Driver's Drowsiness Detection

Challa Yashwanth[1] and Jyoti Singh Kirar[2]

Abstract— **The amelioration of technology from the past 50 years accommodated a good amount of succour to the driver by providing a great level of comfort and safety in the vehicles. The accidents may occur because of many reasons and one of the reason which we are going to portray and solve in this paper is driver fatigue. In this paper, we are going to use Artificial Intelligence-based advanced algorithms to detect driver fatigue and the rate at which the driver is drowsy. We propose an algorithm that uses eye and mouth vertical distances, eye closure, yawning and other engineered facial features to detect driver drowsiness.**

I. INTRODUCTION

Precaution is better than cure, we, by this research are trying to bring that proverb into action. In India, there were 4, 64,674 [1] accidents in 2015 and theyve been increasing strenuously since then. The major cause for these accidents is, either the driver does not follow the traffic rules or he is drowsy. The US National Highway Traffic Safety Administration has gauged about 100,000 accidents every year provoked predominantly because of the drowsiness of driver[2]., we can only imagine what it is like for India which has a higher road accident rate compared to the USA. We have discussed several Machine Learning algorithms that were proposed by other researchers to automatically detect the drowsiness of the driver. However, these algorithms do not utilize all the relevant features for classification such as the vertical distances of eyes and mouth.

Automatic processes invented to know about driver weariness and detect driver sleepiness will be an intrinsic part of the upcoming automatic vehicle inventions. There are many features to consider for detecting driver fatigue and some of them which are used till now are lane detection, pulse and heart rate, steering wheel movement[3] pattern detection, etc. These are the ways in which the driver would not get any obstruction because of the system. As described earlier, researchers also used physical and mental features[4] for detecting fatigue while driving.

In this work, we will be developing a Machine Learning, Pattern Recognition and Computer Vision-based algorithm that will use all the relevant features for fast and accurate classification of drowsiness in a driver. We will also develop an algorithm that will determine the level of drowsiness of the driver called the rate of drowsiness.

*This work was supported by Shiv Nadar University,NH91, Tehsil Dadri,Greater Noida, Uttar Pradesh 201314,India

[1]Challa Yashwanth is with Department of Computer Science Engineering, Shiv Nadar University,NH91, Tehsil Dadri,Greater Noida, Uttar Pradesh 201314, India

[2]Jyoti Singh Kirar is with Faculty of Computer Science Engineering, Shiv Nadar University,NH91, Tehsil Dadri,Greater Noida, Uttar Pradesh 201314, India.

II. RELATED WORK

Almost all the published researches which rest on machine learning approaches are image-based real-time systems for drowsiness supervising using classic facial attributes. Singh[5] invented the machine learning-based system that works on eye blink duration. Saito[6] used the driver's psychological and physical conditions to detect drowsiness. Horng[7] used edge detection for detecting eyes and dynamical matching for eye-tracking and also for driver drowsiness recognition. Smith[8] published a machine algorithm that mainly focuses on trailing a person's head by using optical flow. Kartik Dwivedi[9] invented a new drivers fatigue detection by anchoring multi-layer convolutional neural networks. It is proved in Zutao zang[10] paper that if the person closes his/her eyes for 5 seconds then he is considered as drowsy but in reality, we need more time to decide drowsiness of a person and the rate at which the person is drowsy is also not considered.

III. PRE-PROCESSING

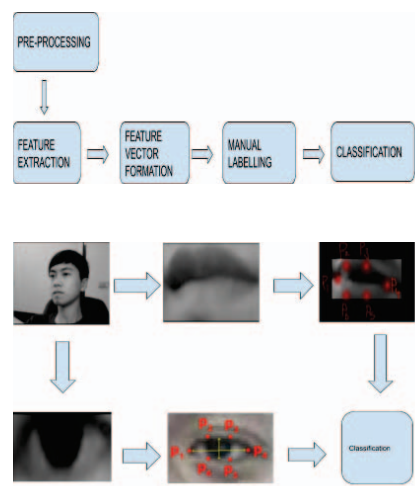

Fig. 1. An Example of Proposed Work.

A. Framing

We have converted video into frames using the OpenCV library of python. We have captured the video using the VideoCapture method of OpenCV and constructed every frame by the read method included in the library, wrote these images into a folder using imwrite method from the OpenCV library. The dataset we had was videos(30fps) of drowsy people. By framing the videos we converted every video into 1500+ frames(videos had varied duration, so we took 1500 samples as limit).

B. Framing Threshold

We have used Eye aspect Ratio(EAR) [11] for identifying the frames in which the subject is drowsy, Mouth aspect ratio(MAR) is considered for mouth region which is another feature in the model proposed.EAR is calculated by adding the two distances from the top end to the bottom end of the eye, which is then divided by the horizontal distance of the eye. Mouth aspect ratio(MAR) is also calculated similarly, where both are elucidated below.

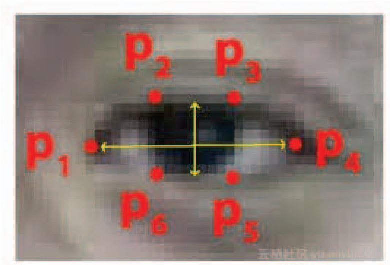

Fig. 2. Detection points for eye

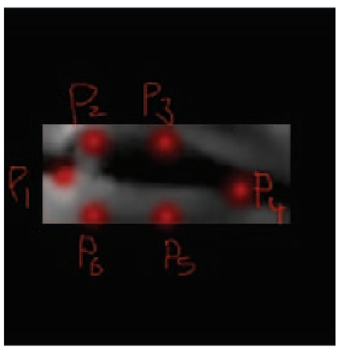

Fig. 3. Detection points for mouth

$$MAR = ||p2 - p6|| - ||p3 - p5||/2||p1 - p4|| \quad (1)$$

$$EAR = ||p2 - p6|| + ||p3 - p5||/2||p1 - p4|| \quad (2)$$

IV. FEATURE EXTRACTION

A. Calculation of vertical distance

We have defined Vertical distance for eyes and mouth, which are discussed below.

1) Eye vertical distance(EVD): The Eye is detected by dlib which is an OpenCV library dlib and a Dlibs 68 Point Facial Landmark Detector. Every frame is cropped to the eye region which makes it easier for future calculations. The length would always be the same for a particular subject, only the width would be changing as stated below by an example. The width of the cropped image is calculated and is defined as eye vertical distance(EVD). In the below figures 250 pixels is the length and 93, 56 are the widths.

2) Mouth vertical distance(MVD): The mouth is detected by the OpenCV library dlib and a Dlibs 68 Point Facial Landmark Detector. Every frame is cropped only to the mouth region which makes it easier for future calculations as done for the eye region. The length would always be the same for a particular subject, only the width would be changing. The width of the cropped frame is calculated and is defined as mouth vertical distance(MVD).

B. Feature vector formation

We calculate the ratio of vertical distances both EVD(eye vertical distance) and MVD(mouth vertical distance) frame by frame to the average value of all EVD(AEVD) and MVD(AMVD) respectively in which the subject is drowsy, we put the values in two different columns for classifying the subject drowsiness state.

$$RATIOeye = EVD/AEVD \quad (3)$$

$$RATIOmouth = MVD/AMVD \quad (4)$$

V. MANUAL LABEL FORMATION

We have defined our own label format which is shown in the fig.4 and fig.5.

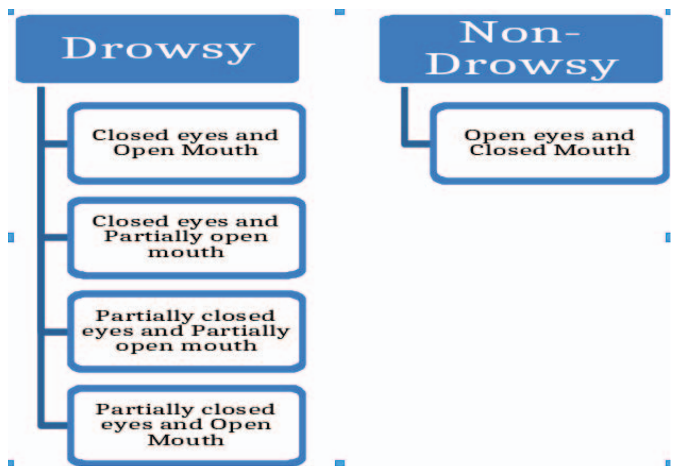

Fig. 4. The sub-states of Drowsy and Non-Drowsy states.

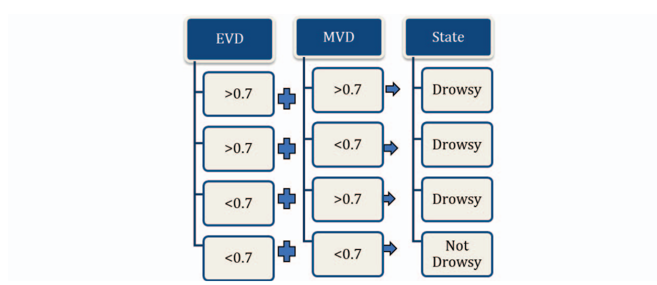

Fig. 5. The EVD and MVD values that are considered for corresponding states.

VI. Experimental setup

A. dataset description and experimental setup

The driver fatigue dataset possessed by NTHU Computer Vision Lab[12] is used. In this dataset, there were 9 subjects who had a different origin which made the dataset more diverse. The subjects were simulated under day and night conditions and videos are trimmed if the subject yawns, feels drowsy or if the blink rate is reduced. The subjects were conditioned to sit on a chair and play a driving simulator with a steering wheel and pedal. The total time of the entire dataset is about half an hour.

The video resolution is 640x480 in AVI format. The frame rate of the recorded videos is 30FPS.

B. Result Analysis

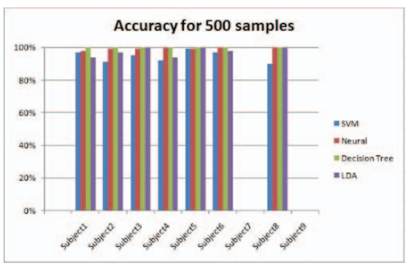

Fig. 6. Accuracy for 9 subjects with 500 Frames each.

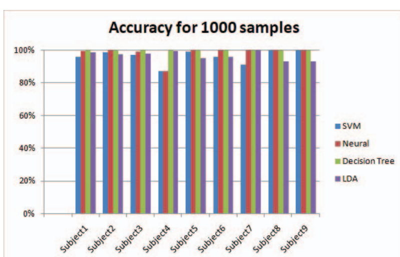

Fig. 7. Accuracy for 9 subjects with 1000 Frames each.

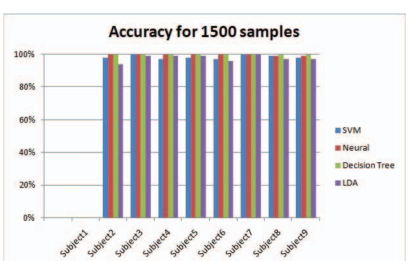

Fig. 8. Accuracy for 9 subjects with 1500 Frames each.

We can infer from the above graphs that every subject has an accuracy of more than 85 percent in all cases. Neural network and decision tree classifiers have shown better accuracy than other classifiers. In the case of 500 samples, subject 1 and subject 3 have shown better accuracy to decision tree classifier than the neural network classifier. In 1000 samples, subject 1 and subject 5 have given better accuracy to decision tree than any other classifier. In 1500 samples, subject 8 and subject 9 showed better accuracy to the decision tree classifier. Except in the above cases, neural network and decision tree classifiers have shown equal results. We can infer that the decision tree classifier is better than all other classifiers for this project but the neural network classifier has given almost equal results. So, We can use any of the two classifiers for classifying the subject to be drowsy or non-drowsy.

- Anomalies:
- 1. For subject 7 and subject 9, we can observe zero accuracies because the subjects did not have drowsy stated frames in the first 500 samples.
- 2. For subject 1, we have zero accuracies in 1500 samples because the subject had less than 1500 samples in total.

  We have also proposed an algorithm for the rate of drowsiness, which is as follows.

VII. Rate of Drowsiness

We took the ratio values of 10 consecutive images and found average of these values, multiplied this set of 10 frames to the next set of 10 frames. We have considered 5 sets of 10 consecutive images for rate calculation as the results were optimal. When the 5th set completes, we go in a Round Robin manner to fill in the next 10 consecutive frames.

$$RATE = P1 * P2 * P3 * P4 * P5 \tag{5}$$

- P1=Average of first ten consecutive images
- P2=Average of second ten consecutive images
- P3=Average of third ten consecutive images
- P4=Average of fourth ten consecutive images
- P5=Average of fifth ten consecutive images

We experimented this algorithm on 10 subjects and the results are follows:

TABLE I
Results for Rate of Drowsiness

Subjects	Non-drowsy (Rate)	Drowsy(Rate)
Subject 1	0.27-0.6	0.6-0.9
Subject 2	0.3 - 0.6	0.6-0.9
Subject 3*	1.0 - 1.28	1.1 - 1.3 , 0.3-0.8
Subject 4	0.3 -0.74	0.74-1.2
Subject 5	0.5-0.7	0.7-1
Subject 6*	0.13- 0.6	0.15- 1
Subject 7	0.23-0.5	0.5-0.7
Subject 8*	0.4 - 1.13	0.67-1.27
Subject 9	0.3-0.6 0.7	0.6-1.24
Subject 10	0.42-0.6	0.6-1.2

* = Subjects for which algorithm couldnt differentiate drowsy and non-drowsy states .

So, we could not find the rate of drowsiness for 3 subjects in 10, as there were many blurred frames in those subjects which disturbed the ratio. We took the average of the values and established the values for non-drowsy and drowsy rate detection. For non-drowsy it is 0.33-0.63 and for drowsy it is 0.63 -1.02.

We have scaled this into 1-10 according to the following scale.

TABLE II
RATE OF DROWSINESS SCALED FROM 1 TO 10

Rate before Scaling	Rate after Scaling	State
0.33-0.39	1	Extremely alert
0.39-0.45	2	Very alert
0.45-0.51	3	Alert
0.51-0.57	4	Rather alert
0.57-0.63	5	Neither alert nor sleepy
0.63-0.708	6	Some signs of sleepiness
0.708-0.786	7	Sleepy, no effort to stay awake
0.786-0.864	8	Sleepy, some effort to stay awake
0.864-0.942	9	Very sleepy
0.942-1.00	10	Excessively sleepy

VIII. CONCLUSION AND FUTURE WORK

This research project proposes an algorithm for driver drowsiness detection and drivers rate of drowsiness. For classifying the driver into drowsy and non-drowsy we conclude from the data of 9 subjects that decision tree and neural network classifiers have given better results than linear SVM and LDA. For the Rate of Drowsiness, we have described an algorithm as explained above. Prior techniques were able to make decisions based on attributes such as eye blinks, eye closure. We have considered both eyes and mouth as features and used modern classifiers for classifying the subject to drowsy or non-drowsy.

Although the proposed classifiers are good enough to give reasonable results, still there is a lot of latitude for improvement in their performance. A more robust drowsiness detection classifier can be applied by checking many other classifiers. For the rate of drowsiness detection, the algorithm can be still be improved by researching on other datasets.

REFERENCES

[1] http://ncrb.gov.in/StatPublications/ADSI/ADSI2015/table-1A.5.pdf

[2] United States Department of Transportation., Saving lives through advanced vehicle safety technology http://www.its.dot.gov/ivi/docs/AR2001.pdf.

[3] Takei, Y., Furukawa, Y. (2005, October). " Estimate of driver's fatigue through steering motion". In 2005 IEEE international conference on systems, man and cybernetics (Vol. 2, pp. 1765-1770). IEEE.

[4] International Federation of Societies for Electroencephalography, Clinical Neurophysiology. (1983). " Recommendations for the practice of clinical neurophysiology ". Elsevier Publishing Company.

[5] Singh, M., Kaur, G. (2012). " Drowsy detection on eye blink duration using algorithm ". International Journal of Emerging Technology and Advanced Engineering, 2(4), 363-365.

[6] Saito, H., Ishiwaka, T., Okabayashi, S. (1994, August)." Applications of driver's line of sight to automobiles-what can driver's eye tell". In Proceedings of VNIS'94-1994 Vehicle Navigation and Information Systems Conference (pp. 21-26). IEEE.

[7] Horng, W. B., Chen, C. Y., Chang, Y., Fan, C. H. (2004, March). " Driver fatigue detection based on eye tracking and dynamic template matching ". In IEEE International Conference on Networking, Sensing and Control, 2004 (Vol. 1, pp. 7-12). IEEE.

[8] Smith, P., Shah, M., da Vitoria Lobo, N. (2000, September). " Monitoring head/eye motion for driver alertness with one camera ". In Proceedings 15th International Conference on Pattern Recognition. ICPR-2000 (Vol. 4, pp. 636-642). IEEE.

[9] Dwivedi, K., Biswaranjan, K., Sethi, A. (2014, February)." Drowsy driver detection using representation learning ". In 2014 IEEE International Advance Computing Conference (IACC) (pp. 995-999). IEEE.

[10] Zhang, Z., Zhang, J. S. (2006, August). "Driver fatigue detection based intelligent vehicle control". In 18th International Conference on Pattern Recognition (ICPR'06) (Vol. 2, pp. 1262-1265). IEEE.

[11] Soukupova, T., Cech, J. (2016, May). "Eye blink detection using facial landmarks". In 21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia.

[12] Weng, C. H., Lai, Y. H., Lai, S. H. (2016, November). "Driver drowsiness detection via a hierarchical temporal deep belief network". In Asian Conference on Computer Vision (pp. 117-133). Springer, Cham.