Vision-Based Yawning Classification System for Real-Time Drowsiness Detection

Dr. Ana María Torres[1], Dr. Carlos Eduardo Rivas[2], Dr. Lucia Fernández[3]*

[1]Department of Environmental Sciences, Universidad de Buenos Aires, Buenos Aires, Argentina
[2]Department of Civil Engineering, Universidad Nacional de Colombia, Bogotá, Colombia
[3]Department of Earth Sciences, Universidad de Chile, Santiago, Chile

ABSTRACT

As the facts depict drowsiness or fatigue is one of the major causes of road accident. Drowsiness impairs driver's senses and the driver is not able to respond quickly resulting in higher risk of road accidents. Being a significant indicator of drowsiness, yawning detection has been explored in this paper. Many geometric and feature based methods have been previously proposed but they are sensitive to illumination, appearance variations such as skin color etc. Hence, effective feature extraction techniques are needed that extract good features in the presence of varying lightning conditions and further improve classification accuracy. Frequency domain feature extraction has not been used extensively in this area. In this paper, Local energy based shape histogram (LESH) has been implemented that extracts the features in the frequency domain, as a result it is insensitive to contrast and illumination. Support Vector machine (SVM) is a widely used machine learning tool because of its binary nature of classification and also efficiently classifying non-linear data. The proposed framework is evaluated on Yawdd database. A sensitivity of 100%, specificity of 94% and accuracy of 95% was achieved.

Keywords: *Drowsiness detection, yawning detection, feature extraction, Support vector machine.*

I. INTRODUCTION

Dangerous road accidents can occur if one drives while feeling sleepy, as there is much higher risk of meeting an accident if a person is not alert. If a risky situation ascends, drowsiness weakens our ability to respond quickly and safely and he less conscious of what is happening on the road. Facts show that, 51% of youngsters feel drowsy while they had driven and even 17% have really fallen asleep conferring to National Sleep Foundation (NSF) [1]. Driver drowsiness can cause numerous physical and economical fatalities. Hence, Driver assistive systems that detect a driver's drowsiness level and alert the fatigued, are required to avoid road accidents. One way to detecting driver's drowsiness is to observe the driver's driving behavior, if driver not paying attention while driving alarm sound alerts are generated.

Correlation between driver fatigue and a noteworthy rise in the likelihood of car accident arises the need to develop reliable detection systems. After detecting the fatigue and drowsiness levels, a variety of steps can be taken to warn the driver, such as alerting alarm sound, vibrating the steering wheel or driver's seat, displaying warning and advice messages, or providing more oxygen to the driver, for example, by paced breathing using a pulse sound synchronized with heartbeats. Driver behavior monitoring systems, which may include a driver fatigue detection, also depend on yawning detection as one the factors of determining the driving behavior.

There are three ways in which driver drowsiness can be detected: Vehicle based measures [2], [3], [4] Behavioral measures [5], [6], [7] Physiological measures. In vehicle based methods, a number of sensors are placed on several vehicle machineries like steering wheel and the acceleration pedal; the level of drowsiness is determined by the signals sent by the sensors. The drowsy driver behavior like head pose, frequent yawning, droopy eyes, eye closure and frequent blinking etc. are observed to detect the driver drowsiness using a camera and alert the driver in case any one of these indications is seen. Physiological based method is a direct measure of the brain activity and it can avoid many road accidents due to fatigue driver. Electro cardiogram (ECG), electro-myogram (EMG), electro encephalogram (EEG) and electro-occulogram (EOG), heart rate (HR) are possible measures for physiological signals.

Driver behaviour monitoring systems, count on yawning detection as one the major factors of determining the driving behaviour. Not only driver drowsiness systems, yawning detection can also be implemented in in-home health care system, Operator attentiveness that uses yawning detection as one of a few deciding factors in determining whether or not an operator of a critical system, such as heavy machines, nuclear reactor controls and air traffic controllers, is concentrating on the operation or not. Yawning detection can also be used in systems that determine the communication intents of a person having tongue disability, specifically to detect incorrect estimation. For all the above systems that require an efficient detection of yawning.

Literature survey is described in Section II. Section III provides method and material for yawning detection. Section IV provides results and discussions. Finally, conclusion is provided in section V.

II. LITERATURE SURVEY

Masrullizam Mat Ibrahim et al [8] proposed a method for yawning detection based on widest area of darkest region between the lips and occlusions due to mouth covered with hand during yawning. Then classification of local binary patterns (LBP) features extracted from the mouth when covered by a hand and evaluated using neural network classifier, support vector machine. It is concluded that SVM shows better performance. Zhuoni Jie et al, [9] proposed a method that analyzed yawning behavior in a dataset of spontaneous expressions of drowsy drivers, and proposed a new method to detect yawning that combines geometric and appearance features of the mouth and eye areas. A set of geometric and appearance features is extracted that can represent both hand-covered and uncovered spontaneous yawns, namely: mouth openness, Histograms of Oriented Gradients (HOGs) and Local Binary Patterns (LBP).Ashish Kumar et al [10] developed a system where a webcam records the video and driver's face is detected in each frame employing image processing techniques. Ensemble of regression trees is used to estimate the landmark positions on face from a sparse subset of pixel intensities. Using this method, the boundary points of eyes, mouth and the central line of the nose are marked. Facial landmarks on the detected face are pointed and subsequently the eye aspect ratio, mouth opening ratio and nose length ratio are computed and depending on their values, drowsiness is detected based on developed adaptive thresholding. Machine learning algorithms have been implemented as well in an offline manner. Manu B.N et al [11] describes an efficient method for drowsiness detection by three well defined phases. These three phases are facial features detection using Viola Jones, the eye tracking and yawning detection. Once the face is detected, the system is made illumination invariant by segmenting the skin part alone and considering only the chromatic components to reject most of the non-face image backgrounds based on skin color. The tracking of eyes and yawning detection are done by correlation coefficient template matching. The feature vectors from each of the above phases are concatenated and a binary linear support vector machine classifier is used to classify the consecutive frames into fatigue and non-fatigue states and sound an alarm for the former, if it is above the threshold time. Anitha C et al [12] proposed a yawning detection system which is based on a two-agent expert system. In the proposed system, as the first part of detection we use the face detection algorithm's skin detection part. The skin region is extracted. For all the skin region blocks detected, their boundaries are defined. The lower half of the face is considered for the mouth region extraction. The presence of yawning would be indicated by a black blob in the mouth region of the binary image. Mona Omidyeganeh et al [13]designed and implemented such automatic system, using computer vision, which runs on a computationally limited embedded smart camera platform to detect yawning. They used a significantly modified implementation of the Viola–Jones algorithm for face and mouth detections and, then, used a backprojection theory for measuring both the rate and the amount of the changes in the mouth, in order to detect yawning. Nawal Alioua et al [14] proposed an efficient and nonintrusive system for monitoring driver fatigue using yawning extraction. The proposed scheme uses face extraction based support vector machine (SVM) and a new approach for mouth detection, based on circular Hough transform (CHT), applied on mouth extracted regions. Narender Kumar et al [15] used Contour Activation algorithm to locate the contour of lips. If the mouth is open, the inside of the mouth is dark which is roughly segmented out. After segmentation contour of the mouth is obtained and height of the mouth is calculated. Height being greater than the threshold for several consecutive frames, means the person is yawning. Tayyaba Azim et al [16] The focus of the paper is on how to detect yawning which is an important cue for determining driver's fatigue. Initially, the face is located through Viola-Jones face detection method in a video frame. Then, a mouth window is extracted from the face region, in which lips are searched through spatial fuzzy c-means (s-FCM) clustering. The degree of mouth openness is extracted on the basis of mouth features, to determine driver's yawning state.In the proposed scheme by Saima Naz, Sheikh Ziauddin et al, they captured videos from a camera mounted inside the vehicle. From the captured video, we localize the eyes using Viola-Jones algorithm. Once the eyes have been localized, they are classified as open or closed using three different techniques namely mean intensity, SVM, and SIFT.

III. METHOD & MATERIAL

In this paper, a yawning detection framework has been proposed that extracts features using LESH feature extraction. The prepared feature vector is fed to the SVM to perform classification between yawning and non-yawning frames. The methodology is explained below.

A. Video Acquisition

Videos of drivers are captured using a webcam or a smartphone installed on the dashboard. The extracted video is then converted into a series of images resulting in an image array. Images are then read at regular gaps. Videos are obtained from yawdd database [17].

B. Face and mouth detection

To detect yawning it is essential that search space is reduced and mouth area is extracted from the images. Viola jones algorithm [18] was implemented for face detection and further for mouth detection.

C. LESH Feature extraction

Finding and extracting consistent and discriminative features and attributes is always a critical step to complete the task of computer vision, image recognition and classification. Feature extraction is the process of converting an input set of images into a set of features. These features are distinctive attributes of input patterns which help in distinguishing between the categories of input patterns. Further machine learning algorithms gain knowledge from these extracted features, a training model is build and a solution is provided i.e. classification of features. Histogram of oriented gradients and local binary pattern are two most widely used feature extraction techniques in appearance based yawning detection. In my work, Local Energy Based Shape Histogram extracts frequency domain based features from the set of input images.

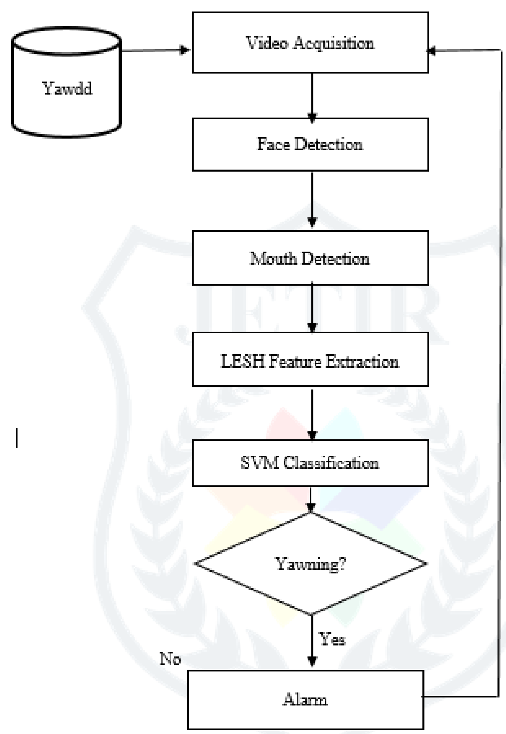

Figure 1: Yawning detection process

LESH [19] is a frequency domain based feature extraction technique that extracts the local energy information of the image. The energy of an image is always higher at edges and corners. This technique first obtains the Fourier transform of the image and Phase congruency measure is calculated. Phase is an underused local image attribute that can convey more information regarding the signal behavior in presence of edges and corners than magnitude does. It is invariant to illumination and contrast. The local energies are computed for each sub-region in different orientations and scales. An accumulated 512-dimensional LESH feature vector is obtained for each image.The results are show in section IV.

D. Support vector machine

A Support Vector Machine (SVM) is a machine learning classifier which, given labelled training data builds an optimal separating hyperplane which categorizes new data instances. A *good margin* or separating hyperplane is the one where this parting is larger for both the classes. A non-linear SVM with radial basis kernel function has been implemented in this paper.

IV. RESULT & DISCUSSION

A. Feature extraction results

A training set and test set consisting of yawning and non-yawning mouth states, was prepared. For each image, a LESH feature vector is obtained and in the end, it is concatenated to get an accumulated feature set consisting features of all images. The image of male and female mouth with their corresponding feature vector is shown in figure 2.

B. Classification results

A non-linear binary classifier support vector machine was trained to develop a training model and classification results were obtained using the accuracy, sensitivity and specificity parameter. Our training set includes a set of mouth images which are is yawning state and non-yawning statesuch as smiling, coughing, laughing. Support vector machine was efficiently able to classify the data with overall accuracy of 95%, sensitivity 100% and specificity 96%.

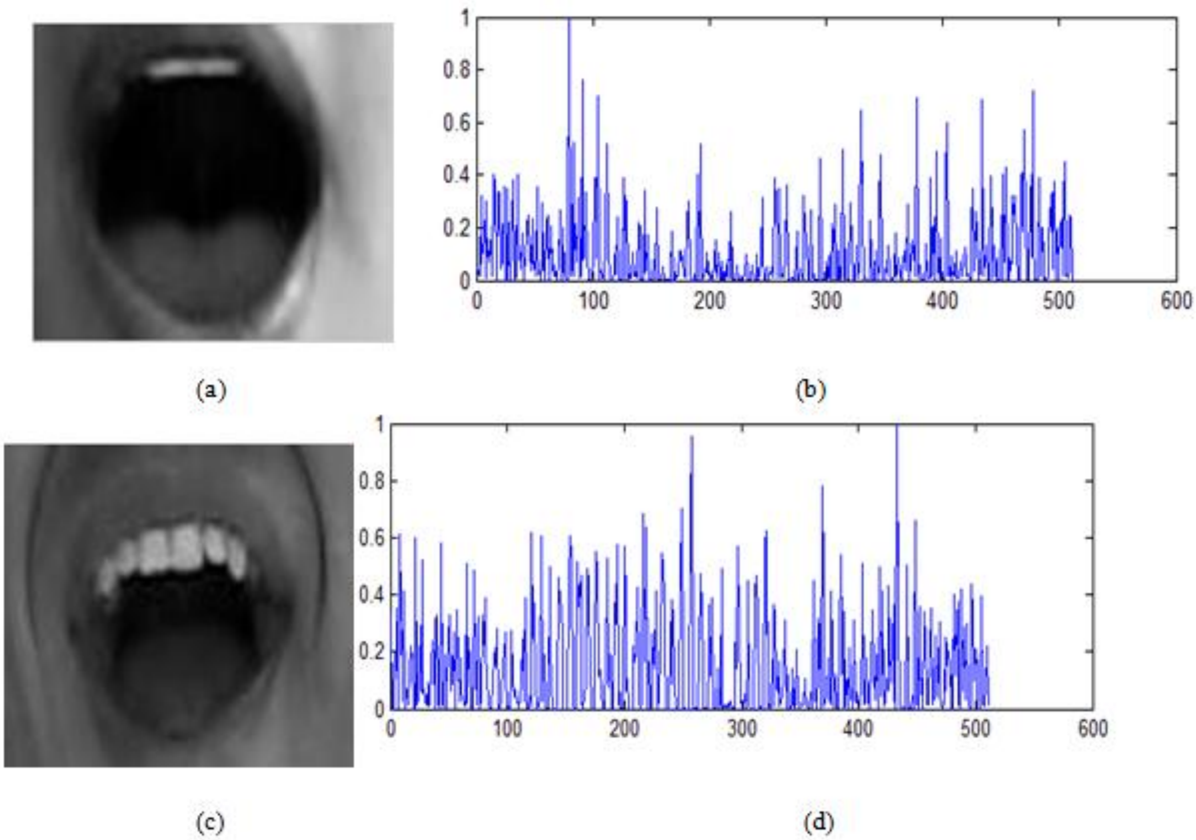

Figure 2: LESH feature extraction results (a) Female yawning mouth (b) LESH feature vector of female yawning mouth (c) Male Yawning mouth (d) LESH feature vector of male yawning mouth

Table 1. Classification results based on different size of feature sets

Features	50	100	All
Accuracy	95.28%	95.15%	95%
Sensitivity	100%	100%	100%
Specificity	96.17%	96.04%	96%

V. CONCLUSION

A frequency domain based feature extraction was proposed in this paper to extract the features of the yawning and non-yawning mouth states images. The performance was evaluated using most widely used data learning tool i.e. Support Vector Machine. This approach is invariant to illumination and facial variations such as skin color and satisfactory classification results were obtained. Unlike feature based methods, this approach is more likely to reduce false positives and negatives in case of laughing and coughing mouth states which are somewhat similar to yawning mouth state.

Further work can be done to improve the efficiency of the system by making it a multimodal approach i.e. addition of more indicating factors such as eyes and distortion on the forehead etc.

REFERENCE

1. *National Sleep Foundation's Sleep in America poll, "Global Status Report on Road Safety [Online] Available:https://www.sleepfoundation.org/sleep-topics/drowsy-driving*
2. *A. D. McDonald, C. Schwarz, J. D. Lee, and T. L. Brown, "Real-Time Detection of Drowsiness Related Lane Departures Using Steering Wheel Angle," Proc. Hum. Factors Ergon. Soc. Annu. Meet., vol. 56, no. 1, pp. 2201–2205, 2012.*
3. *P. Choudhary and G. Rishikesh, "A Survey Paper on Drowsiness Detection & Alarm System for drivers," International Research Journal of Engineering and Technology, vol 3, 2016.*
4. *J. H, Yang, Z.H. Mao, L. Tijerina et al. "Detection of Driver Fatigue Caused by Sleep Deprivation," IEEE transactions on system, VOL. 39, NO. 4, 2009.*
5. *M.H. Sigari, M. Fathy and M. Soryani "A Driver Face Monitoring System for Fatigue and Distraction Detection" International Journal of Vehicular Technology 2013, Article ID 263983, 11 pages.*
6. *B. Mandal, L. Li, G.Wang, and J.Lin, "Towards Detection of Bus Driver Fatigue Based on Robust Visual Analysis of Eye State," IEEE Trans. Intell. Transp. Syst., vol. 18,issue no. 3, pp. 545 - 557, Sep. 2016.*
7. *A. Dasgupta, A. George, S. Happy, and A. Routray, "A vision-based system for monitoring the loss of attention in automotive drivers," IEEE Trans. Intell. Transp. Syst., vol. 14, no. 4, pp. 1825–1838, Dec. 2013.*
8. *Ibrahim, M. Mat et al. "Yawn analysis with mouth occlusion detection", Biomedical Signal Processing and Control 18 (2015), pp.360-369.*
9. *Z. Jie, M. Mahmoudy et al. "Analysis of yawning behavior in spontaneous expressions of drowsy drivers", 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition ,2018*
10. *A. Kumar, R. Patra, " Driver Drowsiness Monitoring System using Visual Behavior and Machine Learning", 2018 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)*
11. *Manu B.N, "Facial Features Monitoring for Real Time Drowsiness Detection", 2016 12th International Conference on Innovations in Information Technology (IIT).*
12. *M.K. Venkateshab, B. Adiga, "A Two-Fold Expert System for Yawning Detection" 2nd International Conference on Intelligent Computing, 2016, pp. 63-71*
13. *M. Omidyeganeh, S. Abtahi, et al. "Yawning Detection using embedded smart cameras," in IEEE Transactions on instrumentation and measurement,2016.*
14. *N. Alioua, A. Amine, M. Rziza, " Driver's Fatigue Detection Based on Yawning Extraction" International Journal of Vehicular Technology, Volume 2014*
15. *N. Kumar, N.C. Barwar, " Detection of Eye Blinking and Yawning for Monitoring Driver's Drowsiness in Real Time", International Journal of Application or Innovation in Engineering & Management, November 2014*
16. *T. Azim, M. Jaffar et al. " Automatic Fatigue Detection of Drivers through Yawning Analysis" International Conference on Signal Processing, Image Processing and Pattern Recognition 2010, pp 125-132.*
17. *S. Abtahi, M. Omidyeganeh, S. Shirmohammadi, and B. Hariri, "YawDD: A yawning detection dataset," in Proc. 5th ACM Multimedia Syst. Conf., Singapore, 2014, pp. 24–28.*

18. *P. Viola and M. J. Jones, "Robust Real-Time Face Detection," Int. J. Computer Vision., vol. 57, pp. 137–154, 2004.*
19. *Sarfraz, S., Hellwich, O.:"Head Pose Estimation in Face Recognition across Pose Scenarios", Proceedings of VISAPP 2008, Int. conference on Computer Vision Theory and Applications, Madeira, Portugal, pp. 235-242, January 2008 (Best Student Paper Award)*