International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Driver's Drowsiness Detection Based on Behavioural Changes using ResNet

A. Jeyasekar, Vivek Ravi Iyengar

Abstract—Recently there has been growing interest in intelligent transportation system because the road accidents become biggest problems of mankind and the casualties of accident also increases rapidly every year. The casualties are very often witnessed in heavy and light motor vehicles. Moreover, the accidents occur mainly due to carelessness and drowsy feeling of the driver. Intelligent transportation systems use deep learning mechanism to detect drowsiness of the driver and alert the same to driver. It results in reduction of accidents. The driver's behaviour during drowsiness is detected by three types of approaches. One approach deploys the sensors in steering wheel and accelerator of the vehicle and analyzes the signal sent by the sensors to detect the drowsiness. Second approach focuses on measuring the heart rate, pulse rate and brain signals etc to predict the drowsiness. Third approach uses the facial expression of the driver such as blinking rate of eye, eye closure and yawning etc. The cause for most of the road accidents is driver's drowsiness. Therefore, in this paper, the behavioural changes of driver is accounted to detect the drowsiness of the driver. Eye movement and yawning are two behavioural changes of driver is considered in this paper. There are many CNN based deep learning architectures such AlexNet, VGGNet, ResNet. In this paper, we propose the drowsiness detection using ResNet because this method works on the principle of passing the output to the next la. The performance of proposed mechanism detects the drowsiness of the driver better than AlexNet and VGGNet.

Keywords— Drowsiness Detection, Convolutional Neural Networks, Activation Functions, Microsoft Resnet

I. INTRODUCTION

Accident is an incident that unexpectedly results in a damage or injury. Accidents cause catastrophic loss to mankind. Many models are proposed, and laws are imposed to curb accidents but still accidents are increasing many folds every year. Accidents are mostly unavoidable situations but by using a detection model, it can be reduced. Driving for lengthy periods of time can lead to accidents if proper rest is not taken. Rash driving, drink and drive are important reasons amongst youngsters which lead to death. With the advent of modernization and modern road, drivers tend to ride fast and cause loss to mankind and money. The World Health Organization (WHO) reported that India and China are prone to more road accidents. The number of world's road traffic deaths remains unacceptably high that is 1.35 million deaths in each year [2]. WHO's report in [2] has 2016 data which was republished again in 2018 wherein research was conducted research on various levels of road safety. There are 17 deaths every hour or 413 deaths every day caused due to accidents in India [5]. The main reasons for road accidents are due to over speeding, drunken driving, distraction to driver, red light jumping, avoiding safety rules [3][4][5]. Table 1 shows the number of road accidents, number of deaths in India and type of vehicle causes the accidents. It is evident that wo wheelers and cars account for 50% of road accidents in India [2].

	Number of Road Accidents		Number of Persons	
	Fatal	Total	Killed	Injured
Two-Wheelers	41608 (30.6)	162280 (33.8)	44366 (29.4)	153060 (30.9)
Auto-Rickshaws	6,095 (4.5)	31440 (6.5)	6767 (4.5)	39680 (8.0)
Cars, Jeeps, Taxis	28746 (21.1)	113267 (23.6)	32599 (21.6)	1257723 (25.4)
Buses	10394 (7.6)	237487 (7.8)	12088 (8.0)	50686 (10.3)
Trucks, Tempos, Tractors	36147 (26.6)	101085 (21.0)	39504 (26.2)	91784 (18.6)

Table 1: Road Accidents in India-2016[2]

The causes for the accidents are broadly classified into:

- Driver Perspective
- Vehicle Perspective
- Environment Perspective [6]

Reasons caused by driver behaviour are the most responsible for accidents compared to vehicle and other environmental conditions.

Manuscript published on 30 September 2019
* Correspondence Author
Dr. A. Jeyasekar*, Associate Professor, Department of CSE, SRMIST, Chennai
Vivek Ravi Iyengar, P.G Student, Internet of Things(IoT), Department of CSE, SRMIST, Chennai.

Drowsy driving occurs when someone is too tired to operate a motor vehicle which in turn puts themselves and their co-passengers in danger. There are many Intelligent transportation systems (ITS) proposed to detect the drowsiness of a driver using deep learning mechanism. Deep Learning is a form of artificial intelligence that uses Artificial Neural Networks (ANN) with multiple hidden layers that learns hierarchical representation of the underlying data in order to make prediction if new data are given. Another type of deep learning neural network is called Convolution Neural Networks (CNN) gained 80% accuracy in prediction [1]. AlexNet, VGGNet and ResNet are popular CNN based deep learning architectures in the literature. Since the ResNet provides better prediction than other architectures, in this paper, we propose a drowsiness detection mechanism using ResNet to prevent or reduce road accidents.

II. RELATED WORKS

MultiLayer Perceptron (MLP) Classifier using artificial neural networks is used to predict the drowsy behaviour of the driver [6]. The dataset contains a video frame which covers driving modes, yawning, blink rate and dizzy dozing. The implementation of the paper includes a Dlib library which extracts landmark coordination from images. Each perceptron is considered as one neuron carrying collection of an image. The activation functions used are tangential and logistic which ranges from -1 to 1 and 0 to 1 respectively. Model is prepared by reducing the error rate of the weighted connections of the perceptron after the data is being processed. It extensively uses Back Propagation. The performance of MLP achieves highest accuracy only if less number of images are given for training.

AlexNet [2] algorithm was founded in 2012 by Alex Krizhevesky and it won ImageNet challenge named as ImageNet Large Scale Visual Recognition Challenge (ILSVRC)-2012. AlexNet contains five layers out of which three are convolution layers and two are fully connected layers. Alexnet is efficient in training million images. The filters are applied on the first three convolution layers of stride 4. The activation function used in Rectified Linear Unit (ReLU) which is a modified linear function having positive advancement. The dataset used to evaluate the efficiency of AlexNet was NTH drowser driver detection video dataset which contains 360 video clips on behaviours of drowsy state like nodding, yawning and blinking. Under different subjects, the accuracy of the dataset on an average is around 65%.

VGGNet [15] was introduced in 2014 and developed by Simonyan and Zisserman. It consists of 19 layers with 2x2 pooling layers and 3x3 convolution layers. It can train 2.6 million images therefore it takes a lot of memory [10]. It contains 13 convolution layers and 3 fully-connected layers. The accuracy of prediction is better, but the memory utilization is high.

Residual Network or Resnet is a ImageNet Large Scale Visual Recognition Challenge(ILSVRC)-2015 winner. Activation function used is Rectified-Linear Unit(ReLU). The layers are stacked and are trained to the task at hand. ResNet-50 is a deeper network which consist of 5x5 pooling layers and 3x3 convolution layers comprising of 50 weight layers. The algorithm takes the image size and the adds the image parameters to it. This is known as Residual Learning. There are many variants of ResNet based on the number of layers. Moving towards deeper networks may cause training degradation and in turn it affects the accuracy of the model. Convolution layer contains hidden layers of *convolutions* which is the main building blocks of CNN. Convolution refers to mathematical combination of two functions. The two functions are input image and weights. The next building block of CNN is *pooling*. Pooling layer reduces the number of computation i.e. reduces the number of parameters to train. *Max-pooling* is one of the types of pooling. Max-pooling takes the maximum values from each window.

Mechanism	**Layer**	**Conv /Pool.**	**Act. Fn.**	**Acc.**
MLP [6]	2	1/1	Sigmoid, Tangent	~80%
AlexNet [2][13][14]	5	3/2	ReLU	~65.93%
VGG Net [15]	16	13/3	ReLU	~78%

Table 2: Comparison of CNN Methods

III.DROWSINESS DETECTION USING RESNET

In this paper, we use two set of dataset: Eye dataset and Yawning dataset. There are 6000 images of size 29x29 pixel with four different parameters in eye dataset: (i) Open Eye (ii) Closed Eye (iii) Left Eye and (iv) Right Eye. The Yawning Dataset contains video frames of drivers with different parameters based on lighting, facial expressions, yawning etc. The dataset is published on University of Ottawa and Nanjing University. The sample data in the form of table is represented in Figure 1.

Male Participant Number	Action	Facial Hair	Background Movement	Type of Glasses	Lighting
1	Talking	No	No	No Glasses	Sunny
1	Yawning	No	No	No Glasses	Sunny
1	Normal	No	No	Sun glasses	Sunny
1	Yawning	No	No	Sun glasses	Sunny
2	Normal	No	Yes	Prescription	Sunny
2	Talking	No	Yes	Prescription	Sunny
2	Yawning	No	Yes	Prescription	Sunny
3	Normal	No	no	Prescription	Sunny
3	Talking	No	No	Prescription	Sunny
3	Yawning	No	No	Prescription	Sunny
3	Normal	No	No	No Glasses	Sunny
3	Yawning	No	Yes	No Glasses	Sunny
4	Normal	No	No	No Glasses	Sunny
4	Talking	No	No	No Glasses	Sunny
4	Yawning	No	No	No Glasses	Sunny
5	Normal	No	No	Sun glasses	Sunny
5	Talking	No	Yes	Sun glasses	Sunny
5	Yawning	No	No	Sun glasses	Sunny

Fig. 1: Yawning Dataset Example

These two datasets are used to train and test the network post model creation. Figure 1 shows different actions a driver can perform e.g.yawning, normal, talking.The architecture shown in Figure 1 is used in this paper to detect the drowsiness of driver using ResNet. *Preprocessing* includes resizing of image to certain number of pixels.

This preprocessing step is done in appropriate dataset (e.g.yawning) where there is need to extract images from videos frames. The images extracting from the video frames can be done in intervals so that at dicrete different facial expressions can be captured.

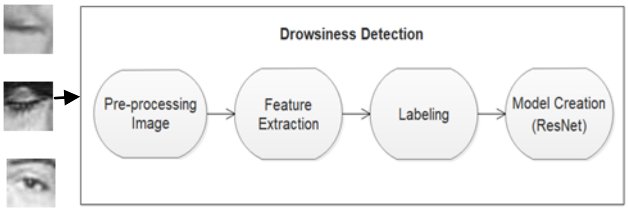

Figure 2: DD High-Level Architecture

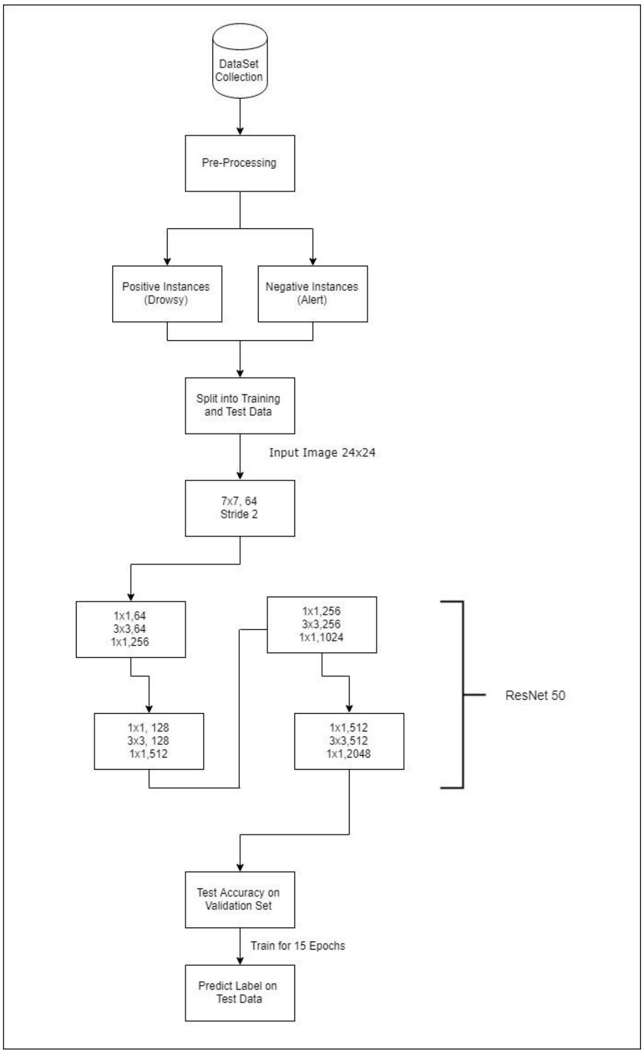

Figure 3: Drowsiness Detection Detailed Architecture

After extraction of images, the images are resized to 48x48 pixels in case of AlexNet and 24x24 pixels in case of VGGNet. But ResNet accepts same input size as that of the original image in the dataset. The feature extraction is performed by applying Dlib library which extracts facial features of the images. *Labeling* module labels all the image features so that while modelling, it is easy for the network to identify the images when passed the test data. *ResNet Model creation* follows the Resnet Architecture. Detailed architecture is explained below as a flowchart:

IV. PERFORMANCE ANALYSIS

This section deals with performance analysis of Resnet model represented in a graph andout test results done and presented in 3 epochs with an interval of 5 each. An epoch is one pass when the entire training dataset is passed to network. As it is difficult to pass the entire dataset in one epoch so images are passed in 5 epochs. Batch size refers to dividing the entire dataset in batches. To calculate performance, ResNet model is passed into 15 epochs with a batch size of 64. We evaluate the proposed drowsiness detection method using four metrics. They are precision, recall, F1 score and support. Precision is the ratio of true positive observations (correct positive value) to the total no. of observations. Recall is the ratio of correctly predicted observations to all observations in actual class (eye + yawning). F1 Score is a harmonic mean of precision and recall. The table shows the comparative analysis done on ResNet, AlexNet and VGGNet considering a batch of 64 images for training.

Model	Top-1 Error (EyesOpen+Closed model)	Top-1 Error (Alert+Yawning model)
VGGnet		
Epoch-5	29.63	11.42
Epoch-10	20.85	10.68
Epoch-15	22.67	8.6
Resnet50		
Epoch-5	26.76	11.25
Epoch-10	20.42	8.57
Epoch-15	16.52	7.75
Alexnet		
Epoch-5	21.95	9.30
Epoch-10	23.17	8.61
Epoch-15	22.21	10.38

Table 3: Top-1% & Top-5% Error of ResNet, AlexNet,VGGNet

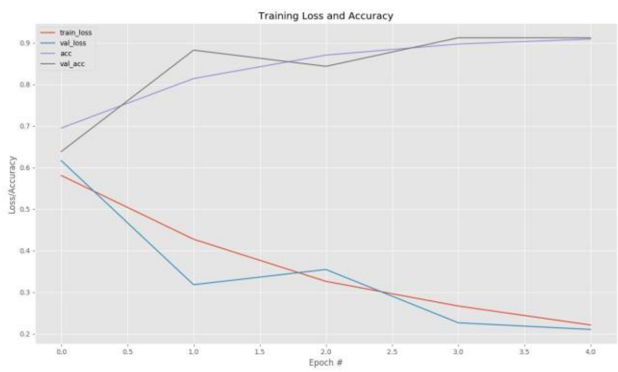

Figure 4: Training Loss/Accuracy vs Epoch-5

The forthcoming graphs and tables are generated after the ResNet model is tested on the given dataset.
The figures represent training loss, validation loss, accuracy vs epochs. The
Training Loss Graph (Batch Size 64Epoch=5)

	Precision	Recall	F1 Score
Awake	0.91	0.88	0.90
Drowsy	0.92	0.92	0.90
Avg/Total	0.91	0.91	0.91

Table 4: Accuracy ResNet Epoch-5

Epoch=10, Batch Size=64

Figure 5: Training Loss/Accuracy vs Epoch-10

Figure 5 represents loss lines(red/blue) decreasing with each epoch.

	Precision	Recall	F1 Score
Awake	0.96	0.80	0.87
Drowsy	0.87	0.97	0.87
Avg/Total	0.91	0.90	0.90

Table 5: Accuracy Resnet Epoch-10

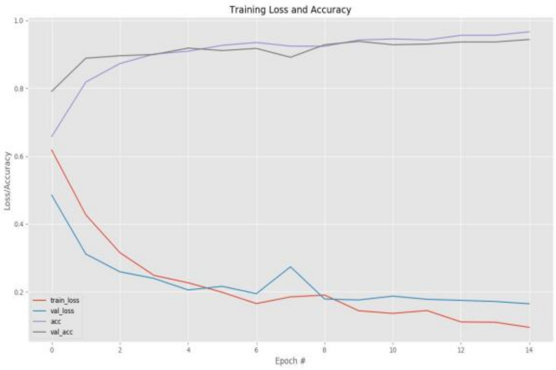

Figure 6 : Training Loss Accuracy vs Epoch-15

	Precision	Recall	F1 Score
Awake	0.94	0.93	0.93
Drowsy	0.95	0.95	0.95
Avg/Total	0.94	0.94	0.94

Table 6: Accuracy Resnet Epoch-15

The graphs and performance analysis of VGGNet and AlexNet is represented below:

VGGNet Epoch 5

	Precision	Recall	F1-Score
Awake	0.82	0.91	0.86
Dwosy	0.93	0.86	0.89
Avg/Total	0.88	0.88	0.88

Table 7: Accuracy VGGNet Epoch-5

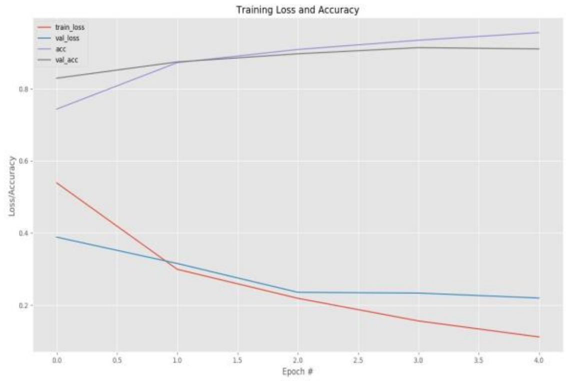

Figure: 7 Accuracy VGGNet Epoch-5

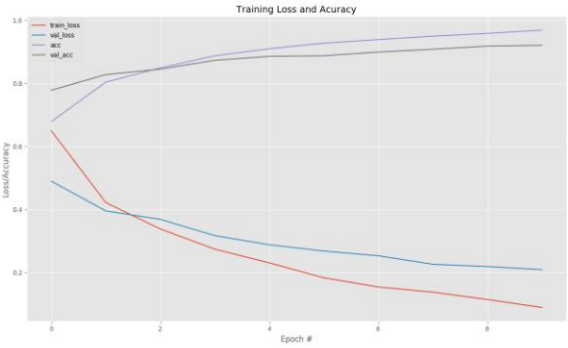

Figure 8: Training Loss/Accuracy vs Epoch-5(AlexNet)

	Precision	Recall	F1-Score
Awake	0.91	0.90	0.90
Drowsy	0.93	0.93	0.93
Avg/Total	0.92	0.92	0.92

Table 8: Accuracy AlexNet Epoch-5

V. DISCUSSION AND CONCLUSION

From the graph and performance results it is evident that Resnet50 model has achieved consistent accuracy of approximately **94%** for a span of 15 epochs. Comparison of Resnet with AlexNet and VGGNet also concludes that Top 5 error around 8% for Resnet compared to Alexnet's 10.38%. This paper has done performance analysis of Resnet and compared it with AlexNet and VGGNet and it proved to better model than other two.

REFERENCES

1. RishuChabra, Dr.SeemaVerma,Dr. C. Rama Krishna,(2017),"A Survey on Driver Behaviour Detection Techniques for Intelligent TransportationTechniques",Conference,36-41.

2. World Health Organization -Road Accidents https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/
3. World Health Organization, "Global Health Observatory (GHO) data", http://www.who.int/gho/mortality_burden_disease/causes_death/top_10/en/
4. Global status report on road safety2013.
5. http://jhtransport.gov.in/causes-of-road-accidents.html.
6. KhTohidul Islam, Ram Gopal Raj, "Performance of SVM, CNN, ANN with BoW, HoG and Image Pixels in Face Recognition", Conference, ICEEE-2017.
7. Rateb Jabbar, Mohammad Kharbeche, "Real-time Driver Drowsiness Detection for Android Application Using Deep Neural Networks Techniques ",IEEE Conference, ANT 2018,400-407.
8. Xuedan Du, Yinghao Cai, Shuo Wang, Lejjie Zhang, "Overview of Deep Learning", IEEE Conference, YAC 2016,159-164
9. Manish Mishra, Monika Srivastava, "A View of Artificial Neural Network", IEEE Conference, ICAETR-2014.
10. Charlotte Jacob, Christophe Bourdin, Anca Stratulat, Jean-Louis Vercher,"Detection and prediction of driver drowsiness using artificial neural network models", Accident Analysis and Prevention , ELSEVIER-2017,https://doi.org/10.1016/j.aap.2017.11.038
11. Karthik Dwivedi, Kumar Biswaranjan, Amit Sethi, "Driver Drowsy Detection using Representation Learning", IEEE Conference, IACC-2014, 995-999
12. MkhuseliNgxande, Jules-Raymond Tapamo, Michael Burke, "Driver drowsiness detection using Behavioral measures and machine learning techniques: A review of state-of-art techniques", IEEE Conference, PRASA-2017, 156-161
13. WEB- https://adeshpande3.github.io/The-9-Deep-Learning-Papers-You-Need-To-Know-About.html
14. Vijayan Asari, Tarek Taha: "The History began from AlexNet: A comprehensive survey on Deep Learning Approaches", Cornell University Library, https://arxiv.org/ftp/arxiv/papers/1803/1803.01164.pdf
15. Sanghyuk Park, Fei Pan, Sunghun Kang and Chang D. Yoo, "Driver drowsiness detection system based on feature representation learning using various deep networks",Springer-Conference,ACCV-2016.
16. https://medium.com/@sidereal/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5
17. Karen Simonyan, Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", ICLR-2015, https://arxiv.org/pdf/1409.1556.pdf