JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Analysis On Driver Distraction Detection And Performance Monitoring System

Hanumanthavajjala Bharat Chandra[1], K. Sree Lakshmi Sravya[2], Bheesetti Eshitha[3],

Boyidi Sai Ram[4], Lenin Sai Priya[5]

Department of Information Technology, Vignan's Institute of Information Technology(Autonomous), Vishakapatnam.

bharat.chandra200@gmail.com[1], sravyakosuri123@gmail.com[2], eshithabheesetti@gmail.com[3], sairamboyidi156@gmail.com[4], lovelysai8035@gmail.com[5]

1. ABSTRACT

According to records, 57% of accidents in India occur on roads. One of the main reasons behind accidents is drivers that gets distracted by various tasks. Chances are getting more for the number of accidents to get increased due to the development of new mobile technologies, which can negatively affect the attention level of the driver.

According to the reports from The National Highway Traffic Safety Administration (NHTSA), over 25% of police reported crashes involved inattentive drivers. This finding is not surprising since it is estimated that about 30% of the time that drivers are in a moving vehicle, they are engaged in secondary tasks. Detecting driver's distraction tasks is an important research problem to prevent accidents and increase the security on the roads. Commonly performed secondary tasks can deviate the attention of the drivers from the primary driving task. These in cab demands can produce visual, cognitive, auditory, psychological and physical distractions. Therefore, it is very important to understand the effect induced by different secondary tasks on the drivers. A key step in the analysis is to define reference metrics or criteria to assess the attention level of the driver. These reference labels can be used as ground truth to train to detect distracted drivers Developing feedback systems that can detect the attention level of the driver plays a key role in preventing accidents by alerting the driver about possible hazardous situations. Monitoring driver distraction is an important research topic, and various forms of technology are available for drivers that can interfere with key driving tasks. An important issue is how to define a reference label that can be used as a ground truth for training and detecting distracted drivers. The answer to this question is not easy, as drivers are affected by visual, cognitive, auditory, psychological and physical distractions. This paper examines three different approaches that characterize driver distraction. Perceptual assessment by an external evaluator, self-assessment from a post-driving questionnaire, and analysis of the differences observed between multimodal characteristics and standard patterns.

The driver usually feels distracted from the road while performing various tasks that can lead to traffic accidents. This document proposes a solution to detect driver distraction and generate reports that provide insight into driver performance. We use different models such as Convolutional Neural Networks (CNN), namely: CNN, VGG16 for the classification of distracted drivers. The deep learning library used for this is TensorFlow. Our best result is 94% accuracy on the validation set and 98% accuracy on the test data.

KEY WORDS: Classification, image segmentation, accuracy, over fitting, dashboard, convolution neural network.

2. INTRODUCTION

According to a report by the Department of Road Transport and Highways, the number of traffic accidents increased by 10%, equivalent to 4,87,421 cases nationwide. Road accidents in the year 2021 is 4,43,110 cases. Our project aims to overcome the previous problem by detecting whether the driver is distracted or not. Software alone or integrated with hardware can warn drivers of distraction and thus prevent accidents. We import camera-based driver tracking images into our model. The model then predicts the class of an image specifying the output class to which this image belongs.
Distracted driving deaths increase faster than deaths from drinking,

speeding and not wearing a seat belt. Drivers are considered distracted when there is an activity that distracts them from their driving duties. There are three types of distracted driving:

Manual distraction: the driver takes his hands off the wheel, e.g., to drink, eat, etc.

Visual distraction: Visual distraction involves taking one's eyes off the road, while manual distraction involves taking one's hands off the wheel, e.g., reading, looking at his phone, etc.

Cognitive distraction: the driver's attention or concentration is not on the driving task, e.g., talking, thinking, etc.

It is important to note that although distracted driving is classified into three different categories, they do not always occur separately. For example, when talking on the phone, two types of distraction occur at the same time: manual distraction and cognitive distraction. Another source that can affect driving performance is phone use. Talking on the phone while driving consumes a significant amount of brain energy. By doing both, a person's brain activity dedicated to driving could be reduced by 37%. Texting while driving can be even more distracting as it not only takes the driver's mind, but also their hands and eyes off the driving task for an average of 4.6 seconds. A recent study indicates that about 78% of drivers use mobile phones while driving, which greatly increases the likelihood of traffic accidents on the road.

To reduce vehicular accidents and improve traffic safety, a system capable of classifying distracted driving is highly desirable and has attracted much research interest in recent years. This research is motivated by the development of a distraction detection system that has the potential to be deployed on real vehicles. Due to the lack of hardware facilities and safety concerns when testing on real roads, we conducted a study on a custom designed electric power assisted driving test rig that is very useful for research studies. preliminary rescue. Therefore, the objective of this work is to develop a driver assistance system capable of detecting distracted driving behaviours and alerting the driver to concentrate on the driving task.

3. RELATED WORK

In the existing system, some of the solutions were based on CNN models built from scratch and some were pre-trained on ImageNet, e.g., VGG16. However, these are of moderate accuracy and have no software for real-time usage.
Techniques absent from some of the above include data sampling, which selects subset of dataset which represents whole dataset, thereby reducing overfitting. Some of them didn't use dropout, a regularization technique that helps reduce overfitting.
Call-backs were used to continuously track the training performance and save the metrics from epoch having good accuracy and fever loss
A distraction log is introduced that tracks all detected distractions and generates visuals, graphs and displays on dashboard that provide feedback-based insights on driver performance.
It is very useful for organizations like Uber, Ola, etc.

4. PROBLEM STATEMENT

Many road accidents are taking place due to the distraction of driver from road while performing other actions such as talking to mobile, chatting on mobile, drinking water, talking to other passengers.
Our motive is to detect such kind of actions and raise alert to driver to ensure safe driving.

4.1 EXISTING SYSTEM

There are few AI based trained models that detect driver distraction but they are of moderate accuracy say 87% and some are of high accuracy but over fitted which leads to faulty results, and there is no proper software for real-time usage.

4.2 PROPOSED SYSTEM

AI model is trained with accuracy up to 94% without any overfitting problems by following some data sampling techniques and carefully designed CNN architecture.

Software is also designed for tracking driver actions in real-time and in addition to this a dashboard was designed which gives valuable insights on driver performance i.e., how often distraction is noted and what kind of actions are frequently being done.

This dashboard consists of visuals giving insights on driver distraction and day wise performance of driver which shows at what time distraction took place and what kind of distraction is detected. It also shows what kind of actions are frequently done by driver of being distracted.

ADVANTAGES:

Safety and business insights: Driver's performance record provides valuable information which helps in talking safety measures.

5. ARCHITECTURE

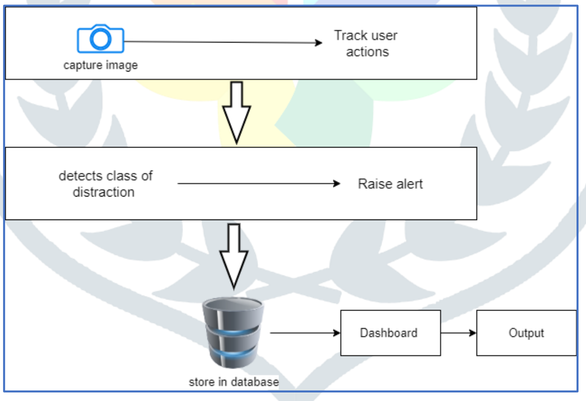

Fig.5.1 architecture

The above figure shows the work flow of our system.

Initially driver will be continuously tracked using web cam or some kind of camera and if any distraction is detected then alert is raised in order to notify the driver.
Tracked actions are stored in database along with date, time, detected action at that particular time to generate visuals and provide insights on driver's performance which is used to take further safety measures.
These insights and visuals are displayed on dashboard which is can be accessed by both driver and administrator at the organization level.
E.g., Uber, Ola

6. LITERATURE SURVEY

In the various studies that have been conducted, many researchers focus on achieving higher accuracy in detecting driver distraction, but this can lead to overfitting problems, and there is also no suitable software and monitoring system to deal with notifying the user of distraction in real-time.

Furthermore, no researcher has focused on how often distraction is noted and what ideas and actions can be taken to reduce it.

Alexey Kashevnik[1,2], Roman Shchedrin[1,2], Christian Kaiser[3], Alexander Stocker[3] main driver in a paper explores various reasons and approaches behind getting distracted. Distraction detection approaches: i) manual distraction, ii) visual distraction, and iii) cognitive distraction.

Pramila M. Chawan[2] in this paper proposed a solution by exploring various models such as several CNN models, our best ensemble was created after averaging the probabilities generated by VGG-16, VGG-19 and InceptionV3. The final log loss which we got was 0.795. Also says that using GPU and algorithms such as K Nearest Neighbor, ResNet50 might increase the performance.

Jimiama Mafeni Mase[1], Peter Chapman[2], Grazziela P. Figueredo[1], Mercedes Torres Torres[1] in a paper proposed model which is trained using various algorithms such as VGG-16, Resnet50, InceptionV3, InceptionV3-LSTM, C-SLSTM and concluded on getting 92% accuracy with C-SLSTM model.

Convolutional Neural Network
Neural networks are computer systems with interconnected nodes that function similarly to neurons in the human brain.
A convolutional neural network is a deep learning algorithm that can take an input image, learns various features or patterns from it, and distinguish one from another.

Input Layer
The input layer consists of our input images of the data set in the desired size i.ie in our case 64x64 which is reduced from 640x480 to speed up the training.

Convolution Layer
Convolutional layer is the main layer of convolutional neural networks. A convolution is the simple application of a filter to an input image, resulting in an activation. Repeatedly applying the same filter to an input result, showing the locations and strength of a recognized trait in an input.

Max Pooling Layer
Max pooling is a pooling operation that selects the maximum element of feature map region covered by the filter. Thus, the output after max-pooling layer would be a feature map containing the most prominent features of the previous feature map.
There are other kinds of pooling layers such as min pooling and average pooling layers.

Rectified Linear Units Layer
ReLu is an activation function applied to each layer to increase the non-linearity of the model. It is a linear function that will output the input directly if it is positive, otherwise, it will output zero.
Formula: max (0, z)

Dropout Layer
The dropout layer is a mask that cancels the contribution of some neurons to the next layer and leaves all others untouched.
Used to reduce the overfitting problem. Sets the input units to 0 with a rate frequency at each step during training time, preventing overfitting. Inputs that are not set to 0 are scaled by 1/ (rate 1) so the sum of all inputs does not change.

Fully Connected Layer
Fully connected layer consists of 10 neurons since we have 10 output classes.

7. APPROACH

Basic CNN

First, we wrote a basic CNN architecture which consists of 3 convolution layers and 3 max pooling layers, one flatten layer, dropout layer and 10 neuron output dense layer, which was trained to 13 epochs. Then we got an accuracy of 94% on training data and 93% accuracy on validation data.

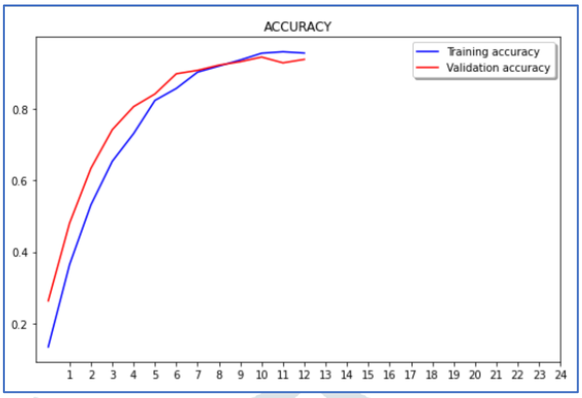

Fig.7.1 Training and Validation accuracy of basic CNN

Above results show that the training accuracy matches the validation accuracy perfectly and shows good model performance.

Now to increase accuracy we tried to train our dataset on some huge CNN architecture, hence used VGG16 pre-trained model.

VGG-16

VGG-16 is a part of VGG network architecture presented in the article Very Deep Convolutional Networks for Large Scale Image Recognition. VGG16 carries 16 weights. Now as VGG16 is very huge network, to reduce training time we freeze the upper layers and made use of last layer weights.

We got an accuracy of around 99% for both training and validation sets.

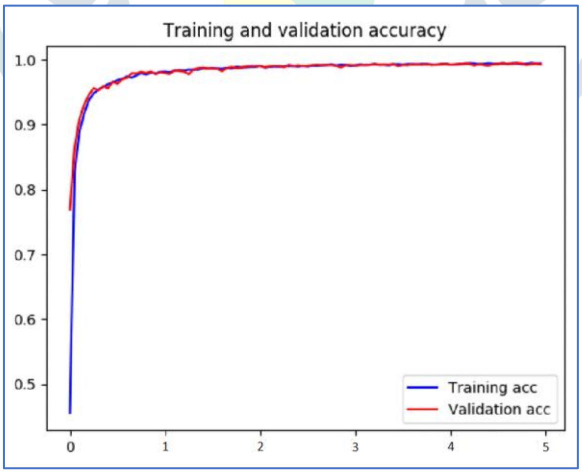

Fig.7.2 Training and validation accuracy of VGG-16

As discussed earlier too much accuracy leads to over fitting problems, as shown below we got some wrong predictions.

Fig.7.3 Wrong prediction

Then we again thought of writing our own CNN architecture which resembles VGG16.

CNN 2.0

This time we designed the CNN architecture in such a way that it has two consecutive conv2d layers followed by one max pooling layer

Therefore, the data fed to model for training is augmented. Data augmentation makes the model become more robust and prevent overfitting. Dropout layers were also added. Dropout helps in ignoring a few random neurons while training of the model on the dataset.

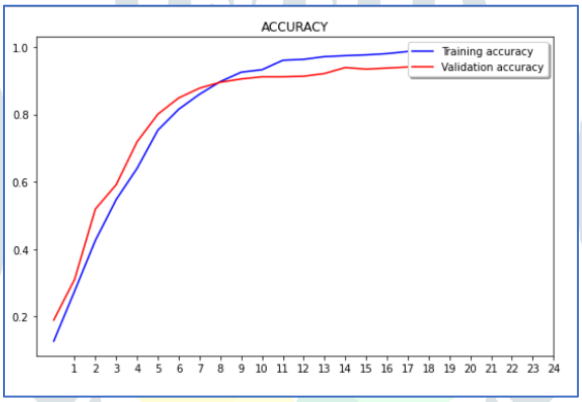

Fig.7.4 Training and validation accuracy of new CNN architecture

8. FUTURE WORK

Used small dataset due to limited computational resources.

1. Adding more features to dataset such as detecting wireless devices like Bluetooth earphones, tracking mouth movement to detect continuous talking can help in detecting more categories of distractions.
2. Using GPU can help in increasing computational power and training large datasets with huge networks like ResNet without causing overfitting.

9. RESULTS

Software tracks driver actions and sends alert when distraction is detected, dashboard will provide insights on driver's performance which can be used for taking further decisions.

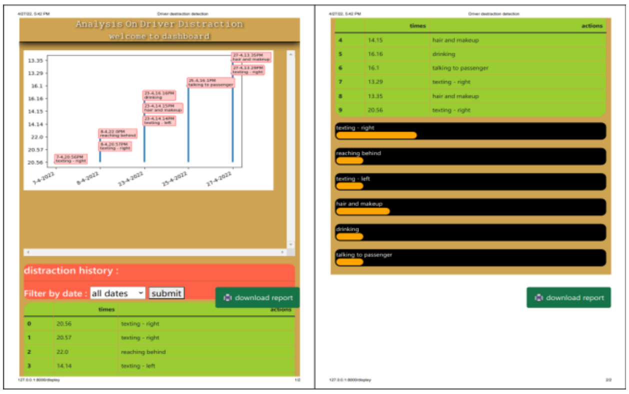

Fig.9.1 Dashboard

10. CONCLUSION

	Epochs	Train Loss	Train Acc	Val Acc	Val Loss
CNN 1.0	13	0.1266	0.9545	0.9437	0.2439
VGG16	5	0.0143	0.9928	0.9852	0.0482
CNN 2.0	18	0.0452	0.9867	0.9405	0.2767

Fig.10.1 – model comparative analysis

Thus, after testing out several CNN architectures like VGG16, ResNet, our final CNN model concluded with the loss of 0.042 and accuracy of 98% for training and loss of 0.27 and accuracy of 94% for testing data.

11. REFERENCES

[1] Driver Distraction Detection Methods: A Literature Review and Framework | IEEE Journals & Magazine | IEEE Xplore

[2] https://www.researchgate.net/publication/326416658

[3] ICTC-2020-final.pdf (horizon.ac.uk)

[4] Report on Road Accidents in India 2016- Ministry of Road Transport & Highways (MoRTH), Government of India pp. 1-2 http://morth.nic.in/showfile.asp?lid=2904

[5] Kaggle. A brief summary https://www.kaggle.com/c/state-farmdistracted-driverdetection

[6] Yehya Abouelnaga, Hesham M. Eraqi, and Mohamed N. Moustafa, “Real-time Distracted Driver Posture Classification”, arXiv preprint arXiv:1706.09498

[7] T. H. N. Le, Y. Zheng, C. Zhu, K. Luu and M. Savvides, "Multiple Scale Faster RCNN Approach to Driver‟s Cell-Phone Usage and Hands on Steering Wheel Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, NV, 2016, pp. 46-53

[8] Hssayeni, Murtadha D; Saxena, Sagar; Ptucha, Raymond; Savakis, Andreas,“Distracted Driver Detection: Deep Learning vs Handcrafted Features”, Society for Imaging Science and Technology, Imaging and Multimedia Analytics in a Web and Mobile World 2017, pp. 20-26(7)

[9] dImageNet: VGGNet, ResNet, Inception, and Xception with Keras https://www.pyimagesearch.com/2017/03/20/imagenet-vggnet-resnet-inception- xceptionkeras/

[10] CS231n Convolutional Neural Networks for Visual Recognition http://cs231n.github.io/convolutionalnetworks/

[11] Karen Simonyan, Andrew Zisserman, “Very Deep Convolutional Networks for LargeScale Image Recognition”, arXiv preprint arXiv:1409.1556

[12] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, “Going Deeper with Convolutions”, arXiv preprint arXiv:1409.4842