Data set 1

Most efficient data set

# **Facial point annotations**

****Description:****

Existing facial databases cover large variations including: different subjects, poses, illumination, occlusions etc. However, the provided annotations appear to have several limitations.

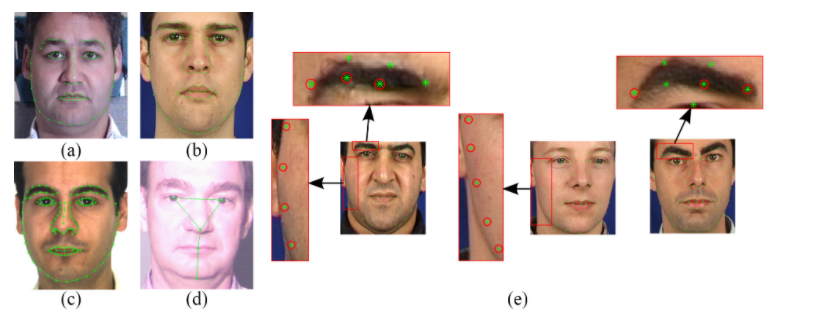
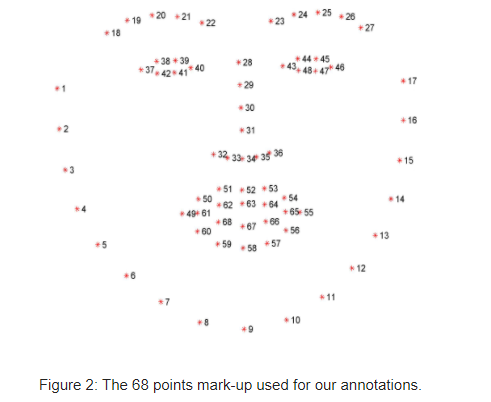


 Figure 1:  (a)-(d) Annotated images from MultiPIE, XM2VTS, AR, FRGC Ver.2 databases, and (e) examples from XM2VTS with inaccurate annotations.

1. The majority of existing databases provide annotations for a relatively small subset of the overall images.
2. The accuracy of provided annotations in some cases is not so good (probably due to human fatigue).
3. The annotation model of each database consists of different number of landmarks.

These problems make cross-database experiments and comparisons between different methods almost infeasible. To overcome these difficulties, we propose a semi-automatic annotation methodology for annotating massive face datasets. This is the first attempt to create a tool suitable for annotating massive facial databases.

All the annotations are provided for research purposes ONLY (NO commercial products).



Data set 2

Simpler one

Using frame rate of video of driver.

The RLDD dataset consists of around 30 hours of RGB videos of 60 healthy participants. For each participant we obtained one video for each of three different classes: alertness, low vigilance, and drowsiness, for a total of 180 videos. Subjects were undergraduate or graduate students and staff members who took part voluntarily or upon receiving extra credit in a course. All participants were over 18 years old. There were 51 men and 9 women, from different ethnicities (10 Caucasian, 5 non-white Hispanic, 30 IndoAryan and Dravidian, 8 Middle Eastern, and 7 East Asian) and ages (from 20 to 59 years old with a mean of 25 and standard deviation of 6). The subjects wore glasses in 21 of the 180 videos, and had considerable facial hair in 72 out of the 180 videos. Videos were taken from roughly different angles in different real-life environments and backgrounds. Each video was self-recorded by the participant, using their cell phone or web camera. The frame rate was always less than 30 fps, which is representative of the frame rate expected of typical cameras used by the general population.

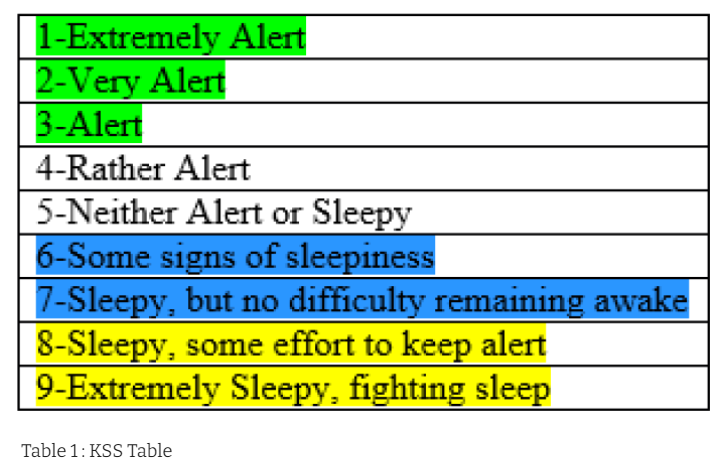
Each video was self-recorded by the participant, using a cell phone or web camera of the participant. The frame rate was always less than 30 fps, which is representative of the frame rate expected of normal cameras used by the general population.

 The three classes were explained to the participants as follows:

1) ****Alert :**** One of the first three states highlighted in the KSS table in Table 1. Subjects were told that being alert meant they were completely conscious so they could easily drive for long hours.

2) ****Low Vigilant :**** As stated in level 6 and 7 of Table 1, this state corresponds to subtle cases when some signs of sleepiness appear, or sleepiness is present but no effort to keep alert is required. While subjects could possibly drive in this state, driving would be discouraged.

3) ****Drowsy :**** This state means that the subject needs to actively try to not fall asleep (level 8 and 9 in Table 1).



[https://uta.box.com/s/e7nsmloj8xmblosvfg98q42fgqnjy6dv](https://www.google.com/url?q=https://uta.box.com/s/e7nsmloj8xmblosvfg98q42fgqnjy6dv&sa=D&sntz=1&usg=AFQjCNGZdOEWnjYJYOX9e95AjyfG1ufDsA" \t "https://sites.google.com/view/utarldd/_blank)

<https://sites.google.com/view/utarldd/home>

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*