TASK 2-COMPUTER VISION

Sanika Ardekar

INTRODUCTION

In a car safety technology, driver drowsiness detection is very essential to prevent road accidents. Now-a-days, many people using automobiles for daily commutation, higher living standards, comfortability, and timing constraints to reach destinations. This trend leads to high volumes of traffic in urban areas and highways. In turn, it will raise number of road accidents with several factors. Driver drowsiness could be the one reason for road accidents. One way to reduce number of accidents is early detection of driver drowsiness and alerting with an alarm before it is too late.

To determine the level of driver drowsiness various measures are used. These measures are Physiological Measures, Behavioural Measures and Vehicle-based Measures. In our proposed method Behavioural Measures are used to detect driver drowsiness.

PROBLEM STATEMENT:

Given a dataset of human face images taken from a dashboard of a car, detect if the driver is drowsy or not. Note that a driver can be classified as drowsy if their eyes are closed more often than open.

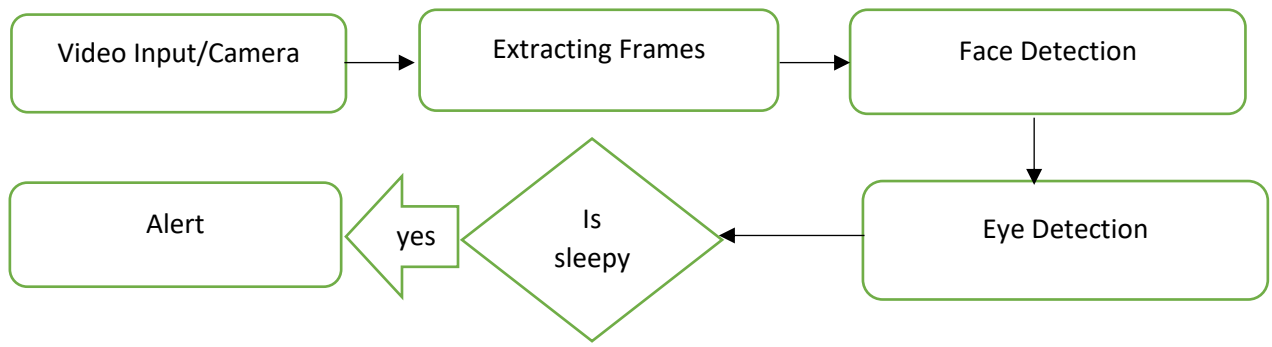
DATASET: Dataset contains images of open eye and closed eye.

OBJECTIVE:

The objective of this project is to build a drowsiness detection system that will detect that a person's eyes are closed for a few seconds. This system will alert the driver when drowsiness is detected. The model will classify whether the person's eyes are 'Open' or 'Closed'.

METHOD: Convolutional Neural Network (CNN)

APROACH



A) Face detection and Eye Region Extraction

Whole face region may not be required to detect the drowsiness but only eyes region is enough for detecting drowsiness. At first step by using the Viola-jones face detection algorithm face is detected from the images. Once the face is detected, Viola-jones eye detection algorithm is used to extract the eye region from the facial images.

B) Feature Extraction

Feature extraction is one type of dimensionality reduction where useful parts of an image represented as a feature vector. In this method features from the eye region images are extracted using a Convolutional Neural Network (CNN).

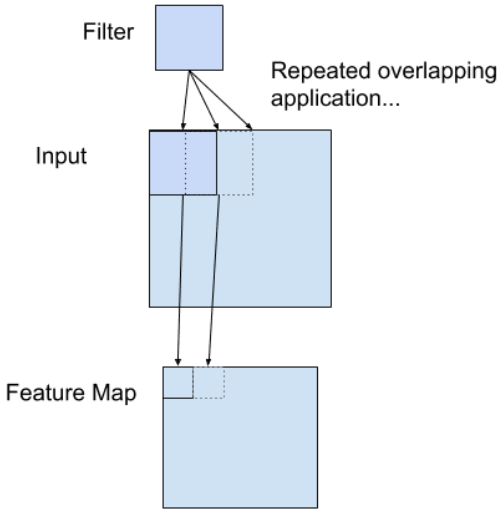
CNN:

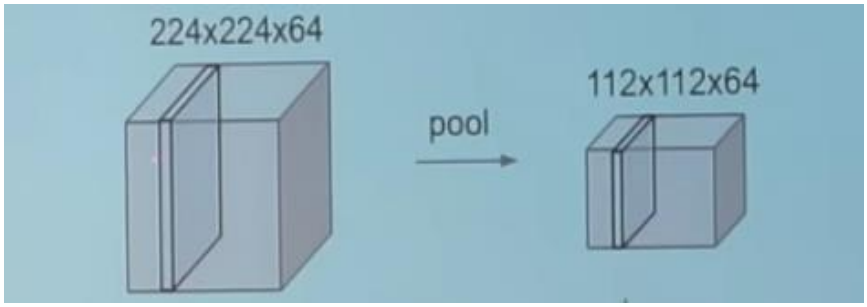
Convolutional neural network (CNN) is used for detection of driver drowsiness.
A convolution is a combined integration of two functions that shows you how one function modifies the other. There are three important items to mention in this process: the input image, the feature detector, and the feature map. The input image is the image being detected. The feature detector is a matrix, usually 3x3 (it could also be 7x7). A feature detector is also referred to as a kernel or a filter.

Since a feature detector is needed for each drowsy image to compare with existing features in a database to detect either drowsy or not. Usually, CNNs requires fixed size images as input so pre-processing is required. The pre-processing includes extracting the key frames from video based on temporal changes and store in database. From these stored images, feature vectors are generated in convolution layers of CNN. These feature vectors are then used for the detecting the driver drowsiness.

A convolution is the simple application of a filter to an input that results in an activation. Repeated application of the same filter to an input result in a map of activations called a feature map, indicating the locations and strength of a detected feature in an input, such as an image. Convolution layer has kernels (filters) and each kernel having width, depth and height.

Spatial invariance is a concept where the location of an object in an image doesn't affect the ability of the neural network to detect its specific features. Pooling enables the CNN to detect features in various images irrespective of the difference in lighting in the pictures and different angles of the images. CNN uses pooling layers (Max or Average) to minimize the size of the feature maps to speed up calculations. Its function is to progressively reduce the spatial size of the representation to reduce the number of parameters and computation in the network.





Once the pooled featured map is obtained, the next step is to flatten it. Flattening involves transforming the entire pooled feature map matrix into a single column which is then fed to the neural network for processing.

The rectified linear activation function or ReLU for short is a piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero. The ReLU layer applies the max/average function on all the values in the input data and changes all the negative values to zero. The following equation shows the ReLU activation function:

$$\mathrm{ReLU'(x)} = \begin{cases} 0, & \text{for } x < 0 \\ 1, & \text{for } x \geq 0 \end{cases}$$

SoftMax classifier is used for classification. On the basis of these features, the SoftMax layer classifier was trained. During testing phase, we capture the video frames through camera and alert with an alarm when the model predicts drowsy output state continuously. Static images are used for training but during testing phase key frames are extracted from continuous video and tested against the trained static images.

A binary signal for each frame in the form of drowsy or non-drowsy face is been obtained. For an alert signal to be delivered to a driver, at least 40 out of 60 frames should be detected as drowsy. A buffer of 60 recent frame outputs is maintained and a warning is sent to the driver in the form of an alerting sound. Thus, the driver is being successfully alerted.

CONCLUSION

In this method used for driver drowsiness detection based on eye state. This determines the state of the eye that is drowsy or non- drowsy and alert with an alarm when state of the eye is drowsy. Face and eye region are detected using Viola-Jones detection algorithm. Convolution neural network is developed to extract features using filter to develop feature map, pooling, flattening and ReLU and used for learning phase. A SoftMax layer in CNN classifier is used to classify the driver as sleep or non-sleep.