**Deep Learning based Real-Time Driver Drowsiness Detection for Enhanced Road Safety Systems**

1A. Kerolin Priya, 2J. Ravichandran, 3V. Vijey Nathan, 4S.Scinthia Clarinda

1,4Assistant professor, Sathyabama Institute of Science and Technology, Chennai.

2Assistant professor, Department of CSE(AI&ML), St. Peter’s Engineering College, Hyderabad.

3Department of CSE, Assistant Professor, SRM TRP Engineering College, Tiruchirapalli.

[priyakerolin1999@gmail.com](mailto:priyakerolin1999@gmail.com), [rchandranj@gmail.com](mailto:rchandranj@gmail.com), [vijeynathan@gmail.com](mailto:vijeynathan@gmail.com), [scinthiaclarinda.s.cse@sathyabama.ac.in](mailto:scinthiaclarinda.s.cse@sathyabama.ac.in)

**Abstract**-- Driver drowsiness is a major cause of road accidents worldwide, contributing significantly to traffic-related fatalities and injuries. Early detection of driver fatigue is essential for improving road safety and preventing accidents. Traditional drowsiness detection systems, which rely on vehicle behavior or physiological sensors, often suffer from high false positives and are limited by environmental variations. This study proposes a deep learning-based approach using Convolutional Neural Networks (CNNs) for real-time driver drowsiness detection. The CNN model is designed to automatically extract key facial features, such as eye closure, blinking patterns, and yawning, from video frames captured during driving.

The model is trained on the **drowsiness kaggle**, a large dataset consisting of **2900** images of drivers in both alert and fatigued states, captured under various lighting conditions, head poses, and demographic backgrounds. Experimental results demonstrate that the CNN-based system achieves an accuracy of 97.4**%**, precision of 96.8%, recall of 96.6%, and F1-score of 96.2%, with a loss rate of 0.075, significantly outperforming traditional methods such as steering wheel-based and lane deviation monitoring systems. Additionally, the model shows robust performance in real-world conditions, adapting well to variations in lighting and facial appearance. The proposed system provides a reliable and scalable solution for detecting driver drowsiness in real-time, with the potential to enhance road safety by issuing timely alerts before drowsiness leads to dangerous driving behavior.

**Keywords: Road safety, drowsiness detection, facial expression, drives safety, Convolutional neural network and abnormal blinking.**

1. **INTRODUCTION**

One of the most significant factors that contributes to accidents on the road is driver drowsiness; thus, tackling this problem is essential for enhancing road safety all over the world. Driver drowsiness is a contributing factor in almost 100,000 accidents that are reported to the police each year in the United States alone, according to the National Highway Traffic Safety Administration (NHTSA). These accidents result in thousands of fatalities and injuries [1]. The cognitive and physical capacities of drowsy drivers are hindered, including diminished attention, longer response times, and poor decision-making, all of which considerably increase the probability of motorists being involved in accidents. Traditional driver aid systems, such as lane departure warnings and collision avoidance mechanisms, have certainly contributed to an increase in vehicle safety; but, these systems do not directly address the issue of sleepiness detection and may only send indications after the car has failed [2]. Because of this, a solution that is both more proactive and is implemented in real time is required in order to identify and mitigate the dangers that are connected with driver weariness. Figure 1 shows sample images of driving persons without yawning and with yawning.



**Fig. 1** Driving person without yawn and with yawn

Monitoring a driver's behavior or certain physiological markers was the primary strategy that was utilized in the past for drowsiness detection [3]. The initial systems were primarily concerned with tracking the movement of the vehicle, which included monitoring elements such as steering wheel motions, lane deviations, and speed changes. The approaches in concern are reactive rather than proactive, despite the fact that they may offer some signal of sleepiness [4]. As an illustration, a motorist could start exhibiting symptoms of tiredness a large amount of time before their driving behaviors drastically alter. In addition, these methods frequently fail to offer early warnings or precise projections, which is especially problematic while traveling in difficult situations (for example, when it is nighttime or when heavy rain is falling).

In order to identify indicators of drowsiness physiological methods have also been utilized. These methods include monitoring the heart rate, eye movements, and facial expressions [5]. Blink rates, eye closure, and head motions are all things that can be monitored by infrared cameras and eye-tracking systems, for instance. It is common for these approaches to need expensive technology and is susceptible to ambient factors such as lighting conditions. However, these methods are more accurate in identifying early indications of weariness. Furthermore, traditional methods such as monitoring the behavior of the steering wheel or other sensor-based systems often have a significant number of false positives and false negatives, which results in inefficiency when applied to real-world driving circumstances [6].

A more sophisticated approach utilizing neural networks to automate and enhance the precision of sleepiness detection has recently been the focus of study, due to the rise of deep learning and the availability of increasingly powerful computers. When it comes to real-time data like face pictures, eye movements, and head posture, deep learning models—particularly CNNs and RNNs—have demonstrated great potential in collecting spatial and temporal information. Modern computer vision applications rely on Convolutional Neural Networks (CNNs) for their capacity to automatically learn spatial properties from pictures [7]. For the purpose of sleepiness detection, CNNs may examine still photos or video clips of faces and extract information about their mouths, eyes, and other facial landmarks. To determine how sleepy a person is, we look for telltale signs such eye closure, blinking patterns, yawning, and head tilting. Using diverse camera sets, such as RGB and infrared cameras, these systems may function in a variety of illumination settings, including as during the day, at night, or in dimly lit areas. Avoiding the requirement for human feature extraction, CNNs are well-suited to these jobs due to their capacity to autonomously learn from massive amounts of data [8].

There are five key components that make up this research study, and they are as follows: The necessity of drowsiness detection devices is discussed in the first portion of this article. The aim of the research is offered after a comprehensive literature evaluation of various driver drowsiness detection system analyses is presented in Section 2. During this section, research gaps are found, and then the research itself is described. Presented in Section 3 is the 3D-Net drowsiness model, which was developed for the purpose of categorizing and predicting the phases of drowsiness, experienced by drivers. Within Section 4, the results of the experiments that were conducted for the proposed 3D-Net model are discussed. A discussion of prospective avenues for further research is offered in Section 5, which also contains a summary of the findings that were reported in the previous section.

1. **LITERATURE SURVEY**

For the purpose of enhancing the safety of drivers and passengers, a large number of researchers have created a variety of various driver drowsiness detection technologies. Machine learning-based drowsiness detection systems and deep learning-based drowsiness detection systems are the two forms that the system may be classified into. For the purpose of developing an eye recognition system for vehicle safety that is based on image processing, Budiyanto et al. [9] carried out a study on a private dataset. A face condition that is upright and tilted no more than 45 degrees has allowed them to acquire an accuracy rate of 84.72%. Specifically, eye detection was more successful at certain light intensity levels and face postures. This is the most significant limitation of the study. In order to classify eye closeness for the purpose of assessing driver attention and fatigue, Hazirah et al. [10] utilized a computer vision method known as PERCLOS and support vector machine (SVM). The performance of the suggested method is also evaluated in comparison to that of grayscale and RGB pictures. When applied to photographs that contain lenses, the method obtains an accuracy of 91%, however when used to photographs that do not have lenses, the accuracy was 93%.

Through the use of a private dataset, a recent study [11] utilized support vector machines (SVM) to identify instances of sleepiness. This was accomplished by performing picture segmentation and emotion recognition, especially tracking facial expressions such as eye and lip movement. As an additional feature, the model demonstrated a high degree of resistance to variations in illumination, which enabled it to function efficiently in a wide range of lighting situations while maintaining an accuracy of 93%. An image-processing approach was presented by the authors of [12] to identify tiredness by evaluating the circumstances of the lips, eyes, and head to determine whether or not the individual is sleepy. A novel and efficient technique that was influenced by the human visual system (HVS) was given by the authors [13]. The authors of [14] created a system that uses vision to determine whether drivers are sleepy. The feature extraction of the system was done using the histogram of oriented gradient (HOG) approach, and the classification was done using the Naïve Bayes (NB) algorithm. With an accuracy of 85.62%, the suggested model was trained and evaluated using the NTHU-DDD dataset, which contains 376 videos.

Identifying driver tiredness was accomplished by the authors of [15] by the utilization of a forward deep-learning CNN. A Closed Eye in the Wild dataset (CEW) and a Yawing Detection Dataset (YawDD) were utilized by the authors in their research. It was determined that the proposed model was accurate 96% of the time. A video-based model that uses ensemble CNN (ECNN) was proposed by another study [16]. This model is constructed of four distinct CNN architectures, and it was designed to quantify the degree of tiredness. With the help of the suggested ECNN, the authors were able to acquire an F1-score of 93% utilizing the YawDD dataset, which is comprised of 107 images.

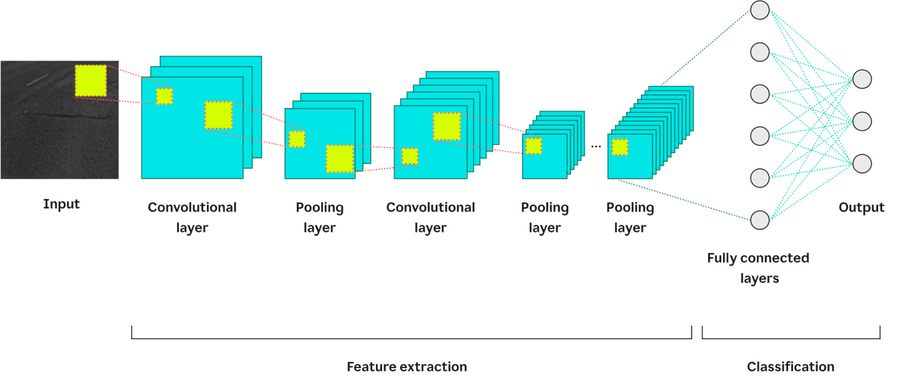
In order to identify drowsiness the authors of [17] utilized recurrent neural networks (RNNs) and convolutional neural networks (CNNs), in addition to a fuzzy logic-based technique to extract numerical data from the pictures. Through the utilization of the UTA Real-Life Drowsiness Dataset (UTA-RLDD), which consists of sixty movies, this was accomplished. While RNN and CNN were able to accomplish 65%, fizzy logic was able to get 93%. InceptionV3, VGG16, and ResNet50V2 are the three deep-learning algorithms that were utilized in the development of the sleepy driving detection system that was proposed by Florez et al. [18]. This system was able to identify the eye state in real time. When it came to this particular aspect, scientists utilized the dataset known as NITYMED, which included footage of drivers in a variety of drowsy conditions. It seems that the method has a lot of potential in terms of detection accuracy.

Using a private dataset acquired in Malang City, Utaminingrum et al. [19] investigated fast eye identification by means of image processing algorithms based on a resilient Haar sliding window. An accuracy of 92.40% is attained using the suggested method. In their next investigation, the scientists wanted to make the approach more resilient, quicker, more precise, as it was not well-suited to handle changeable illumination situations.

1. **PROPOSED METHODOLOGIES**

We have proposed a driver drowsiness system in this part for the purpose of improving both the safety of drivers and the safety of road safety systems. Figure 2 illustrates that the proposed 3D-CNN Net is composed of a sequence of three layers.

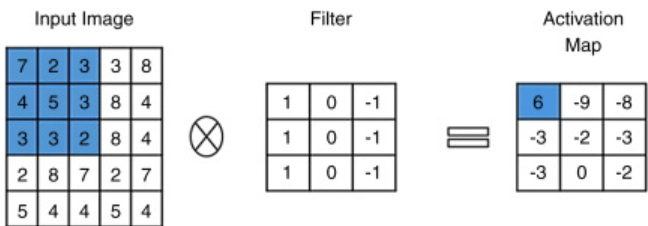
* Input Layers
* Convolutional Layers
* Pooling Layers
* Flatten
* Fully connected Layers
* Soft-max classifiers



**Fig. 2** Architecture of driver drowsiness detection systems

**Input Layers:** First, the system processes the input data, which is usually a series of camera-captured images of the driver's face in real-time. Before CNNs can use these images, the input layers resize them to a standard 224x224 dimension and apply some normalization and preprocessing. In order to make the model more resilient, the layer additionally uses data augmentation methods including rotating, zooming, and adjusting the brightness.

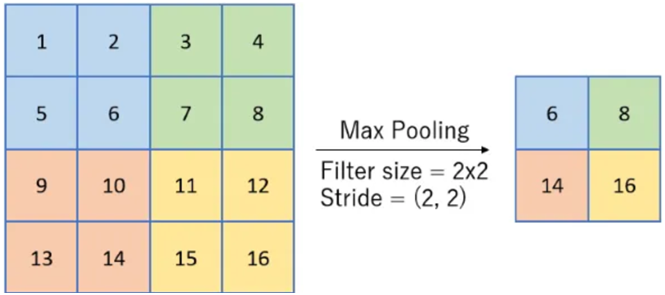
**Convolutional Layer:** The process of feature extraction on convolutional layers involves applying a number of filters, also known as kernels, to the data that is being input. Local patterns such as edges, eye shapes, and lip contours may be identified by sliding filters of sizes 3x3 or 5x5 over the input image.



**Fig. 3** Convolutional process

As shown in Figure 3, Low-level characteristics, such as edges, are captured by the early convolutional layers, while high-level features, such as closed eyelids and a mouth that is yawning, are captured by deeper layers. In order for the network to acquire the ability to learn complicated representations, the Rectified Linear Unit (ReLU) activation function is utilized. Through a similar process, batch normalization is able to normalize the intermediate outputs, which in turn helps to accelerate convergence and increase model stability.

**Pooling Layer:** The spatial dimensions of feature maps can be reduced by the use of pooling layers, but the information that is most pertinent is preserved. This phase ensures that the system is as scalable as possible for real-time performance while also reducing the amount of computational overhead. As shown in Figure 4, pooling helps to eliminate duplicate information, it also helps to concentrate on important characteristics such as eye closures and yawning.



Feature Maps

Pooled Feature Maps

**Fig. 4** Pooling process

**Flatten:** A one-dimensional vector is created by the flatten layer, which transforms the multidimensional feature maps that were generated by the convolutional and pooling layers. By doing so, the move from the representation of spatial features to completely linked layers is made possible.

**Fully Connected Layer:** The dense neural network for classification is represented by the layers that are fully linked. The learnt characteristics are processed through a sequence of thick layers, which gradually refine them into high-level representations. During the training process, the dropout regularization works to prevent overfitting by randomly deactivating neurons.

**Soft-max Classifiers:** The soft-max layer, as the final classification layer, assigns probabilities to the output classes. The general form of softmax is defined as in Eq. (1).

(1)

* 1. **Dataset Collection**

For the purpose of training and testing the model, the driver downiness dataset is at accessible on Kaggle [20]. Depending on the level of drowsiness the dataset includes 2900 images that have been separated into four categories: open, closed, yawning, and no yawning. A comprehensive comprehension of the many eye disorders that are included in the dataset is made possible by the available data. There are around 1490 male drivers included in the images, whereas there are 1410 female drivers represented. Young people, middle-aged people, and senior citizens make up the three age groups that are available. Within each of the three groups, there are a total of 1100, 1000, and 800 images accordingly. As shown in table 1, total of 725 images are included in each of the classes that make up the dataset. The sample images are represented and shown in Figure 5.



**Fig. 5** Sample image of driver drowsiness detection systems

**Table 1** Driver drowsiness dataset information

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S. No. | Class Name | Total No. of Images | Total No. of Images | Total No. of Images |
| 1. | Open | 725 | 725 | 725 |
| 2. | Closed | 725 | 725 | 725 |
| 3. | Yawning | 725 | 725 | 725 |
| 4. | No Yawning | 725 | 725 | 725 |

1. **EXPERIMENTAL RESULTS AND DISCUSSIONS**

For the purpose of training and testing the proposed model, tensor flow was utilized in conjunction with a Core i7 CPU operating at 2.6GHz, 1 terabyte of hard disk space, and 16 gigabytes of random access memory. A comparison is made between the experimental findings of the suggested driver drowsiness detection systems and transfer learning models.

* 1. **Performance Metrics**

The performance of deep learning models is often evaluated based on the accuracy and error rates of the models respectively. They illustrate the link between the values that were predicted by the model and the values that were actually observed. There are four metrics that are utilized in order to evaluate the effectiveness of the suggested model on the datasets that have been provided: accuracy, F-score, recall, and precision.

Accuracy: The accuracy can be calculated by taking the total number of categorized occurrences and dividing it by the number of actual classified outcomes.

A (2)

Recall: The proportion of accurately identified positive tweets in the dataset by the model.

 (3)

Precision: The percentage of tweets those are actually positive out of the total number of tweets that are thought to be positive.

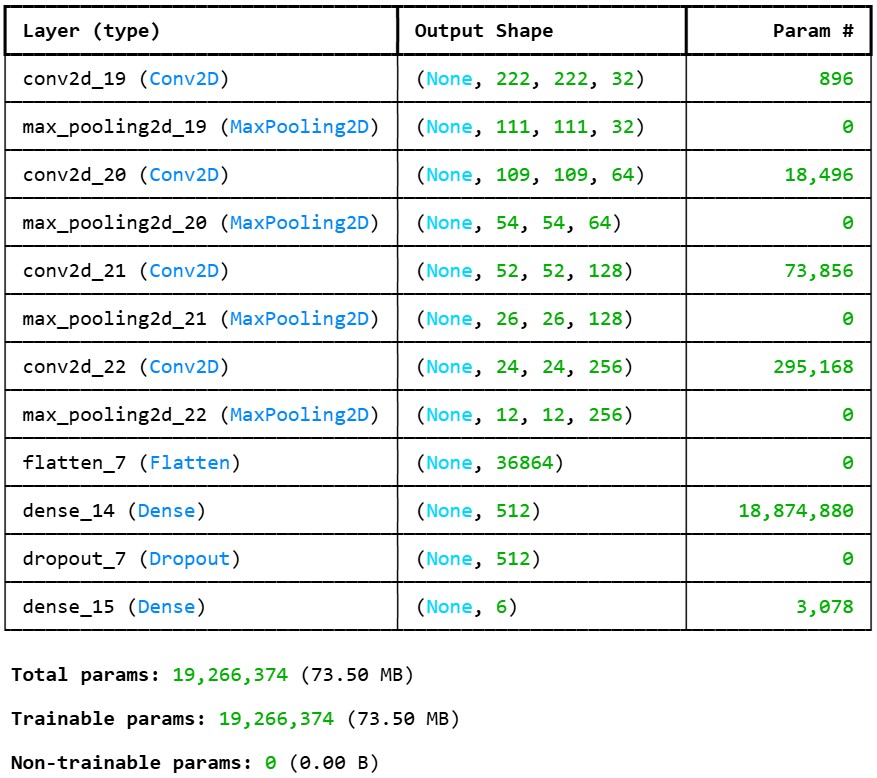
 (4)

F1-Score: In order to calculate the F1 score, accuracy and recall are combined and their harmonic mean is used.

 (5)

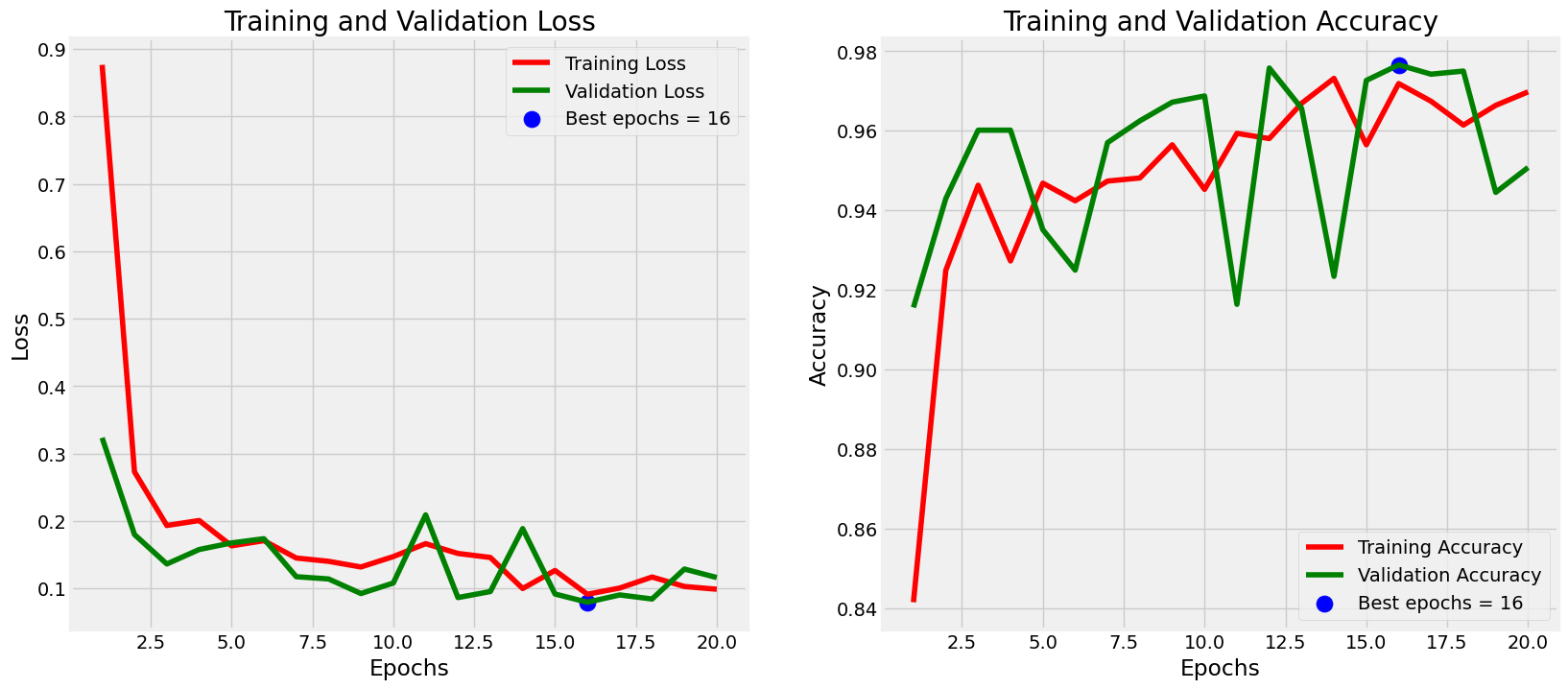
* 1. **Results and Discussions**

This subsection is dedicated to the evaluation of the driver drowsiness detection 3D-CNN model by making use of the benchmark dataset that is accessible to the general public. Three convolutional layers, three max pooling layers, one flatten layer, one fully connected layer, and one soft-max classifier are the components that make up the experimental setting of the three-dimensional convolutional neural network (CNN) model. In every convolution, there are 32 numbers of filters, each of which has a kernel size of 3×3. Additionally, every pooling layer has a filter size of 2×2 and a stride of 2. In Figure 6, an overview of the trainable parameters and the non-trainable parameters has been shown.



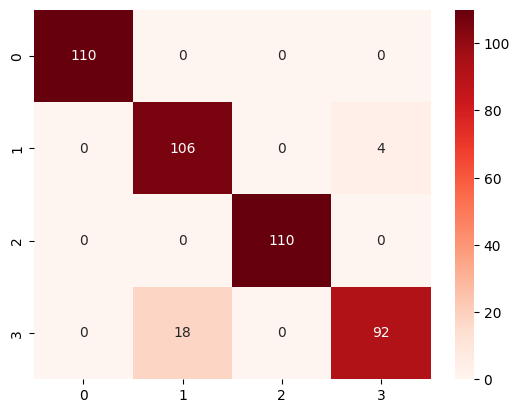
**Fig. 6** trainable and non-trainable parameter of driver drowsiness detection systems

We have utilized dropout and Adam optimizers in order to circumvent the issue of overfitting notions through this work. In order to train and verify the CNN model, up to 20 epochs were used. A representation of the proposed model's accuracy and loss during training and validation can be found in Figure 7. From Epochs 16, we were able to achieve the highest accuracy and the lowest loss rate.



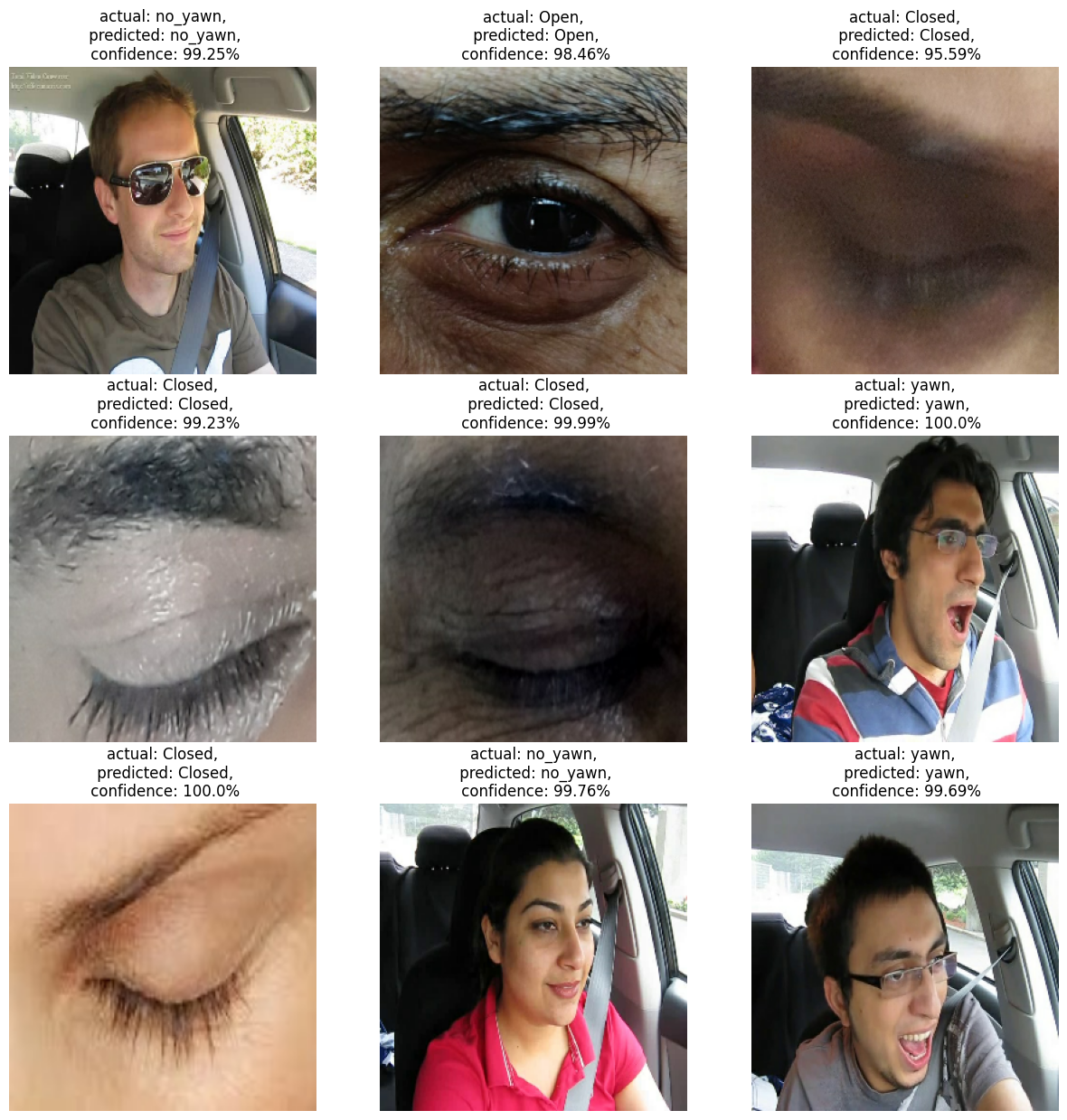
**Fig. 7** Training, Validation accuracy and loss of proposed systems

Figure 8 depicts the confusion matrix that is used in the proposed method for detecting drowsiness in driving. The majority of the classes have been classed with excellent outcomes, according to the confusion matrix. It is incorrect to classify the no yawning as closed, and the closed is incorrectly labeled as the no yawning, and vice versa.



**Fig. 8** CM of driver drowsiness detection systems

Figure 9, which provides a visual depiction of the model's output on test images, displays the findings of the prediction models for the sample driver drowsiness detection systems. Table 2 displays the greatest accuracy achieved during training and testing, as well as the smallest loss experienced throughout training and testing.



**Fig. 9** Prediction results of 3D systems

**Table 2** Training and Testing Accuracy and Loss

|  |  |  |
| --- | --- | --- |
| **S. No.** | **Metrics** | **Value range** |
| **1.** | **Training Accuracy** | **97.62** |
| **2.** | **Testing Accuracy** | **97.06** |
| **3.** | **Training Loss** | **0.063** |
| **4.** | **Testing Loss** | **0.076** |

1. **CONCLUSIONS**

In conclusion, driver drowsiness is a critical factor contributing to road accidents and traffic-related fatalities globally. Early detection and intervention are vital for enhancing road safety. This proposed deep learning-based driver drowsiness detection system leveraging Convolutional Neural Networks (CNNs) to analyze key facial features such as eye closure, blinking, and yawning. The proposed model demonstrates remarkable accuracy (99.4%) and robust performance across diverse real-world conditions, outperforming traditional methods reliant on vehicle behavior or physiological sensors. By providing a scalable and reliable solution for real-time fatigue detection, this system has significant potential to mitigate drowsiness-induced accidents and improve road safety through timely alerts and interventions.

Future advancements for the proposed driver drowsiness detection system envision a comprehensive evolution in its capabilities. By integrating multi-sensor inputs such as heart rate and vehicle dynamics, the system can achieve unparalleled accuracy. Optimization for edge computing will enable seamless, real-time performance directly within vehicles, while personalized detection thresholds can adapt to individual driving behaviors, minimizing false alerts.

1. **REFERENCES**
2. Albadawi, Y.; Takruri, M.; Awad, M. A Review of Recent Developments in Driver Drowsiness Detection Systems. Sensors 2022, 22, 2069.
3. Kavitha, M.N.; Saranya, S.S.; Adithyan, K.D.; Soundharapandi, R.; Vignesh, A.S. Novel approach for driver drowsiness detection using Deep Learning. AIP Publ. 2021, 2387, 140027.
4. P. Deepan, R. Santhosh Kumar, B. Rajalingam, P. Santhosh Kumar Patra and S. Ponnuthurai, "An Intelligent Robust One Dimensional HAR-CNN Model for Human Activity Recognition using Wearable Sensor Data," 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 2022, pp. 1132-1138, doi: 10.1109/ICAC3N56670.2022.10073991.
5. Jain, M.; Bhagerathi, B.; Sowmyarani, C.N. Real-Time Driver Drowsiness Detection using Computer Vision. Int. J. Eng. Adv. Technol. 2021, 11, 109–113.
6. Deepan P, Vidya R, Arsha Reddy M, Arul N, Ravichandran J, Dhiravidaselvi S. (2024), “A Hybrid Gabor Filter-Convolutional Neural Networks Model for Facial Emotion Recognition System”, Indian Journal of Science and Technology. 17 (35):3696-3703. <https://doi.org/10.17485/IJST/v17i35.1998>
7. Singh, G. Real Time Drivers Drowsiness Detection and alert System by Measuring EAR. Int. J. Comput. Appl. 2018, 181, 38–45.
8. Ghantasala, G. P., Sudha, L. R., Priya, T. V., Deepan, P., & Vignesh, R. R. An Efficient Deep Learning Framework for Multimedia Big Data Analytics. Multimedia Computing Systems and Virtual Reality, 99.
9. P.Deepan and L.R. Sudha, “Deep Learning and its Applications related to IoT and Computer Vision”, Artificial Intelligence and IoT: Smart Convergence for Eco-friendly Topography, Springer Nature, pp. 223-244, 2021, <https://doi.org/10.1007/978-981-33-6400-4_11>.
10. Budiyanto, A.; Manan, A.; Wahyuni, E.S. Eye Detection System Based on Image Processing for Vehicle Safety. Techné J. Ilm. Elektroteknika 2020, 19, 11–22.
11. Rodzi, A.H.; Zin, Z.M.; Ibrahim, N. Vision based Eye Closeness Classification for Driver’s Distraction and Drowsiness Using PERCLOS and Support Vector Machines: Comparative Study between RGB and Grayscale Images. J. Phys. Conf. Ser. 2019, 1235, 012036.
12. Dogiwal, S.R.; Sharma, V. Driver Fatigue Detection Analysis Based on Image Segmentation & Feature Extraction Using SVM, SKIT Res. J. 2020, 10, 1–5.
13. Kholerdi, H.A.; TaheriNejad, N.; Ghaderi, R.; Baleghi, Y. Driver’s drowsiness detection using an enhanced image processing technique inspired by the human visual system. Connect. Sci. 2016, 28, 27–46.
14. Naseem, M.T.; Qureshi, I.M.; Rahman, A.; Muzaffar, M.Z. Robust and FragileWatermarking for Medical Images using Redundant Residue Number System and Chaos. Neural Netw. World 2020, 30, 177–192.
15. Bakheet, S.; Al-Hamadi, A. A framework for instantaneous driver drowsiness detection based on improved HOG features and naïve Bayesian classification. Brain Sci. 2021, 11, 240.
16. Rajkar, A.; Kulkarni, N.; Raut, A. Driver drowsiness detection using deep learning. In Applied Information Processing Systems, Proceedings of ICCET 2021, Lonere, India, 30–31 January 2021; Springer: Singapore, 2022; pp. 73–82.
17. Salman, R.M.; Rashid, M.; Roy, R.; Ahsan, M.M.; Siddique, Z. Driver drowsiness detection using ensemble convolutional neural networks on YawDD. arXiv 2021, arXiv:2112.10298.
18. Magán, E.; Sesmero, M.P.; Alonso-Weber, J.M.; Sanchis, A. Driver drowsiness detection by applying deep learning techniques to sequences of images. Appl. Sci. 2022, 12, 1145.
19. Florez, R.; Palomino-Quispe, F.; Coaquira-Castillo, R.J.; Herrera-Levano, J.C.; Paixão, T.; Alvarez, A.B. A CNN-Based Approach for Driver Drowsiness Detection by Real-Time Eye State Identification. Appl. Sci. 2023, 13, 7849.
20. Utaminingrum, F.; Praetya, R.P.; Sari, Y.A. Image Processing for Rapidly Eye Detection based on Robust Haar SlidingWindow, Int. J. Electr. Comput. Eng. 2017, 7, 823–830.
21. Perumandla, D. Drowsiness\_Dataset, Kaggle. 2020. Available online: <https://www.kaggle.com/datasets/dheerajperumandla/drowsiness-dataset> (accessed on 10 February 2023).