iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** IV **Month of publication:** April 2020

DOI:

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IV Apr 2020- Available at www.ijraset.com

Sleep Detection using Image Processing Techniques

Sarthak[1], Vanshika Chaudhary[2], Nancy Gupta[3], Mohit Sarraf[4]
[1, 2, 3, 4]*Student, Department of Computer Science Engineering, Raj Kumar Goel Institute of Technology, Ghaziabad, Uttar Pradesh, India*

Abstract: Nowadays, concentration is a basic need in all the professions. Driving is one of them, drivers must keep their eyes on road so that they can act accordingly. A number of accidents on road occur due to driver's drowsiness. Hence, there must be a system that can tell a driver about his/her mental and physical condition, so designing this kind of system can reduce the number of road accidents which are caused due to drowsiness of a driver. Though, it's not easy to make a system like this as it needs a very fast and active awareness of a driver's drowsiness characteristics. Image Processing methods can be used to implement this kind of system.
Keywords: Drowsy Driving, Drowsiness Detection, Image Processing, Dlib, Haar Features, OpenCV.

I. INTRODUCTION

Fatigue driving is becoming a major cause of road accidents these days. Identification of drowsiness symptoms of a driver is not an easy task, as there are no readymade tests that can be performed on driver. Extenuation is a way of reducing the number of these accidents. The right way to measure a driver's fatigue is to invigilate anatomical characteristics like heart pace, skin communication and mind's activity. Sometimes to gauge this,it is required to have a direct physical contact between the driver and electrodes but it may be uncomfortable and can cause distraction also. Another method which is generally used is observing of direction patterns and alley driving blueprint, deviations from which will be shown as an indication of drowsiness of the person who is driving. One of the most efficient approach is the implementation of image processing to observe the physical transition that happens to a drowsy person. These transitions include sagging of eyelids, sagging of head and oscitancy. Here, sagging of eye is that what we observed. Many techniques can be applied on unique eye coordinates and the results can be used to conclude if eyes are in a closed state or an open state. To alarm the driver, we decided to use a buzzer that will make a sound whenever both eyes of driver will be in a closed state for the time longer than the decided time. First we will learn to distinguish between a shut eye and an open eye after that we will note the time of an eye blink.

Soukupova and Cech used the EAR(Eye Aspect Ratio) in their research in 2016 "Real-Time Eye Blink Detection Using Facial Landmarks" [1]. It was great to use this technique at that time. Dlib was opted for the solution of the problem as it is a modern C++ toolkit that contains machine learning algorithms and tools for developing complex software in C++ to solve real world problems. Dlib identifies an eye by its six coordinates. For a shutting eye, EAR is 0.3. The static frame, if EAR less than 0.3 then the eyes are considered as closed. If EAR is more than 0.3, the eyes are considered to be open. This fundamental combines with the threshold value(time for which an eye of the person is shut or closed),blinks are noticed.

II. LITERATURE REVIEW

Drowsy means sleepy, which means when someone is about to sleep. There are various stages of sleep and these stages are:

A. Non Rapid Eye Movement [NREM]
B. Rapid Eye Movement[REM]
C. Awake

NREM is divided further into following categories:
1) Change from awake to drowsy
2) Heavy Sleep
3) Petty Sleep

The main kick in designing cost effective sleep detection system is to gather proper data about drowsiness. Due to safety issues, this system cannot be implemented in a real time scenario so this system first needs to be designed and approved in a lab. In lab ,the crucial part in the designing of this type of system is the approach followed by the subject to get drowsy.
Driver's drowsiness basically depends on three factors:

a) The depth of last sleep
b) Time at which he/she is driving
c) Duration of the driving time

In experiments which were performed, it was seen that some of the persons were completely deprived of sleep and in some other experiments people were deprived of a part of sleep. Some of the researchers took people from the night shift into consideration, in result they found that these people were completely deprived of sleep. Kokonozi,et al. Performed an experiment in which the subjects were monitored for a period of 24 hours to ensure that they were entirely sleep deprived[2].In some of the experiments ,subjects were allowed to sleep for half the period of the time they use to sleep in their daily routine. Peters,et al. Also performed an experiment, in that he studied the effect of sleep deprivation on the drowsiness level of a person for four days and obtained results for no sleep deprivation, partial sleep deprivation and total sleep deprivation[3]. They found that even for the cases of half sleep deprivation, subjects became drowsy after some time .Therefore, it was noted that the quality of last sleep has an impact on drowsiness. Otamani,et al. Took some other factors into consideration which directly have an impact on signals that create drowsiness, in which the one factor is duration of the task[4]. Researches also claim that prolonged driving generates drowsiness.

III.DETECTION TECHNIQUES

A. LAB Color Space

The CIE LAB color space which is also called CIE L*a*b*, is a color space defined by the International Commission on Illumination (CIE). It expresses color as three values: L* for the lightness from black i.e.0 to white i.e.100, a* from green (−) to red (+), and b* from blue (−) to yellow (+) are the 2 vibrant components, which range from -120 to 120.It describe all the colours visible to the human eye and was created to function a device-independent model to be used as a reference. Since the L*a*b* model is a three-dimensional model, it can be represented properly only in a 3-D space. As within the computer, the image is saved in RGB color model. To use Lab color model, the primary step is convert the image from RGB to Lab color model within the experiment. To complete the conversion, the crucial part is to change the image to XYZ color model. After the conversion, L, a, b all belong to [0,128],so we can opt cluster method to classify the colours within the object with Lab color space.

B. Thresholding

Thresholding is one of the most easiest technique of image segmentation. Thresholding is performed when you order the pixel strengths in an image. Thresholding is performed on grayscale images where images have pixel strength which is ranging from 0 to 255. Once you restrict an image you classify these pixels into bunches setting an upper and lower bound to each group. From a grayscale image, thresholding can be utilized to make binary picture. The easiest thresholding technique supersede each pixel in an image with a black pixel, if the image intensity Ix,y is less than some fixed value constant T (that is, Ii,j<T), or a white pixel if the image intensity is more than that of constant(T).For a thresholding algorithm to be authentically efficacious, it should retain logical and semantic content. The types of thresholding algorithms are:

1) Global thresholding algorithms
2) Adaptive thresholding algorithms

In global thresholding, a single threshold for all the photo pixels is used. At the point ,when the pixel estimation of the elements and that of background are fairly anticipated in their individual values over the complete image, global thresholding might be used. In adaptive thresholding, unique threshold values for unique nearby areas are used.

a) Global Thresholding

i) Initial estimate of T
ii) Segmentation using T:
1. G1 are pixels that are brighter than T;
2. G2 are pixels that are darker than or equal to T.
iii) Computation of the average intensities r1 and r2 of G1 and G2.
iv) New threshold value: Tnew = (r1 + r2)/2
v) If |T − Tnew| > _T, back to step 2, otherwise stop.

b) Local properties based Thresholding

i) Statistics based criteria can be used for adjusting the threshold.
ii) For example:

Tij = r_ij + cnij

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IV Apr 2020- Available at www.ijraset.com

Tij = r_ij + cnG

iii) The segmentation is operated the use of a appropriate predicate, Rij:
h(i, j) = {1, if Rij
0, otherwise
where Rij can be, for instance:
g(i, j) >Tij
g(i, j) >r_ij AND g(i, j) >cnij

iv) This method can be easily generalized to a couple of thresholds segmentation.

C. Eye Aspect Ratio(EAR)

To extract the facial landmarks of drivers, Dlib library is imported. The library utilizes a pre-trained face identifier, which is based on a conversion to the histogram of oriented gradients and uses linear SVM (support vector machine)method for object detection. Then the facial landmark indicator is initialized and the captured facial landmarks by using the software are used to calculate distance between points. These separations were utilized to calculate EAR value. EAR is described as the proportion of peak and width of the eye and is computed with the use of equation. The numerator indicates the height of the eye and the denominator signifies the width of the eye and the details of all the landmarks of eye are represent by Fig. 1.

$$EAR = (|p2 - p6| + |p3 - p5|) /2*|p1-p4|$$

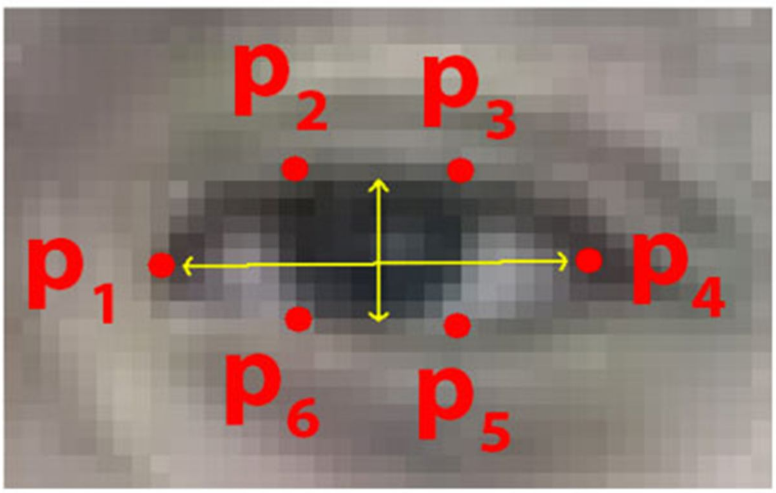


Fig. 1 Landmarks of Eye in EAR

Referring above equation, the denominator represents the horizontal distance of the eye. The numerator calculates the separation between the upper eyelid and the lower eyelid. At the point when the eyes are open, the estimation of numerator increases, for that reason the EAR value grows and when the eyes are closed the estimation of numerator diminishes, in this way the EAR value decreases. In this situation, EAR values are used to detect driver's tiredness. The value of eye aspect ratio is calculated of each eye(left and right eye) and after that mean is taken. In our drowsiness identifier case, the Eye Aspect Ratio is monitored to take a look at if the value falls under threshold value and also it does no longer will increase again above the threshold value in the subsequent frame. The above condition implies that the individual has shut his/her eyes and is in a sleepy condition.On the contrary, if we find the increment in EAR value, it infers that the individual has quite recently squinted the eye and there is no instance of sleepiness.

IV.METHODOLOGY

This section contains the details about the proposed approach for detecting driver's drowsiness. The process starts with a real time webcam as face detection can be done only on an image that is divided into frames . The images are then converted to grayscale image. Dlib library is employed to detect facial landmarks. Using the facial landmarks and Haar cascade sample we detect the eyes and then calculate its's EAR(Eye Aspect Ratio).Now we compare the number of frames for a particular value of EAR to the threshold value of no. of frames. If the value is greater than threshold value the alarm starts beeping.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IV Apr 2020- Available at www.ijraset.com

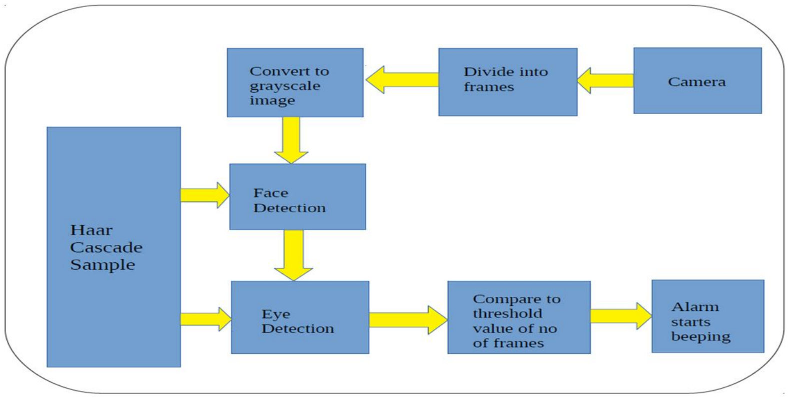


Fig. 2 Process Stages

A. Dividing Into Frames

When detecting drowsiness here we work on a real time situation and the image processing for face detection and eye detection can only be done on an image therefore we divide the real time capturing into frames to work on.

B. Conversion To Gray Scale Image

To convert the images captured into grayscale, we use the function cvtColor function of OpenCV . This function is used for the conversion of an image from one color space to another. As we are dealing with RGB color space , we convert RGB color space to Grayscale. The function syntax is :-

cv2.cvtColor(src,dst,code)->None

cv2.cvtColor(frame,gray,cv2.COLOR_BGR2GRAY) //used in our code

src –the input image

dst -the output image of same size and depth as of src

code – color space conversion code

Formula – RGB[A] to Gray : Y<-0.299.R+0.587.G+0.114.B

C. Face Detection

To determine the facial landmarks of the driver, Dlib library is imported and deployed. The library uses a pre-trained face detector of OpenCV(Haar cascade), which is based on extracting features from images using a kind of 'filter', similar to the concept of convolutional kernel(these filters are called Haar features) . Only facial related structures or features are detcted and all other objects are ignored.

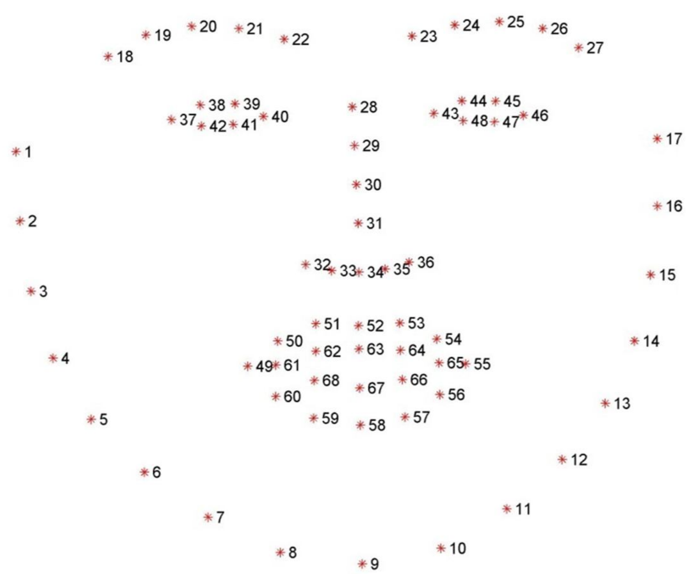


Fig. 3 Haar Cascade sample for facial landmarks

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IV Apr 2020- Available at www.ijraset.com

D. Eye Detection

Similarly as the face landmarks are detected using Haar cascades, the eye is also detected. And when the eye is detected the position of eyes are used to calculate the EAR(Eye Aspectt Ratio) for the respective frame.

EAR is defined as the ratio of height and width of the eye :-

$$EAR=(|p2-p6|+|p3-p5|)/(2*(|p1-p4|))$$

The numerator denotes the height of the eye and the denominator denotes the width of the eye.

The average eye aspect ratio is 0.339 when the eyes are opened and 0.141 when the eyes are closed.

Comparison with threshold value of no. of frames

A threshold value of 50 frames is preset. A counter initialized with 0 is incremented by 1 for the EAR of closed eye until the threshold value is reached or it detects the EAR of open eyes. As soon as the counter surpasses the threshold value, the alarm goes off beeping.

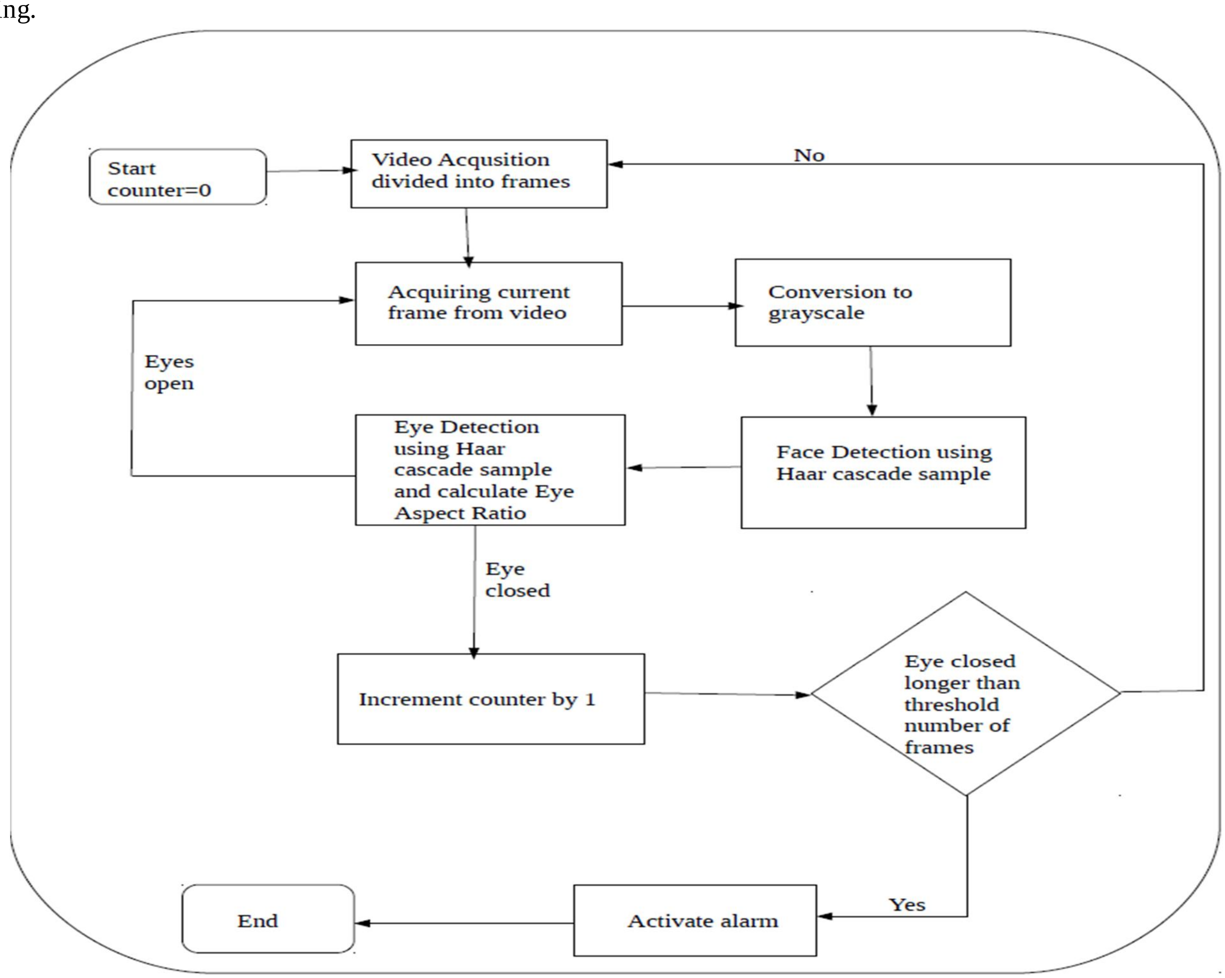


Fig. 4 Flowchart for whole process

V. HARDWARE AND SOFTWARE SETUP

A. Haar Features

Haar features are digital imaging characteristics that are used for detecting object. These are similar to rectangles having dark and light areas which resembles features of our face.Just deploy those features throughout our face to get the output of each feature identified. Since, all faces share some general features like region near eye is darker than the region near upper-cheeks and the region of nose bridge is brighter than the eye region. So these kind of characteristics of face are used for developing haar like features. Each feature is related to a particular spot on the face. To obtain different values rectangular frames are moved over the face.

Value = ∑ (pixels in dark area) **-** ∑ (pixels in light area)

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IV Apr 2020- Available at www.ijraset.com

B. Raspberry Pi

Raspberry Pi is a low priced, card sized computer that is used for the implementation of small projects. A monitor must be connected with it externally in order to visualize its operating system and operate it. A keyboard and a mouse can be used to provide the input. An external memory has to be used to load operating system of the Raspberry Pi. Several languages like C++, Python etc. can be used for programming. Raspberry Pi has the following components :

1) Quad core Cortex-A72(1.5 GHz) processor.
2) 2GB RAM.
3) USB ports for external devices.
4) Micro SD card slots.
5) Ethernet port.
6) HDMI port.
7) 40 GPIO pins.
8) Camera interface.
9) Display interface.
10) Power supply port.

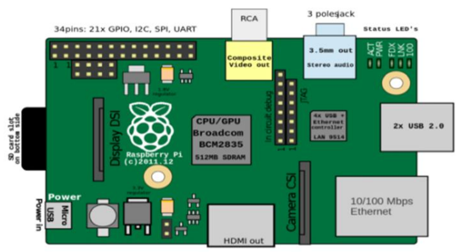


Fig. 5 Different components of RaspberryPi

C. Camera

We can use a web camera and mount it on the dashboard of the automobile in order to capture the picture of the driver . In spite of the fact that the camera can create a video clip, we need to apply developed algorithm at each edge of the video stream. This paper is only focused on the application of the proposed mechanism on single frame only. The camera used in the model is a low priced web camera with 30 fps frame rate in VGA mode.

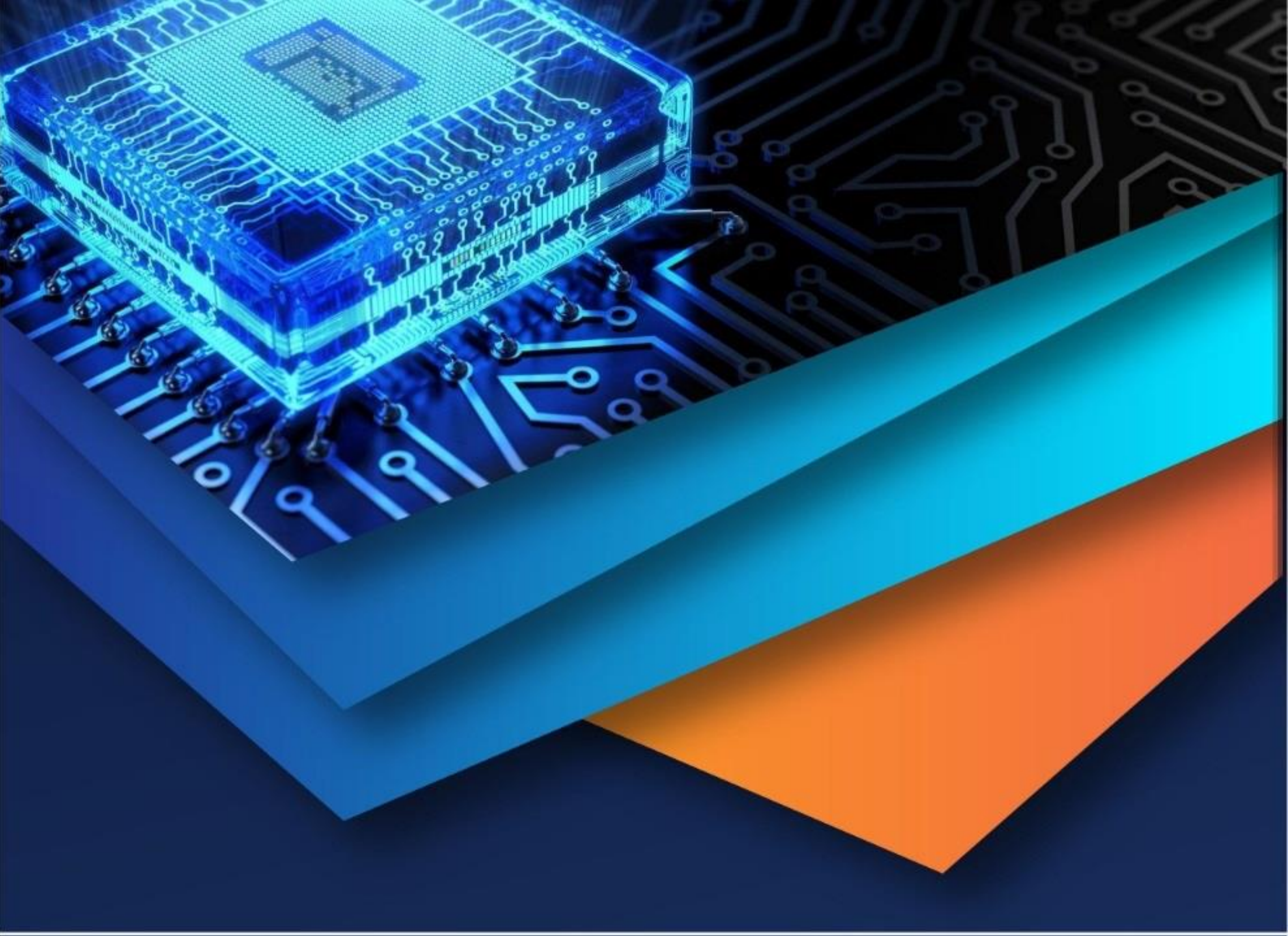
Fig. 6 Camera used for implementing sleep detection system

VI. CONCLUSION

Sleep Detection System is proposed to help a driver to stay awake while driving so that the possibility of car accidents due to drowsiness can be reduced. This paper is concerned with designing a system that uses a camera which points directly at the driver's face to monitor the eyes for the detection of fatigue by image processing algorithm.

REFERENCES

[1] T. Soukupova and J. Cech, "Real-time eye blink detection using facial landmarks," Computer Vision Winter Workshop (CVWW), 2016.
[2] A Study of Heart Rate and Brain System Complexity and Their Interaction in Sleep-Deprived Subjects. Kokonozi A.K, Michail E.M., Chouvarda I.C., Maglaveras N.M. Bologna, Italy. : Computers in Cardiology, 2008.
[3] Effects of partial and total sleep deprivation on driving performance. Peters R.D., Wagner E., Alicandri E., Fox J.E., Thomas M.L., Thorne D.R., Sing H.C., Balwinski S.M. 1999.
[4] Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers. Otmani S, Pebayle T, Roge J, Muzet A. 2005.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY