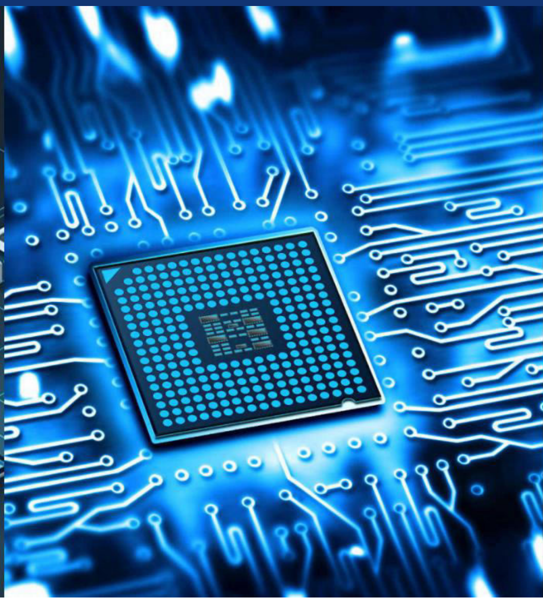
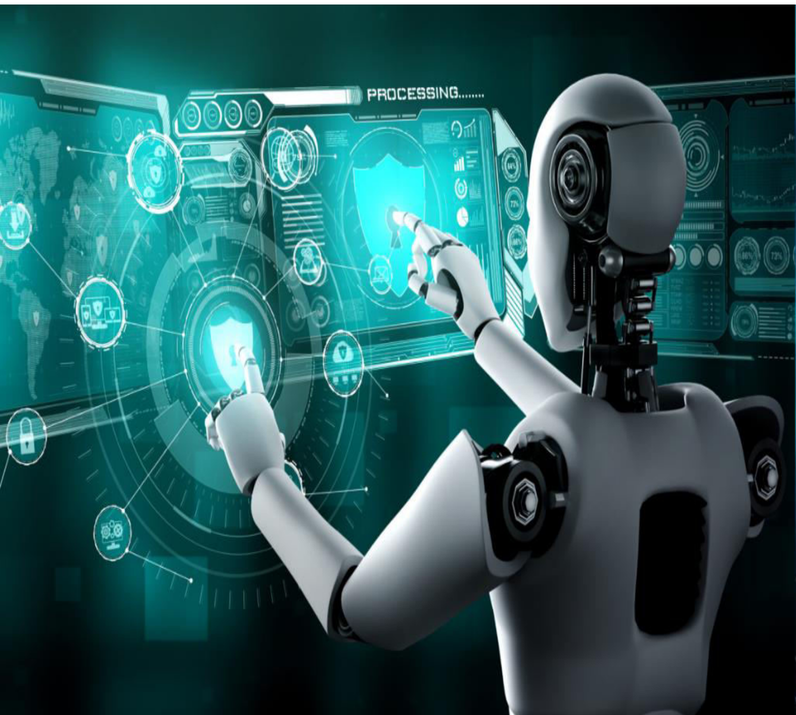
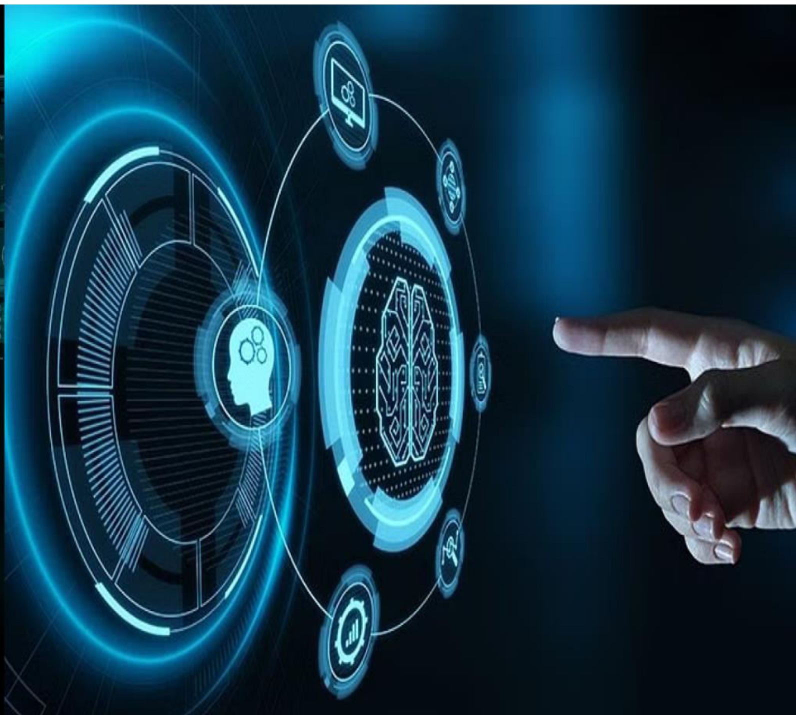
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 9.927

Volume 14, Issue 5, May 2026

www.ijircce.com ijircce@gmail.com +91-9940572462 +91 63819 07438

© 2026 IJIRCCE | Volume 14, Issue 5, May 2026 DOI: 10.15680/IJIRCCE.2026.1405109

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Intelligent Automatic Driver Distraction Detection and Classification Using CNN Deep Learning Model in Vehicles

Aditi Gholap[1], Aadesh Rawas[1], Rutuja Kurhade[1], Mr. Swapnil Bangal[2]

UG Student, Dept. of Electronic & Computer Engineering, PREC, Loni, Maharashtra, India[1]

Assistant Professor, Dept. of Electronic &Computer Engineering, PREC, Loni, Maharashtra, India[2]

ABSTRACT: currently, one of the main factors thought to be responsible for traffic accidents are distraction of driver. Intelligent automobile driving technologies are therefore becoming increasingly crucial. It is clear that driver-assistance systems, which recognize drivers' behaviors and aid them in driving safely, have become more and more popular in recent years. Although there are other forms of data as well, such as the driver's physical conditions, aural and visual features, and automobile information, the main source of data for these research is a camera in the car that captured pictures of the driver's hands, arms, and face. This study proposes a convolution neural network (CNN)-based architecture for the classification and detection of driver distraction. A novel architecture based on the EfficientNetV2 architecture was introduced to run heavy convolutional networks for large-scale image recognition. It makes use of an efficient CNN with excellent accuracy. The State Farm dataset was used to assess the proposed architecture for detecting driver attention. This kind of study often uses this data, which may be accessible publicly on Kaggle. The recommended architecture has a 99.4% accuracy rate.

KEYWORDS: Convolution neural networks, deep learning, and driver-distraction detection component.

I. INTRODUCTION

Road accidents are the tenth most prevalent cause of death worldwide, taking 1.3 million people's lives annually, per a World Health Organization report Twenty to fifty million more people are hurt or disabled. According to a research by the Indian government's National Crime Research Bureau (NCRB), Indian roadways are the world's leading cause of fatalities India has seen a steady increase in the number of fatal road accidents since 2006. The study found that the most common cause of traffic accidents, which claimed 1.46 lakh lives in total in 2015, was driver error.

Distracted driving is defined by the National Highway Traffic Safety Administrator of the United States as any activity that prevents a motorist from carrying out their driving responsibilities. Cognitive distraction is essentially a driver's incapacity to focus on driving, according to the Center for Disease Control and Prevention. In other words, the driver is not concentrating on the task (driving) even while he is in the correct driving position. Instead, he could be daydreaming, lost in contemplation, or otherwise preoccupied due to exhaustion, weariness, drowsiness, or inattention. Visual distraction is the term used to describe when a driver's eyes are taken off the road. Manual distractions are several techniques used while the driver's hands are not on the wheel. These distractions can be communicating through cell phones or text messages, eating and drinking, communicating with the other car occupants, etc.

Identifying the cases of manual distraction when a driver is not engaged in the process of driving The paper at hand is dedicated to the process of safely and also establishing the source of course of distraction. Our approach to this problem is a Convolutional Neural Network based one. The principal contribution of this work are: [1] development of a CNN-style classification algorithm to detect distracted drivers. [2] Qualitative results about the development of common metrics in a recognizable image-based automatic detection system, including accuracy, precision, recall, and f1 score.

II. RELATED WORK

A CNN-based method for identifying distracted drivers using the VGG-16 architecture was presented by Oberoi et al. [1] Their systems maintained real-time performance while achieving an accuracy of 82.5%., Kusuma et al. [2]

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|
International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

presented a DL and computer vision-based method that uses facial characteristics like eye movement, hand position, and head orientation to identify driver distraction. By combining several visual cues, their system lowers false positives and increases detection accuracy. Omerustaoglu et al. [3] proposed an additional hybrid deep learning approach, which uses LSTM networks to combine vision-based CNN models with sensor data. AlShalfan et al. [4] introduced a deep learning approach using VGG-16 architecture on the given dataset. The system developed an accuracy of 96.95% and highlighted the importance of image-based driver monitoring systems. Baheti et al. [5] developed a improved VGG-16 model. Using regularization methods like batch normalization and dropout. Their model greatly decreased computational complexity and attained an accuracy of 96.31%.In order to classify driver distraction.

A deep convolutional neural network-based technique that uses driving signals like acceleration, speed, and RPM by transforming them into image representations was presented by Shahverdy et al. [6] Their method successfully classified a variety of driving behaviours and showed that CNN is not limited to image-based inputs.

Masood et al. [7] proposed a CNN-based architecture that can use dashboard camera photos to determine whether a motorist is distracted. Their system was approximately 99 percent accurate and revealed that CNN is effective in computing the behaviour of drivers.

Hossain et al. [8] designed a dl based system that can detect driver distraction and used different architectures like CNN, VGG-16, ResNet50, and MobileNetV2. They found that MobileNetV2 was the most efficient and the one that was most effective to classify things. In a comparison study, Jagadale et al. [9] combined a custom model with the pre-trained CNN models of AlexNet, VGG-16, and ResNet50. The effectiveness of their proposed model was shown DL, with 93.28 accuracy. Bouhsissin et al. [10] gave a thorough review of the literature on ML and DL methods for classifying driver behavior. Sajid et al.[11] put forward an effective deep learning framework for detecting distracted drivers in order to make the roads safer by accurately identifying different driver behaviors.A thorough survey on machine learning-based driver distraction detection from 2014 to 2021 was presented by Koay et al. [12] The study classifies techniques according to sensors like visual, physiological, and external sensors and draws attention to issues like the limitations of intrusive sensors and the absence of real-world validation, Qi et al. [13] presented a DL-based distracted behaviour detection system. For real-time the system improved accuracy and attained a faster detection speed of 27 frames per second. Using an enhanced YOLOv4 model with a MobileNet backbone and depth wise separable convolution, Mishra et al. [14] suggested an AI-based way to tell if a driver is tired by looking at their eye movements and facial expressions.By identifying drowsiness conditions, their method,Dong et al. [15] suggested a hybrid methodology that integrated Random Forest and CNN techniques for the detection of driver fatigue and distraction. This model utilizes facial feature extraction methods such as S3FD and FAN, which provide high accuracy and real-time capability in embedded systems.

Abouelnaga et al. [16] proposed a distracted driver detection system in real-time using deep learning models with large datasets such as the State Farm dataset. They used convolutional neural networks for driver activities, which showed high accuracy in detecting distracted drivers, A Convolutional Neural Network (CNN)-based model for identifying distracted driving behaviors was proposed by Ruthuparna at el. [17] The technology sorts several driver behaviors, such as chatting, texting, and driving safely, using visual data. According to the study, CNNs can effectively learn spatial information from driver photos. Additionally, it discusses how preprocessing techniques like normalization and resizing might improve model performance, Huang et al [18] Presented a Hybrid CNN Framework for detecting distracted driving behavior in. Their method uses several feature extraction methods and CNN models to make categorization better. The hybrid architecture combines deep learning characteristics with hand-crafted features, making it more robust and accurate than classic CNN-only techniques. This technique works better in complicated real-world situations when the illumination and the driver's position change. In Kashevnik et al. [19] provided an extensive study and framework for approaches of detecting driver attention. The research classifies methodologies into vision-based, sensor-based, and hybrid systems. It indicates that vision-based CNN models are extensively employed due to their excellent accuracy in image classification applications.

III. PROPOSED ARCHITECTURE

Currently, a typical CNN is enhanced and changed by increasing processing power and adding substantial amounts of labeled data. The architectures developed for computer vision tasks, including Alex Net, ZFNet, VGGNet, Google Net,

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

and ResNet, have been used in several CNN adaptations. The Visual Geometry is used in this investigation. In order to identify inattentive drivers, the Group (VGG-16) architecture, as shown in the Fig. 1

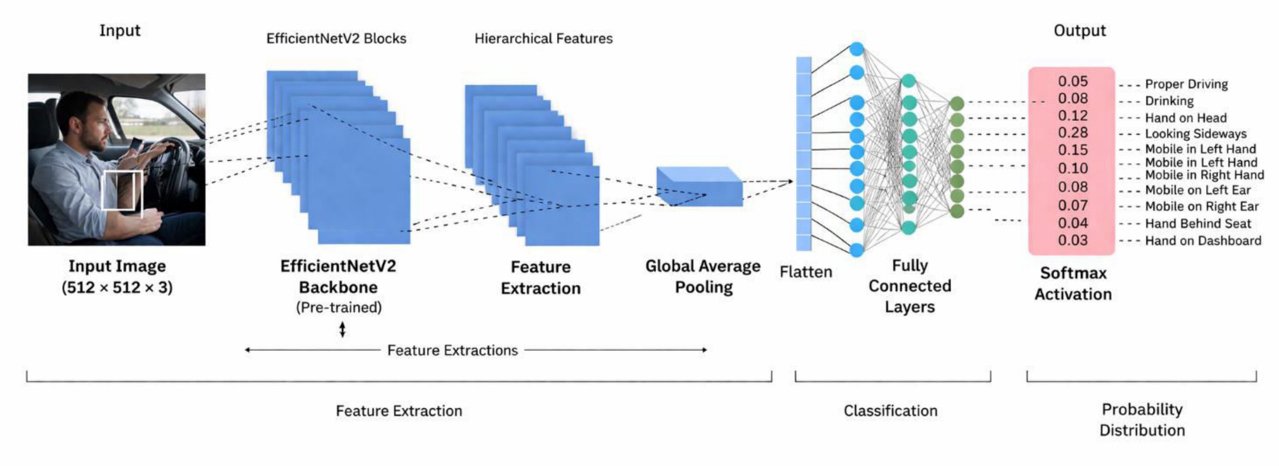


Fig. 1 Proposed EfficientNet-V2 Architecture

Input Layer: An image of a driver taken inside a car is accepted by the input layer. The image is shown as an RGB-formatted numerical matrix that corresponds to pixel intensity values. For example, a 224 × 224 × 3 input image has three colour channels (Red, Green, and Blue) in addition to height and width.

Convolution Layer: Convolutional layers, which are employed for the automatic extraction of important picture information, are the fundamental component of the suggested driver distraction detection system. The layers employ a number of filters, sometimes referred to as kernels, for the convolution operation along the image's spatial dimensions. The filters are employed to identify particular patterns in the picture. Low-level elements including edges, corners, gradients, and textures are detected by the network's lower layers. The mathematical representation of the convolution operation is given by in equation (1)

$$\text{Feature Map} = (\text{Input Image} * \text{Kernel}) + \text{Bias} \quad (1)$$

Activation Function: The network can learn intricate patterns thanks to activation functions, which give it non-linearity. To improve feature learning, convolution operations are followed by the application of GELU or ReLU.

Pooling Layer: These layers preserve the most crucial information while decreasing the spatial size of feature maps. This lessens the complexity of computation and avoids over fitting.

Efficient-NetV2 Backbone: The model uses EfficientNetV2 blocks that consist of fused and MBConv layers instead of the traditional CNN layers. From low-level patterns (edges) to high-level semantic features (driver behaviour), these blocks effectively capture hierarchical features.

Flatten layer: The feature maps that have been identified are transformed into a one-dimensional feature vector by a flattening process. This prepares the data to be sorted.

Fully Connected (FC) Layer:The FC (dense) layer is trained on the relationships between extracted features with various forms of driver behavior. It does this by making decisions, depending on the position of your hands, where you are gazing and how you are manipulating something.

Output Layer (Softmax Activation): The output layer, which is the last layer, generates probability values for each driver behaviour class using the Softmax activation function

The model believes that the most possible alternative in this scenario is "Talking on Phone. This probabilistic output causes the driver monitoring system to be more effective, on the whole, and it assists people to make good decisions. For Example: Drinking: 0.05 Texting: 0.15 Turning away: 0.08

Since "Talking on Phone" has the highest probability in this instance, the model predicts it. This probabilistic output enhances the overall efficacy of the driver.

IV. EXPERIMENTATION

The State Farm dataset was utilized for this investigation, 22424 photos of persons driving safely or engaging in any of the nine types of distracted driving behaviors outlined in Table I are included in the dataset. All of the 640 x 480 pixel photos are RGB. Eighty percent of the total photographs are utilized for training, while twenty percent are used for testing. Below is a discussion of each type of distraction

A. Dataset

For this study, the State Farm dataset was used. The dataset contains 22424 images of people driving safely or participating in any of the nine categories of distracted driving behaviors shown in Table 1. The 640 × 480 pixel images are all RGB. Twenty-five percent of the total images are used for testing, and seventy-five percent are used for training. A description of each kind of distraction is shown in the Fig. 2.

Fig. 2 Classes of images in dataset.

Classes	DRIVER ACTION	FRAME COUNT
C0	safe driving	2489
C1	texting with right hand	2267
C2	talking on the phone with right hand	2317
C3	Texting with left hand	2346
C4	talking on the phone with left hand	2326
C5	operating the radio	2312
C6	Drinking	2335
C7	reaching behind	2002
C8	hair and makeup	1911
C9	Talking to passenger	2129
	TOTAL	22424

TABLE I: TOTAL IMAGES IN DATASET

B. Hardware and software requirement

Hardware Requirements:

1. GPU: NVIDIA RTX3050
2. RAM: 32 GB
3. Storage :1 TB

Software Requirements:

1. Programming language: Python (3.10
2. Libraries: Pandas, Numpy, Matplotlib, Seaborn, OpenCV, PIL
3. Frameworks: Tensor flow
4. Development Environment: VS Code

V. RESULTS AND DISCUSSION

The results in Table II show that traditional methods like VGG-16 and basic CNN models work well, but they have problems with optimization and generalization. The updated VGG-16 shows how architectural improvements can make

© 2026 IJIRCCE | Volume 14, Issue 5, May 2026 DOI: 10.15680/IJIRCCE.2026.1405109

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

a difference by increasing accuracy through regularization. The suggested EfficientNetV2-M model, on the other hand, gets the best training and testing accuracy (99.4%) on the State Farm dataset. This shows that it learns better and lasts longer.

Ref.	**Model**	**Dataset**	**Training Accuracy (%)**	**Testing Accuracy (%)**
[1]	VGG-16 (Transfer Learning)	State Farm Dataset	85	82.50
[5]	Modified VGG-16	State Farm Dataset	97	96.31
[18]	CNN Model	Driver Distraction Dataset	96	95
Our Model	**EfficientNetV2-M**	**State Farm Dataset (Kaggle)**	**99.4**	**99.4**

Table II: Comparative Analysis of Training and Testing Performance

In Below given Fig. 4 the EfficientNetV2-M model was good at classifying the task of recognizing driver distraction. The model was able to identify nearly all of the instances of the 10 types of driver distraction as is evidenced by the preponderance of the data clustering along the diagonal. The model is able to acquire the ability to differentiate among different properties in the data as demonstrated by its accuracy of approximately 99.4%. The confusion matrix above shows how well the proposed driver distraction detection model works for different types of behavior. The true class label is in each row, and the predicted class label is in each column Loss Plot.

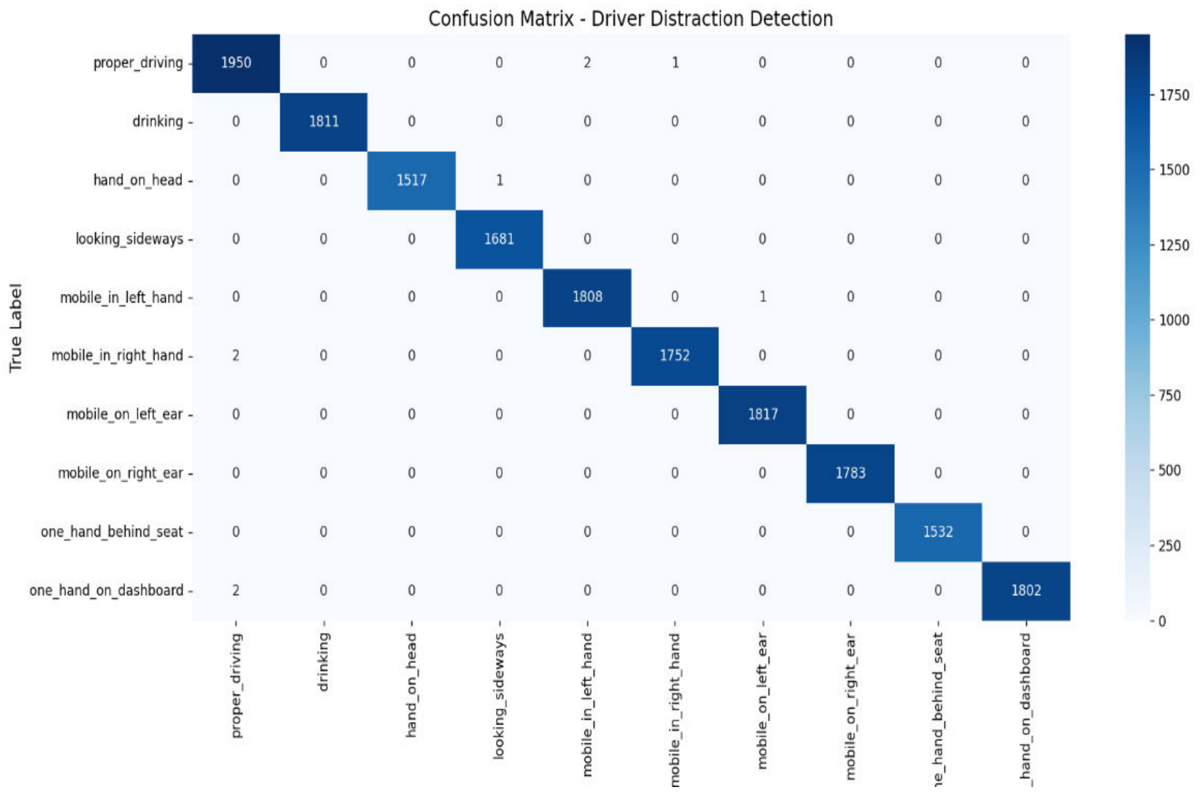
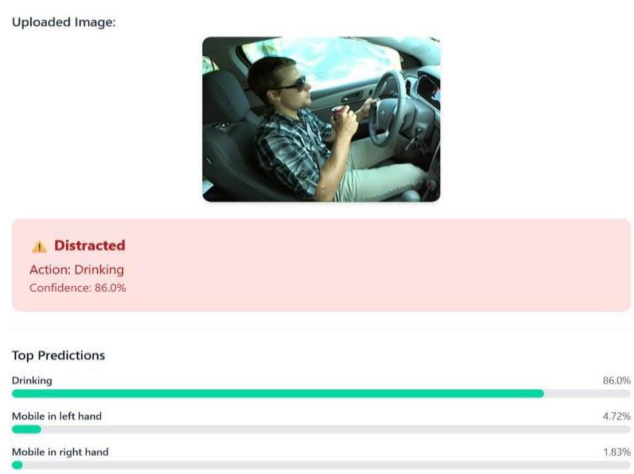
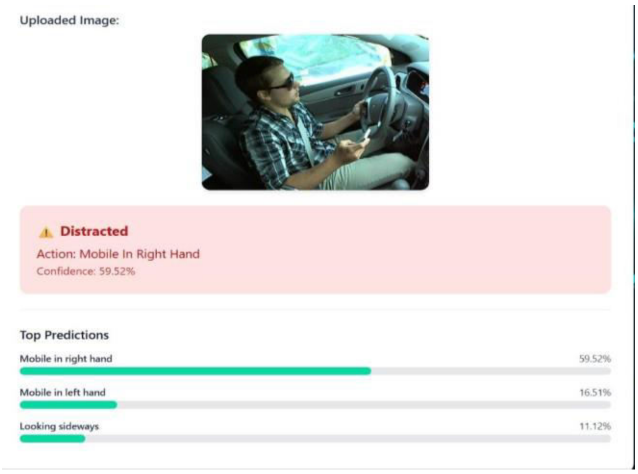


Fig 3. Confusion Matrix of the Proposed Model

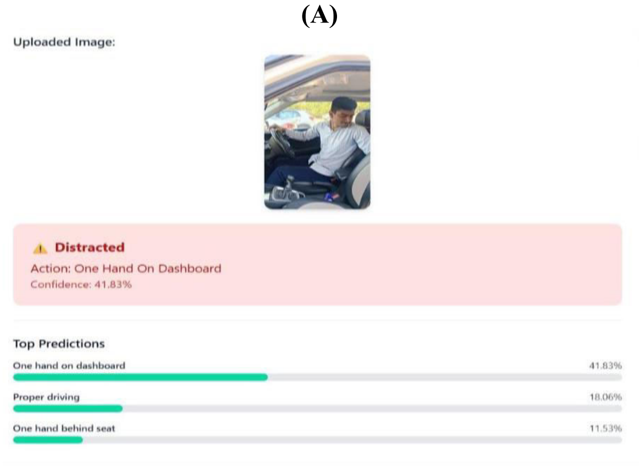
The Fig. 4 in the findings below show the comparisons of the photos to the real images and Kaggle (StateFarm) dataset. Due to the clean and arranged data, the model works exceptionally well with the photos on the Kaggle dataset. It has a high level of confidence with a range of 80-99. It does not have many downsides, and the difference between safe and distracted driving is easy to draw. Nevertheless, the algorithm is only able to identify distracted activities in real-life photographs 40-60 per cent of the time. The reason is that the set of training data might not be similar to real-life conditions in illumination, camera angles, background noise, and other aspects. Nonetheless, this model proves that it can be generalized because it can identify distracted driving behaviors in real time



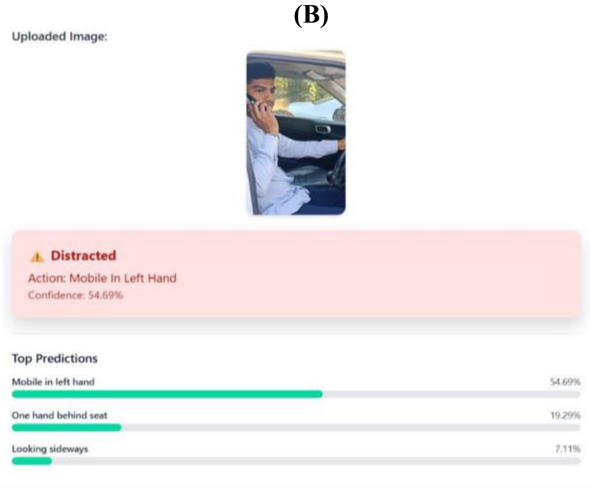
(A)



(B)



(C)



(D)

Fig. 4 Diseacted driver classes result: (A) Drinking, (B) Mobile in Right hand, (C) One hand in dashboard, (D) Mobile in Right hand

VI. CONCLUSION AND FUTURE WORK

This project shows that deep learning can greatly improve road safety by automatically finding when a driver is distracted. The system uses a cutting-edge CNN-based model to accurately identify different driving behaviors and tell the difference between safe and unsafe actions in real time. The model is very accurate, with an accuracy rate of about 99.4%. This shows that it can learn complicated visual patterns, like how to move your hands, how to turn your head, and how to use objects like phones. The model still gives good results even though things like changes in lighting, camera angles, and background noise can have a small effect on performance. This system can be added to smart cars to warn drivers and lower the risk of accidents if it is further optimized and deployed in real time. Future research can concentrate on strengthening the model's resilience to changing environmental factors like illumination and camera angles as well as tailoring it for real-time deployment in smart cars. Road safety can also be improved by connecting the system with advanced driver assistance systems (ADAS), which offer prompt alerts and preventive measures.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

REFERENCES

[1] Manpreet Oberoi, "Driver Distraction Detection using Transfer Learning," International Journal of Engineering Research and, vol. V9, no. 05, Jun. 2020, doi: 10.17577/ijertv9is050862

[2] S. Kusuma, J. Divya Udayan, and A. Sachdeva, "Driver Distraction Detection using Deep Learning and Computer Vision," 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), pp. 289–292, Jul. 2019, doi: 10.1109/icicict46008.2019.8993260

[3] F. Omerustaoglu, C. O. Sakar, and G. Kar, "Distracted driver detection by combining in-vehicle and image data using deep learning," Applied Soft Computing, vol. 96, p. 106657, Nov. 2020, doi: 10.1016/j.asoc.2020.106657

[4] K. A. AlShalfan and M. Zakariah, "Detecting Driver Distraction Using Deep-Learning Approach," Computers, Materials & Continua, vol. 68, no. 1, pp. 689–704, 2021, doi: 10.32604/cmc.2021.015989.

[5] B. Baheti, S. Gajre, and S. Talbar, "Detection of Distracted Driver Using Convolutional Neural Network," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1145–11456, Jun. 2018, doi: 10.1109/cvprw.2018.00150.

[6] M. Shahverdy, M. Fathy, R. Berangi, and M. Sabokrou, "Driver behavior detection and classification using deep convolutional neural networks," Expert Systems with Applications, vol. 149, p. 113240, Jul. 2020, doi: 10.1016/j.eswa.2020.113240.

[7] S. Masood, A. Rai, A. Aggarwal, M. N. Doja, and M. Ahmad, "Detecting distraction of drivers using Convolutional Neural Network," Pattern Recognition Letters, vol. 139, pp. 79–85, Nov. 2020, doi: 10.1016/j.patrec.2017.12.023.

[8] Md. U. Hossain, Md. A. Rahman, Md. M. Islam, A. Akhter, Md. A. Uddin, and B. K. Paul, "Automatic driver distraction detection using deep convolutional neural networks," Intelligent Systems with Applications, vol. 14, p. 200075, May 2022, doi: 10.1016/j.iswa.2022.200075.

[9] S. H. Jagadale and Dr. P. R. Attar, "Distracted Driver Detection: A Comparative Study Using CNN," Proceedings of the International Conference on Industrial Engineering and Operations Management, pp. 2449–2458, Aug. 2022, doi: 10.46254/in02.20220579.

[10] S. Bouhsissin, N. Sael, and F. Benabbou, "Driver Behavior Classification: A Systematic Literature Review," IEEE Access, vol. 11, pp. 14128–14153, 2023, doi: 10.1109/access.2023.3243865.

[11] Sajid F, Javed A R, Basharat A, et al. An efficient deep learning framework for distracted driver detection [J]. IEEE Access, 2021, 9: 169270-169280.

[12] H. V. Koay, J. H. Chuah, C.-O. Chow, and Y.-L. Chang, "Detecting and recognizing driver distraction through various data modality using machine learning: A review, recent advances, simplified framework and open challenges (2014–2021)," Engineering Applications of Artificial Intelligence, vol. 115, p. 105309, Oct. 2022, doi: 10.1016/j.engappai.2022.105309.

[13] X. Qi, "Driver Distracted Behavior Detection Based on Deep Learning," Proceedings of the 2022 6th International Conference on Electronic Information Technology and Computer Engineering, pp. 786–790, Oct. 2022, doi: 10.1145/3573428.3573569.

[14] U. Mishra, S. Tyagi, P. Goyal, and S. Gupta, "A Novel Approach Distracted Driver Detection System Using ML Techniques," SSRN Electronic Journal, 2024, doi: 10.2139/ssrn.4811988.

[15] B.-T. Dong, H.-Y. Lin, and C.-C. Chang, "Driver Fatigue and Distracted Driving Detection Using Random Forest and Convolutional Neural Network," Applied Sciences, vol. 12, no. 17, p. 8674, Aug. 2022, doi: 10.3390/app12178674.

[16] Y. Abouelnaga, H. M. Eraqi, and M. N. Moustafa. Real-time distracted driver posture classification. CoRR, abs/1706.09498, 2017.

[17] M. P. Ruthuparna and M. Mallikarjuna, "Distracted driver detection using convolutional neural network," 1ST INTERNATIONAL CONFERENCE ON EMMA-2021, vol. 2742, p. 020058, 2024, doi: 10.1063/5.0191591

[18] Huang C, Wang X, Cao J, et al. HCF: a hybrid CNN framework for behavior detection of distracted drivers [J]. IEEE access, 2020, 8: 109335-109349.

[19] Kashevnik, A.; Shchedrin, R.; Kaiser, C.; Stocker, A. Driver distraction detection methods: A literature review and framework. IEEE Access 2021, 9, 60063–60076. Distract CNN. Available online: https://github.com/nkkumawat/Driver-Distraction-dtection/branches (accessed on 5 January 2020).