International Journal on Emerging Technologies **11**(3): 502-506(2020)

ISSN No. (Print): 0975-8364
ISSN No. (Online): 2249-3255

Driver Drowsiness Detection System using Convolutional Neural Network

Komal[1], P. Sharma[1], A. Lamba[2], B. Nagpal[2] and S. Chauhan[2]
[1]Assistant Professor, Department of Computer Science, Amity University Haryana, Gurgaon (Haryana), India.
[2]Student, Department of Computer Science, Amity University Haryana, Gurgaon (Haryana), India.

(Corresponding author: Komal)
(Received 04 March 2020, Revised 27 April 2020, Accepted 29 April 2020)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Rapid increase in the number of automobiles plying on the road has not only increased the traffic jams but the probability of having more road accidents. Sleep deprivation remains the primary cause regarding loss in concentration levels and has progressively increased in modern times. As per the statistics of NHTSA, Drowsiness during driving resulted in 44,000 crashes leaving nearly 800 dead. The methods proposed for detecting the drowsiness are very complex and time consuming due to usage of large mathematical computations. In this paper, an unprecedented Deep learning based approach towards real-time driver drowsiness and distraction detection using object recognition is proposed. The work is done making use of a vision-based cost-effective system. It uses 68 point facial Landmarks to find aspect ratios for eyes and mouth- EAR and MAR as these remain the most prominent features that get affected by drowsiness, to identify the probability of driver to fall asleep. It implements object recognition built using various deep learning algorithms simultaneously to get the most promising results to recognize objects that leads to distraction. On identifying drowsiness or distraction, the system will sound a caution alarm along with a message to notify the driver. The proposed model is capable of achieving accuracy of 90%.

Keywords: Convolutional neural network, deep learning, drowsiness detection, object detection, tensor flow.

I. INTRODUCTION

Deprived Concentration levels while driving are becoming more and more common nowadays due to the hectic work schedules, lack of sleep and getting engaged into multiple things while driving. Such situations increase the probability of road accidents that can be fatal for the driver. Based on the statistics gathered from National Sleep Foundation (NSF) and National Highway Traffic Safety Administration (NHTSA), it can be clearly stated that feeling drowsy and getting distracted because of external objects (food, phone call) while driving – are the primary causes of road accidents [3]. The drowsiness increases the reaction time of the driver and reduces the control of driver on the vehicle. On the other side, the object distraction leads to lack of attention on road and can be as dangerous as drowsy driving.

Another shocking statistics from AAA Foundation for Traffic Safety predict that the annual count of road accidents due to drowsiness driving is around 328,000. Which is much higher than the actual reported number by the authority. It also revealed that one-third of these cases had severe injuries and 6% people succumbed to their injuries. In recent years, lot of research has been taking place to develop early warning systems to combat the issues related to driver drowsiness and vehicle safety features to reduce the number as well as severity of accidents happening on road.

Other than the researcher community at large, the R & D divisions of top-notch car manufacturing companies are also working towards the development of such systems. One such development is the Ford's driver assistant system. It analyses rapid steering movements, crossing lanes separating roads, abnormal acceleration and speedy braking. Based on the processing of previously collected data, the driver is graded from 5 (highest concentration level) to 1 (least concentration level). The driver is warned through beeps and warning signs on the panel located inside the vehicle. Another such system has been developed by Skoda which keeps a record of steering movements in normal condition (starting minutes after car starts) and analyses the movements of steering at any later stage and abnormality in the behavior of steering movements is reported to driver with the help of warning system [10].

Bosch has developed a system known as the Driver Drowsiness Detection system [2]. The system uses information gathered from the sensor (fitted into the steering wheel) about the sudden movements of the steering vehicle, analyzes the abnormal trajectory of the vehicle and generates the warning signals for the driver. Other driver tiredness detection techniques use brain wave analysis and heart rate analysis to detect signs of tiredness or fatigue [6]. Such systems are highly accurate but impractical to use and these features are only available in luxury vehicles.

Also the systems which are designed using normal features like movements of facial components, suffers from the problem of mathematical computations and hence are very slow. However, deep learning is making groundbreaking advances in these fields. The faster graphical processing units in the deep learning networks and their high learning capabilities make them the best choice for designing such systems.

In this paper we have implemented a technique using convolution neural network to identify the signs of drowsiness in the driver. The proposed method has shown very good results as compared to other state of the art methods.

II. LITERATURE REVIEW

Drowsiness detection has been a popular research topic

among the researchers not for a very long time. Earlier the researchers were using simple SVM based classifications for doing the same. But these algorithms were very complex in mathematical computation and also very slow as manual feature extraction was required in these methods for further classification.
With the advent of deep learning there has been a huge shift in terms of object detection [12] and classification problems. In drowsiness detection problem also these techniques have been used recently and have shown very good results. Some of the research done in this area with their advantages and disadvantages is shown in Table 1.
The main challenge to create a robust system to detect drowsiness seems unaccomplished till now. Some technologies used by the above authors seem more accurate but have less practicality, some are slow some are fast but not accurate. Others are not cost-effective again reducing their practicality in various conditions. So to handle all these downfalls proposed system uses pretrained dlib shape_predictor_68_face_landmarks trained on IBUG300w Dataset to detect facial landmarks and is quite fast with good accuracy after getting the landmarks EAR and MAR are calculated to calculate the drowsiness.

The proposed system has a perfect balance between accuracy and speed and is smaller in size, needs less computational ability again increasing its practicality on embedded systems. The proposed system uses Convolution neural networks which work very fast for object detections. Some of the object detection problems in literature are summarized in Table 2.
Various Object detection techniques are represented in graph below (Fig. 1) using their accuracy percentage and speed.

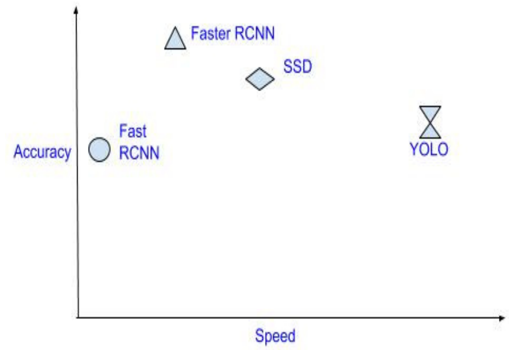

Fig. 1. Accuracy of object detection techniques.

Table 1: Drowsiness Detection techniques in literature.

Author	Methodology	Method cons	Data Set	System Weakness	Accuracy
Xie *et al.*, (2018)	CNN(Inception V3), LSTM(Recurrent Neural Networks) [11]	CNN requires more computational ability and is slow and difficult to apply in embedded systems.	YawDD, NTHU DDD dataset	Only one Drowsiness Trait, Accuracy decreases under weak Illumination.	97.36%
(PRASA-RobMech)	HOG SVM, HMM CNN [7]	Review of existing methods	YawDD, NTHU DDD dataset	Accuracy traits were limited	
Choi *et al.*, (2018)	Viola and Jones facial landmark algorithm with physiological traits(BPM) [1]	Physiological sensors are expensive and have low practicality.	Dataset of 5 volunteers on Virtual Driving Environment.	Dataset is small, wearable's are necessarily decreasing the practicality.	90%
Jabbara *et al.*, (2018)	Facial Landmark, neural networks [5]		NTHU(Drowsy Driver Database)	Working on only a single drowsiness trait, Accuracy calculated on high-performance GPU	80%
Nikolskaia *et al.*, (2019)	CNN, Trained neural network(shape predictor face landmark) [8]	CNN requires more computational ability.	CEW dataset containing 2500 Photographs	Memory performance should be high, Training time is more, costly, Message not enough to alert	89%

Table 2: Object Detection Using CNN literature.

Author	Methodology	Dataset	Weakness	Accuracy
Hayat *et al.,* (2018)	CNN On top of Tensorflow framework [4]	Caltech-101 9,144 images	Fixed no. of objects	90.12%
Sudharshan and Raj (2018)	CNN on top of TensorFlow [9]	CIFAR-10 60,000 images	Done on Cpu version of TensorFlow, more processing time, performed on only 1 category of object	96%

Tensorflow deep learning framework has been used by earlier researchers for object detection and same is used to implement this model. SSD (single shot detector) provides the right balance between speed and accuracy that's is the reason why the proposed model uses SSD. SSD seems to be a good choice as we are able to run it on a video and the accuracy trade-off is very limited.

III. PROPOSED METHODOLOGY FOR CNN BASED DROWSINESS DETECTION

The proposed system for drowsiness detection consists of various steps using deep learning libraries and use of convolution networks. The flowchart of the methodology is presented in Fig. 2. The details of each step are provided below.

STEP 1: In the first step, the video is extracted through webcam using openCV and then the video is divided into frames.

STEP 2: Process of conversion of image to greyscale has been held for faster processing using open CV. After converting into greyscale, region of interest is calculated i.e., Face.

STEP 3: For facial landmarks, the pre-trained Dlib shape predictor is used to extract landmarks on the face. After detecting the facial landmarks aspect ratios of eye and mouth are calculated using Eqns. 1 and 2.

The state of eye (open/close) is determined with the calculations using Eqn. 1. It represents a constant value when opening state of eye and it transitions quickly to 0 in case eye gets closed. Similarly to determine the yawning parameter the aspect ratio of the mouth (MAR) is calculated. When closed state of mouth represents zero & open state of mouth represents value greater than zero and a very high value suggests that mouth is wide opened then it will be considered as yawning.

STEP 4: For object detection, object detection technique SSD (single shot detector) for detecting unwanted objects the driver uses while driving is used. Tensor flow deep learning framework is used to implement this model.

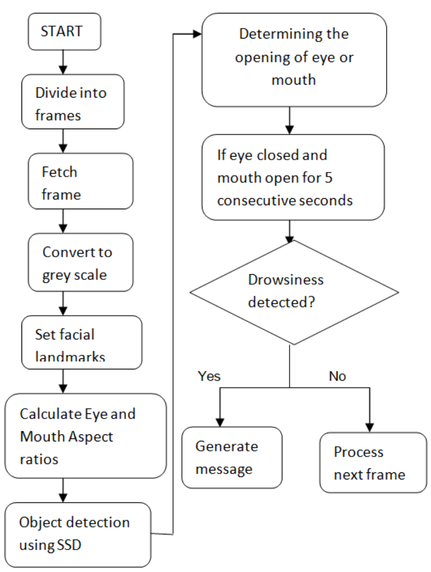

Fig. 2. Flowchart of the proposed system.

SSD uses single shot to detect multiple objects. The algorithm provides a better mix of accuracy and speed having better productivity. It eliminates certain steps used by many similar algorithms making it faster than them. It comprises of a gradually decreasing convolution filter for calculating object classes and offsets in bounding box locations.

STEP 5: After calculating EAR and MAR and after detecting the unwanted objects, the system will determine whether there are any signs of drowsiness- closed state of eyes for continuous frames and/or yawning state of mouth.

STEP 6: If eyes closed and mouth open for 5 consecutive frames and if EAR and MAR are beyond set threshold and any unwanted object detected then system will sound an alarm using the play sound library to alert the driver, else more frames to process.

IV. COMPUTATION RESULTS

Here we used the pre-trained Dlibshape_predictor to add landmarks on the face, it is trained on the IBUG-300W dataset and has approx.. 100% accuracy in applying facial landmarks under well-illuminated conditions considering the fact that the face is in the correct position.

After detecting the facial landmarks, aspect ratios of eye (EAR) and mouth (MAR) are calculated by using NumPy (Difference between the vertical and horizontal coordinates) using Eqns. 1 and 2 respectively (the mechanism is shown in Figs. 3 and 4). Eyes and Mouth are the most critical in drowsiness classification in any circumstances. If the EAR and MAR are beyond the set threshold system will sound an alarm using the play sound library to notify the driver.

EAR=|p2-p6|+|p3-p5|/2||p1-p4|| (1)

MAR=||p2-p8||+||p3-p7||+||p4-p6||/2||p1-p5|| (2)

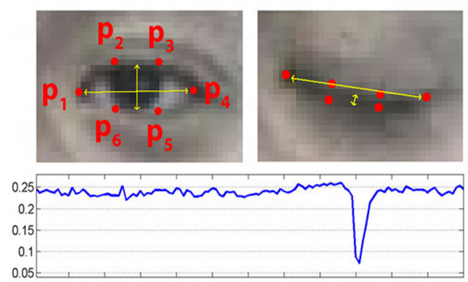

Fig. 3. Landmark detection for eyes.

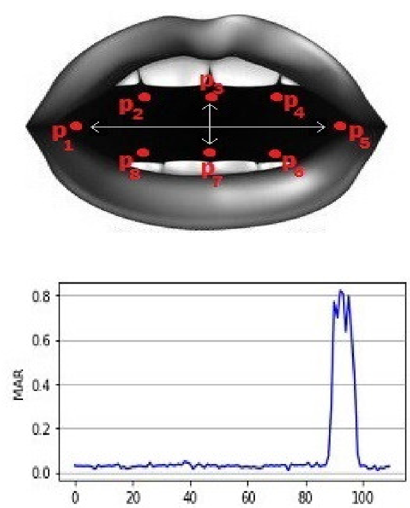

Fig. 4. Landmark detection for mouth.

To register the blinking state of eye, while a real-time video is being captured, the calculated EAR value

(using Eqn. 1) is compared with a pre-set threshold value (i.e. 0.3). A sudden change of EAR to a higher value than the threshold value is accounted for blinking. If this observation repeats for 5 consecutive frames, it is considered a sign of drowsiness and hence, the system will sound an alarm using predefined play sound library.
For distraction detection deep learning technique SSD(single shot detector) is used for detecting unwanted objects in real-time that the driver uses while driving. For this pre-trained model, ssd_mobilenet_v1_coco is used that is trained on coco dataset that consists of 330k images of 80 object categories. Proposed model can detect up to 10 objects in a single image. Tensorflow deep learning framework is used to implement this model.
The model used for object detection has been trained to detect any external object presence using following descriptors- class of object (Mobile, pen, keys etc.), location of object in the image (dimensions of bounding box) and value of confidence score (depicting the accuracy of detection).
Output Model for Object Distraction Detection: The output of the model for various classes is shown in Table 3.

Table 3: Model output for different classes.

Class of Object	Confidence Score Values	Location of Identified Object
Mobile	0.91	[18, 23, 58, 64]
Laptop	0.75	[21,45,69,84]
Food	0.85	[44,22,15,65]
Pen	0.21	[41, 21, 36, 85]

Confidence Score: To interpret the results of the model, we take into account the score as well as the location of the identified objects. The values of Score range from 0 to 1, the model is considered to be more confident if the values are close to 1, in determining the class of the detected object. So here Pen will be discarded due to low confidence score.
Input: The input is provided in the form of an Image with dimensions of 300x300 pixels using RGB channel representation for each pixel. It is provided in the model with a buffer value of 270,000 bytes (300 × 300 × 3). As the model undergoes expansion, each value represents one byte (values ranging from 0 to 255).
Output: The output of the model consists of four arrays, drawn in to indices 0-3 described below in Table 4, inserting one item each in corresponding list. Maximum of 10 objects will be detected in a single image.

Table 4: Description of Array index locations.

Index	Name	Description
0	Location	Multidimensional array of values ranging from 0 to 1, representing region of bounding box.
1	Class	Array of size 10 with integer values depicting the key of a class label.
2	Score	Array of size 10 with floating point values depicting probability of a detected class.
3	Numbers & Detections	Array of size depicting count of detected objects.

The Accuracy of the proposed method was calculated using True Detection Rate (TDR). True detection rate is calculated using ratio of correctly identified samples and total number of samples.

$$TDR = \frac{\text{Correctlyidentifiedsamples}}{\text{Totalnumberofsamples}} \quad (3)$$

It was observed that the accuracy of the system was about 90 percent considering all the factors and illumination conditions.

V. DISCUSSION

The proposed method was compared with many state of the art methods. In some of the methods the accuracy results were very good [3, 4]. But there were having problems related to cost and special infrastructure within the vehicles. Also these methods were developed using very small datasets and were tested on the same. High illumination conditions were also a challenge for these methods. On the other hand, the proposed method does not require any special monitoring device in the vehicle and as a result can be used with every vehicle. Also, separate data set (100 real-time videos including all illumination conditions)was used for testing the system.
Some of the other methods [5, 6] have taken landmarks into consideration for drowsiness detection but these methods were also suffering from the problem of high memory requirement and costly message transmission.
The pre-trained models used in the method are highly specific for object detection and hence give good results at a very fast rate using limited memory capabilities. As a result though the method is not giving very high accuracy results but has taken all factors into considerations to design a system which can be used with all the vehicles in all the conditions.

VI. CONCLUSION

The paper presents a driver drowsiness and object distraction detection system using facial landmarks by calculating the aspect ratio of eyes and mouth. The calculated aspect ratios i.e. EAR and MAR are compared with the threshold values to identify drowsiness. To identify object distraction, presence of multiple objects in video frame is detected using object detection technique SSD (single shot multi-box detector) based on convolution neural networks. The purpose of using these methods is to reduce the complexity of the system therefore leading to increase in practicality. The proposed method has provided good balance between speed and accuracy and is also capable of being used on embedded systems due to less complexity. The system is capable of achieving an accuracy of 90 percent under different conditions. However, there is lot of scope for further increasing the accuracy and speed of the system. Rather than only using the facial landmarks, head movements and other body traits may also be taken into consideration for further improvement in the proposed method.

Conflict of Interest. The authors declare no conflict of interest.

REFERENCES

[1]. Choi, H., Back, M., & Lee, K. (2018). Driver Drowsiness Detection based on Multimodal using Fusion of Visual-feature and Bio-signal. *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, 1249-1251

[2]. *Driver Drowsiness Detection*. (2018). Retrieved from BOSCH: https://www.bosch-mobility-solutions.com/en/products-and-services/passenger-cars-and-light-commercial-vehicles/driver-assistance-systems/driver-drowsiness-detection/

[3]. *Drowsy Driving.* (2018). Retrieved from NHTSA, United States Department of Transportation: https://www.nhtsa.gov/risky-driving/drowsy-driving#2206.

[4]. Hayat, S., Kun, S., Tengtao, Z., Yu, Y., Tu, T., & Du, Y. (2018). A Deep Learning Framework Using Convolutional Neural Network for Multi-Class Object Recognition. *2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC)*, 194-198.

[5]. Jabbara, R., Al-Khalifaa, K., Kharbechea, M., Alhajyaseen, W., Jafari, M., & Jiang, S. (2018). Real-time Driver Drowsiness Detection for Android Application Using Deep Neural Networks Techniques. *The 9th International Conference on Ambient Systems, Networks, and Technologies (ANT 2018)*, 1-8.

[6]. Khan, M. Q., & Lee, S. (2019). A Comprehensive Survey of Driving Monitoring and Assistance Systems. *Sensors*, 1-32.

[7]. Ngxande, M., Tapamo, T., & Burke, M. (2017). Driver drowsiness detection using behavioral measures and machine learning techniques: A review of state-of-art techniques. *2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech)* , 156-161.

[8]. Nikolskaia, K., Bessonov, V., Starkov, A., & Minbaleev, A. (2019). Prototype of Driver Fatigue Detection System Using Convolutional Neural Network. *2019 International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS)*, 82-86.

[9]. Sudharshan, D. P., & Raj, S. (2018). Object recognition in images using convolutional neural network. *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, 718-821.

[10]. *Superb Safety Assists*. (2018). Retrieved from Skoda: https://www.skoda-auto.co.in/models/superb/superb/superb-safety-assists

[11]. Xie, Y., Chen, K., & Murphe, Y. L. (2018). Real-time and Robust Driver Yawning Detection with Deep Neural Networks. *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, 532-538.

[12]. Zhiqiang, W., & Jun, L. (2017). A review of object detection based on convolutional neural network. *2017 36th Chinese Control Conference (CCC)*, 11104-11109.

How to cite this article: Komal, Sharma, P., Lamba, A. Nagpal, B. and Chauhan, S. (2020). Driver Drowsiness Detection System using Convolutional Neural Network. *International Journal on Emerging Technologies*, *11*(3): 502–506.