralysed and others with special needs may benefit from this.
Keywords: Haar cascade, CNN, EAR (Eye Aspect Ratio)

I. INTRODUCTION

The rapid advancement of computer technologies has significantly highlighted the importance of Human-Computer Interaction (HCI). For disabled persons, Eye Ball Movement Control is often employed as an accessible means of using computers. Incorporating this eye-controlling system into computers allows for independent operation without the need for assistance from another individual. HCI focuses on providing an interface between a computer and a human user. Thus, there is a need to develop suitable technology that facilitates effective communication between humans and computers. Human Computer Interaction plays a pivotal role in this regard, and thus there is a requirement to identify an alternate method to enable communication between persons with impairments and computers, allowing them to become part of the Information Society.

Given the presence of certain illnesses which impede the ability of certain individuals and groups to operate conventional computing devices, it is essential to develop a method of computer operation that is accessible to those with disabilities. The human eye may serve as an effective substitute for traditional computer operating hardware.

A Human-Computer Interaction (HCI) system is being implemented, allowing users to control computers with their eyes. This development will be particularly beneficial for individuals with disabilities who lack the ability to use their hands, enabling them to operate computers independently. This type of control is made possible through the detection of pupil movement relative to the center of the eye, which is accomplished using OpenCV technology.

II. RELATED WORK

A thorough literature review has been conducted in the field of healthcare applications pertaining to eye tracking systems; some of these methods are detailed below.

Firstly, Khare, Vandana, S. Gopala Krishna, and Sai Kalyan Sanisetty (2019) assert that personal computers have become an essential element of modern life; however, their usage is limited by physical abilities. To address this issue, a Raspberry Pi and OpenCV were utilized in conjunction with an external camera to capture images for facial feature detection. This approach provides a straightforward correlation between eyeball movement and cursor movement; however, it is relatively expensive due to the necessary additional hardware. [1] Z. Sharafi et al. [2] conducted an evaluation of metrics pertaining to eye-tracking in software engineering. The authors consolidated disparate metrics into a unified standard to facilitate the analysis of eye-tracking experiments for researchers. S. Chandra et al. [3] proposed an application based on eye-tracking and its interaction with humans. The research concentrated on the determination of direction, position, and sequence of gaze movement. R. G Bozomitu et al.[4] proposed an eye tracking system based on the circular Hough transform algorithm for the purpose of providing benefits to neuromotor disabled patients. RamshaFatima and AtiyaUsmani [5] proposed a method related to Eye Movement-based Human Computer Interaction in IEEE 2016. This paper aimed to outline and execute a Human Computer Interface framework that tracks the direction of the human eye using EAR ratio for cursor movement implementation. In their 2016 paper, Venugopal and D'souza [6] proposed a hardware-based system utilizing an Arduino Uno microcontroller and a Zigbee wireless device.

©IJRASET: All Rights are Reserved

First, real-time images were captured and processed via Viola Jones Algorithm to detect faces, then Hough Transform was used to detect pupil movement, which was used as input to the hardware containing the Arduino Uno microcontroller and Zigbee device for data transmission. O. Mazhar et. al [7] developed an ingenious eye-tracking system, designed to provide a revolutionary solution for those with neuromotor disabilities. M. Kim et. al [8]proposed a radical system for detecting and tracking eye movement based on cardinal direction. In the experiment, participants were instructed to look towards the north in a specific cardinal direction and their reaction times were then observed. In 2012, Gupta, Akhil, Akash Rathi and Dr Y. Radhika unveiled their revolutionary invention: "Hands-free PC Control", [9] which enabled users to control their mouse cursor with eye movement. The Image-based method relied on Support Vector Machines (SVM) to detect the face in streaming video; once the face was detected using SVM and SSR (Scale Space Representation), an integral image was tracked moving forward. K. Kuzhals et al. [10] unveiled a groundbreaking eye tracking system, which promises to revolutionize the application of Personal Visual Analytics (VSA). This research paper explores the challenges of VSA in real-time scenarios and identifies potential areas for further study.

III. METHODOLOGY

This section provides the description of proposed model.

A. Haar Cascade Classifier

The Haar Cascade Classifier is a machine learning-based method for finding objects in pictures or movies. To identify the presence of a particular object, such as eyes, faces, or cars, it employs a series of classifiers. So, in our study, this classifier was utilized to identify both the right and left eyes. We must first train the classifier on a sizable dataset of both positive and negative images to detect eyes using the Haar Cascade Classifier. Negative images lack eyes while positive images do. The program then picks up on the patterns that set the two classes apart .Here are the Five steps for utilizing the Haar Cascade Classifier to find eyes:

1) Get the pre-trained Haar Cascade Classifier for eye detection. It is accessible online or in the data folder of OpenCV. which is an XML file.
2) Use the OpenCV library to load the classifier into your Python code.
3) Read the image you wish to use to find the eyes and make it grayscale.
4) To identify the eyes in the image, use the classifier's detectMultiScale function. Each rectangle in the list that this method returns is an identified eye.
5) Each eye that was found in the image should have a rectangle drawn around it.

B. Deep Learning and CNN

Deep learning is a subfield of machine learning that has revolutionized the field of artificial intelligence (AI) in recent years. Deep learning models are capable of learning and solving complex problems by automatically discovering hierarchical representations of the input data. They are particularly well-suited for tasks such as image and speech recognition, natural language processing, and playing games, among others.

At the core of deep learning are neural networks, which are computational models inspired by the structure and function of the human brain. Neural networks consist of layers of interconnected nodes or neurons, each of which computes a weighted sum of its inputs and applies a non-linear activation function to produce an output. The weights and biases of the neurons are learned through a process called backpropagation, in which the error between the predicted outputs and the true outputs is propagated backwards through the network and used to update the weights.

There are several types of deep learning architectures that are commonly used today. One of the most popular is the convolutional neural network (CNN), which is used for image and video recognition tasks. A CNN consists of .several layers of convolutional and pooling operations, which enable it to learn spatial features in the input image.

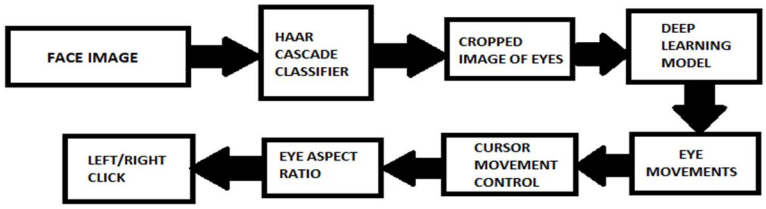


Fig 1 : flow chart

C. *Eye Aspect Ratio*

Eye tracking and eye movement research use the eye aspect ratio (EAR) measurement to gauge how much of the eye is open or closed. It is determined by the location of the eye's landmarks, which are normally the top and bottom, as well as the inner and outer corners. When a user blinks or closes their eyes, the EAR is utilized in eye tracking and cursor movement systems to detect it. To avoid accidental clicks or motions, the cursor movement is stopped or disabled when the user's eyes are closed.

A threshold value of 0.2 is established in our software to employ EAR in eye tracking and cursor movement to identify whether the user's eyes are open or closed. The system assumes that the eyes are closed if the EAR is below the threshold value and stops or inhibits cursor movement. The system assumes that the eyes are open and permits cursor movement if the EAR is higher than the threshold value. To distinguish between a typical wink of the eye and a wink of the eye used to accomplish clicking actions, we set the threshold at three consecutive frames.

$$\text{EAR} = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

IV. IMPLEMENTATION

The following is the process how we implemented this project.

A. *Dataset*

We have generated our own dataset which is having the following process.

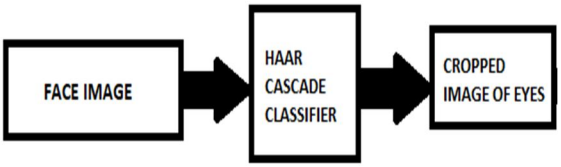


Fig 2 : flow chart for dataset

We have used the support of Haar cascade classifier which is used for object detection. In this algorithm the model trains with many positive and negative values based on the shadowing and edge feature, line feature and four rectangle feature. There is an open source xml file which can be downloaded from the github link of OpenCV.

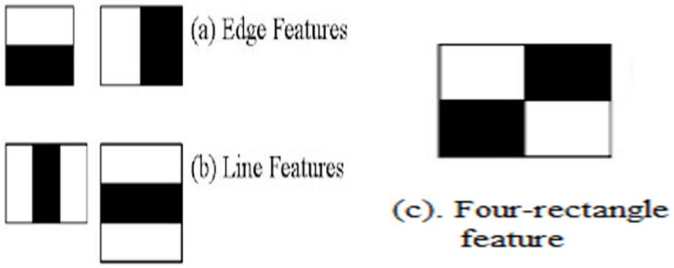


Fig 3 : haar cascasde

For creating the dataset we have looped with the frames of our face using OpenCV and generated an event listener which will capture the image when an click operation is done. While clicking it is necessary to look at the cursor because it will be the picture storing with a file name of “X_pixel Y_pixel_buttonPressed”. The intention behind it is that how out eyes will be when we are looking at the particular pixel on screen. Once the button is clicked the haar cascade classifier will crop the image with the borders around the 2 eyes and at last we will be concatenating the images and save it as the name mentioned. We have to generate around 10000 images to make our model train enough and guess the exact location where we are looking at. And the image is fixed to have 32 pixels. Capturing images is quite a task because only if the eyes are detected it will capture a picture. There has to be enough light to capture images and specs will make the system understand that here are no eyes in the image. Hence without specs we have to capture the images. The following is the picture which depicts out dataset.

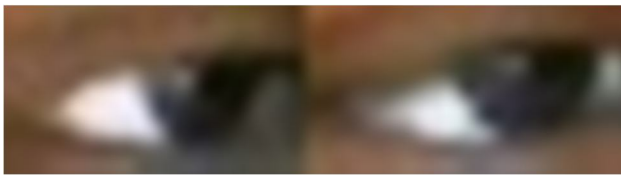


Fig 4 : data sample

B. Model Training

We have used a sequential model which consists of 3 convolution layers and a max pooling layer. We have run around 200 epochs and loos value was satisfying. After getting good accuracy we have deployed the model. The sequence of the layers are 3 convolution layers, succeeded with one max pooling layer for getting a sensible feature map. It is flattened into normal numeric layers with 2 dense layers. You can see the flow of layers for the model training used in the process in the below diagram(fig 5) of model summary.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 5, 21, 32)         896

 conv2d_1 (Conv2D)           (None, 2, 10, 64)         8256

 conv2d_2 (Conv2D)           (None, 2, 10, 32)         51232

 max_pooling2d (MaxPooling2D  (None, 2, 10, 32)        0
 )

 flatten (Flatten)           (None, 640)               0

 dense (Dense)               (None, 32)                20512

 dense_1 (Dense)             (None, 2)                 66

=================================================================
Total params: 80,962
Trainable params: 80,962
Non-trainable params: 0
_________________________________________________________________
```

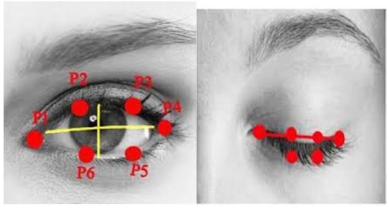
Fig 5: model summary

C. Eye Tracking

Once the model is trained it is now time to control our coputer with our eyeball movements.As the model will be able to recognize the direction in which we are looking and can map our cursor to that position of the screen.

Using the image in the frame the face landmarks are located and we will pass a function on the points of left and right eye.

The function will return the EAR (Eye aspect ratio) value which will be verified with the threshold value. If the value of EAR drops below the threshold value for 3 consecutive frames, it will be assumed as a click operation.



Open eye Closed eye

Fig 6 : Eye aspect ratio

Now if we want to perform any click operation we just need to wynk our eyes. If we wynk our left eye it will be assumed as a left click operation, if we wynk our right eye it will be assumed as a right click operation. We can invoke the left and right clicks using pyautogui package in python. For left click we will invoke pyautogui.click() funtion and for right click pyautogui.rightClick() function.

D. Streamlit Interface

We have used Streamlit an interface for this model exploration.

V. EXPERIMENTAL RESULTS

A. Results

The below figure demonstrates the detection of eye for the cursor movement.

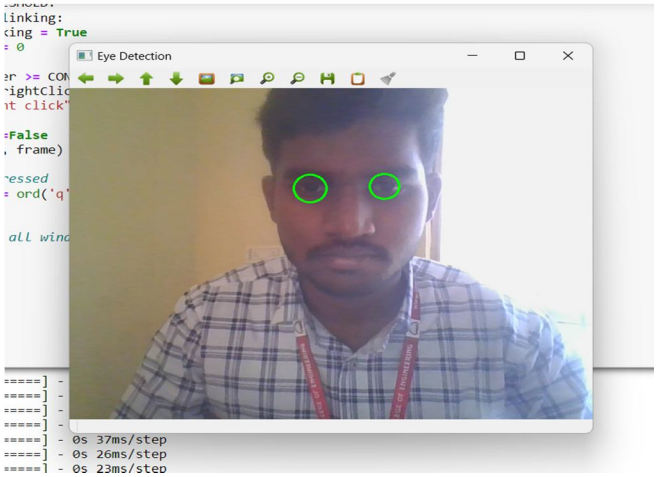


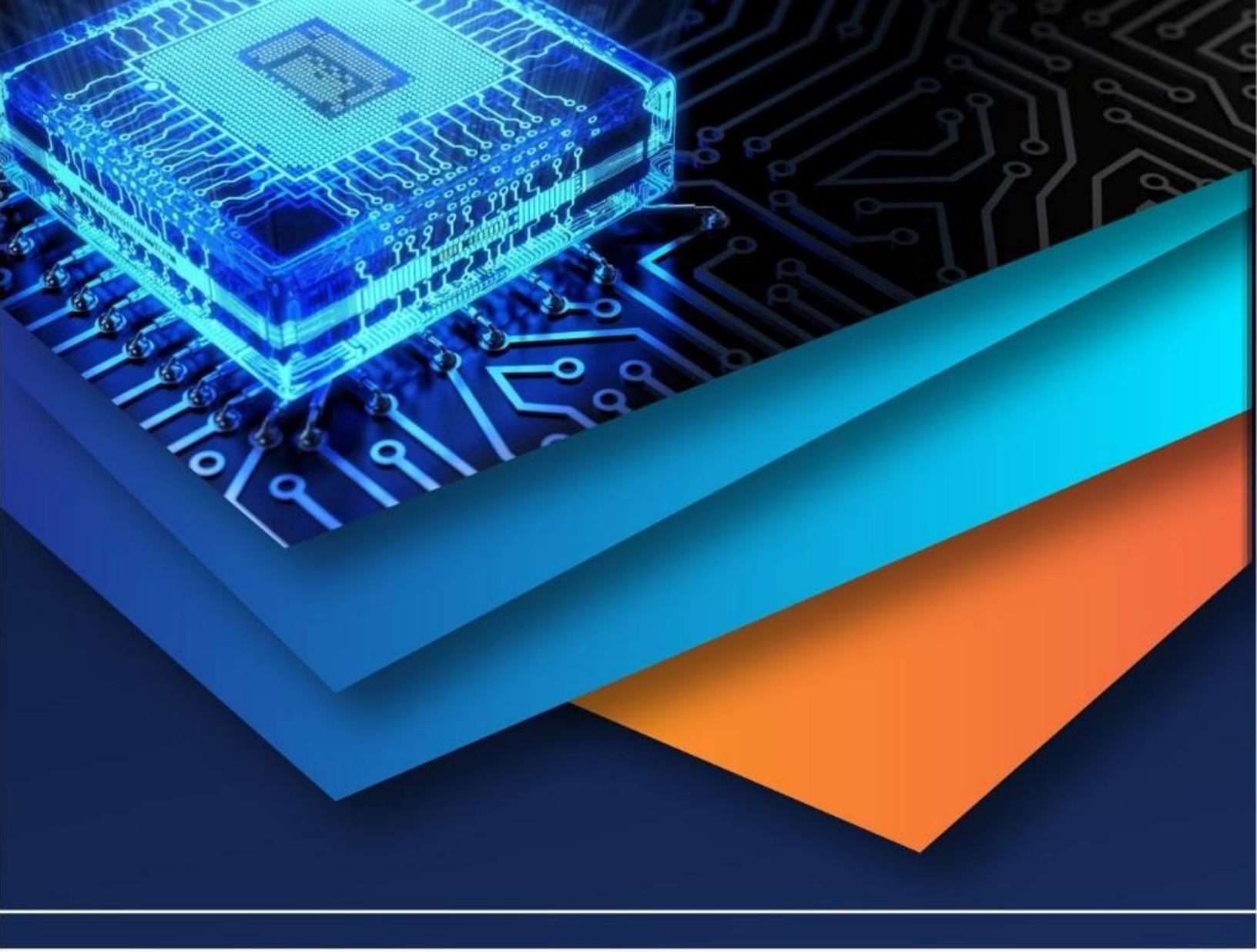
Fig 7 : eye detection

VI. CONCLUSION AND FUTURE SCOPE

As or aim is to control computer with eyes. We are looking to implement few other operations which can be done by mouse. Other than click operations there are scrolling operation, selection of text and typing. For typing we are trying to invoke a virtual keyboard and the letters that are clicked will be displayed on the notepad or other blank screen. We are looking forward to generalize the cursor control to the remote controlling of televisions or any other software. For this we need an hardware device which will do the whole thing at once and perform the task at ease. A specs or something with a camera inbuilt and the whole process running can make the daily life easier and interesting.

REFERENCES

[1] Khare, Vandana, S. Gopala Krishna, and Sai Kalyan Sanisetty. "Cursor Control Using Eye Ball Movement." In 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), vol. 1, pp. 232-235. IEEE, 2019.
[2] Sharafi, Zohreh, Timothy Shaffer, and Bonita Sharif. "Eye-Tracking Metrics in Software Engineering." In 2015 AsiaPacific Software Engineering Conference (APSEC), pp. 96-103. IEEE, 2015.
[3] Chandra, S., Sharma, G., Malhotra, S., Jha, D. and Mittal, A.P., 2015, December. Eye tracking based human computer interaction: Applications and their uses. In 2015 International Conference on Man and Machine Interfacing (MAMI) (pp. 1-5). IEEE.
[4] Bozomitu, R.G., Păsărică, A., Cehan, V., Lupu, R.G., Rotariu, C. and Coca, E., 2015, November. Implementation of eye-tracking system based on circular Hough transform algorithm. In E-Health and Bioengineering Conference (EHB), 2015 (pp. 1-4). IEEE.
[5] RamshaFatima,AtiyaUsmani(2016) Eye movement based human computer interaction, IEEE.
[6] Venugopal, B. K., and Dilson D'souza. "Real Time Implementation of Eye Tracking System Using Arduino Uno Based Hardware Interface."in 2016
[7] Mazhar, O., Shah, T.A., Khan, M.A. and Tehami, S., 2015, October. A real-time webcam based Eye Ball Tracking System using MATLAB. In Design and Technology in Electronic Packaging (SIITME), 2015 IEEE 21st International Symposium for (pp. 139-142). IEEE.
[8] Kim, M., Morimoto, K. and Kuwahara, N., 2015, July. Using Eye Tracking to Investigate Understandability of Cardinal Direction. In Applied Computing and Information Technology/2nd International Conference on Computational Science and Intelligence (ACIT-CSI), 2015 3rd International Conference on(pp. 201-206). IEEE.
[9] Gupta, Akhil, Akash Rathi, and Dr Y. Radhika. "Hands-free PC Control" controlling of mouse cursor using eye movement." International Journal of Scientific and Research Publications 2, no. 4 (2012): 1-5.
[10] "An eye tracking algorithm based on Hough transform" https://ieeexplore.ieee.org/document/8408915/ (12.5.2018)
[11] Eye gaze location detection based on iris tracking with web camera: https://ieeexplore.ieee.org/document/8404314/ (5.7.2018)
[12] Kurzhals, K. and Weiskopf, D., 2015. Eye Tracking for Personal Visual Analytics. IEEE computer graphics and applications, 35(4), pp.64-72.
[13] Muhammad Usman Ghani, Sarah Chaudhry "Gaze Pointer: A Real Time Mouse Pointer Control Implementation Based On Eye Gaze Tracking"/http://myweb.sabanciuniv.edu/ghani/files/2015/02/GazePoint er_INMIC2013.pdf [6 February 2014].
[14] Mohammad Shafenoor Amin, Yin Kia Chiam, Kasturi Dewi Varathan. Identification of significant features and data mining techniques in predicting heart disease.
[15] "Face and eye tracking for controlling computer functions" https://ieeexplore.ieee.org/document/6839834/ (2014)
[16] "An Adaptive Algorithm for Precise Pupil Boundary Detection" by Chihan Topel and Cuneyt Akinlar (2012)
[17] "Eye tracking mouse for human-computer interaction" https://ieeexplore.ieee.org/document/6707244/ (2013).



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)