Drowsiness Detection

Anirudh Bhupalarao

Abstract :

For a given video stream, detect whether the eyes have been closed for a particular amount of time to tell if the person is drowsy. This is done by using real-time facial landmark detection. If the eyes have been closed for a certain amount of time, we'll assume that they are starting to doze off and play an alarm to wake them up and grab their attention. This can help reduce the number of accidents that happen every year on the road.

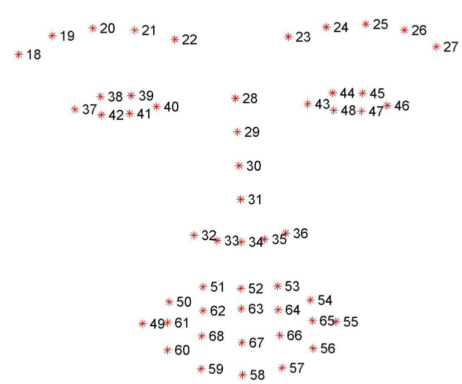
Approach Used :

As the project statement requires us to detect drowsiness we use image processing techniques like facial landmark detection to detect whether the eyes are closed or not. Facial landmarks are used to represent salient features of the face such as eyes, nose, mouth, ears etc. They have been successfully applied to face alignment, blink detection, head pose estimation and much more. Detecting facial landmarks is a subset of the shape prediction problem. Given an input image (and normally an ROI that specifies the object of interest), a shape predictor attempts to localize key points of interest along the shape. In the context of facial landmarks, our goal is to detect important facial structures on the face using shape prediction methods. This is done in two steps :

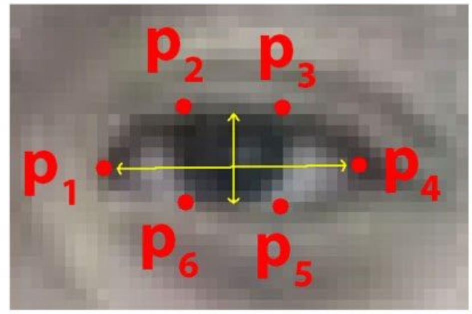
- Localize the face in the image
- Detect key facial structures on the face ROI (Region of Interest)

What's important is that we obtain the face bounding box (i.e., the *(x, y)*-coordinates of the face in the image). 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 uses a training set of labeled facial landmarks on an image. These images are manually labeled specifying specific x-y coordinates for each facial structure and also uses the distance between pairs of input pixels. Given this training data, an ensemble of regression trees are trained to estimate the facial landmark positions directly from the pixel intensities themselves.

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 are taken from the 68 point iBUG 300-W dataset.



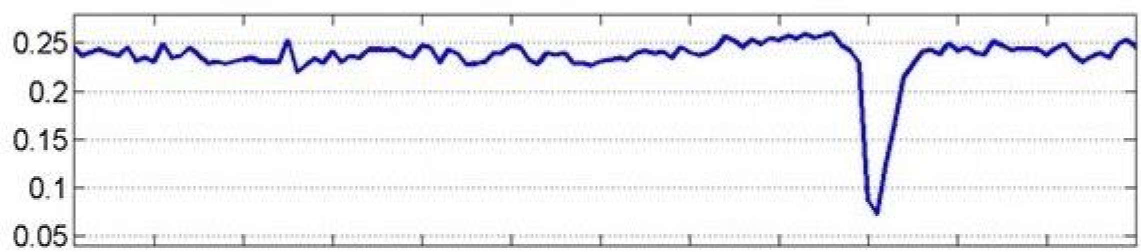
To build our drowsiness detection algorithm we will be computing a metric called the Eye Aspect Ratio (EAR). It involves a very simple calculation based on the ratio of distances between facial landmarks of the eyes. Each eye is represented by 6 (x, y)-coordinates, starting at the left-corner of the eye and then working clockwise around the remainder of the region.



Based on this image, we see that there is a relation between the width and height of these coordinates. We use this to derive an equation to compute the EAR :-

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

This eye aspect ratio is approximately constant while the eye is open, but will rapidly fall to zero when a blink is taking place.



Now to implement this into a drowsiness detection algorithm, we check if the eye aspect ratio indicates that the eyes have been closed for a sufficiently long enough amount of time and if so sound an alarm to wake up the driver.

Recommendations :

This method has proven to be successful with accurate results shown. However due to noise in a video stream, subpar facial landmark detections, or fast changes in viewing angle, a simple threshold on the eye aspect ratio could produce a false-positive detection. To make our detector even more robust we could compute the EAR for the N-th frame along with the eye aspect ratios for *N – 6* and *N + 6* frames, concatenating these eye aspect ratios to form a 13 dimensional feature vector and train a Support Vector Machine (SVM) on these feature vectors.

References :

[1] Kazemi Vahid, Sullivan Josephine, “One Millisecond Face Alignment with an Ensemble of Regression Trees”, IEEE Conference on Computer Vision and Pattern Recognition, 2014.
[2] Soukupova Tereza, Cech Jan, “Real Time Eye Blink Detection using Facial Landmarks”, Czech Technical University, Prague, 2016.