Dimensionally improved residual neural network to detect driver distraction in real time

M Balamurugan* and R Kalaiarasi
School of Computer Science, Tamilnadu Open University, Chennai, Tamilnadu, India.
E-mail: *techbalamurugan@gmail.com

Abstract. Bountiful reasons are there for crop up of road accidents. Predominantly most of the road accidents are turn out by the humans. In compare with other sources of road accidents, distracted driving is the foremost one. Driver gets distracted ascribed to many factors such as texting using phone, seeing outside of the car. The regime can grasp the road accidents by enforcing the traffic regulations moreover the regime seeks the assistance from the technology to reduce the car accidents due to driver distraction. There are numerous research undergoing to reduce the chance of the driver to get distracted by monitoring the physical activities of the driver while driving the four wheeler. Automatic detection of driver distraction can support to develop an apprise system by observing the activities of the driver that provide better result and evading the car accidents. Deep learning is a leading technology in automated industries to detect the traffic light signs and pedestrians. The novel ResNext is the proposed work and inherited from its parent model called Residual Network (ResNet) to train and test the model. The “Next” in the word ResNext represents next dimension of ResNet called as “cardinality”. In this paper, with the help of Convolutional Neural Network (CNN), the system detects and captures the distracted moments of the driver from proper driving gestures automatically in the form of 2D image. The test results of the model outperforms and proved its significance over the existing driver detection algorithms by bagging higher efficacy with 97.6% accuracy in classifying the live driver posture.

1. Introduction

Road accidents are rapidly increasing these days due to increasing number of vehicles across the globe. Globalization plays a major role in the increase of vehicles. Densely populated countries like India and China have very less proportion of public transportation when compared to the population. So, people over there are more likely to use their own vehicles to travel than using a public transport to avoid crowd and to be on time. Increase in average speed directly increases the chance of accident occurrence and also increases the severity of the crash. In practical, 1% of increase in speed will lead to 4% increase in fatal crash risk and 3% increase in severe crash risk. Due to road traffic crashes, around 13lakh people die every year as per WHO (World Health Organization) report. Road accidents are the main cause of death for people around 5-29 years of age. More than 90% of the road traffic death occurs on low - middle economic countries and the death rates are higher in the African region. According to male-female death ratio, 73% of male are more likely dead in road accidents when compared to female and they were also under the age of 25, which is almost male have 3 times higher death risk rate in accidents than female population. Intake of drugs, psychoactive substances and alcohol also elevates the chance of fatal crash or may lead to severe injuries. Following the traffic protocols like wearing helmets, seat belts and strictly following road signs will drastically reduce the risk of accidents. Wearing the helmet in a proper way can reduce the risk of death by 42% and reduce the risk of major injuries by 69% according to a study. Accidents also occur due to distraction of vehicle driver due to various reasons.
According to NHTSA (National Highway Traffic Safety Administration) a study report on U.S (United States) road crash in 2018 alone, 2,841 people were killed in road accident involving distracted drivers. Distraction due to mobile phones becomes very common among drivers and also a growing concern for traffic safety. Texting becomes the more common distraction these days. Reading or chatting on phone can take the drivers eyes off the road for almost 3-5 seconds which is compared to travel a football ground at the speed of 80-90 KM/H with eyes closed. Drivers using mobile devices while driving increase the chance of accident by 4 times than the drivers who were not using mobile phones/devices. Because it considerably reduces the reaction time for the driver to act suddenly in some critical conditions, notably it delays braking action. The government also needs to enforce the traffic law effectively by establishing the effective law (rules/protocols may vary country to country), updating the law on regular intervals and enforcing at all levels without any excuses and fails. Unsafe road infrastructure, unsafe vehicles also contributes to road accidents. But all other aspects cannot be fully controlled by the driver, but the driver can minimize the risk of met with an accident by be attentive and concentrated on the road by not letting himself distracted over other things.

This paper mainly focuses on four wheeler/car driver’s distraction detection mechanism. Some expensive cars now a day have built in driver alert system to keep the driver more focus and conscious on driving. Most listed driver distractions are narrowed down into to four major categories, Visual Distraction, Auditory Distraction, Manual Distraction and Cognitive Distraction. The proposed work functions on a deep learning model to detect the driver and capture the distracted moments as images using a camera. Deep Learning methods works better in object detection and image recognition. It mimics the decision making function of the brain by analyzing and processing the data and creating patterns. The obtained Artificial Intelligence is used in the applications like detecting objects, recognizing speech etc. To detect driver’s unusual movements while driving, 2 dimensional camera will be placed on the front cabin/dashboard of the vehicle. The eye movement, head position, hand movements, and body position will be taken in account and be captured if any distracted gesture is observed by the system. The detective performance of the proposed algorithm effectively outperforms the existing driver detection models.

The remaining work of the paper is ordered as follows. The existing research works related to this study is discussed in section 2. The next section briefly describes the methods and mechanism used in the proposed system. Results, comparisons and discussions will be scrutinized and discussed in Section 4. The last section concludes the work with its importance and significance over other mechanisms.

2. Background study

Accidents are come to be a usual one in contemporarily life ascribed to health issues, by obtaining the distractions from driving. Drivers may has a frazzle as a result of long drive. The author of this paper illustrate about the detection of the frazzle condition of the driver with the help of real time using computer vision. Real time using computer vision composed of hereunder components, Raspberry Pi, video camera, and for alerting the driver it uses the signal system. This paper focus on collecting the data of gaze identification [1][2] gleaned from the data it can send the alert signal to driver. In an effort to recognize the face of the driver, the upcoming algorithms are typically used. Bayesian Filtering and Gaussian processor, Convolutional Neural Network and hyper spectral imaging algorithms. Viola Jones Face Recognition Approach is implemented here to reach the more accuracy and high speed. Viola Jones Face Recognition approach detects the face which is in front of camera. It differentiate the face part from non-face parts. For selecting the features it used Haar feature selection and for training and classifiers those features AdaBoost and Cascading classifiers algorithm are used. Moreover, this approach provides 95 per cent of veracity when compares to prevailing algorithms [3].

This research paper describes the algorithm to determine the driver distraction with the assistance of the machine learning algorithms (supervised and unsupervised) and in vehicle signal processing. Most of the accidents are occur because of driver distractions. Driver may distract due to tiredness, hearing the music,
speaking on the phone, doing some activities while driving [5]. Detecting the distraction of driver done by two things one is detection of distraction and distraction of driving. Distraction Detection can be achieved by unsupervised machine learning algorithms, in that Gaussian Mixture Model (GMM) is used [6]. GMM Model couldn’t generate the difference between the normal driving and distracted driving because if a vehicle speed is over 50 Km/hr, with the help of the data the GMM couldn’t make a decision because of that additional algorithm is introduced as detecting the driver distraction along with driving distraction by pattern learning. Artificial Neural Network received the inputs of both normal pattern and pattern of distracted driving, 1 indicates as distracted driving pattern and 0 indicates as normal driving pattern and train those inputs till achieve the expected result. When compare with the previous algorithm, the new algorithm achieves the better result with better accuracy. [8].

Detecting the distraction of driver by virtue of conducting the validation of deep Convolutional Neural Network trained models on the embedded graphic processing system. VGG-16, AlexNet, GoogleNet and ResNet[9-12] are the trained models of the CNN. The author describes the distraction into 3 ways, Distraction occur by manual (Reading), Distraction by visual (Seeing the roadside) and Distraction by cognitive (Thinking). Distraction of drivers can be captured by using video camera [25-29]. The above models are trained and implemented in the assisted driving testbed which holds the four components a) Simulator b) Software tool c) Script d) embedded system. NVIDIA Jetson TX1, NanoPi M3 are the embedded systems used to implement the CNN models, every models implemented on each embedded system to get the result faster and more accurate. Moreover all those trained models couldn’t discover the misclassification behavior of the driver. Eg., put make up etc., its one of the drawback of all the models. Following of the individual implementation on the embedded system, the results are compared, comparing with the other trained models, ResNet produces the better accuracy with minimum frequency of 8 Hz but has more complexity and too slow. Therefore, based on the accuracy, frequency, complexity and speed, GoogleNet performs better than rest of the model for detecting the driver distraction [14].

Improvised architecture of VGG – 16 designed to detect the distraction of drivers while driving the car. Major accidents of car occurs due to the driver distraction. Nearly 13 lakh people die in accidents, Most of the accidents happens because of cognitive distraction. This paper explains one of Convolutional Neural Network model to detect the manual type of distractions of the driver. Kaggle provides the dataset, it holds 10 classes like eating, drinking, texting on the phone.[15][16]. Transfer learning is applied to detect the distraction with more accurate and in better performance. A solution of a problem can be applied to the different problem but that should be related to it. This system uses the VGG-16 model with ReLU Activation function as a classifier [18] for initializing the weight ImageNet model is used and for optimizing the dataset Adam Optimizer is applied. The model generate the result with the accuracy of 82.5 percent. Moreover, this system can be used only for the manual distraction, not for cognitive and visual type of distractions.

Distraction of driver creates a major path for the accidents, to avoid these kinds of accidents many researches proposed many algorithms to detect the distraction of drivers and alert the drivers using machine learning models. Predominantly, the camera captures the image of the driver or some researchers taken the data from dataset but nowadays the data collects with the help of smartphones in non – intrusive approach[34]. Based on the data collected from the smartphones. But it is challenging one to identify that the phone is either used by the passenger or the driver [19]. To overcome the issue, GPS, gyroscope sensors of the smartphone, accelerometer and magnetometer are used to gather the data. The collected dataset undergoes to the training, validation, testing. 10 folds cross validation is performed for training and validation and the model is tested using Gradient Boosting Machine. A Built in application is used in smartphone, when a driver or a passenger needs to use the phone, the sensor sends the message to the phone. If a driver receives the message and driver starts to use the phone there should be a change in the driving based on the data receives from accelerometer and magnetometer, the phone will be locked to
avoid the distraction of driver\cite{22} \cite{24}. This paper produces the efficient analysis and performs better in identifying the driver and passenger.

3. Materials and methods

This segment discusses the procedure of ResNeXt-101 in detailed together with the comprehensive explanation of dataset and algorithms.

3.1 Dataset Description

To ascertain the driver distraction, the initial step is analyze the dataset. The dataset holds 12977 training images, 4331 test images and the images are retrieve from videos in the form of frames are given in table 1. The preparation of dataset have been done with the help of 4 cars and 31 subjects in different form of environment \cite{34}. The images in the dataset classified into 10 unique classes for instance c0: safe driving, c1: texting-right, c2: talking on the phone-right, c3: texting-left, c4: talking on the phone left, c5: operating the radio, c6: drinking, c7: reached behind, c8: hair and makeup, c9: talking to a passenger.

| Parameters               | Details                        |
|--------------------------|-------------------------------|
| Source                   | Distracted Driver Dataset      |
| Data Type                | 2D Images                     |
| Training Test Data       | 12977                         |
| Testing Test Data        | 4331                          |
| Number of Classes        | 10                            |
| Number of Channels       | 3 (RGB Image)                 |

3.2 Residual Neural Network (ResNet)

Microsoft Research introduced a new architecture called Residual Network which was proposed in the year of 2015. Residual Neural Network is a type of convolutional neural network for performing the image recognition.

Complexity of the function gets increased when adding the more number of layers, at that same time the structure of the network is eloquent. In an effort to train the model by grasping the input features from complex image, neural network needs several layers. Moreover there is a vanishing gradient problem \cite{vanishing paper} occurs when the number of layers increased more than 30 layers in VGG Net. The solution for the vanishing gradient problem offered by the ResNet architecture by skipping the two or more layers and by skipping the layer the output wouldn’t have major changes.

3.2.1 Structure of ResNet

A building block of ResNet is known as Residual Block. Each residual block holds 3*3 convolutional layer, batch normalization and follow up with ReLU activate function and this residual block proceed again with 3*3 convolutional layer and batch normalization. The entire block considered as 1*1 convolutional layer and performs activation function.
In the above ‘figure 1’, the two 3*3 convolutional operations are skipped and the input x is directly perform the addition operation with ReLU. The output of two 3*3 convolutional layer may produce the same input, so instead of performing the operations in the residual block, 3*3 is changed to 1*1 convolutional layer and added before to ReLU function as R(x)= f(x)+x, where R(x) is residual layer.

3.2.2 ResNet

ResNet holds several residual blocks. The architecture of ResNet has 7×7 convolutional layer, 64 output channels with a stride of 2 and it is followed by the 3×3 maximum pooling layer. ResNet uses 4 modules and each modules is formed by residual blocks. The number of channels in input layer is same as the number of channels in the first module. ResNet performs best with the lowest error rate 5 percent when comparing with the other architectures of CNN.

3.3 ResNeXt-101

ResNeXt refers to the next dimension of ResNet. Such name of the next dimension called as “Cardinality” dimension. The dimension of Cardinality referred as “C” and it refers to the total number of paths present in the ResNeXt block and it has responsibility to controls the number of complex transformations. 101 describes the number of layers used in this algorithm. It is sometime called as ‘Network in Neuron’ instead of Network in Network. ResNet gradually increases the dimension from 4 to 128 and from 128 to 256 but ResNeXt 101 requires the negligible additional effort to create the path. Every path has same configuration. ResNeXt decreases the validation error due to high cardinality. The neurons in the path would not connected to the other neurons located at some other paths. In the below architecture, 3 convolutional operations is performed at each path i.e., 1*1 Convolutional, 3*3 Convolutional and 1*1 Convolutional. The interior dimension of each path is represented by the symbol ‘d’ (here d=4) and Cardinality (C) of the path is 32. When we add the dimension of each 3*3 Convolutional, then the dimension will be 128 which means d*C, but the dimension of ResNeXt is directly increased from 4 to 256. However ResNeXt-101 achieves 21.2% of top -1 error while ResNet 101
achieves only 22.0% top-1 error rate. This result shows the cardinality is vitally important to improvise the accuracy of image classification.

Figure 2. Architecture of ResNeXt

Let us take the image of a driver who is texting on phone. The system train the model with the images and stored in the database. The principle of the ResNeXt 101 shown in 'figure 2' is to identify the actions of driver and capture it when the driver distracted from driving. The system continuously monitoring the actions of the driver, it sends the input to the layers. First, it extracts the features from the image and perform the above composite operations such as convolution, batch normalization and ReLU. This operations performs repeatedly to get the output image that approximately matches with the image. Once the image gets matched it capture the particular action of the driver.

Algorithm: ResNeXt 101 Model

| DefResNeXt101 (image_shape, depth, cardinality, width) |
|-------------------------------------------------------|
| /* Layers =101 */ |
| image_shape = (64, 64, 3) |
| class=6 |
| depth=4 |
| width=4 |
| Y_image=Input(image_shape) |
| Y=ZeroPadding((3, 3))(Y_image) |
| /* Phase - 1 */ |
| Y=Conv(64,(7,7), stride(2,2), pro_name= 'convo1' kernel_init = gl_uni(seed=0))(Y) |
| Y= Batch_Normalization(row= 3, pro_name = 'B_convo1')(Y) |
| Y = Activation('ReLU')(Y) |
| Y= Max_Pooling((3, 3), stride=(2, 2))(Y) |
| /* Phase - 2 */ |
| Y = conv_block(Y, f_s= 3, fil = [64, 64, 256],C=32, stage_process= 2, block='a1', p=1) |
| Y= iden_block(Y, 3, [64, 64, 256], stage_process=2, block='b1') |
| Y= iden_block(Y, 3, [64, 64, 256], stage_process=2, block='c1') |
| /* Phase - 3 */ |
| Y = conv_block(Y, f_s= 3, fil = [128, 128, 512],C=32, stage_process= 3, block='a1', p=2) |
Y = iden_block(Y, 3, [128, 128, 512], stage_process=3, block='b1')
Y = iden_block(Y, 3, [128, 128, 512], stage_process=3, block='c1')
Y = iden_block(Y, 3, [128, 128, 512], stage_process=3, block='d1')

/* Phase - 4 */
Y = conv_block(Y, f_s= 3, fil = [512, 512, 2048],C=32, stage_process= 4, block='a1', p=2)
Y = iden_block(Y, 3, [256, 256, 1024], stage_process=4, block='b1')
Y = iden_block(Y, 3, [256, 256, 1024], stage_process=4, block='c1')
Y = iden_block(Y, 3, [256, 256, 1024], stage_process=4, block='d1')
Y = iden_block(Y, 3, [256, 256, 1024], stage_process=4, block='e1')
Y = iden_block(Y, 3, [256, 256, 1024], stage_process=4, block='f1')

/* Phase - 5 */
Y = conv_block(Y, f_s= 3, fil = [512, 512, 2048],C=32, stage_process= 5, block='a1', p=2)
Y = iden_block(Y, 3, [512, 512, 2048], stage_process=5, block='b1')
Y = iden_block(Y, 3, [512, 512, 2048], stage_process=5, block='c1')
Y = Average_Pooling()(Y)
/* Output_Layer */
Y = Flatten()(Y)

/* To reduces the input to the number of classes by applying softmax activation function use Fully Connected Layer (FCL). */
Y= FCL(classes, activation='softmax', pro_name='fc' + str(classes), kernel_init=gl_uni(seed=0))(Y)
/* output image */
Out_image= Output_Image(image= Y_image, output = Y, pro_name='ResNeXt101')
return Out_image

4. Results and Discussion

4.1 System Requirement

In this section, the proposed model is going to be evaluated using some standard quality metrics to measure its performance and compares it with the existing works to prove its significance. Graphical Processor Unit (GPU) and system implemented with NVIDIA GeForce GTX 950M 4GB Graphic processor is the minimum hardware required to run this model without hassle.

4.2 Results

The distracted driver dataset were employed to train and test the effective performance of the model and 5-fold cross validation technique is used to evaluate the model. The dataset needs to complete 10 epoch cycles to train and validate the model which is fixed manually to standardize the system to calculate the accuracy and loss scores of the model during the cycle. The number of cycles of the training dataset the model takes to train the system effectively is referred to as epochs. For every increasing epoch, the model progressively gives better performance and attains its best at 10th epoch. At 1st epoch, the model fetches comparatively low accuracy with maximum loss, but at 10th epoch, it concludes with high accuracy and having minimal loss. Error on the training dataset is said to be Training loss. Validation loss is calculated after running the validation dataset through already trained network. As epoch increases, both loses will tend to drop thus increases the system accuracy are shown in table 2.
Table 2. Experimental outcome of the proposed model at various epoch levels in %.

| Quality Metrics | 1st Epoch | 5th Epoch | 10th Epoch |
|-----------------|-----------|-----------|------------|
| Training Loss   | 0.15      | 0.12      | 0.06       |
| Validation Loss | 0.19      | 0.1       | 0.005      |
| Accuracy        | 78.48     | 89.37     | 98.1       |

4.3 Comparison and Discussion

This model's performance will not only be determined by its sole results but by comparing it with the scores of other existing deep learning methodologies with same dataset values will distinctly manifest the importance of the proposed work. For that, some of the existing image detection methodologies have been employed and the comparison analysis table has given below. The below table 3 represent the distracted driver detection capability of various models with their accuracy rate at first and last epoch in percentage.

Table 3. Comparative performance analysis of the existing models with the proposed ResNeXt 101 in %.

| Model         | Accuracy @ 1 | Accuracy @ 10 | Training Loss | Validation Loss |
|---------------|--------------|----------------|---------------|-----------------|
| VGG 19        | 70.89        | 89.69          | 0.22          | 0.36            |
| DENSENET 121  | 74.67        | 93.04          | 0.15          | 0.19            |
| RESNET 101    | 76.58        | 94.1           | 0.14          | 0.17            |
| INCEPTION V3  | 77.55        | 95.42          | 0.12          | 0.1             |
| Reg-DenseNet  | 78.09        | 98.06          | 0.07          | 0.005           |
| RESNEXT 101   | 78.48        | 98.1           | 0.06          | 0.005           |

Figure 3. Comparison of Training and Validation Loss of Various models
Figure 4. Comparison of Accuracy of Various models

From the above illustrations and comparisons of various models, the evaluation results were stated in the form of tables and graphs as shown in 'figure 3' and 'figure 4' it clearly indicates that the proposed method detects the distracted driver gestures and captures it as image with 98.1% precision with low training and validation loss of 0.06% & 0.005% respectively and is better than the compared models thus proved the significance of the work. Also the proposed system performs a bit better than the previous model (Reg-DenseNet) and proved the importance of this work by comparatively gives better test results with a slight higher precision. The efficacy of the deep learning models is again exhibited in this experimental study, which is not only limited to image classification, also in bioscience [30-33], self-learning models, real-time monitoring system, etc.

5. Conclusion

The proposed framework classifies the distracted driver with the help of the obtained dataset to train the model and effectively detects the distracted driver by capturing the moments with 97.6% classification accuracy and stores it in the database in the form of image. The performance of the model fetches better score as per the quality metrics examinations and proved its significance and importance among the other existing techniques with comparatively more efficiency. In future, to improve the robustness of the model, the driver facial expression detection can also be used combined and implemented to monitor the eye movement, dizziness and consciousness of the driver to furthermore improve the system and alert the driver or automate the vehicle to act accordingly to take control of the vehicle over the driver in abnormal situations to minimize the occurrence rate of accidents drastically. As a next step, in order to enhance the proposed work and to complete the entire system, an alert system needs to be built further along with the proposed model to alert the driver in real time whenever the driver gets distracted. This will instantly help the driver to get back to proper stance and gain the control over driving, thus extensively reduces the chance of occurrence of road accidents caused by four wheelers and also lower the probability of fatality rate.
References

[1] Malin Lundgren, Lars Hammarstrand, and Tomas McKelvey 2016 Driver-gaze zone estimation using bayesian filtering and gaussian process IEEE Transactions on intelligent transportation systems.

[2] Mohamad-Hoseyn Sigari, Mahmood Fathy and Mohsen Soryani 2013 A driver face monitoring system for fatigue and distraction detection International Journal of Vehicular Technology.

[3] Ankita S, Kulkarni Sagar, Shinde B 2017 A review paper on monitoring driver distraction in real time using computer vision system International Conference on Electrical, Instrumentation and Communication Engineering.

[4] Mohamad Hoseyn Sigari, Muhammad-Reza Pourshahabi, Mohsen Soryani and Mahmood Fathy 2014 A review on driver face monitoring systems for fatigue and distraction detection.

[5] Dong Y, Hu Z, Uchimura K and Murayama N 2011 Driver inattention monitoring system for intelligent vehicle: a review IEEE Transactions on Intelligent Transportation Systems.

[6] Ersal T, Fuller H J A, Tsimhoni O, Stein J L and Pathy H K 2010 Model-based analysis and classification of driver distraction under secondary tasks IEEE Transactions on Intelligent Transportation.

[7] Liang Yulan, Michelle L, Reyes and John D Lee 2007 Real-time detection of driver cognitive distraction using support vector machines IEEE Transactions on Intelligent Transportation Systems.

[8] Seongsu Im, Cheolha Lee, Seokyoul Yang, Jinhak Kim and Byungyong You Driver distraction detection by in-vehicle signal processing.

[9] Simonyan K, Zisserman A 2014 Very deep convolutional networks for largescale image recognition.

[10] Krizhevsky A, Sutskever I, Hinton G E 2012 ImageNet classification with deep convolutional neural networks.

[11] Szegedy C, Liu W, Jia Y et al. 2015 Going deeper with convolutions Proc. IEEE Conf. on computer vision and pattern recognition, Boston, Massachusetts, USA, pp 1–9.

[12] He K, Zhang X, Ren S, et al. 2016 Deep residual learning for image recognition Proc. IEEE Conf. on computer vision and pattern recognition, Las Vegas, Nevada, USA, pp 770–78.

[13] Carnetsoft Inc. Research driving simulator.

[14] Weihua Sheng, Duy Tran, Ha Do, He Bai 2018 Real-time detection of distracted driving based on deep learning.

[15] Center for disease control and prevention.https://www.cdc.gov/motorvehiclesafety/distracteddriving/.

[16] Sultana H Parveen, Shrivastava, Nirvishi, Dominic, Dhanapal Durai, Nalini N and Balaje J M 2019 Comparison of machine learning algorithms to build optimized network intrusion detection system, Journal of Computational and Theoretical Nanoscience 16(5-6), pp 2541-49.

[17] Yan Y, Coenen F and Zhang B 2016 Driving posture recognition by convolutional neural networks.

[18] Basheer Shakila, Bivi S, Mariyam Aysha, Jayakumar S, Rathore Arpit, Jeyakumar Balaje 2019 Machine learning based classification of cervical cancer using k-nearest neighbour, random forest and multilayer perceptron algorithms Journal of Computational and Theoretical Nanoscience 16(5-6), pp 2523-27.

[19] Liu X, Cao J, Tang S, He Z and Wen J 2017 Drive now, text later: nonintrusive texting-while-driving detection using smartphones IEEE Trans. Mob. Computing.

[20] Lipovac K, Derić M, Tešić M, Andrić Z and Marić B 2017 Mobile phone use while driving-literary review Transp. Res. Part F Traffic Psychol. Behav.

[21] Eraqi H M, Abouelnaga Y, Saad M H and Moustafa M N 2019 Driver distraction identification with an ensemble of convolutional neural networks.

[22] Ye M, Osman O A, Ishak S and Hashemi B 2017 Detection of driver engagement in secondary tasks from observed naturalistic driving behavior.

[23] Liu T, Yang Y, Huang G B, Yeo Y K and Lin Z 2016 Driver distraction detection using semi-supervised machine learning.

[24] Renato Torres, Orlando Ohashi and Gustavo Pessin 2019 A machine-learning approach to distinguish passengers and drivers reading while driving.
[25] Munif Alotaibi and Bandar Alotaibi  Distracted driver classification using deep learning.

[26] Abdel-Hamid O, Mohamed A R, Jiang H and Penn G 2012 Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition In Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference IEEE. pp 4277-80.

[27] Abdel-Hamid O, Mohamed A R, Jiang H, Deng L, Penn G and Yu D 2014 Convolutional neural networks for speech recognition IEEE/ACM Transactions on audio, speech, and language processing 22(10), pp 1533-45.

[28] Jeyakumar, Balajee, MA Saleem Durai and Daphne Lopez  2018 Case studies in amalgamation of deep learning and big data In HCI Challenges and Privacy Preservation in Big Data Security pp 159-74.

[29] Garcia-Garcia A, Orts-Escolano S, Oprea S, Villena-Martinez V and Garcia-Rodriguez J 2017 A review on deep learning techniques applied to semantic segmentation arXiv preprint :1704.06857.

[30] Karthik S and Sudha M 2020 Predicting bipolar disorder and schizophrenia based on non-overlapping genetic phenotypes using deep neural network Evolutionary Intelligence pp 1-16.

[31] Sekaran K and Sudha M 2020 Predicting drug responsiveness with deep learning from the effects on gene expression of obsessive–compulsive disorder affected cases Computer Communications 151, pp 386-94.

[32] Sekaran K and Sudha M Predicting autism spectrum disorder from associative genetic markers of phenotypic groups using machine learning Journal of Ambient Intelligence and Humanized Computing pp 1-14.

[33] Karthikeyan T, Sekaran K, Ranjith D, Vinoth Kumar V and Balajee J M 2019 personalized content extraction and text classification using effective web scraping techniques International Journal of Web Portals 11(2), pp 41–52.

[34] https://www.kaggle.com/c/state-farm-distracted-driver-detection.