Driver Behavioural Detection & Alert System

Khaja Moinudeen AK, Kruthik J, Madhu Sudhan S,Mohammed Saif, Mr. Shreenidhi BS

Student, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, Dayananda Sagar Academy of Technology and Management, Bangalore, India

Abstract

Driving requires a series of actions that call for significant concentration. Sometimes these behaviours are put on the back burner in favour of other activities including smoking, eating, drinking, talking, making phone calls, adjusting the radio, or falling asleep. These are also the main causes of today's traffic accidents. As a result, developing apps

It is essential to notify drivers in advance. To help the warning system convey accurate information and reduce traffic collisions, a lightweight convolutional neural network architecture is developed in this study.

This network is produced by combining feature extraction and classifier modules.
By utilising the advantages of average pooling layers, depthwise separable convolution layers, standard convolution layers, and suggested adaptive connections, the feature extraction module extracts the feature maps. The benefit of the convolution block attention module is utilised by the feature extraction module, which guides the network in learning the important features. The classifier module makes use of a global average pooling and softmax layer to calculate the probability of each class. Throughout the architecture, classification accuracy is kept while network parameters are optimised.

The entire network is trained and tested using three benchmark datasets: the State Farm Distracted Driver Detection, the American University in Cairo version 1, and the American University in Cairo version 2. Since there are ten classes, the overall class accuracy is 99.95%, 95.57%, and 99.61%, respectively. A number of video tests were also conducted in HD (High Definition), FHD (Full High Definition), and VGA (Video Graphics Array) resolutions; you can view them at https://bit.ly/3GY2iJl.

Introduction

There are many more and more intricate road traffic systems in use today. Accidents increasingly increased in number as a result. The World Health Organization estimates that 50 million automobile accidents result in 1.35 billion fatalities per year [1]. Accidents are increasing due to a number of variables, including driver behaviour. Additionally, it was mentioned in the previous clause that driving while paying close attention could reduce accidents by half. Approximately 2,895 fatal road accidents involving distracted driving are estimated to have happened in 2019 (US), accounting for 8.7% of all fatal traffic accidents that year [2]. These reports indicate that, between 2010 and 2019, distracted driving continued to cause 8% to 10% of all fatalities and accidents, and 14% to 16% of all accidents. The exact number of distracted driving fatalities and crashes during the preceding ten years is shown in Figures 1 and 2.

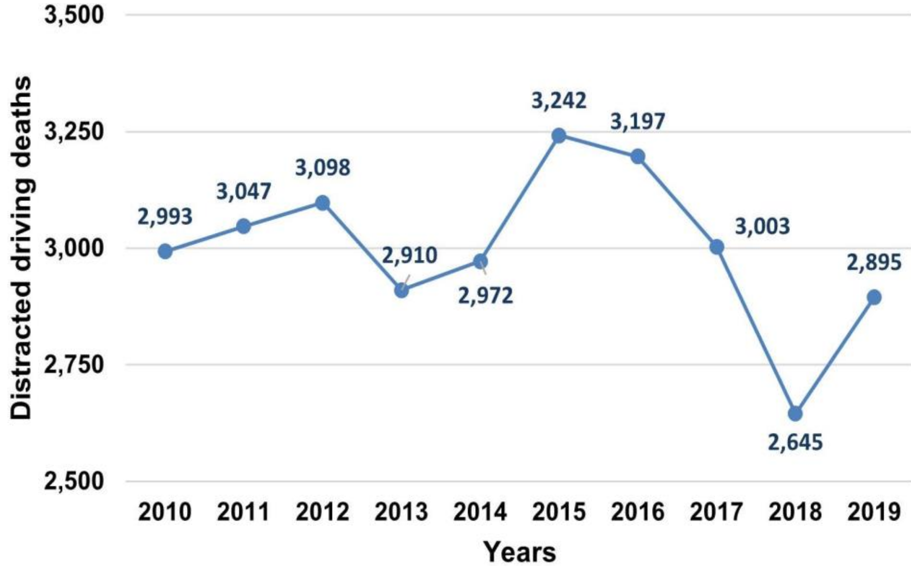

FIGURE 1. **Statistics on fatalities from distracted driving in the US during the next ten years.**

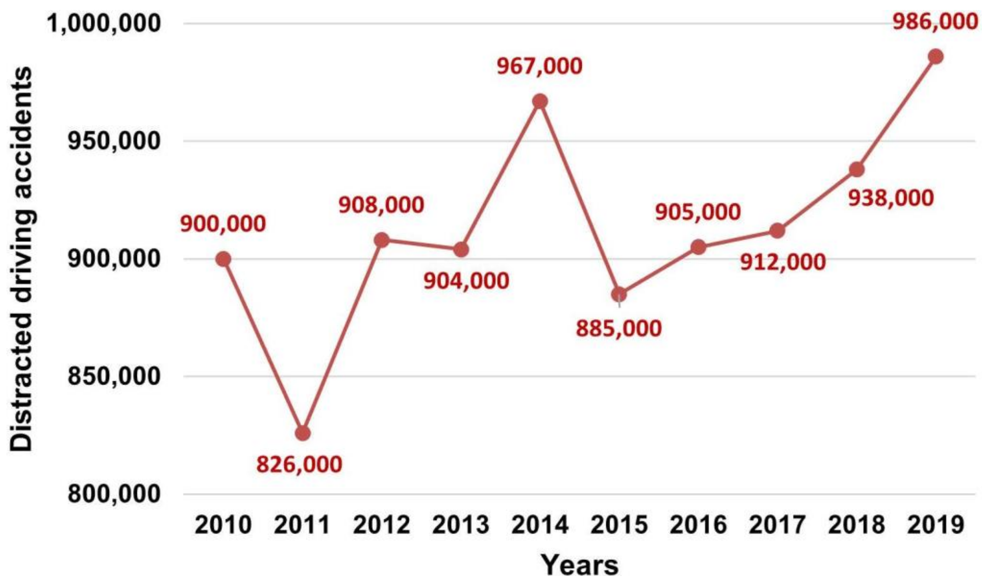

FIGURE 2. **Statistics on distracted driving collisions in the US during a ten-year period.**

The scientists focused on investigating the reasons behind traffic collisions, coming up with fixes, and defining distracted driving. Distracted driving, according to the authors of [2], is the act of a driver utilising a telephone (talking or texting), eating, drinking, or manipulating the entertainment system while operating a vehicle (radio, stereo). Distracted driving, according to a separate definition in [3], is any activity that keeps the driver's focus off the road. In this study, there are three different forms of distracted driving: visual, manual, and cognitive. The major objective of visual distraction is the examination of eye position and head posture. The main equipment for measuring visual distraction is a variety of sensors and cameras that are physically connected to the driver or installed on vehicles to gather, analyse, and record information. In the event of physical distraction, the devices are designed to watch a driver's hand or foot motions on the stop and gas pedals. Cognitive distraction can predict the psychophysiological condition of drivers, including their heart rate, blood pressure, and body temperature. Due to the discoveries presented above [4]–[6], automakers have integrated analytical and driver warning devices in modern automobiles.

The bulk, nevertheless, are still undergoing testing. However, the majority of these devices are pricey, and the older vehicles make deployment difficult. Additionally, safe driving practises and various naturally occurring human body structures may cause wearable electronics to suffer from signal interference. In order to simplify the devices, make them less intrusive for drivers, and save money, this research proposes a way to recognise driving behaviours using a rudimentary convolutional neural network (CNN) architecture combined with the attention methodology.

In order to extract feature maps, the proposed network combines adaptive connections, average pooling layers, depthwise separable convolution layers, and standard convolution layers. The standout features are then learned using the attention mechanism using the convolution block attention module (CBAM). The global average pooling (GAP) layer and softmax function are used by the classifier module to calculate 10 probabilities of associated driver behaviours in the datasets. The primary contributions of this study are as follows:

1) The driver warning system is supported by a lightweight convolutional neural network for driving behaviour identification. Modules for feature extraction and classifiers make up this network. The approach uses fundamental CNN building blocks, suggested adaptive connections, and a convolution block attention module to learn crucial feature map information. Additionally, it replaces all fully connected layers in typical classification networks using global average pooling. As a result, it maintains high speed and accuracy while optimising the network settings. This architecture may be used without incurring any additional installation or modification expenditures on low-cost and low-computing equipment, including deployment on older cars. On the other hand, it doesn't interfere with the driver's psyche because of how the camera's picture signals are used and processed.

2) All three benchmark datasets were used to thoroughly train, assess, and report the proposed network. Additionally, this team created the application for testing videos on various hardware, including the Jetson Nano, GPU, and CPU.

II. RELATED WORK

This section will describe numerous methods for identifying driver behaviour along with its advantages and disadvantages.
When comparing these approaches, both conventional machine learning and CNN-based approaches are taken into account.

A. Traditional Machine Learning Methodology

The first research aimed to determine how often people use cell phones while driving. Ref. [7] uses the Supervised Descent technique, a Histogram of Oriented Gradients (HOG), and an Adaboost classifier to identify smartphone usage with a 93.9% accuracy rate. The constraints of this study include the mobile phone area extraction from facial landmark method, lighting, and occlusion situations. By measuring the distances between four components, such as the face, mouth, hands, and cellphone, further study uses Hidden CRF [8] and Support Vector Machines (SVM) [9] to categorise cellphone use.

These techniques still strongly rely on the skin's illumination even though they may achieve accuracy of 91.2% and 91.57%, respectively. With an accuracy of 86.19%, the research in [10] also uses the SVM approach to detect smartphone usage in images acquired with cameras deployed on roads and traffic signals. The study's accuracy is quite poor because it only included a small sample of 1,500 images. The authors of [11] recommended utilising Hidden Markov models and an Adaboost classifier to categorise pictures collected from the RGB-D sensors of Kinect devices that imitate smartphone use.

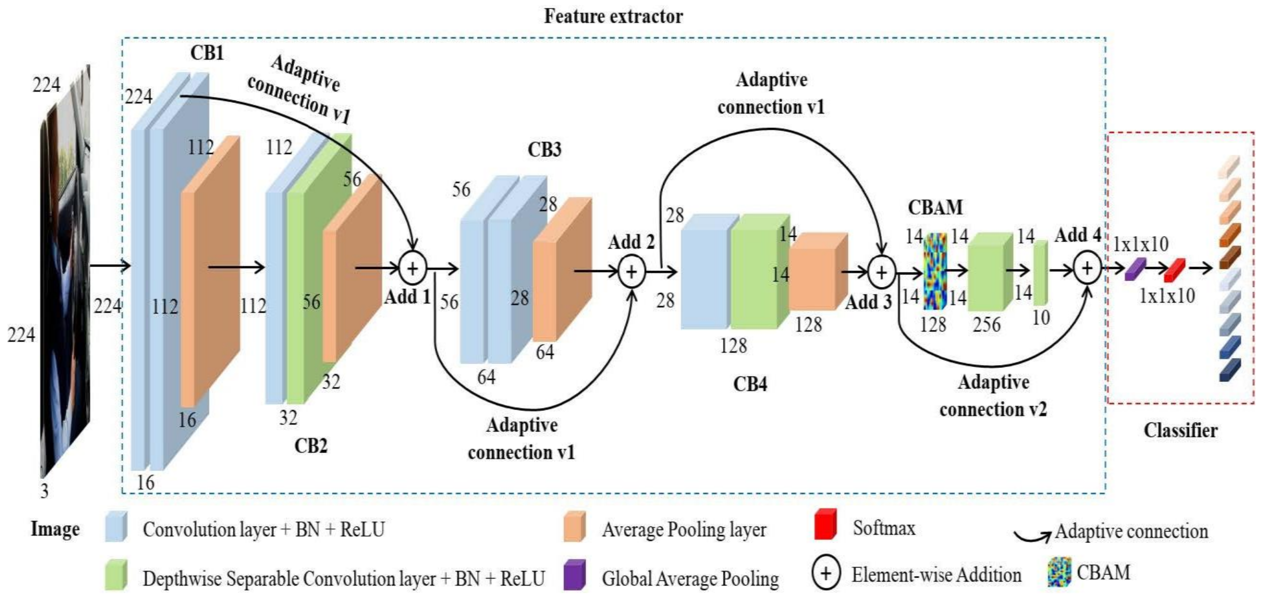

FIGURE 3. The suggested network for classifying driving behavioural.

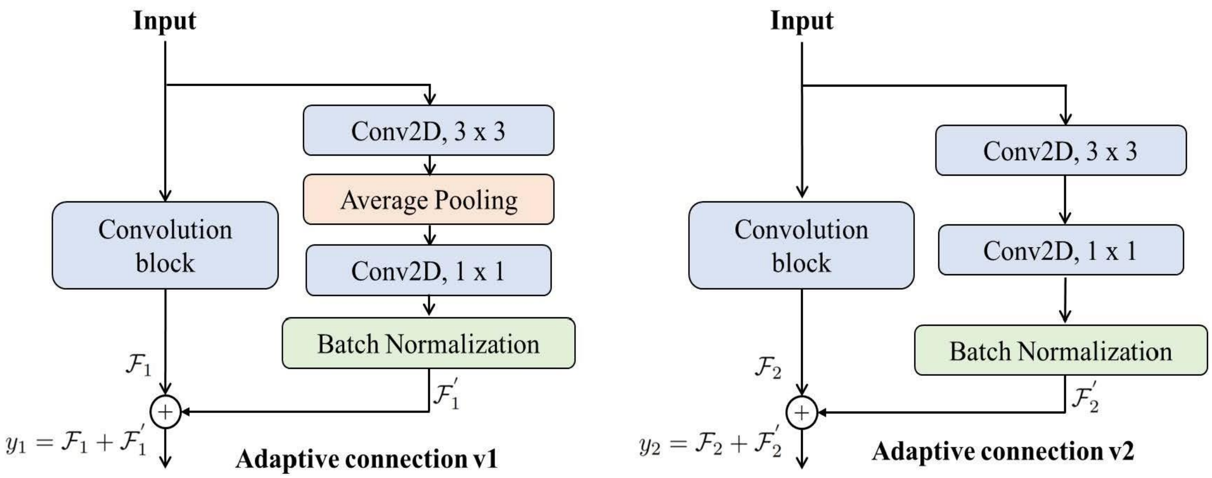

FIGURE 4. The design of the flexible connectors.

Distracted driving can be recognised 90% of the time, however the system is made up of many complex parts. Using the Aggregate Channel Features (ACF) method, [12] focuses on detecting the driver's hand movement; the best prediction result was 70.09% of AP (Average Precision), and [13] uses the Contourlet Transform in conjunction with Random Forests classification to classify multiple actions; the eating action had the highest accuracy, 88%. In conclusion, while standard machine learning algorithms are easy to use, their classification accuracy is subpar.

B. CNN_Based Methodology

Recently, computer vision areas have made substantial use of convolutional neural networks. Convolutional neural networks have also been utilised to develop applications for monitoring and warning in research on human behaviour in general and driver behaviour in particular. Picture segmentation, image classification, and image detection are examples of applications. The work in [14] employs a Faster R-CNN network as a detector to infer the hand motions on the steering wheel. The results show that this method reliably detects instances of using a mobile phone while driving and having one's hands on the wheel, with accuracy rates of 92.4% and 91%, respectively. In Ref. [15], the dashboard, gear lever, and steering wheel are localised using the photo segmentation approach.

They then propose a network architecture to identify the driver's hand position on previously segmented regions with 74.3% accuracy. This combo strategy can deal with the problem of light changes despite being computationally challenging. The first dataset for the picture classification task that [16] recommended was the Southeast dataset, which had four classes—smoking or eating, talking on the phone, driving safely, and shifting gears. With an overall accuracy of 99.78%, [17] employed a variety of convolutional neural network methods on this dataset, which is used in typical machine learning methodologies, to categorise these four categories. Later, [18]–[20] presented ten-class enhanced datasets for distracted driving (the dataset part includes a thorough description of each class).

On the basis of these datasets, several research have trained and evaluated various CNN network topologies. The authors also provide suggestions in [19] and [20]. The accuracy was 94.29% and 93.65%, respectively, using an ensemble training technique with five different CNN networks. Driving activities have also been classified using other popular classification neural networks, such as VGG [21], [22], DenseNet [23], and GoogleNet [24], with accuracy ranging from 95% to over 99%. In order to decrease network parameters and use low-computation devices, [25, [26] proposed convolutional neural network designs with depthwise separable convolution operation and a residual network to recognise 10 driving behaviours. With a very high accuracy (over 95%), these algorithms create a very small number of network parameters (less than 0.5M parameters).

Although the research listed above demonstrate high accuracy, they either looked at individual datasets with fewer classes or only concentrated on categorising driver behaviour using a small number of classes (four classes) (ten classes). On the other hand, several of the solutions that have been suggested are complex and difficult to integrate into real-time systems. In order to test the effectiveness of typical convolutional and depthwise separable convolution layers, as well as Inception and Residual networks, this study proposes a lightweight driver behaviour classification convolutional neural network. Despite having just 0.43M parameters, the network gives high accuracy when compared to several other approaches.

III. PROPOSED METHODOLOGY

The suggested network design consists of two modules: a feature extractor and a classifier. The feature extractor module is constructed utilising the stem module, adaptive connections (AC), and a CBAM to extract the feature maps. The classifier module calculates the probability of 10 driving behaviours before categorising them using the final feature map and the GAP and softmax functions. Figure 3 provides an extensive description of the planned architecture.

A. FEATURE EXTRACTOR MODULE

The majority of well-known convolutional neural networks can extract high-level feature maps from raw images without any further human processing steps. It will be easy to use and achieve high precision for subsequent tasks including picture segmentation, object recognition, and classification. Since traditional machine learning algorithms mostly rely on feature extraction and picture preprocessing, the received accuracy is unreliable.

This work focuses on developing the feature extractor utilising a number of novel methodologies in order to provide the most usable feature maps. The feature extractor consists of two depthwise separable convolution layers, four ACs, one CBAM, and four convolutional blocks (CBs). The CBs have two different architectural designs. The first design consists of two common convolution layers, a batch normalising layer, a rectified linear unit activation function, and an average pooling layer (in CB1, CB3). The alternative design uses a depth-separable convolution layer, a ReLU activation function, a standard convolution layer, an average pooling layer, and a BN layer (in CB2, CB4).

The kernel widths and channel counts of convolutional blocks range from 7 7 16 (CB1) through 5 5 32 (CB2), 3 3 64 (CB3), and 3 3 128 (CB4) (CB4). Early high kernel sizes can increase the network's receptive fields, allowing the feature extractor to more precisely extract the crucial object data from the picture. However, this strategy considerably increases the cost of computation on the network. After then, depthwise separable convolution layers are utilised to reduce the cost of the prior computation. After four convolution blocks, the 224x224x3 input picture will be reduced by 16 times, yielding a 14x14x128 feature map.

A significant amount of important information is lost throughout this process. As a result, it is crucial to combine the information from the previous and current feature map levels. This maintains and enhances the crucial information for each feature map's level. As seen in Figure 4, our study provides adaptive connections using two different approaches, leveraging ideas from the ResNet [27] and Inception [28] networks. Adaptive connection version 1 has a 3 3 standard convolution layer, an average pooling layer, and a 1 1 standard convolution layer followed by a BN layer (ACv1). Adaptive connection version 2 (ACv2) is nearly equivalent to ACv1 but does not use the average pooling layer between the two common convolution layers.

The suggested network makes use of four adaptable connections, each at a different level. In particular, the first ACv1 applies to the output of the second convolution layer in CB1 and adds the output of CB2. The second one (ACv2) increases the output of the first addition by the output of CB3 (Add 1). A process that is equivalent for the third adaptive connection (ACv1) and the fourth adaptive connection is used to create the output of Adds 3 and 4. (ACv2). The equations for these relationships are as follows:

$$y_i = F_i(x) + F_i^r(x), \qquad (1)$$

where x and y are the corresponding input and output feature maps. Version I of the adaptive connection is either 1 or 2. Each convolutional block's output feature map is Fi(x) RW H C.

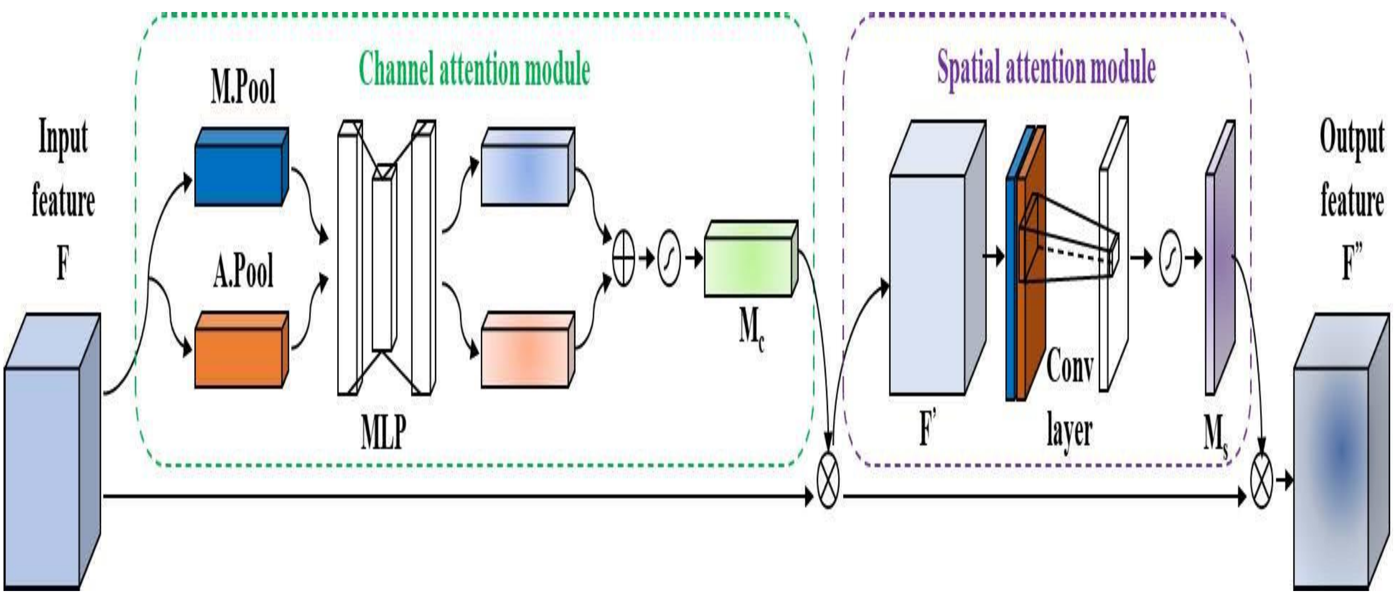

FIGURE 5. The convolution block attention module's construction.

Fir(x) RW HC is the feature map that results from the adaptive connections. Within each iteration:

F1r (x) = BN (f 11(A.Pool(f 33(x))),
F2r (x) = BN (f 11(f 33(x))),

where f 11 and f 33 represent 1 1 and 3 3 standard convolu- tion layers, respectively. The batch normalisation layer is called BN. The standard pooling layer is A.Pool.

C. Classifier Module

At the end of the categorization network, totally connected layers have always been a common practise. However, when used to low-powered computer devices, this technique significantly increases network parameters, increasing the computational load and delaying processing. This article recommends a method for replacing all entirely connected layers in well-known categorization networks with a single GAP layer. For this method, the spatial features are extracted along each channel, and the 14 by 14 by 10 feature map produced by the extractor will quickly be reduced to 1 by 1 by 10, saving a substantial amount of network parameters. A softmax function is then used to calculate the probability of each item class appearing in the input picture.
The difference between the predicted value and the target value is calculated during training using the categorical cross-entropy loss function for simplicity.

D. VIDEO TESTING SYSTEM

Figure 6 outlines the whole video testing setup. (Stage of testing). The system is composed of the input, trained model, and output. A selection of videos in different resolutions, including VGA, HD, and FHD, are accepted as input. The model is trained using the State Farm dataset and the stored weight file. The output consists of on-screen message text signals with prediction class, accuracy, and speed in FPS. The output of this system, which may instal audio signals to the speaker to warn the driver, can also be used in place of the input by a standard camera. The organisational structure of the real-time driver warning system is shown below.

IV. EXPERIMENTS

A. DATASETS

The following three datasets for categorising driver behaviour were utilised in this study's training and assessment phases: The American University in Cairo version 1 of the State Farm Distracted Driver Detection (State Farm) [18], and

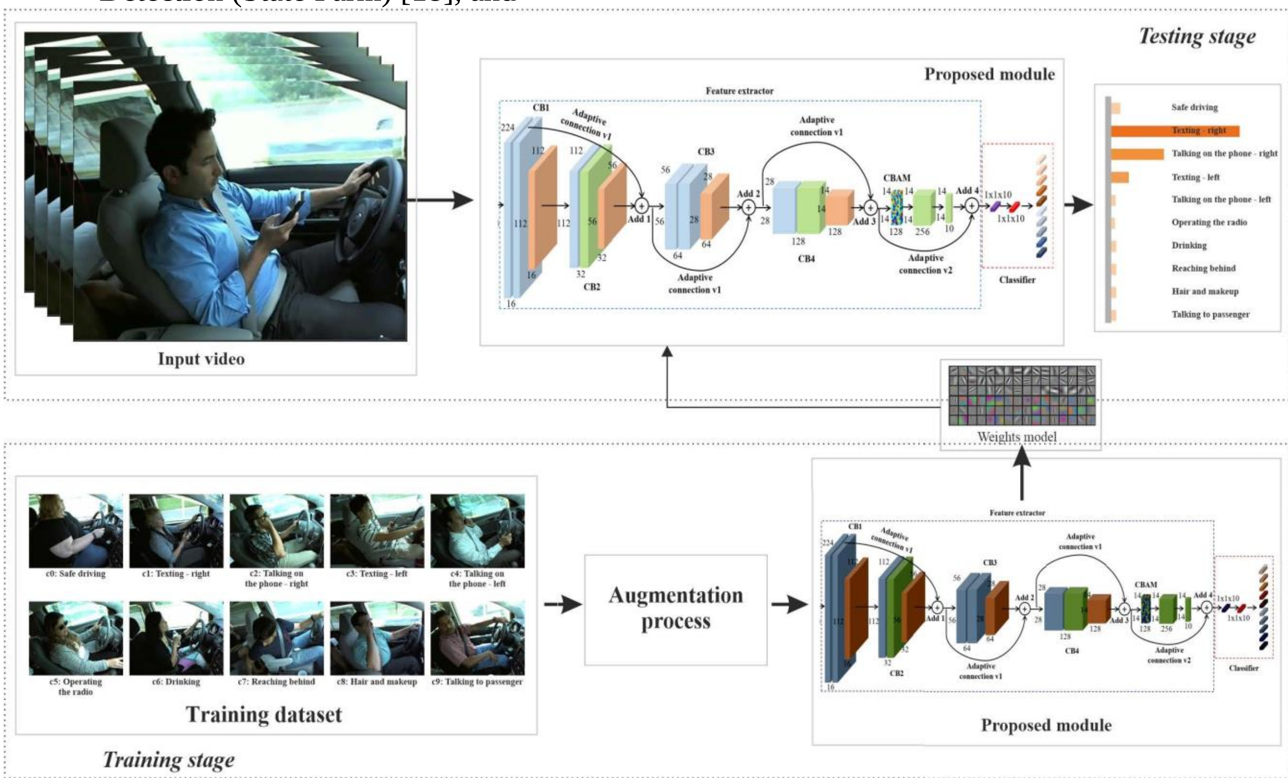

FIGURE 6. The system used to test videos generally.

(AUC version 1) [19], and the American University in Cairoversion 2 (AUC version 2) [20].

1. STATE FARM DATASET

This Kaggle competition's dataset was downloaded. It consists of 22,424 colour photos at a resolution of 640 by 480 pixels. Ten folders with names like "safe driving" (c0), "texting-right," "talking-right," "texting-left," "managing the radio," "drinking," "reaching behind," "hair and makeup," and "talking to passenger" are used to organise the photos (c9).

2. AUC VERSION 1 DATASET

This dataset contains videos of 31 individuals' simulated driving behaviours from seven different nations—22 males and 9 women. The AUC version 1 dataset was recorded in video format by the SUS ZenPhone's back camera. After that, 17,308 modified high-resolution (1080 x 1920) images were chosen. 10 classes made up the

State Farm dataset, and ten classes make up this dataset as well, albeit with different titles (Drive Safe, Text Right, Talk Right, Text Left, Talk Left, Adjust Radio, Drink, Reach Behind, Hair & Makeup, Talk Passenger).

3. AUC VERSION 2 DATASET

The DS325 Sony DepthSense camera was utilised to capture five distinct types of cars utilising both the ASUS ZenFone smartphone back camera and the AUC version 2 dataset. There are a total of 44 athletes from seven different nations, comprising 29 men and 15 women. These movies were shot in a variety of settings, including various lighting, weather, and topic and road conditions. 14,478 images with a resolution of 640 x 480 pixels or 1080 x 1920 pixels are retrieved. The labels and accompanying folders of the classes are separated, much like the AUC version 1 dataset.

B. EXPERIMENTAL SETTING

The proposed network is implemented using Python programming and the Keras framework. This network is examined and trained on a GPU (GeForce GTX 1080Ti). A separate CPU (Intel Core I7-4770 CPU @ 3.40 GHz, 32GB of RAM), one Jetson Nano, and a system for assessing movies with VGA (640 480 pixels), HD (1280 720 pixels), and FHD (1920 1080 pixels) resolutions also used it (Nvidia Maxwell GPU, 4GB of RAM). 300 epochs are completed during the training phase with a batch size of 16. The Adam optimization technique is employed throughout the network's weight updating process. After 20 epochs, if the accuracy has not improved from the previous phase, the learning rate approach starts off at 103 and then reduces to 0.55 times. For each dataset, a training set (80%) and an evaluation set (20%) are kept apart. In order to increase accuracy and avoid overfitting issues, this experiment uses a number of data augmentation techniques, including random brightness, random zoom, and shift.

TABLE 1. The outcomes of the comparisons using various techniques on the State Farm, AUC version 1 and AUC version 2 datasets.

Model	Parameters (Million)	Acc (%)
State Farm dataset		
Simple CNN [24]	0.65	99.51
Light-weight CNN [24]	0.46	99.95
Mobile VGG [22]	2.20	99.75
Our network	**0.43**	**99.95**
AUC version 1 dataset		
VGG with Regularization [21]	140	96.31
Original VGG [21]	140	94.44
GA weighted ensemble [19]	120	95.98
Majority Voting ensemble [19]	120	95.77
AlexNet [19]	62	93.65
InceptionV3 [19]	24	95.17
Modified VGG [21]	15	95.54
DenseNet+Latent Pose [21]	8.06	94.20
NasNet Mobile [22]	5.30	94.69
MobileNet [22]	4.20	94.67
MobileNetV2 [22]	3.50	94.74
Mobile VGG [22]	2.20	95.24
SqueezeNet [22]	1.25	93.21
Light-weight CNN [25]	0.46	95.36
Our network	**0.43**	**95.57**

AUC version 2 dataset		
AlexNet [20]	62	94.29
InceptionV3 [20]	23.91	90.07
Resnet50 [20]	23.77	81.70
	15.25	76.13
	0.43	**99.61**

C. RESULT ANALYSIS

The proposed network was trained and assessed on the three aforementioned datasets, and then tested on movies utilising a GPU, a CPU, and a Jetson Nano device. These tests are provided based on the accuracy and frames per second (FPS) measurements. This network achieved accuracy scores of 99.95% on the State Farm dataset, 95.57% on the AUC version 1 dataset, and 99.61% on the AUC version 2 dataset using just 426,785 parameters. The whole network parameters are established in the typical way by adding the weights and biases of the convolutional and fully connected layers. In order to optimise network parameters, this article totally replaced the fully linked layers with a GAP layer and the four convolution layers with four depthwise separable convolution layers. This considerably reduces the network parameter while keeping the network's classification and feature extraction accuracy. The outcome above shows that the findings using State Farm and AUC version 2 are nearly absolute since the images in both datasets are clearly divided into folders after the class labels. The AUC version 1 dataset, in comparison, has some photographs that are scattered across the two behaviours, confounding the network's learning process. For the State Farm dataset, the suggested network performs better than the Simple CNN [24] and the Mobile VGG [22]. Although it has around 30K less network parameters, it is still comparable to the Light-weight CNN [24]. Only the VGG with Regularization [21], the GA weighted ensemble [19], and the Majority Voting ensem- ble [24] perform worse than the suggested network, which beats the majority of already used networks for the AUC version 1 dataset. However, the network parameters of the Majority Voting ensemble, the GA weighted ensemble, and the VGG with Regularization ensemble are all 325.58 times larger than those of the proposed method. For the AUC version 2 dataset, the proposed network totally surpasses the well-liked classifier networks in [20], with an accuracy range of 5.32% to 23.48%.

The three confusion matrices in Figures 8, 9, and 10 show the recommended network's classification skills in each class. The State Farm and AUC version 2 datasets' class prediction rates, which range from 98% to 100%, are extremely stable. The prediction rates were frequently between 95% and 97%, with the exception of the labels "Adjust radio," "Reach behind," and "Hair & cosmetics," which had lower prediction rates that ranged from 92% to 93% in the AUC version 1 dataset. With regard to that conclusion, it is evident that the driver behaves in a manner that is less closely related to other items (such as the steering wheel, cup, and bottle) and that his or her ability to categorise them is less than that of other activities. This study also tested the speed of films with VGA, HD, and FHD resolutions using a video testing system with a trained model. For real-time deployment, a normal camera with a speaker can be used in place of the system's video inputs. However, this study only looks at driving safety in real-time simulation films. Figure 12 also shows how all devices' speeds decline when video resolution rises from VGA to HD and FHD.

As a result, utilising VGA resolution is suggested to offer the fastest processing possible and eliminate delay. In order to offer warning systems for autos, the network may be implemented on embedded systems and low-power computer devices.

With the provided image data, as was previously described, the suggested network functions superbly. However, occasionally, while playing with movies or live-stream videos, the internet exposed a number of shortcomings. These combine similar behaviours such as safe driving and texting while using just one hand to grip the steering wheel, all-two-handed and one-handed driving, wearing cosmetics and applying hair, drinking, etc.

With the provided image data, as was previously described, the suggested network functions superbly. However, occasionally, while playing with movies or live-stream videos, the internet exposed a number of shortcomings. These combine similar behaviours such as safe driving and texting while using just one hand to grip the steering wheel, all-two-handed and one-handed driving, wearing cosmetics and applying hair, drinking, etc. The lighting conditions have a considerable influence on the network's classification accuracy. Excessive illumination will make it difficult to tell foreground items from background ones. The camera angle is also a critical component for optimising the network and successfully integrating the aforementioned components to boost accuracy.

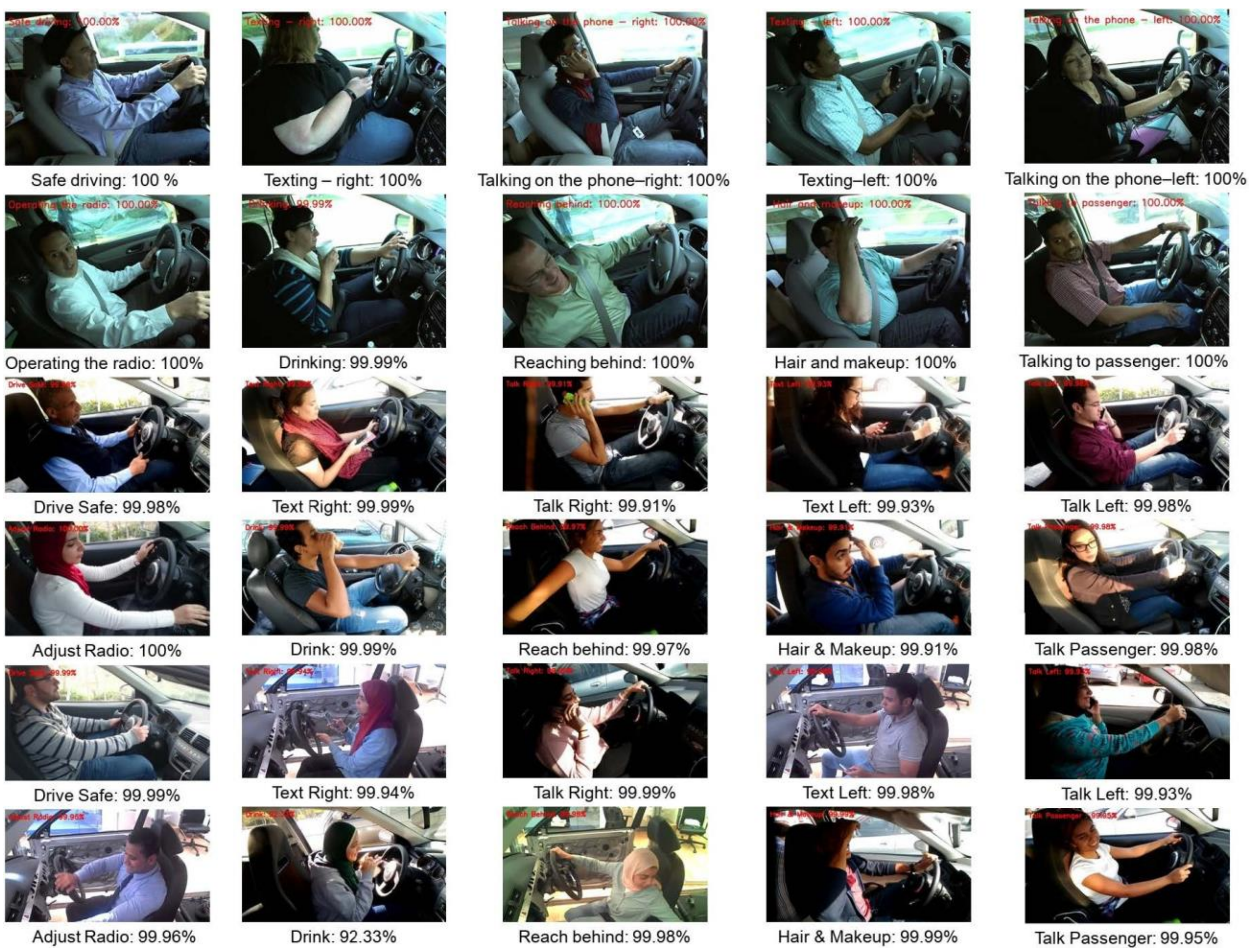

FIGURE 7. AUC version 1, version 2, and State Farm datasets' qualitative findings. The State Farm dataset appears in the top two rows, followed by the AUC version 1 dataset in the middle, and the AUC version 2 dataset in the bottom two rows.

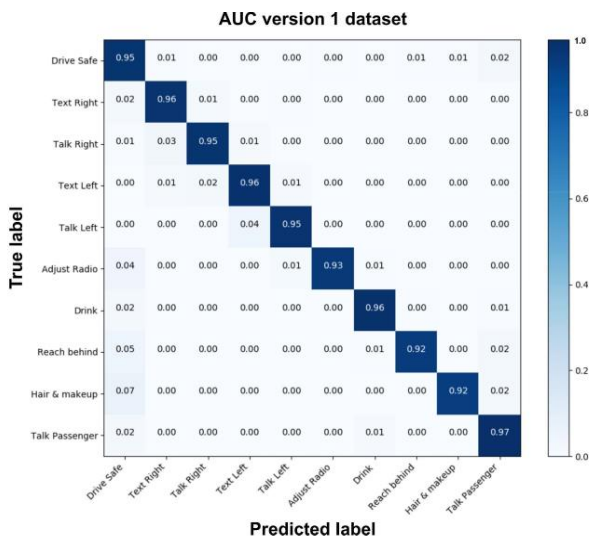

FIGURE 8. **On the State Farm dataset, there is a confusion matrix..**

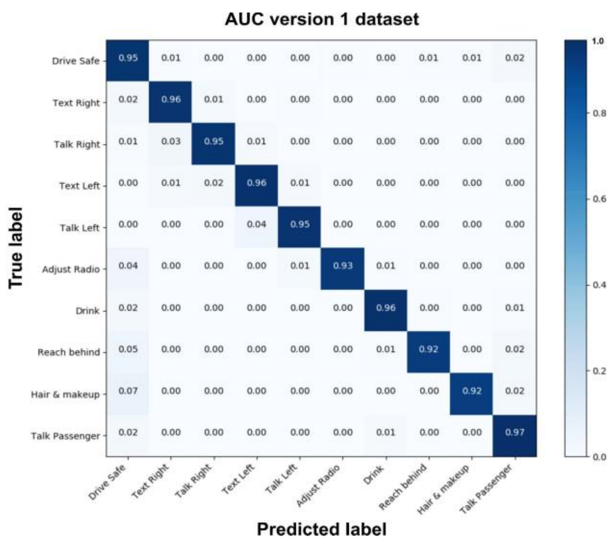

FIGURE 9. **The AUC version 1 dataset's confusion matrix.**

CONCLUSION

This research describes a driver behaviour recognizer that relies on a lightweight convolutional neural network and an attention mechanism. To extract feature maps, this architecture employs suggested adaptive connections, depthwise separable convolution operation, and ordinary convolution. The network then instructs the CBAM attention mechanism to focus on taking up the most crucial information. Finally, the classifier identifies ten driving behaviours. In this work, a number of different tactics were employed to boost the precision and quantity of network parameters. The proposed network made use of all three benchmarks for training, evaluating, and presenting the outcomes in the accuracy measure. However, movies with different resolutions and processing speeds were also used to test it. Future advancements of this method will be based on a two-stage driver behaviour warning system. The proposed network will be added into the system following the driver body detection stage. When using real-time applications, it is feasible to considerably increase the accuracy of the behaviour categorization by first extracting the driver's body postures.

REFERENCES

[1] Data from NCRB, Government of India, available at http://ncrb.nic.in/StatPublications/ADSI/ADSI2014/ADSI2014.asp

[2] "Internet of Things Global Standards Initiative". ITU. Retrieved 26 June 2015. - http://www.itu.int/en/ITU-T/gsi/iot/Pages/default.aspx

[3] Z. Xiaorong et al, ―The Drunk Driving Automatic Detection System Based on Internet of Things‖, International Journal of Control and Automation.

[4] J. Dai, J. Teng, X. Bai, Z. Shen, and D. Xuan. "Mobile phone based drunk driving detection." In 2010 4th International Conference on Pervasive Computing Technologies for Healthcare, pp. 1-8. IEEE, 2010

[5] A. R. Varma, S. V. Arote, C. Bharti, and K. Singh. "Accident prevention using eye blinking and head movement." IJCA Proceedings on Emerging Trends in Computer Science and Information Technology-2012 (ETCSIT2012) etcsit1001 4 (2012)

[6] V. Savania, H. Agravata and D. Patela , ―Alcohol Detection and Accident Prevention of Vehicle‖, International Journal of Innovative and Emerging Research in Engineering, Volume 2, Issue 3, 2015, pp 55-59

[7] http://en.wikipedia.org/wiki/Global_Positioning_System.

[8] http://www.dadss.org/breath-based-technology.

[9] http://www.dadss.org/touch-based-technology.

[10] S sahabiswas, S sourav, ―Drunken driving detection and prevention models using Internet of things‖, Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2016 IEEE 7th Annual, pp.1-16, IEEE,2016

[11] Open CV 2.4.11.0 documentation; Open CV Tutorials; objdetect module;Object Detection.

[12] Viola, Jones: Robust Real-time Object Detection, IJCV 2001.

[13] Open CV Computer Vision with Python; Tracking Faces with Haar Cascades

[14] S. Motorist, Driver Fatigue is an important cause of Road Crashes, Smart Motorist.

[15] Driver Fatigue and Road Accidents: A Literature Review and Position Paper

[16] Royal Society for the Prevention of Accidents, 2001.

[17] Road and Traffic Authority (RTA) Annual Report, Sidney, 2008.

[18] B. Praveenkumar, K Mahendrakan. "Prevention of Accident Due To Drowsy By Using Eye Blink." International Journal of Innovative Research in Science, Engineering and Technology(IJIRSET)Volume 3, May 2014.

[19] Stan, Ovidiu, Liviu Miclea, and Ana Centea. "Eye-Gaze Tracking Method Driven by Raspberry Pi Applicable in Automotive Traf¿c Safety. "Arti¿cial Intelligence,

Modelling and Simulation (AIMS), 2014 2nd International Conference on IEEE, 2014.

[20] Friedrichs, Fabian, and Bin Yang. "Camera-based drowsiness reference for driver state classi¿cation under real driving conditions." Intelligent Vehicles Symposium (IV), 2010,IEEE. IEEE, 2010.

[21] "Eye detection and recognition in the fatigue warning system." Intelligent Networks and Intelligent Systems (ICINIS), 2010 3rd International Conference on IEEE, 2010.

[22] Turan, Garima, and Sheifali Gupta. "Road Accidents Prevention system using

Driver‘s Drowsiness Detection." International Journal of Advanced Research in Computer Engineering and Technology (IJARCET).