VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

School of Information Technology and Engineering

Digital Assignment-II, MARCH 2021

B.Tech., Winter-2020-2021

Driver Drowsiness Detection System

COURSE CODE	ITE3999
COURSE NAME	Technical Answers to Real World Problems (TARP)
SLOT	TE1
FACULTY	Prof. VIJAYAN E.

TEAM 7 MEMBERS:

REG. NO.	NAME
18BIT0027	SAKINA HUSSAIN BANDOOKWALA
18BIT0231	KUSHAGRA AGARWAL
18BIT0272	PRIYAL BHARDWAJ

Description of the Proposed Methodologies and Architecture

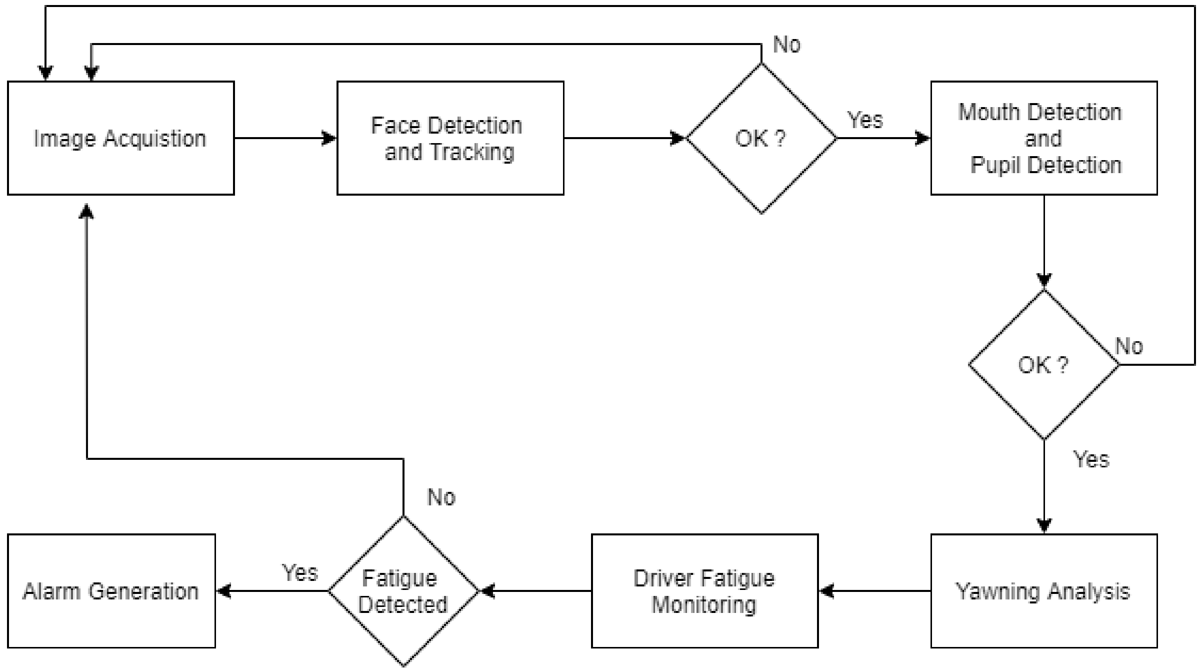


Image Acquisition:

Utilizing a web camera introduced inside the automobile we can get the picture of the driver. Despite the fact that the camera creates a video clip, we have to apply the developed algorithm on each edge of the video stream.

Face Detection:

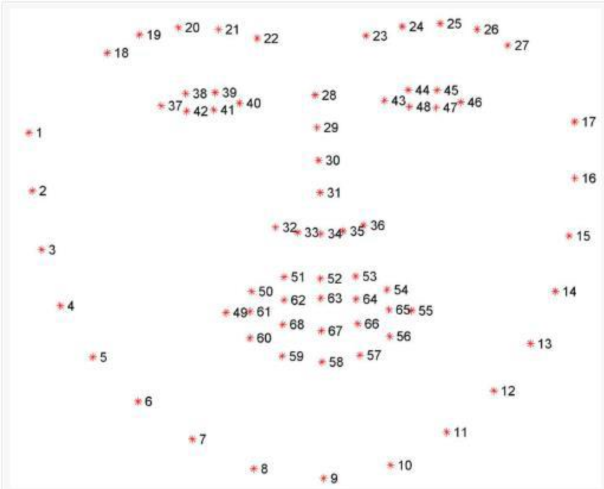
In this stage we detect the region containing the face of the driver. A specified algorithm is implemented for detection of face in every frame. By face detection we mean finding location of facial characters through machine learning model. The frame may be any random frame. Only facial related structures or features are detected and all others types of objects are ignored.

Process:

- In the system we have used facial landmark prediction for face detection.
- The facial landmark detector included in the dlib library is an implementation of the One Millisecond Face Alignment with an Ensemble of Regression Trees paper by Kazemi and Sullivan (2014).
- This method starts by using:

- ➢ A training set of labeled facial landmarks on an image. These images are manually labeled, specifying specific (x, y)-coordinates of regions surrounding each facial structure.
- ➢ Priors, of more specifically, the probability on distance between pairs of input pixels.

- The pre-trained facial landmark detector inside the dlib library is used to estimate the location of 68 (x, y)-coordinates that map to facial structures on the face.



- These annotations are part of the 68-point iBUG 300-W dataset which the dlib facial landmark predictor was trained on.
- On applying dlib library, we can extract face from the image frames captured by the camera.

Eye and Mouth Detection:

After successful detection of face, eyes as well as mouth needs to be detected for further processing. At this point our model performs the detection of these facial features in the required particular region with the use of detection of several features. When eye and mouth detection is done then the result is matched with the reference or threshold value for deciding the state of the driver.

Process:

- We will use the same facial landmark detector to detect eyes and mouth that we used to detect the whole face.
- From the above facial landmarks image, facial regions can be accessed via simple Python indexing.
- The right eye using [36, 42].

- The left eye with [42, 48].
- The mouth can be accessed through points [48, 68].
- After successfully detecting eyes and mouth, we will be calculating respective Aspect Ratio and compare them to the threshold value to determine drowsiness.

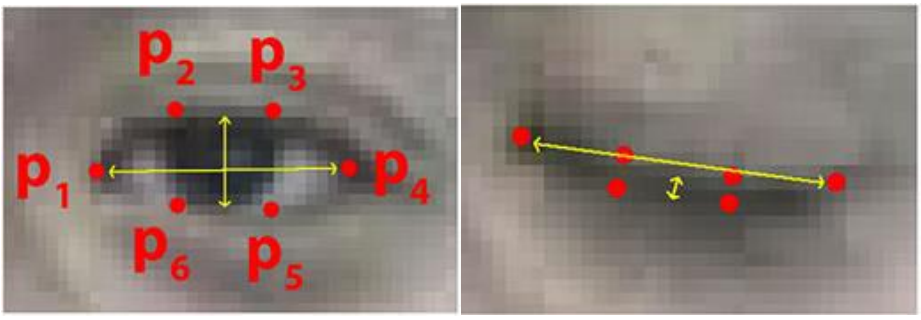
Eye Aspect Ratio (EAR)

For every video frame, the eye landmarks are detected. The eye aspect ratio (EAR) between height and width of the eye is computed.

$$EAR = \frac{||p2 - p6|| + ||p3 - p5||}{2||p1 - p4||}$$

where p1, . . ., p6 are the 2D landmark locations. The EAR is mostly constant when an eye is open and is getting close to zero while closing an eye.

Aspect ratio of the open eye has a small variance among individuals, and it is fully invariant to a uniform scaling of the image and in-plane rotation of the face. Since eye blinking is performed by both eyes synchronously, the EAR of both eyes is averaged.



Mouth Detection

The next step towards yawning detection is to find the location of mouth and the lips. To do so, mouth area will be segmented in the face. The strong difference between lips colour and face colour is used in our method. In the mouth region, the red colour is the strongest component while the blue component is the weakest. Usually, the mouth area is detected based on colour information, after the face is located. The following equations are used to generate the mouth map.

$$\text{Mouth\_map} = (C_r)^2 \times \left((C_r)^2 - \frac{\eta \times C_r}{C_b} \right)^2$$

$$\eta = 0.95 \frac{\frac{1}{n} \sum_{(x,y)} C_r(x,y)^2}{\frac{1}{n} \sum_{(x,y)} \left(C_r(x,y) / C_b(x,y) \right)}$$

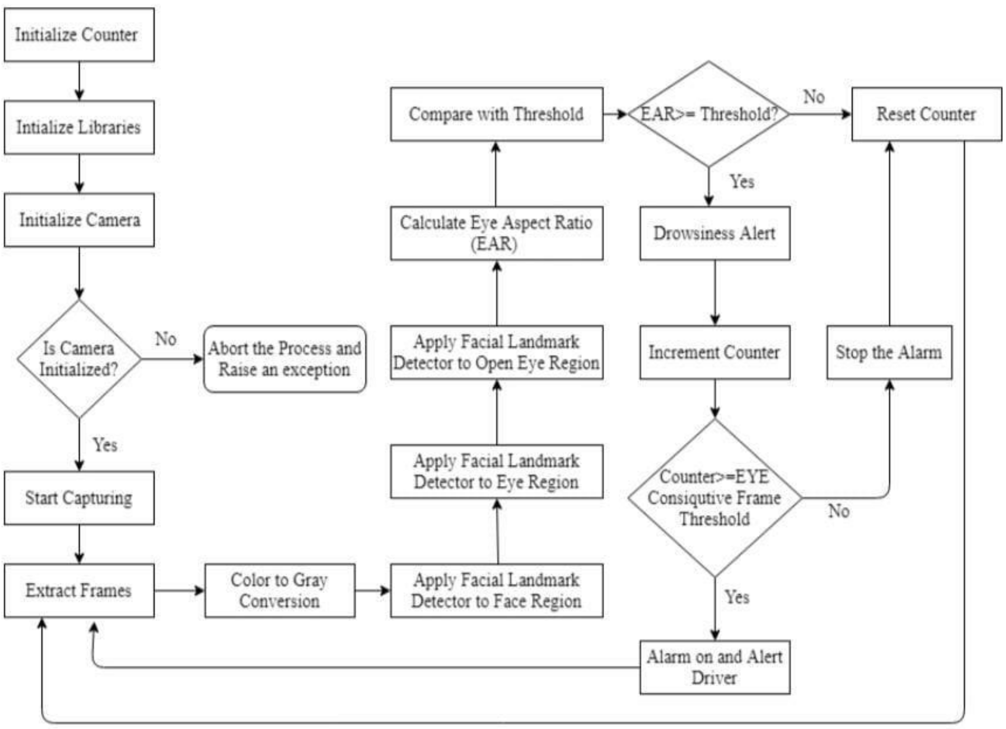
The mouth map will then go through some post processing steps such as black and white conversion, erosion, dilation and finding the biggest connected components in the same way as the eye detection scheme. The geometrical features of the face and relative location of the mouth with respect to eyes can be exploited in this step to verify the validity of the detected mouth.

Driver Fatigue Monitoring:

After we have successfully located eyes and mouth, we will be calculating eye blink frequency and yawning pattern. If both the values cross their respective threshold value then we will sound the alarm else we will repeat the process.

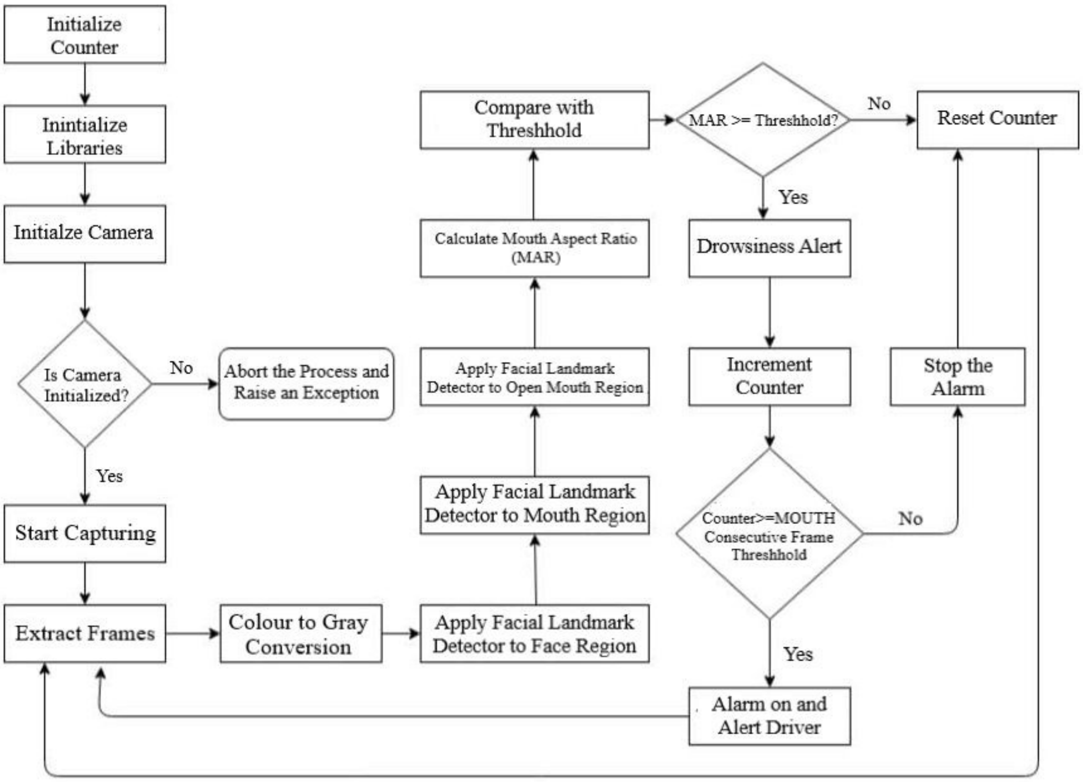
Process:

<u>Eye State Detection</u>



- The decision for the eye state is made based on EAR calculated in the previous step. If the distance is zero or is close to zero, the eye state is classified as “closed” otherwise the eye state is identified as “open”.
- The last step of the algorithm is to determine the person’s condition based on a pre-set condition for drowsiness. The average blink duration of a person is 100-400 milliseconds (i.e. 0.1-0.4 of a second).
- Hence if a person is drowsy his eye closure must be beyond this interval. We have set an interval of 48 frames which is nearly 3 seconds. If the eyes remain closed for three or more seconds, drowsiness is detected and alert regarding this is triggered.

<u>Yawn detection</u>



Yawning is characterized by a widely opened mouth. Like the eye closure detection, the facial landmarks are used to detect an open mouth. The flowchart given above describes the process of detecting the event of yawning from the instance a face is detected. Mouth Aspect Ratio (MAR) is the parameter used to determine if the subject’s mouth is open. If MAR calculated from the frame is greater than the threshold, the subject is determined to be yawning. An alarm is raised if the subject has yawned more than the set boundary value consecutively. Small openings that in reality are taken as a result of talking, eating is ignored.

Drowsiness Detection

The last step of the algorithm is to finally determine if the driver is drowsy or not. We will add up results of both the parameters, i.e. result from Eye state detection and Yawn Detection. If both the parameters cross the threshold value, we will sound an alarm to alert the driver. Along with alarm, we will also send a message to the emergency contact given by the driver which will contain the location of driver. If none of the parameter crosses the threshold value, then the system will keep repeating the process till the end of drive.

****