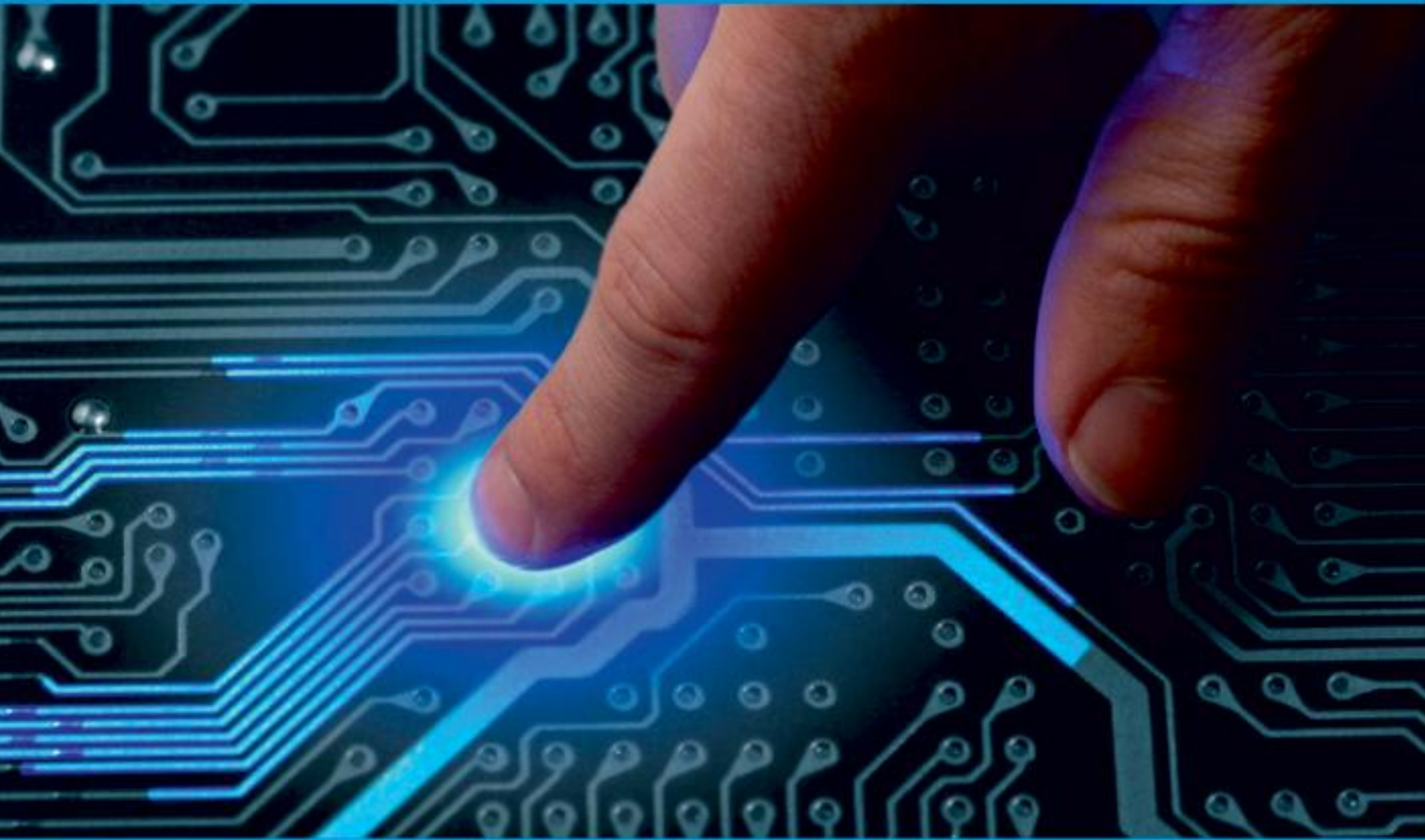
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 7, July 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907196 |

Real Time Detection of Driver Distraction using CNN

Madhumathi T, Deepika P, Kavitha R, Mrs.Sumathi G

UG Student, Department of CSE, Velammal Engineering College, Chennai, India

UG Student, Department of CSE, Velammal Engineering College, Chennai, India

UG Student, Department of CSE, Velammal Engineering College, Chennai, India

Assistant Professor, Department of CSE, Velammal Engineering College, Chennai, India

ABSTRACT: Distracted driving is the main cause for large number of motor vehicle accidents across the globe. Detecting a distracted driver is considered as the significant research area for reducing the road accidents. This paper focuses on a methodology to reduce the accidents caused by distracted driver with deep neural networks. CNN based method is used to develop the actions of driver from driver image dataset, which is used to classify the distracted driver into different categories. The increasing use of in-vehicle information systems (IVISs), such as navigationdevices and MP3 players, can jeopardize safety by introducing distraction into driving.One way to address this problem is to develop distraction mitigation systems, whichadapt IVIS functions according to driver state. In such a system, correctly identifyingdriver distraction is critical, which is the focus of this dissertation. Visual and cognitivedistractions are two major types of distraction that interfere with driving most comparedwith other types. Visual and cognitive distraction can occur individually or incombination. The research gaps in detecting driver distraction are that the interactions ofvisual and cognitive distractions have not been well studied and that no accuratealgorithm/strategy has been developed to detect visual, cognitive, or combineddistraction.

I. INTRODUCTION

Every year more than 1 million people die due to the unsafe and dangerous driving all over the world. Distracted driving is one of the main causes for these road accidents. This is mainly due to the negligence of the drivers and secondary tasks done by the drivers while driving, such as answering a phone call, talking to co-passenger, texting, controlling the radio etc.

Furthermore 86% of the drivers admit that they use any of the following: attending the calls, replying to the texts, checking GPS, talking to the fellow passenger, eating or drinking while driving, checking maps, watching video, grooming, and browsing web.

These statements are further motivated to identify the ways to reduce the distracted driving. As pointed out, these inattentions can lead to road accidents. Accidents can be reduced to a high extend if the drivers are warned at the moment the distraction is observed.

To bridge these gaps, the dissertation fulfilled three specific aims. The first aimdemonstrated the layered algorithm developed based on data mining methods couldimprove the detection of cognitive distraction from my previous studies. The second aimdevelopedestimationalgorithms for visual distraction and demonstrated a strongrelationship of the estimated distraction with the increased risk of real crashes using thenaturalistic data. The third objective examined the interaction of visual and cognitivedistractions and developed an effective strategy to identify combined distraction.Together these aims suggest that driver distraction can be detected from performanceindicators using appropriate quantitative methods. Data mining techniques represent apromising category of methods to construct such detection algorithms. When combinedin a sequential way, visual distraction dominates the effects of distraction while cognitivedistraction reduces the overall impairments of distraction on driver performance.

The proposed paper attempts to identify the distraction of drivers by the actions recorded in the camera. A camera is placed in the right side of the driving seat and record all the actions made by the driver. These body movements are categorized into

10 classes and if the actions or body movements falls in any of these classes, it is said that the driver is distracted. CNN based approach will help us to classify the driver as distracted or non-distracted based on the body movements.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907196 |

II. RELATED WORK

Finding the distracted driver using machine learning technique is one of the problems that undergoes continuous research. Many researches has been conducted on this topic. Table 1 shows the various researches conducted on driver distraction and their results.

Table 1. Literature Survey

Literature Survey 1	
Title	An Adaptive Forward Collision Warning Framework Design Based on Driver Distraction
Authors	Seyed Mehdi Iranmanesh , HosseinNourkhizMahjoub , HadiKazemi , and Yaser P. Fallah
Published Year	2018
Efficiency	A higher accuracy was claimed in comparison with the correspondent SupervisedExtreme Learning Machine (SELM) approach.
Drawbacks	car damages which impose undesired costs and traffic jams which result in mental stress and considerable waste of time and energy are the other important drawbacks of these crashes.
Description	Forward Collision Warning (FCW) is a promising Advanced Driver Assistance System (ADAS) to mitigate rear-end collisions. The deterministic FCW approaches may occasionally lead to the issuance of annoying false warnings, as they cannot be individualized for different drivers. This application oversight, which may cause the driver to deactivate the system, has been tackled with some adaptive methods. However, driver distraction, which is one of the most influential driver-specific factors on FCW warnings acceptability, has not been considered yet and is analyzed in this paper for the first time. Specifically, the adaptive FCW method proposed in this paper generates the warnings by continuously comparing Time Headway with a flexible threshold. The core of the proposed threshold updating mechanism is a realtime monitoring of the driver reactions against the previously generated warnings using the available indicators such as braking history. This method considers the driver distraction in parallel to fine-tune the calculated threshold in accordance with driver cognitive state. In order to incorporate the driver distraction in the system framework, a learning-based approach is designed which continuously estimates driver distraction by the virtue of different available Controller Area Network (CAN) bus time series, such as throttle pedal position, velocity, acceleration, and yaw rate. Neural network, as a widely adopted classification method, is nominated to detect driver distraction. The framework performance is evaluated over two realistic driving datasets. An approximately 80% false warning reduction is observed in analyzed safe scenarios, while no critical warning is missed in the dangerous ones.

Literature Survey 2	
Title	Incorporation of Driver Distraction in Carfollowing model based on Driver's Eye Glance Behavior
Authors	Yeeun Kim ,Seongjin Choi ,Daejeon
Published Year	2018
Efficiency	Variables affecting glance behavior is divided into road environmental variables and vehicle movement variables and analyzed. According to the results of analysis, a hierarchical decision tree is generated considering both road environmental

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907196 |

	variables and vehicle movement variables
Drawbacks	CDF of TTC distributions for a range of 0 to 5 seconds to analyze more risk situation. Distracted OFFA shows a closer distribution to NGSIM, however, causes more unsafe situations with TTC value of less than 1.5 seconds
Description	This paper aims to incorporate driver distraction into car-following model to demonstrate hazardous situations such as vehicle collision. Many reports point out a driver distraction as one of the major causes of automobile collisions and driver's eye glance behavior is representative indicator for quantitatively evaluating distraction. To analyze distraction, a 100-car Naturalistic Driving Study data including eye glance data is used. This study classified several variables affecting the glance behavior into two different groups and analyzed in different ways. Based on the results of the analysis, decision tree analysis is conducted to derive driving scenarios according to the eye glance behavior and total of eleven scenarios are derived. Driver's glance behavior is modelled by scenarios and incorporated into the existing car-following model. As one of the existing car-following model, Oversaturated Freeway Flow Algorithm (OFFA) is extended to distracted OFFA. We show that distracted OFFA describes real-world driving more closely in terms of vehicle safety by comparing a distribution of time to collision (TTC) with existing car-following models.

Literature Survey 3	
Title	Driver Distraction Detection Using Semi-Supervised Machine Learning
Authors	Tianchi Liu ; Yan Yang ; Guang-Bin Huang ; Yong Kiang Yeo ; Zhiping Lin
Published Year	December 2015
Efficiency	This shows that by exploring the data structure without actually labeling them, we gain extra information to improve the performance of driver models.
Drawbacks	The future study will investigate another solution, i.e., to develop distraction detection system that adapts its model to new scenarios in an online manner.
Description	Real-time driver distraction detection is the core to many distraction countermeasures and fundamental for constructing a driver-centered driver assistance system. While data-driven methods demonstrate promising detection performance, a particular challenge is how to reduce the considerable cost for collecting labeled data. This paper explored semi-supervised methods for driver distraction detection in real driving conditions to alleviate the cost of labeling training data. Laplacian support vector machine and semi-supervised extreme learning machine were evaluated using eye and head movements to classify two driver states: attentive and cognitively distracted. With the additional unlabeled data, the semi-supervised learning methods improved the detection performance (G-mean) by 0.0245, on average, over all subjects, as compared with the traditional supervised methods. As unlabeled training data can be collected from drivers' naturalistic driving records with little extra resource, semi-supervised methods, which utilize both labeled and unlabeled data, can enhance the efficiency of model development in terms of time and cost.

Literature Survey 4	
Title	Detection of Driver Cognitive Distraction: A Comparison Study of Stop-Controlled Intersection and Speed-Limited Highway

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

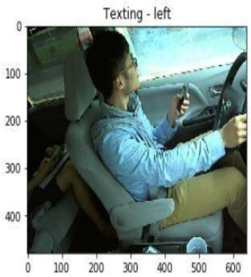
|| Volume 9, Issue 7, July 2021 ||

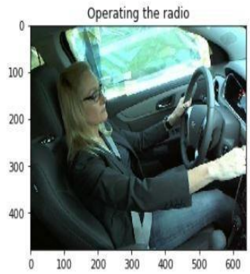
| DOI: 10.15680/IJIRCCE.2021.0907196 |

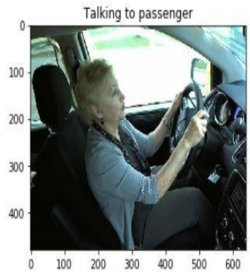
Authors	Yuan Liao ; ShengboEben Li ; Wenjun Wang ; Ying Wang ; Guofa Li ; Bo Cheng
Published Year	January 2016
Efficiency	This paper presents a method for the detection of driver cognitive distraction happened at stop-controlled intersections, and compares its feature subsets and classification accuracy with that on speed limited highway.
Drawbacks	Future works will focus on the overcoming of limited data in short time event and develop real-time detection of cognitive distraction adaptive to both urban and highway scenarios.
Description	Driver distraction has been identified as one major cause of unsafe driving. The existing studies on cognitive distraction detection mainly focused on high-speed driving situations, but less on low-speed traffic in urban driving. This paper presents a method for the detection of driver cognitive distraction at stop-controlled intersections and compares its feature subsets and classification accuracy with that on a speed-limited highway. In the simulator study, 27 subjects were recruited to participate. Driver cognitive distraction is induced by the clock task that taxes visuospatial working memory. The support vector machine (SVM) recursive feature elimination algorithm is used to extract an optimal feature subset out of features constructed from driving performance and eye movement. After feature extraction, the SVM classifier is trained and cross-validated within subjects. On average, the classifier based on the fusion of driving performance and eye movement yields the best correct rate and F-measure (correctrate = 95.8 ± 4.4%; for stop-controlled intersections and correct rate = 93.7 ± 5.0%; for a speed-limited highway) among four types of the SVM model based on different candidate features. The comparisons of extracted optimal feature subsets and the SVM performance between two typical driving scenarios are presented.

III.PROPOSED WORK

A.Dataset State Farm Dataset [10] is used for training and testing the model. State farm, an insurance company released dataset, which contains driver images for Kaggle competition. The dataset consists of 102,150 labelled images of 26 subjects which includes different colour, ethnicity, action, age etc. these subjects were used to perform classification of the subjects into 10 classes such as normal/safe driving, talking in the phone by both left and right hands, controlling the radio, text messaging by left hand, test Messaging by right hand, talking to the co passenger, drinking while driving, hair and makeup, reaching behind, etc. each image is labelled with their action class. State farm data set is aimed for frame by frame classification.







| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907196 |

Figure 1. All 10 classes of Driving Categorized by the Dataset

Category distribution graph of distracted driver's class is given in figure 2. All the ten classes are represented from c0 – c9. There are total of 102150 images are there in the dataset, out of which 17939 are taken as training data and rest as test data.

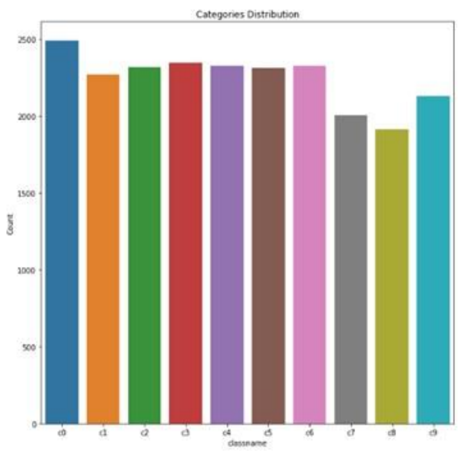


Figure 2. Category Distribution Graph

B.METHODOLOGY

Deep learning networks learn from multiple levels of representation, which helps to extract meaningful features from raw data. In this work all the implemented strategies are either similar to convolutional neural networks or slight variation from CNN architecture. First general methods of CNN are described and later focus on the variations done in the CNN model. CNN consist of many layers including convolutional layer, max pool layer, ReLU layer, drop out layer and fully connected layers. State farm data set is used for training the model which includes images of drivers while performing a number of actions including drinking, talking to co-passenger, texting, talking on phone while driving, etc. Keras and tensor flow libraries are used to create convolutional neural network. The model is built using Google colab research platform, which provides high performing setup to train and test the model. Images in the data set are cleaned to a same specification and 10 classes are defined. Cleaning includes cropping the image into equal size, correcting the pixels and removing the unwanted data. Name of the image and their corresponding classes are

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907196 |

maintained in the excel sheet. From the CSV file, class names are taken as the labels for the images and use the image names to match the labels with the correct images. We develop three models and compare their results for accuracy.

1.Improved Vanilla CNN Model

Vanilla CNN model is constructed with a total of 3 convolutional layers, Then a flatten layer and 3 dense layer.
Convolutional layer 1 consist of 60 filters, with each kernel of size 3x3. We use relu activation function with same padding and weights are initialized from 0.0001. A dropout of 0.3 is also added to this layer.
In the second convolutional layer, a total of 90 filters each with size of 3x3 is used. A dropout of 0.3 is added in this layer also.
Similar to this, in the third convolutional a total of 200 filters are used with a dropout of 0.5. After 3 convolutional layers, a flatten layer and 3 dense layers are also used with 512, 128 and 10 filters each. *RMS-Prop* optimizer and *categorical cross entropy* are also used for loss. Activation function for 3 convolutional layers are ReLU and activation function for the first dense layer is ReLU and the last dense layer is softmax. Model configuration is shown in table 2.

Table 2. Model Configuration for Improved Vanilla CNN

Layer	Output Shape
Convolutional Layer	(None, 62, 62, 32)
Max pooling	(None, 31, 31, 32)
Convolutional Layer	(None, 31, 31, 64)
Max pooling	(None, 16, 16, 64)
Convolutional Layer	(None, 16, 16, 128)
Max pooling	(None, 8, 8, 128)
Flatten	(None, 8192)
Dense	(None, 512)
Dense	(None, 128)
Dense	(None, 10)

2.Vanilla CNN with data Augmentation

The previous model is augmented to generate more images. This data is used to train the model. Vanilla CNN model with data augmentation is constructed with a total of 3 convolutional layers, then a flatten layer and 3 dense layer. *Adam* optimizer and *categorical cross entropy* are used for loss. Activation function for 3 convolutional layers are ReLU and activation function for first two dense layer is ReLU and last dense layer is sigmoid.

Once the model is ready, train the model using the generated data. The images in the dataset are classified into 2 categories, training and testing. 80% of the total datasets are used for training and the rest is for testing the model accuracy. More images are generated using *ImageDataGenerator*.
The model is trained for 5 epochs and saved.

Table 3: Model Configuration for Vanilla CNN with Data Augmentation.

Layer	Output Shape
Convolutional Layer	(None, 238, 238, 128)
Max pooling	(None, 119, 119, 128)
Convolutional Layer	(None, 117, 117, 64)
Max pooling	(None, 58, 58, 64)

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907196 |

Convolutional Layer	(None, 56, 56, 32)
Max pooling	(None, 28, 28, 32)
Flatten	(None, 25088)
Dense	(None, 1024)
Dense	(None, 256)
Dense	(None, 10)

3.CNN with Transfer Learning(VGG, MobileNet)

To reduce training time without losing the accuracy, CNN using transfer learning is used. A total of 4 convolutional layer and 2 dense layer are used for this model. Model configuration of CNN with transfer learning is given in table 4.

Table 4. Model configuration for CNN with transfer learning.

Layer	**Output Shape**
Convolutional Layer	(None, None, None, 64)
Max pooling	(None, None, None, 64)
Convolutional Layer	(None, None, None, 128)
Max pooling	(None, None, None, 128)
Convolutional Layer	(None, None, None, 256)
Max pooling	(None, None, None, 256)
Convolutional Layer	(None, None, None, 512)
Max pooling	(None, None, None, 512)
Dense	(None, 1024)
Dense	(None, 10)

IV. EXPERIMENTS

A.Results and Analysis

Accuracy and Loss are calculated for all 3 models, in order to find which model performed better in different scenarios. Accuracy can be calculated by forming the confusion matrix. Equation for accuracy is,

Classification Accuracy= $\frac{TP+TN}{TP+TN+FN+FP}$

Where FP= False Positive (where we predicted YES and the actual output was NO), TP= True Positive (where we predicted correct classification), FN= False Negative (where the predicted output was a NO but actual output was YES), TN= True Negative (predictions where correct).

Categorical Cross Entropy is calculated for all the three models. This loss function is most popularly used for multiclass classification. Categorical Cross Entropy is defined as,

Categorical cross entropy = $-\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} 1y_i\epsilon C_c \log P_{model}[y_i\epsilon C_c]$

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907196 |

Where N is the number of observation, C the number of categories and P is the probability predicted by the model. Three different models are trained with 5 epochs each. From the reports it is obvious that optimized vanilla CNN give the maximum accuracy and minimum loss. The detailed analysis for each model is given below.

1. Improved Vanilla CNN Model.

Optimized vanilla CNN Model shown highest accuracy of 97.66% and lowest loss of 9.6%. Detailed analysis is shown in the figure 3.

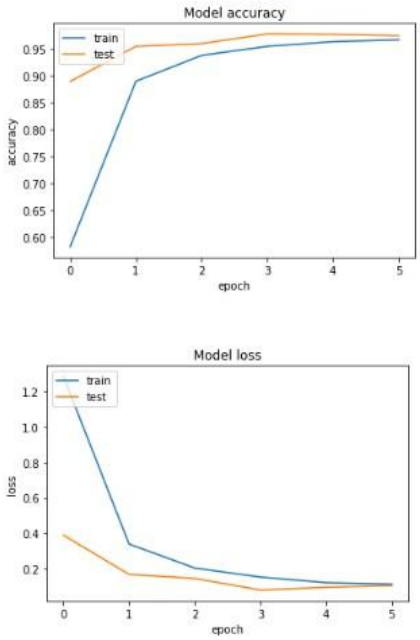


Figure 3. Accuracy and Loss of Improved Vanilla CNN model

2. Vanilla CNN with Data Augmentation.

Vanilla CNN model with data augmentation has shown 97.05% accuracy and 11.18% loss. Detailed analysis is shown in figure 4.

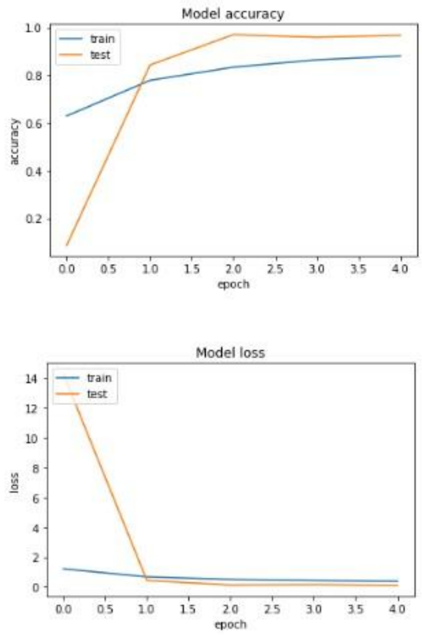


Figure 4. Accuracy and Loss of vanilla CNN Model with Data Augmentation.

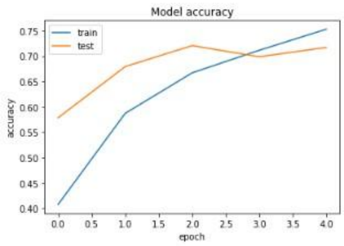
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907196 |

3. CNN with Transfer Learning

CNN with Transfer learning reduces training time without sacrificing accuracy. This model has shown an accuracy of 71.72 % and loss of 75.78 %. Detailed analysis is shown in figure 5.



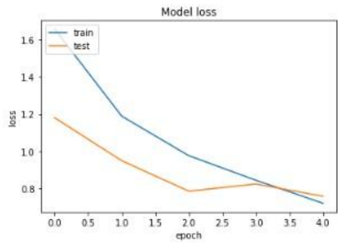


Figure 5. Accuracy and Loss of CNN with Transfer Learning.

Table 5. Effect of Varying Parameters on Overall Test Accuracy

No. of Layers	**No of Filters**	**Testing Accuracy% of CNN model**		
		Improved Vanilla CNN	**Data Augmentation**	**Transfer Learning**
2	conv1=32, conv2=64	53.32	62.96	40.75
2	conv1=64, conv2=128	58.36	68.00	42.56
3	conv1= 40, conv2=60, conv3=80	88.95	83.39	54.63
3	conv1=100, conv2=200, conv3=300	**96.63**	86.46	56.12
3	conv1= 100, conv2=250, conv3=350	96.28	88.09	66.70
4	conv1= 50, conv2=100, conv3=150, conv4=200	93.75	95.93	**75.31**
4	conv1=100 conv2=140, conv3=180, conv4=200	94.62	**96.78**	72.13
5	conv1= 50, conv2=100, conv3=150, conv4=200, conv5=250	89.4	87.42	72.08

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907196 |

5	conv1=100, conv2=200, conv3=400, conv4=450, conv5=500	74.7	79.51	67.94

V.CONCLUSION

Three different models are trained to solve the problem of drivers' distraction in this paper.
Proposed model has achieved 97% accuracy and tried a new model with very less training time without compromising much on the accuracy. The transfer learning model has also achieved 75% accuracy with less amount of data. This method can still be improved with more data. Major concern regarding this implementation is the privacy of the passengers. Most of the drivers has found it difficult to maintain a camera to monitor their driving. However, by considering the benefits of this system, the negative aspects can be neglected.

VI. FUTURE WORK

As a futuristic method, skeleton images can be explored as input. This will reduce the noise and other background influence in the system.
Other sensor devices can be used for obtaining better results. The proposed model can also add more action classes to avoid discrimination and misclassifying the actions.

REFERENCES

[1] C. Yan, F. Coenen, and B. Zhang, "Driving posture recognition by convolutional neural networks," IET Computer Vision, vol. 10, no. 2, pp. 103–114, 2016.
[2] Y. Abouelnaga, H. M. Eraqi, and M. N. Moustafa, "Realtime distracted driver posture classification," arXiv preprint arXiv:1706.09498, 2017.
[3] Leonel Cuevas Valeriano, Paolo Napoletano,RaimondoSchettini. "Recognition of driver distractions using deep learning", 2018 IEEE 8th International Conference on
Consumer Electronics - Berlin (ICCE-Berlin), 2018
[4] ManjulaKumari, C.V Hari, Praveen Sankaran."Driver Distraction Analysis Using Convolutional Neural Networks", 2018 International Conference on Data Science and Engineering(ICDSE), 2018
[5] S. Masood, A. Rai, A. Aggarwal, M. Doja, and M. Ahmad, "Detecting distraction of drivers using convolutional neural network," Pattern Recognition Letters, 2018.
[6] Y. Yun, I. Y. Gu, M. Bolbat, and Z. H. Khan, "Videobased detection and analysis of driver distraction and inattention," in Signal Processing and Integrated Networks (SPIN), 2014 International Conference on. IEEE, 2014, pp. 190–195.
[7] A. Rangesh and M. M. Trivedi, "Handynet: A one-stop solution to detect, segment, localize &analyze driver hands," arXiv preprint arXiv:1804.07834, 2018.
[8] Z. Mu, J. Hu, and J. Yin, "Driving fatigue detecting based on eeg signals of forehead area," International Journal of Pattern Recognition and Artificial Intelligence, vol. 31, no. 05, p. 1750011, 2017.
[9] P. Napoletano and S. Rossi, "Combining heart and breathing rate for car driver stress recognition," in International Conference on Consumer Electronics - Berlin (ICCE-Berlin). IEEE, 2018, pp. –.
[10] S. Farm, "State farm distracted driver detection - data," https://www.kaggle.com/c/state-farm-distracted-driverdetection/data, 2016, accessed: 2018-05-21.
[11] NHTSA, "Distracted driving," https://www.nhtsa.gov/risky-driving/ distracted-driving/, 2015, accessed: 2018-05-17.
[12] MORTH "Ministry Of Road Transport and Highways" https://morth.nic.in, 2018,accessed : 2020-03-04.
[13] Pandian, M. D. (2019),"Sleep Pattern Analysis And Improvement Using Artificial Intelligence And Music Therapy"Journal of Artificial Intelligence, 1(02), 54-62. [14]Bashar, A. (2019),"Survey On Evolving Deep Learning Neural Network Architectures",Journal of Artificial Intelligence, 1(02), 73-82.