iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** II **Month of publication:** February 2023

DOI: https://doi.org/10.22214/ijraset.2023.49187

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue II Feb 2023- Available at www.ijraset.com

Deep Learning Approaches to Detect Distracted Driving Activities: A Survey

K Sudheevar Reddy[1], Kiran N Deshmukh[2], Lokesh M[3], Prathik Raj RC[4], Dr. Kavitha K S[5]
[1, 2, 3, 4, 5]*Department of Computer Science, Dayananda Sagar College of Engineering,* Bengaluru, India

Abstract: ***With increase in the use of advanced systems on vehicle dashboards such as modern infotainment systems and mobile phones, we have come to observe a proportional increase in the distraction of drivers. Distracted driving has proven to be a vital cause for most modern-day road crashes due to lane variation, and inattention on the road. Hence, efficient mechanisms to tackle this problem have to be studied and developed. So far, many attempts have been made on the identification of distracted driving in the form of computer vision techniques and ML/DL approaches. This research effort suggests numerous approaches to assess driver behaviour and identify alleged distractions, thereby offering robust solutions to reduce distracted driving-related traffic accidents.***

Keywords: Deep learning, distracted driving, driver behaviour analysis, road safety, transfer learning

I. INTRODUCTION

Road accidents are a serious public safety and health hazard and one of the most common sources of premature deaths among people of all age groups, most prevalently among 15-29 year olds.

Road accidents claim the lives of 1.2 million drivers and occupants each year and render another 50 million permanently disabled, according to the WHO. Road accidents have risen to the tenth-leading source of fatalities globally over the past ten years, and by 2030, they are expected to move up to the fifth spot.

Coincidentally, India forms the largest portion of the worldwide fatality rates related to traffic accidents. Driver distraction, in its simplest form, is the act of operating an automobile while engrossed in another activity, most frequently using a smart phone or other devices, such as the stereo while driving or other related operations.

47% of respondents to a countrywide poll organised by TNS India Pvt Ltd reported taking calls while driving. The poll also infers that the dangers of using a phone while driving are viewed negatively by 94% of respondents. When a motorist is on the phone, 96% of occupants feel insecure, while 41% of drivers admitted to using their handsets for work or business related purposes. Furthermore, 60% of individuals are unwilling to pull over in a secure place and answer their phones, and 20% people have had close calls while operating on their cell phones. This goes to show that the advent of smartphones has worsened the already existing issue of drivers getting distracted and has led to the compromise of passenger safety as well. Moreover, distracted driving would go against the law in many jurisdictions and is subject to fines, licence revocation, and other sanctions.

Distraction while driving can be broadly placed in three categories,

1) *Visual:* It is a type of distraction where motorists appear to be disoriented from the primary task of driving because there is conspicuous information away from their line of sight, such as the usage of multimedia devices.
2) *Manual:* Releasing the steering wheel while engaging with something irrelevant to the driving task.
3) *Cognitive:* This refers to being absent-minded or preoccupied with other tasks like communing or daydreaming.

The above mentioned classifications can lead to lane variations, lack of control over the vehicle, slower responses to hazards which, in turn, contribute to the ever-increasing incidents of road crashes and accidents. Hence, it is important for drivers to avoid all forms of distraction while driving in order to maintain their focus and stay safe on the road.

A. Motivation

Given the increased effort and focus on unimportant things when driving while distracted, the driver's consciousness and judgement capabilities may be diminished. Developing a robust system to perform this task is beneficial to multiple parties, i.e. drivers and passengers in terms of their safety and insurance companies can adapt premiums according to the driver behaviour statistics collected.

©IJRASET: All Rights are Reserved

Coming up with a solid mechanism to detect distracted driving could prove to be highly advantageous.

1) Distraction detection can be coupled with ADAS in modern-day cars that could trigger features like CAS to manoeuvre the car around obstacles.
2) Concerned authorities can be alerted when the codes of conduct are breached by the driver.
3) Although self-driving cars or semi-self driven cars make the drivers' jobs easier, they still require them to be paying attention and ready to manually control the steering wheel in times of emergency. This is one of the reasons why a distraction detection system could prove effective.
4) Promotes a safe environment in driving conditions by mitigating the number of road accidents.

B. *Contribution*

This study aims to analyse various approaches to detecting and classify distracted driving activities. Our contributions through this research work include

1) A survey of multiple research works that discuss various algorithms and approaches to tackle the problem at hand and a brief overview of them.
2) An analysis of multiple approaches in terms of their performance and a summary of the best works.

II. LITERATURE SURVEY

Distracted driver classification primarily incorporates two classical approaches. The first approach involves using wearable sensors that measure brain signals, heartbeat data and data from muscular activities. However, this approach could sometimes prove not to be cost effective or complex hardware with high involvement from users. The other approach involves using camera vision systems to detect and classify the type of distraction. They commonly make use of deep learning techniques to perform the feature extraction and classification tasks. While the first approach can detect cognitive distractions, the second one can detect manual and visual distractions. Hence, it makes sense that the two methods could be used in tandem to ensure efficient detection. In this paper, we discuss various methodologies that involve both the above mentioned approaches vividly.

A. Ezzouhri et al. [1] propose using a segmentation module before performing the classification task to reduce noise in the image eg. background that could contribute to misclassification. In order to do this, they use a human body part segmentation technique called Cross Domain Complementary Learning (CDCL) on raw RGB images obtained from an onboard camera. This segmentation module outputs body part maps of the driver. The authors have designed their own dataset called the Driver Distraction Dataset to train the classification module that makes use of the VGG-19 network that is by itself, trained on the ImageNet dataset. The procedure followed to train the network is transfer learning, by unfreezing some multiperceptron layers for fine-tuning to match the dataset at hand. The authors have also experimented with the Inception V3 network and found it to output similar accuracy scores. Experimentation shows that segmenting the images appreciably boosts the accuracy rates of classification. The authors were able to obtain accuracy rates of 96% and 95% on their dataset and the AUC dataset respectively.

Shokoufeh, et al. [2] used driving data to identify driver distraction using a stacked LSTM network enhanced with an attention layer. To demonstrate the benefit of using an attention mechanism on the model's performance, they contrasted this model with stacked LSTM and MLP models. They tested eight different driving situations. Here is a quick summary of the steps involved: First, the original dataset was divided into training and test datasets, and an MLP model was employed to detect distracted driving (80:20 split). The system's intelligence was then increased using an LSTM network. The LSTM network was chosen because it uses driving data to recall the driver's prior behaviours to forecast the driver's behaviour, rather than simply using the driver's instantaneous actions. Additionally, LSTMs have the benefit of having both long-term and short-term memory. They also have the ability to forget irrelevant information that does not advance the system's intelligence. These LSTM characteristics contribute to a significant performance improvement over traditional RNN models. Third, by including an attention layer in the LSTM network, the model's performance has increased. The attention layer enables the model to take into account the weighted impact of each input sequence step on the output. It was discovered that the LSTM's train and test errors were 0.57 and 0.9, respectively. Over the MLP model, this demonstrated a substantial improvement. The train and test errors were lowered to 0.69 and 0.75 by the attention layer.

A new algorithm for identifying manual driver distraction has been put forth by L. Li et al. [3]. The method was broken into two components by the researchers. The driver's right hand and right ear boundary boxes are simultaneously anticipated and found by the first module utilising RGB images taken by the driver's onboard camera. Using the bounding box as input, the second module guesses the type of distraction. In the first module, YOLO, a DNN for object detection, was utilised. Right ear and right hand detection are accomplished simultaneously with image input.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

A multilayer perceptron in the second module uses regions of interest (ROIs) as input to forecast the type of distraction. The two modules' outputs and their overlap are taken into account in this process, which recommends the driver's condition and hence forecasts the type of distraction. The findings demonstrate that the suggested method is capable of identifying common driving situations, touch screen usage, and phone calls with F1 values of 0.84, 0.69, and 0.82, respectively. The overall F1 score for distraction detection was 0.74.

A comprehensive strategy that relies on PERCLOS has been described by V. Uma Maheshwari, et al. [4]. Similar to EAR (Eye Aspect Ratio) and MAR (Mouth Aspect Ratio), the authors have also suggested the calculation of a new vector, FAR, which stands for Facial Aspect Ratio. This vector aids in determining the status of an opened mouth, such as when yawning, closed eyes, and other frames that suggest hand movements or gestures, such as covering the mouth with the hand. The system also has ways for locating drivers' faces from different angles, for recognising sunglasses on a driver's face, and for situations in which hands might cover the driver's mouth or eyes. The following summarises the suggested methodology: The system receives periodic inputs of images of the driver taken by a camera. Additionally, the image with the produced ROI will be cropped to remove the recognised face. The driver's eyes will then be identified from the ROI, and the CNN algorithm will use this information to categorise the driver's level of tiredness. The model was created by the researchers and evaluated on various datasets, including NTHU-DDD and YawDD, using Keras and CNN. Additionally, they have put out their own dataset, dubbed EMOCDS, which compiles every conceivable instance of sleepiness. With this model and dataset, the suggested method generated an overall accuracy of 95.67%.

Similar to the above, Mohammad-Hoseyn Sigari, et al. [5] have developed a novel method for identifying driver distraction and weariness based on signs in the face and eye areas. The first level of processing involves image capture and face identification. The system is made to adaptively extract the characteristics of the driver's state. Then, the face image is used to derive hypovigilance signs. The authors identified the driver's head rotation using a pattern matching technique. The Fuzzy Expert System is used to calculate the driver's level of exhaustion. The authors of this study employed an adaptive boosting approach and Haar-like characteristics to detect faces. Face detection was followed by the extraction of symptoms from the eye region, including PERCLOS, ELDC, and CLOSNO, and the face region, including head rotation (ROT). A fuzzy expert system that employs fuzzy logic processes these collected features to identify driver tiredness. A shorter amount of time was spent training the fuzzy expert system using average values of extracted symptoms. The level of fatigue and distraction is estimated by this system after training. Using this method, the authors were able to develop a reliable and dynamic driver eye/face monitoring system.

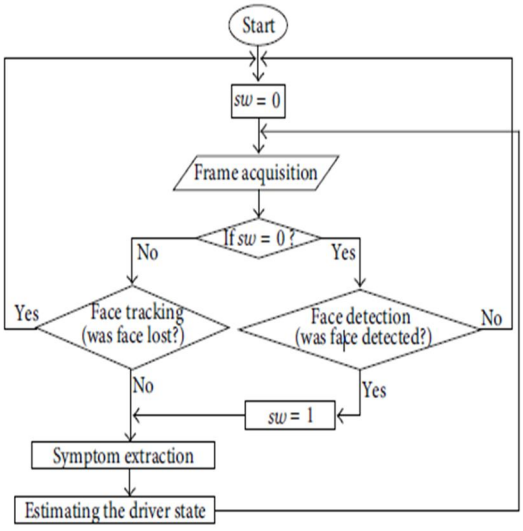


Fig. 1. Flowchart of the system [5]

Jing Wang, et al. [6] have discussed a framework that uses a data augmentation approach to artificially expand the dataset as an attempt to enhance the model performance and work with a smaller dataset. The flow of events in the proposed framework is given as follows – The first step involves the analysis and identification of driving operation key areas using the class activation mapping method. The next step incorporated the augmentation of the dataset using the faster R-CNN model to generate a new sub-dataset called the driving operation area (DOA) dataset constructed using the AUC dataset.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

The R-CNN model was found to give the best accuracy of 0.6271 among the other two models, YOLO and SSD. In the third step, a classification model was designed to process data from both the AUC and the newly formed DOA datasets. The researchers have experimented with multiple classification models such as the InceptionV4, Xception and AlexNet to compare the results and choose the best model. Lastly, the authors have tested these trained models against their own dataset to analyse the classification accuracy. The results show that the classification accuracy was as high as 96.97% with the Xception model as a result of using the data augmentation approach.

The necessity for an optimised framework that uses few parameters, has great accuracy, and runs quickly has been addressed by Qin et al. [7]. The D-HCNN model, which has been proposed by the authors, is a new model that incorporates a decreasing filter size. The needs outlined above are precisely met by this model. In comparison to other models, it only utilises 0.76M parameters, a much smaller quantity. The D-HCNN model includes dropout and batch normalisation in addition to L2 weight regularisation to prevent the overfitting issue, HOG feature pictures to preserve only the driver's posture and remove background noise. All the above elements help increase the performance of the model. The suggested network starts out with a higher convolution filter size to accommodate a wider receptive field. As the network gets deeper, the receptive field gets narrower and the convolution filter gets smaller to operate with fewer network parameters, which allows for finer tuning and faster processing times. Accuracy ratings of 95.59% and 99.87% were obtained by experimental evaluations on the two datasets AUCD2 and SFD3.

Ben Ahmed, et al. [8] propose utilising the capabilities of contemporary smartphone sensors like accelerometers and gyroscopes to identify driving distractions like texting, reading, and talking on the phone. The experiment was carried out employing a driving simulator that included a steering wheel with A,B,C pedals, a broad screen to imitate background vehicular traffic, and features to replicate different weather conditions, such as daytime, nighttime, fog, and rainy circumstances. The participants engaged in a variety of tasks while operating the simulator while simultaneously logging data using the smartphone's sensors. Utilizing the Random Forest Classifier and the 10-fold cross-validation procedure, a machine learning approach was used to solve the classification problem using the sensory data. When the data from both sensors were pooled, the experiment's outcomes were 85% Precision, 84% Recall, and 87% for F-Measure.

A semi-supervised learning strategy to identify distraction was investigated by Liu et al. [9]. The goal of this research project was to reduce the expense of labelling training data. The authors propose a method for identifying cognitive distractions that takes advantage of eye and hand motions. For this, they suggested a methodology that makes use of a semi-supervised extreme learning machine and the Laplacian SVM. Also tested in comparison to the preceding method were Static Bayesian Network Supervised Clustering (SBN-SC), SVM, Transductive SVM (TSVM), and ELM. They divided the dataset for each subject into three subsets (labelled, unlabelled, and test set) for the purposes of this comparison. The supervised learning models were trained using only the labelled data, while the semi-supervised learning models were trained using both labelled and unlabelled data. Performance evaluations are displayed in Fig. 3.

A hybrid CNN Framework, referred to as HCF, has been presented by C. Huang et al. [10]. HCF capitalises on using three modules to considerably improve the detection accuracy. ResNet50, Inception V3, and Xception are three pretrained models that were combined to create the first module, a cooperative CNN module. These models were trained using the MTRO algorithm, a transfer learning-based approach. Concatenating these characteristics is required in order to gain crucial information because the features retrieved by these three models are independent. In order to combine the extracted features into a one-dimensional vector that could aid in further classification, the authors suggest a concatenation module. This concatenated feature vector serves as the foundation for categorization in the fully connected layers of the HCF neurons. The experiment's findings demonstrate that the suggested HCF architecture is significantly more accurate than utilising a single pretrained network. When used to identify distracted driving behaviour, the model has demonstrated accuracy of 96.74%.

PERFORMANCE COMPARISON OF SSL METHODS AND SUPERVISED LEARNING METHODS

Measures	SBN-SC	SVM	TSVM	LapSVM	ELM	SS-ELM
Accuracy (%)	83.8 ± 8.5	95.0 ± 4.2	95.5 ± 4.0	97.3 ± 2.9	95.6 ± 3.9	97.2 ± 3.0
Sensitivity (%)	89.2 ± 9.1	97.3 ± 3.6	97.3 ± 3.1	98.8 ± 2.8	97.1 ± 3.7	98.2 ± 3.2
Specificity (%)	69.6 ± 20.2	88.6 ± 12.0	90.5 ± 8.0	92.8 ± 9.4	91.2 ± 8.9	93.9 ± 7.9
G-mean	0.773 ± 0.144	0.926 ± 0.073	0.938 ± 0.053	0.956 ± 0.054	0.940 ± 0.053	0.959 ± 0.045
Training Time (ms)	291.1	0.5	273.0	227.9	29.7	150.4
Prediction Time (ms)	35.2	0.2	8.7	1.8	4.4	4.5

Fig. 3. [9]

Peng Mao, et al. [11] have discussed a deep learning approach to detect distraction behaviour while driving, i.e. smoking and talking on the phone. The authors have proposed using PCN and DSST algorithms for face detection and tracking. On top of this, the YOLOV3 algorithm, classically used for object detection, has been used to identify objects like cigarettes and mobile phones around the driver's face to thus predict distraction while driving. This simple approach outputs a recall score of 85.4% and the precision scores for cigarettes and mobile phones are 95.7% and 97.3% respectively with a speed of 36fps.

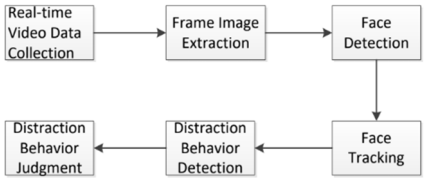


Fig. 2. Flowchart of proposed system [11]

Atran, et al. [12] propose a method based on computer vision to detect cell-phone usage while driving. The experiment apparatus consists of a near infrared camera system that is directed at the front windshield of the car. The method is divided into two modules. The first module is made use of to localise the driver's face region using the deformable part model (DPM) and the second module helps classify a region of interest (ROI) based on a local aggregation model. This is done so to detect mobile phone usage around the face region. The features extracted from the fusion of these modules is provided as input to an SVM Classifier. The proposed model is found to achieve an accuracy score of 86% in detecting mobile phone usage, which hence, plays a vital role in predicting distraction.

A method for once more detecting smartphone activity while driving has been explored by T. Hoang Ngan Le, et al. [13] by determining how many hands are on the wheel while utilising deep learning algorithms. The authors present a MS-FRCNN strategy for ROI pooling that uses RPN generation and feature maps using shallow convolutional feature maps such as conv3 and conv4. The MS-FRCNN suggested in this study aids in the detection of things like the steering wheel, mobile phone, and the motorist's arms, while genetic information is used to assess whether the operator is using his mobile phone or how many hands are on the steering. It has been demonstrated that this model outperformed the regular R-CNN model in tests using two different data sets, VIVA and SHRP-2.

Aksjonov, et al. [14] discussed the methodology for the detection and evaluation of driver distraction while performing secondary tasks like texting on a phone. The suggested approach assesses the impact of driver attention on driving performance while also detecting it. Two models make up the system: one is used for detection and the other is for performance evaluation of the driver. The authors also suggested a model for measuring driver distraction using fuzzy logic. The authors' methods for detecting driver distraction compare normal driving parameters with those obtained when performing secondary tasks.The system experience platform (SEP), which includes a steering wheel, two pedals for the accelerator and brakes, and a virtual world that is displayed on a display in front of the driver, was used to train and test the proposed model. There are three steps in this. In order to forecast a driver's ability to maintain the centerline and the speed limit, the model first includes a predictor that is trained initially without secondary activity for each driver. The performance with secondary activity is contrasted with the projected driving performance. Euclidean distance and a fuzzy logic evaluator that uses linguistic criteria to indicate the percentage of driver distraction are used to calculate the difference between them.

Mohammed S. Majdi, et al. [15] have proposed an automated supervised learning method called Drive-Net for driver distraction detection. For classifying images of a driver, Drive-Net combines a convolutional neural network (CNN) and a random decision forest. The authors adopted the U-Net architecture as a basis of the CNN model because it captures the context around the object in comparison to other architectures like Alex-Net. They modified the U-Net model to suit their application. The authors compared the performance of Drive-Net with two other common machine-learning approaches: a Recurrent Neural Network (RNN) and a Multi-Layer Perceptron (MLP). Two convolution layers, one after the other, are followed by a maxpooling layer and a final ReLU layer in the CNN model.This model was trained by replicating every conceivable driving scenario, including distracted driving scenarios. In order to predict driver distraction, the Random Forest model is fed the output of the CNN model as input. To prevent the CNN model from being overfit, Random Forest is used. The suggested model (Drive-Net) was tested using a publicly available dataset of photos used in the Kaggle competition, and its accuracy (97%) was compared to that of a Recurrent Neural Network (RNN) and a Multi-Layer Perceptron (MLP).

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

An overview of the use of computer vision technologies in the creation of monitoring systems to identify distractions was published by Alberto Fernández, et al. [16]. The main purpose of this paper is detecting driver distraction using vision-based systems. The authors have used various kinds of sensors for face detection, eye tracking sensor, EyeAlert sensor, SensoMotoric instruments used for measuring head position and orientation,Delphi Electronics Driver Status monitor for tracking driver's facial features etc.The authors have also developed models for cognitive distraction. The authors have used different kinds of machine-learning algorithms like SVM, AdaBoost, LDC, DBNs etc. for detecting different kinds of driver distractions. The output of the sensors in the system is fed as an input to machine-learning algorithms to detect drivers' distraction. With this proposal, the authors have developed a system where it can detect various kinds of visual and cognitive distractions accurately.

Zhao et al. [17] introduced a unique system that employs convolutional neural networks to automatically learn and predict pre-defined driving postures. The authors' major goal was to monitor head position and forecast safe/unsafe driving postures. In this approach, the authors built a CNN architecture that seeks to automatically build high-level feature representation from low-level input with limited domain knowledge of the problem and concentrates on characterising the driving postures. The sparse filtering method of unsupervised feature learning, which was used to pre-train the CNN model, was refined using classification. By avoiding explicit modelling of data distribution, sparse filtering maximises the sparsity in the feature distribution.The SEU dataset, which contains images with low lighting and various road conditions, was used to evaluate the proposed model. The results showed that it performed better than traditional approaches that used hand-coded features, with an accuracy of 99.47% on the SEU dataset.

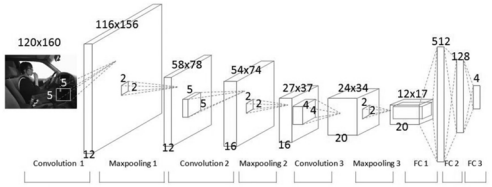


Fig.4. CNN model architecture[17]

According to Ersal et al. [18], earlier studies analysed the average effect of secondary activities on the population of drivers in order to identify general trends. As a result, the authors presented a model to investigate the individual effects of secondary tasks and categorise driving behaviour. The authors have aggregated the model with an SVM to classify distracted and non-distracted driving with very small measurements. The model developed by the author is a Radial-basis Neural Network. The goal of this model is to demonstrate that a model-predictive framework in which normal behaviour of the driver is captured and the model predictions were used as a baseline for predicting driving with secondary tasks.The proposed model captures the nominal behaviour, is trained using the experimental data, validates, and is used to predict hypothetical nominal action of the drivers when they are actually driving with secondary tasks. As a result, it provides a driver specific effects of secondary tasks that uncover effects that are not observed in the average behaviour of the population. The model proposed was successful in gaining insights that were difficult to obtain from analysing the average behaviour of the population. The differences between the predicted nominal and actual distracted actions are analysed using SVM, which shows the high-accuracy classification of driver distraction.

An innovative computational approach for the early identification of driver distraction utilising electroencephalographic (EEG) signals to assess brain activity has been introduced by Wang et al. [19]. Since most research are primarily focused on classifying distracted and non-distracted times, the framework's primary goal is to predict the beginning and conclusion of a driver distraction. Electroencephalographic (EEG) signals were used to record the driver's brain activity while they were requested to drive normally. The authors then created their own prediction model that uses the EEG signal data as an input. A navigation system that can give vocal route instructions in advance has been developed by the author if the model predicts that the driver is inattentive.By doing this, the driver can avoid looking at the map and instead concentrate on the road. The proposed model's overall accuracy for predicting when a map viewing session would begin and end was 81% and 70%, respectively.

For the purpose of identifying the type of driver distraction and detecting it, Céline Craye, et al. [20] have created an effective module using computer vision tools and the Kinect.The proposed model is broken down into four sub-modules: face expression, head orientation, arm position, and eye behaviour. The collected images go through pre-processing, where the background is eliminated, and the image's key elements are extracted. The outputs of these sub-modules are aggregated and fed into the Hidden Markov Model and AdaBoost classifiers, two further classification methods. These classification models were trained using a set of movies that included both typical driving tasks and five distracting tasks.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

HMM was utilised to detect voice recognition tasks and gesture recognition, while AdaBoost was used to significantly identify the retrieved features.Eight drivers, ages ranging from 24 to 40, from various nations, were asked to engage in a driving simulation that included a monitor, steering wheel, and riding software called City Car driving as part of the authors' experiment with the model. The author recorded colour video, a depth map, and audio using a Kinect sensor. Following testing, the authors obtained an overall testing accuracy for AdaBoost of 85.05% and for the HMM model of 84.78%. The proposed approach, according to the authors, is simple to integrate into actual automobile systems.

Tahir Abbas, et al. [21] have proposed a tweaked DNN architecture using Openpose library for the binary classifier problem. Openpose library detects faces and draws 43 points on the face skeleton. The Openpose library is fed to the DNN along with the input images. The authors named the modified neural network model as OpNet-50. This model was obtained by optimising the existing model named ResNet-50. The authors used the SFD3 dataset for training OpNet-50.To test the model, the authors partitioned the dataset into a training and testing dataset. After experimentation, it was proven that the proposed model achieved higher accuracy (98%) and runs 3.44 times faster than the existing ResNet-101 (91%) model at lower running costs.

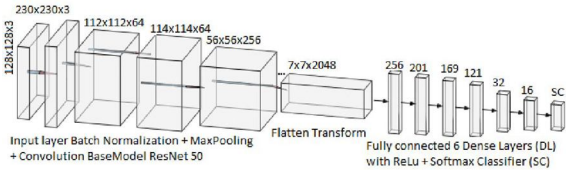


Fig. 5. Proposed Network Diagram[21]

A basic architecture was designed by Peng-Wei Lin, et al. [22] for building a driver alerting that takes into account vehicle condition and driver behaviour. In the projected dataset, the motive force behaviour from the SFD3 dataset and the results of 3D object detection using KITTI's modified KFPN are integrated. They have suggested a system that combines driver status and time to collision (TTC) to generate the protection level call and deliver the necessary warnings to the drivers. A modified CNN model is used for driver perception and driver behaviour detection, alerting the motive force immediately. They have suggested a simple development process for a driver warning system that integrates vehicle state and driver behaviour. They employed MobileNetV3 for recognition to obtain high accuracy, and to boost MobileNetV3's accuracy to roughly 92%, they used the RWSE module to strengthen its channel wise attention module.

Driver assistance technologies like Lane-Departure Warning and Blind Spot Detection systems have been proposed by Sean L. Gallahan et al. [23] as a way to reduce the dangers associated with driving. The Microsoft Kinect works well for full-body 3D motion capture; to identify driving motions, they mainly use skeletal tracking and facial tracking.The algorithm was created by the team in C++, and it was tested by attaching a Kinect to a VDSL simulator. Whenever a driver distraction was identified, an auditory alert sound was played.According to their simulation, they were able to obtain success rates of 100% for drivers reaching for an object, 33% for using a phone while driving, and 66% for looking at an object outside the vehicle.

Using a powerful vision-based system, Hesham M. Eraqi et al. [24] have demonstrated how to identify distracted driving postures. To create and test the system, they used a dataset on distracted driving that was available to the general public. The CNN models employed for detection are the AlexNet network, an InceptionV3 network, a ResNet network with 50 layers, and a VGG-16 network, as skin segmentation is a difficult problem to handle, mostly because of changing lighting conditions while driving. They created a pixel-by-pixel skin segmentation model by employing a Multivariate Gaussian Naive Bayes classifier. They found that in experiments utilising several models, they had a 90% success rate in determining the type of driver distraction.

Duy Tran et al. [25] have presented a system built on the VGG-16, GoogleNet, AlexNet, and ResNet deep learning architectures. Using an aided driving testbed, they were able to create their own dataset on distracted driving. Tests on the assisted driving testbed were conducted in order to evaluate the trained models. The proposed approach outperforms the conventional one. They employed the GoogleNet model, which reaches a frequency of 11 Hz at an accuracy of 89%, to balance the performance in terms of accuracy and frequency. The distraction detection system works in real-time on a Jetson TX1 embedded computer board with an accuracy between 86 and 92 percent and a frequency between 8 and 14 Hz.. The integrated speakers play a sound alert if any distraction is found..

In order to better detect driver distraction, Benjamin Wagner et al. [26] have presented a new method that involves employing two cameras to capture the various driving situations. Using the two cameras, the researchers created their own data set with up to 5 parameters.

The accuracy of three CNN models—VGG-16, VGG-19, and ResNeXt-34—was assessed by the researchers. For each camera, the models were tested individually to identify various driving postures. In the initial phase of their experiment, the ResNeXt-34 model surpassed the two VGG modules with accuracy values for the right camera of 92.88% and the left camera of 90.2%.The researchers achieved the maximum test accuracy of 92.88% for the left camera, 90.36% for the right camera, and an overall accuracy of 92.54% by using the ResNeXt-50 model for the right camera and the ResNeXt-34 model for the left camera in the final stages of development.

Ju-Chin Chen et al. [27] have built a driver behaviour analysis system using a spatial stream ConvNet, which is comparable to the VGG-16 model, to extract the spatial features, a temporal stream ConvNet to record the driver's motion information, and a Fusion network. These recordings of activities from the UCF-101 dataset were used to pre-train this two-dimensional (2D) ConvNet, which can clearly be seen moving its hands. The massive ImageNet served as its pretraining. A feature map with a size of 7x7x256 pixels is taken from the final convolution layer in both ConvNet and processed further in the fusion network.The accuracy rates obtained by the researchers were 49.21% for the Spatial Stream ConvNet and 65.08% for the Temporal Stream ConvNet, and the fusion result was approximately 68.25% correct.

In order to capture the spectral-spatial characteristics of the images, Jimiama Mafeni Mase et al. [28] developed a novel driver distraction posture detection system using layered Bidirectional Long Short-Term Memory (BiLSTM) Networks and CNNs. The CNN-InceptionV3 and CNN-InceptionV3 Plus 1-directional stacked LSTM models were used to train and test the model. Using the AUC distracted driver dataset, which is available to the public, the current setup was evaluated. Because it was able to interpret the spatial and spectral features of the images, the researchers' system outperformed the most sophisticated CNN model with an accuracy rate of 92.7%.

By Matti Kutila et al. [29], a study proposing a tool for tracking driver distraction has been offered. By merging stereo vision and lane tracking data and using rule-based and support-vector machine (SVM) classification approaches, the module determines the visual and cognitive burden of the driver. In particular, the results for identifying cognitive distractions in passenger cars (86%) are encouraging. Less impressive is the truck application result (68%) though. In subsequent experiments, the hit rate increased when the detection of cognitive distraction in a city context was removed because it is much harder to identify and possibly less common in a metropolis where driving demand is higher due to manoeuvring.

III. CONCLUSION

Over the years, there have been many attempts proposed by researchers across the world to detect distracted driving activities. Most of these approaches involve using deep learning techniques or incorporating sensory data with machine learning models. This research work has proposed a review of some of these proposals with factual and numerical data along with a brief on the methodology incorporated in each research work. We have discussed various features that contribute to efficient classification, while also mentioning various datasets that the researchers have used along with the deep learning/classification modules that complement them.

REFERENCES

[1] Ezzouhri, Z. Charouh, M. Ghogho and Z. Guennoun, "Robust Deep Learning-Based Driver Distraction Detection and Classification," in IEEE Access, vol. 9, pp. 168080-168092, 2021, doi: 10.1109/ACCESS.2021.3133797.

[2] S. Monjezi Kouchak and A. Gaffar, "Detecting Driver Behavior Using Stacked Long Short Term Memory Network With Attention Layer," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 6, pp. 3420-3429, June 2021, doi: 10.1109/TITS.2020.2986697.

[3] Li Li, Boxuan Zhong, Clayton Hutmacher, Yulan Liang, William J. Horrey, Xu Xu,Detection of driver manual distraction via image-based hand and ear recognition,Accident Analysis & Prevention,Volume 137,2020,105432,ISSN 0001-4575,
https://doi.org/10.1016/j.aap.2020.105432.
(https://www.sciencedirect.com/science/article/pii/S0001457519309029)

[4] U. Maheswari, R. Aluvalu, M. P. Kantipudi, K. K. Chennam, K. Kotecha and J. R. Saini, "Driver Drowsiness Prediction Based on Multiple Aspects Using Image Processing Techniques," in IEEE Access, vol. 10, pp. 54980-54990, 2022, doi: 10.1109/ACCESS.2022.3176451.

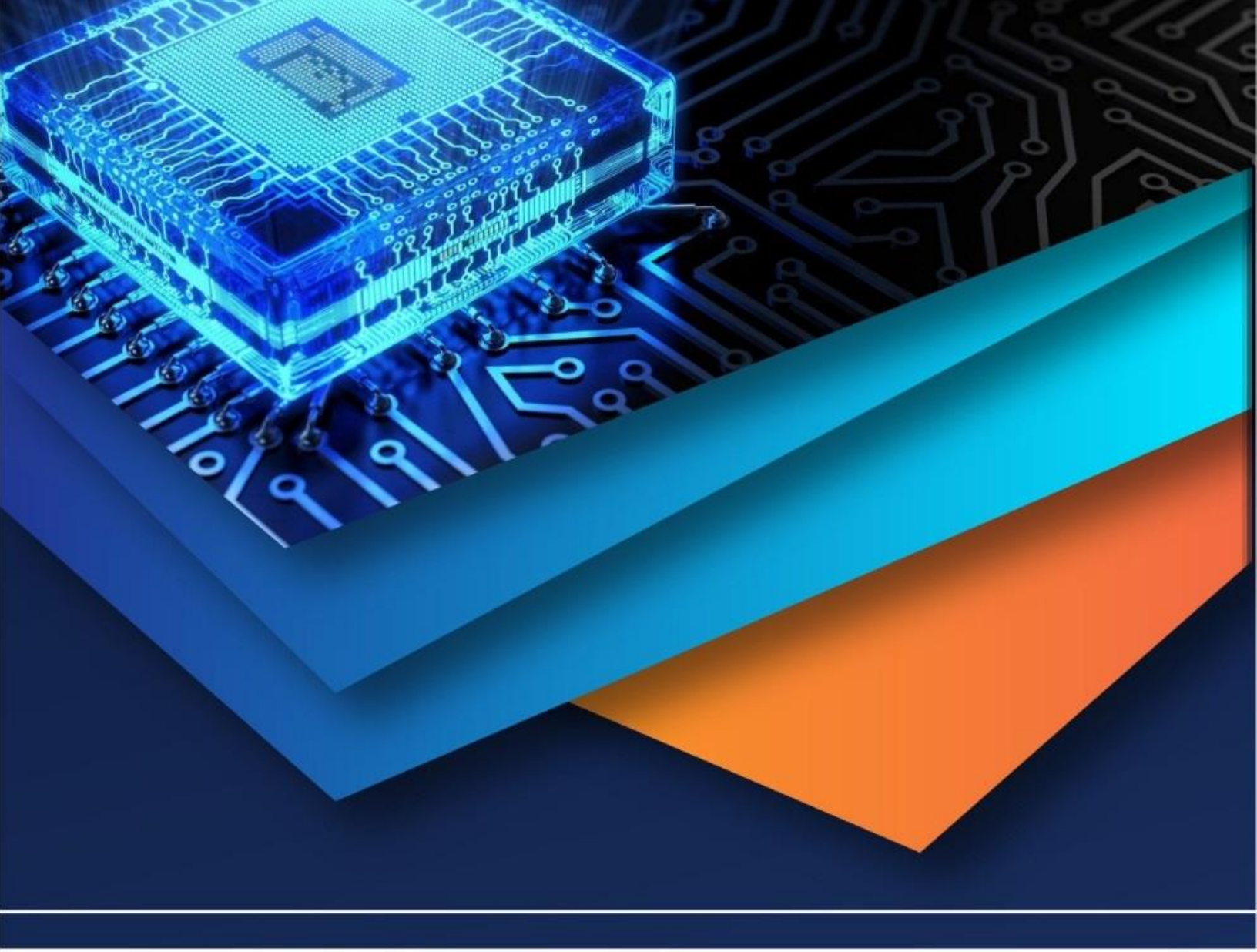
[5] Mohamad-Hoseyn Sigari, Mahmood Fathy, Mohsen Soryani, "A Driver Face Monitoring System for Fatigue and Distraction Detection", International Journal of Vehicular Technology, vol. 2013, Article ID263983,11pages,2013.https://doi.org/10.1155/2013/263983

[6] Wang, Jing, ZhongCheng Wu, Fang Li, and Jun Zhang. 2021. "A Data Augmentation Approach to Distracted Driving Detection" Future Internet 13, no. 1: 1. https://doi.org/10.3390/fi13010001

[7] B. Qin, J. Qian, Y. Xin, B. Liu and Y. Dong, "Distracted Driver Detection Based on a CNN With Decreasing Filter Size," in IEEE Transactions on Intelligent Transportation Systems,vol.23,no.7,pp. 6922-6933,July2022,doi:10.1109/TITS.2021.3063521.

[8] Ben Ahmed, Kaoutar & Goel, Bharti & Bharti, Pratool & Chellappan, Sriram & Bouhorma, Mohammed. (2018). Leveraging Smartphone Sensors to Detect Distracted Driving Activities. IEEE Transactions on Intelligent Transportation Systems. PP. 1-10. 10.1109/TITS.2018.2873972.
[9] T. Liu, Y. Yang, G. -B. Huang, Y. K. Yeo and Z. Lin, "Driver Distraction Detection Using Semi-Supervised Machine Learning," in IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 4, pp. 1108-1120, April 2016, doi: 10.1109/TITS.2015.2496157.
[10] C. Huang, X. Wang, J. Cao, S. Wang and Y. Zhang, "HCF: A Hybrid CNN Framework for Behavior Detection of Distracted Drivers," in IEEE Access, vol. 8, pp. 109335-109349, 2020, doi: 10.1109/ACCESS.2020.3001159.
[11] Peng Mao et al 2020 IOP Conf. Ser.: Mater. Sci. Eng. 782 022012
[12] Y. Artan, O.Bulan, R. P. Loce and P. Paul, "Driver Cell Phone Usage Detection from HOV/HOT NIR Images," 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2014, pp. 225-230, doi: 10.1109/CVPRW.2014.42.
[13] 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), 2016, pp. 46-53, doi: 10.1109/CVPRW.2016.13.
[14] A. Aksjonov, P. Nedoma, V. Vodovozov, E. Petlenkov and M. Herrmann, "Detection and Evaluation of Driver Distraction Using Machine Learning and Fuzzy Logic," in IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 6, pp. 2048-2059, June 2019, doi: 10.1109/TITS.2018.2857222.
[15] M. S. Majdi, S. Ram, J. T. Gill and J. J. Rodríguez, "Drive-Net: Convolutional Network for Driver Distraction Detection," 2018 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), 2018, pp. 1-4, doi: 10.1109/SSIAI.2018.8470309.
[16] Fernández, Alberto, Rubén Usamentiaga, Juan Luis Carús, and Rubén Casado. 2016. "Driver Distraction Using Visual-Based Sensors and Algorithms" Sensors 16, no. 11: 1805. https://doi.org/10.3390/s16111805.
[17] Chao Yan, B. Zhang and F. Coenen, "Driving posture recognition by convolutional neural networks," 2015 11th International Conference on Natural Computation (ICNC), 2015, pp. 680-685, doi: 10.1109/ICNC.2015.7378072.
[18] T. Ersal, H. J. A. Fuller, O. Tsimhoni, J. L. Stein and H. K. Fathy, "Model-Based Analysis and Classification of Driver Distraction Under Secondary Tasks," in IEEE Transactions on Intelligent Transportation Systems, vol. 11, no. 3, pp. 692-701, Sept. 2010, doi: 10.1109/TITS.2010.2049741.
[19] S. Wang, Y. Zhang, C. Wu, F. Darvas and W. A. Chaovalitwongse, "Online Prediction of Driver Distraction Based on Brain Activity Patterns," in IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 1, pp. 136-150, Feb. 2015, doi: 10.1109/TITS.2014.2330979.
[20] arXiv:1502.00250 [cs.CV] (or arXiv:1502.00250v1 [cs.CV] for this version) https://doi.org/10.48550/arXiv.1502.00250
[21] T. Abbas, S. F. Ali, A. Z. Khan and I. Kareem, "optNet-50: An Optimized Residual Neural Network Architecture of Deep Learning for Driver's Distraction," 2020 IEEE 23rd International Multitopic Conference (INMIC), 2020, pp. 1-5, doi: 10.1109/INMIC50486.2020.9318087.
[22] P. -W. Lin and C. -M. Hsu, "Innovative Framework for Distracted-Driving Alert System Based on Deep Learning," in IEEE Access, vol. 10, pp. 77523-77536, 2022, doi: 10.1109/ACCESS.2022.3186674.
[23] Gallahan, S. L. et al. "Detecting and mitigating driver distraction with motion capture technology: Distracted driving warning system." 2013 IEEE Systems and Information Engineering Design Symposium (2013): 76-81.
[24] Hesham M. Eraqi, Yehya Abouelnaga, Mohamed H. Saad, Mohamed N. Moustafa, "Driver Distraction Identification with an Ensemble of Convolutional Neural Networks", Journal of Advanced Transportation, vol. 2019, Article ID 4125865, 12 pages, 2019. https://doi.org/10.1155/2019/4125865
[25] Tran, Duy, et al. "Real-time detection of distracted driving based on deep learning." *IET Intelligent Transport Systems* 12.10 (2018): 1210-1219.
[26] B. Wagner, F. Taffner, S. Karaca and L. Karge, "Vision Based Detection of Driver Cell Phone Usage and Food Consumption," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 5, pp. 4257-4266, May 2022, doi: 10.1109/TITS.2020.3043145.
[27] Chen, Ju-Chin, Chien-Yi Lee, Peng-Yu Huang, and Cheng-Rong Lin. "Driver behavior analysis via two-stream deep convolutional neural network." Applied Sciences 10, no. 6 (2020): 1908.
[28] J. Mafeni Mase, P. Chapman, G. P. Figueredo and M. Torres Torres, "A Hybrid Deep Learning Approach for Driver Distraction Detection," 2020 International Conference on Information and Communication Technology Convergence (ICTC), 2020,pp.1-6,doi: 10.1109/ICTC49870.2020.9289588.
[29] M. Kutila, M. Jokela, G. Markkula and M. R. Rue, "Driver Distraction Detection with a Camera Vision System," 2007 IEEE International Conference on Image Processing, 2007, pp. VI - 201-VI - 204, doi: 10.1109/ICIP.2007.4379556.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)