Driver Drowsiness Detection useful for logistics

Introduction and relevance:

- 1 in 4 vehicle accidents are caused by drowsy driving and 1 in 25 adult drivers report that they have fallen asleep at the wheel in the past 30 days
- Drowsy driving can be as small as a brief state of unconsciousness when the driver is not paying full attention to the road
- Drowsy driving results in over 71,000 injuries, 1,500 deaths, and $12.5 billion in monetary losses per year

Formulating the problem statement:

- Most of the research around this area use the term “prediction” and refers to estimating labels of the data point not used in training. However, actually predicting the ground truth in subsequent few minutes is equally, if not more critical.
- Some studies are able to predict the time when the microsleep will occur in between 15s to 5 min in advance.
- Hence look at two problems
    - Estimating the drowsiness level of the driver in a current set up
    - Predicting the subsequent micro sleep

Approach

Traditionally, this problem can be approached in 3 different levels-

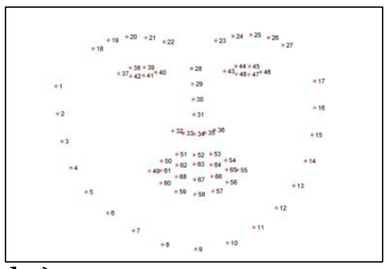
- Rule based solution (Assigning thresholds on defined features): This is the first level of solution, where we define rules based on human expertise on the features.
- Machine Learning based solution (Learning the rules from the defined features): Here we apply statistical techniques for feature selection and learn the rules via machine learning models.
- Deep Learning based solution (Learning the features): Feature extraction requirement is minimal, if not negligible. This approach also has an edge over the previous two since it can handle temporal aspects of the solution as well (in the case of RNNs)

Data and proposed features

Hypotheses-

- It is hypothesized that it is possible to predict when the impaired state will arise by using the sensors, physiological and performance indicators used to detect drowsiness, and or labelling from expert human observer.
- It is hypothesized that adding information such as driving time and participant information will improve the accuracy of the model
- It is hypothesized that adding facial key features such as eyelids and mouth will increase model performance.

Facial Key Feature Extraction-
Facial Key features can be extracted from the pretrained models using open cv, face-net etc.



Proposed Features-

- Blink duration
- Blink frequency
- Average standard deviation of closed eye time (Can be calculated using facial landmarks)
- Average standard deviation of head position in X direction
- Average standard deviation of head position in Y direction
- Head rotation in X direction
- Head rotation in Y direction
- Eye Aspect Ratio – Height of eye divided by width of eye normalized for each person
- Mout Aspect Ratio – Height of mouth divided by width of mouth normalized for each person

Error definition:

- For first problem statement, we use categorical cross entropy for classification
- For the second problem statement, we propose - The error to be the difference between the time remaining from the current frame before the target level is really reached

Modelling technique:

Basic Classification Methods

- Machine Learning approach - After we extracted and normalized our features, we try a series of modelling techniques, starting with the most basic classification models like logistic regression and Naive Bayes, moving on to more complex ensemble models. It’s important to note the trade-off between interpretability and complexity of the model.
- Deep Learning Approaches
    - 1-D CNN – Convolutional Neural Networks (CNN) are typically used to analyse image data and map images to output variables. However, we can build a 1-D CNN and send in numerical features as sequential input data to try and understand the spatial relationship between each feature for the two states.
    - LSTM- Another method to deal with sequential data is using an LSTM model. LSTM networks are a special kind of Recurrent Neural Networks (RNN), capable of learning long-term dependencies in the data.