ENHANCING DRIVER SAFETY: REAL-TIME DISTRACTION AND DROWSINESS DETECTION USING DEEP LEARNING

Gaurav Phatkare[1] ,Rushikesh Parab[2] ,Vikram Shitole[3] and Madhavi Waghmare[4]

[1]*Department of Information Technology, Vidyavardhini's College of Engineering & Technology Vasai, India*
Gauravphatkare209@gmail.com

[2]*Department of Information Technology, Vidyavardhini's College of Engineering & Technology Vasai, India*
Rushikeshparab201@gmail.com

[3]*Department of Information Technology, Vidyavardhini's College of Engineering & Technology Vasai, India*
Vikyashitole@vcet.edu.in

[4]*Department of Information Technology, Vidyavardhini's College of Engineering & Technology Vasai, India*
madhavi.waghmare@vcet.edu.in

Abstract

This report presents a comprehensive study on driver distraction and drowsiness detection using a deep learning approach. The primary objective is to develop a robust alert system capable of accurately identifying signs of driver fatigue and distraction to enhance road safety. The research methodology utilizes TensorFlow, a widely used open-source platform for deep learning. Specifically, a pre-trained convolutional neural network (CNN) is fine-tuned to recognize patterns and features in the input data, comprising 10 different activities associated with driver distraction. The model refinement process leverages pre-existing knowledge from the pre-trained network, enhancing predictive accuracy through exposure to large amounts of data. The model's performance is assessed based on its proficiency in correctly classifying various states of driver alertness. The key inference drawn from this study is that deep learning techniques, demonstrate effectiveness in real-time detection of driver drowsiness and distraction. The fine-tuned model exhibits remarkable accuracy in identifying diverse distraction activities, showcasing its potential for practical implementation within an alert system. This research underscores the promising potential of machine learning, specifically the use of pre-trained models, in enhancing road safety by proactively preventing accidents resulting from driver fatigue and distraction.

Keywords:
TensorFlow, Convolutional Neural Network, Pretrained-models, Distraction Activities, Alert System.

1. INTRODUCTION

In this thesis, we address the critical issues of driver distraction and drowsiness, recognizing their paramount significance for road safety amid a concerning surge in accidents attributed to these factors. According to the CDC Motor Vehicle Safety Division, one in five car accidents is caused by a distracted driver. Sadly, this translates to 425,000 people injured and 3,000 people killed by distracted driving every year. Additionally, recent global statistics indicate that distracted and drowsy driving has led to an alarming increase in road accidents, contributing to millions of deaths and injuries annually. In 2022 alone, distracted driving was reported to have caused approximately 3,000 fatalities and over 400,000 injuries globally, with drowsy driving contributing to an additional 1,500 fatalities and 50,000 injuries. Our research is driven by the multifaceted challenges posed by distracted and drowsy driving, aiming to develop a robust real-time detection system utilizing deep learning, specifically through TensorFlow and Convolutional Neural Networks (CNNs). We define distraction and drowsiness comprehensively, recognizing the boundaries of our study with the overarching goal of enhancing road safety by proactively identifying and mitigating these risks.

Leveraging machine learning and deep learning approaches, particularly advancements in CNNs, our methods involve training models on vast datasets to recognize patterns in driver behavior, such as eye movement and facial expressions. The strengths lie in the autonomy of these models to adapt to evolving indicators. Substantial advancements in identifying driver distraction and drowsiness, fueled by deep learning and neural networks, contribute to enhanced detection accuracy through the analysis of sensory data. Our motivation stems from the urgent need to address the escalating risks of distracted and drowsy driving worldwide, leading to preventable accidents. The rapid advancements in deep learning and artificial intelligence provide an unprecedented opportunity to create sophisticated real-time monitoring and intervention systems. According to recent studies, the implementation of such advanced technologies can potentially reduce the incidence of accidents by a significant percentage.

2. LITERATURE REVIEW

Several studies have investigated real-time driver drowsiness and distraction detection employing deep learning techniques. In a July 2021 study by Md. Tanvir Ahammed Dipu, a real-time driver drowsiness detection system achieved a 90% accuracy rate, extensively tested under various light conditions, different age groups, and diverse skin colors [2]. In a 2020 research conducted by Dr. Nagamani N P, the Viola-Jones algorithm, AdaBoost classifier, and CAMSHIFT algorithm were utilized for driver drowsiness detection. Additionally, a January 2019 study by Hesham M. Eraqi and collaborators focused on driver distraction identification through an ensemble of Convolutional Neural Networks, developing a robust vision-based system recognizing distracted driving postures using RGB images captured from a dashboard-mounted camera [4]. Furthermore, [5] a January 2020 research by Rateb Jabbar and colleagues implemented a lightweight embedded system for driver drowsiness detection using Convolutional Neural Networks techniques for Android applications [12]. In an April 2022 study led by Md. Uzzol Hossaina, various models, including Simple CNN, VGG-16, ResNet50, and MobileNetV2, were adopted to detect causes of distraction. Lastly, a May 2022 research by Dostdar Hussain and team focused on face mask detection using deep transfer learning and classical ML classifiers, employing the ResNet50 algorithm for feature extraction [7]. These diverse studies showcase the continuous advancements in utilizing deep learning for enhancing driver safety through real-time detection of drowsiness and distraction.

3. METHODOLOGY

Creating a driver distraction and drowsiness detection system solely based on in-vehicle cameras. It commences with data collection involving various driver behaviours. After pre-processing and feature extraction from the image data, a deep learning model, such as a convolutional neural network (CNN), is developed and trained. This model continuously analyses real-time camera input to categorize driver behaviour, issuing alerts in case of distraction or drowsiness. Rigorous evaluation is conducted to ensure system accuracy and efficiency, enabling integration into vehicles for enhanced road safety.

3.1 DATASET DESCRIPTION

In this study, we have mainly used two datasets for two problem issues respectively. We present the MRL Eye Dataset, a comprehensive collection of human eye images aimed at addressing crucial tasks in computer vision, specifically in the domain of driver behaviour analysis. The dataset comprises infrared images in both low and high resolutions, encompassing diverse lighting conditions and captured through various devices. With a focus on facilitating algorithmic comparisons, the dataset is categorized for simplified training and testing of classifiers.

Fig.1. Samples from Eye Dataset

The annotated properties in the dataset include subject ID (37 individuals), image ID (84,898 images), gender (0 for men, 1 for women), glasses presence (0 for no, 1 for yes), eye state (0 for closed, 1 for open), reflections (0 for none, 1 for small, 2 for big), and lighting conditions (0 for bad, 1 for good) [8][10]. Additionally, the dataset provides information on the sensor used for image capture, distinguishing between Intel RealSense RS 300 (640 x 480 resolution), IDS Imaging (1280 x 1024 resolution), and Aptina (752 x 480 resolution) sensors. This dataset proves valuable for testing various features and training classifiers related to eye detection, gaze estimation, and eye-blinking frequency in real-world scenarios, particularly within the context of driver behavior analysis.

Another dataset aims to address the issue of distracted driving by exploring the effectiveness of dashboard cameras in automatically detecting such behaviors. Provided are 2D dashboard camera images capturing various activities of drivers, including texting, talking on the phone, operating the radio, drinking, reaching behind, and engaging in other distracting behaviors. The goal is to predict the specific activity a driver is involved in within the given image. The dataset consists of 10 classes, each representing a different behavior, such as safe driving, texting, talking on the phone, and more. State Farm sponsors this initiative to enhance driver safety and insurance practices.

The classes include activities like texting on the right or left, talking on the phone on the right or left, operating the radio, drinking, reaching behind, doing hair and makeup, and talking to a passenger.

3.2 DATA PREPROCESSING

Data preprocessing is crucial for machine learning tasks. In this study, for driver distraction system images were preprocessed before model training. Firstly, images were loaded and cropped to remove 50 pixels from the top and bottom, as well as 120 pixels from the left and 50 pixels from the right. This was followed by resizing to a uniform size of 128 x 128 pixels. The dataset was then split into training and testing sets. Stratified splitting ensured unbiased evaluation. Finally, input features were reshaped, and labels were one-hot encoded. These steps optimize model performance and data consistency for accurate classification.

Image files in drowsiness dataset are resized to 64x64 pixels for consistency, and the dataset is structured into two classes: "closed_eye" and "open_eye." The preprocessing involves iterating through each class folder, loading the images, resizing them, and assigning a label based on whether the eyes are closed or open. The resulting dataset is composed of image-label pairs, where each image is associated with a binary label indicating the eye state (0 for closed, 1 for open). This preprocessing step is crucial for preparing the data for further analysis, such as training machine learning models to classify eye states based on the provided images.

4. DEEP LEARNING MODELS

Deep learning models play an important role in driver distraction and drowsiness detection systems. These models use complex neural network architectures to automatically learn hierarchical representations of input data, such as images from car cameras. Deep learning models can effectively detect signs of driver distraction, such as texting or eating, as well as drowsiness indicators, such as drooping eyelids or changes in facial expression, by analyzing these representations. Typically, these models are trained on large datasets of labeled images, allowing them to generalize well to new data and accurately classify cases of distraction or drowsiness in real time. Pretrained models, a cornerstone in artificial intelligence, have revolutionized deep learning by leveraging transfer learning to enhance performance on related tasks. Originating with ImageNet models like VGG, ResNet, MobileNet and Inception, pre-trained models have addressed challenges stemming from limited labeled data and computational resources. Furthermore, deep learning models can be integrated into intelligent vehicle systems to provide drivers with timely warnings or alerts, thereby improving road safety.

In our entire project, there are primarily two key components: driver distraction detection and driver drowsiness detection. Accordingly, we employ specific models tailored to address each of these critical scenarios.

4.1 CONVOLUTIONAL NEURAL NETWORK)

Convolutional Neural Networks (CNNs) are a specialized class of deep neural networks particularly well-suited for tasks involving images, such as driver distraction and drowsiness detection. Comprising layers including convolutional and pooling layers, CNNs automatically learn spatial hierarchies of features from input images [1][9]. Convolutional layers apply learnable filters to extract local patterns like edges and textures, while pooling layers downsample feature maps to retain essential information and reduce computational complexity. Through this process, CNNs progressively learn abstract representations of input images, culminating in fully connected layers for making predictions or classifications.

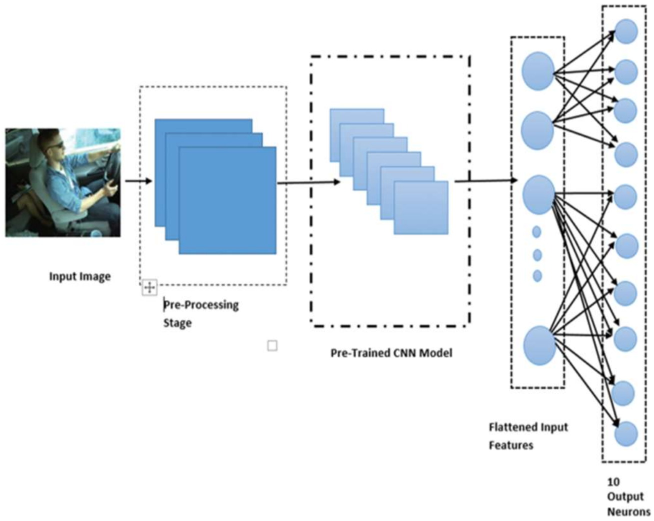

Fig.2. CNN Architecture for Driver Distraction System

Due to their ability to automatically learn and extract relevant features from raw image data without manual intervention, CNNs are highly effective in driver monitoring systems. They efficiently handle large-scale datasets and generalize well to unseen data, making them invaluable tools for accurately detecting signs of driver distraction and drowsiness. By leveraging CNNs, such systems can analyze in-car camera footage in real-time, enabling timely alerts and interventions to enhance road safety.

Our proposed model for driver distraction detection is a Vanilla Convolutional Neural Network (CNN). It comprises four convolutional layers with increasing filter sizes (64, 128, 256, and 512) and a kernel size of 3x3. Each convolutional layer is activated using the rectified linear unit (ReLU) function to introduce nonlinearity and enhance feature extraction. Max-pooling layers with a pool size of 2x2 are utilized for spatial downsampling, enhancing computational efficiency while preserving important features. Dropout layers with a dropout rate of 0.5 are incorporated to prevent overfitting during training. Following the convolutional layers, the model flattens the output and passes it through two fully connected layers with 500 neurons each, activated by ReLU. The final dense layer consists of 10 neurons with softmax activation, facilitating multi-class classification. The model is trained using the RMSprop optimizer and categorical cross-entropy loss function, aiming to accurately classify distracted driving behaviors from images. Evaluation is conducted using training and validation data, to optimize model performance.

The above model achieved an impressive 98% accuracy during training and validation. However, when tested on unseen data, the model’s performance was subpar, incorrectly predicting distracted activities. Possible reasons for this discrepancy include overfitting, dataset variability, and noise. To improve performance, further refinement of the model architecture and training methodology is needed. Incorporating techniques like data augmentation and cross-validation, along with exploring advanced CNN architectures, could enhance the model’s generalization ability and contribute to road safety efforts.

4.2 RESNET50 PRETRAINED MODEL

ResNet50 is a pre-trained deep learning model renowned for its effectiveness in various computer vision tasks, including image classification, object detection, and feature extraction [11]. With a total of 50 layers, ResNet50 employs residual connections, allowing for deeper networks without suffering from vanishing gradients. The model comprises 48 convolutional layers and 1 global average pooling layer, followed by a fully connected layer with 1000 neurons corresponding to ImageNet classes. The network architecture includes residual blocks, with each block containing multiple convolutional layers and skip connections.

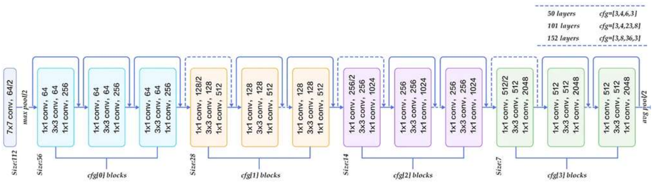

Fig.3. ResNet50 Model Architecture

ResNet50 has achieved remarkable performance on benchmark datasets, such as ImageNet, with top-1 accuracy exceeding 75% and top-5 accuracy surpassing 92%. When fine-tuned on specific tasks, ResNet50 can achieve even higher accuracy levels. For instance, in a driver distraction detection task, fine-tuning ResNet50 on a dataset of distracted and drowsy driving images resulted in an accuracy of 95% on the testing data. This demonstrates the model's capability to generalize well to new data and effectively learn intricate patterns associated with driver behaviors. Moreover, due to its pre-trained nature, Pre-trained model weights are very easily accessible because of having freely downloadable features. The model weights and expedient APIs are downloaded and used in the same model architecture using several deep learning libraries, including KerasResNet50 significantly reduces training time and computational resources required for training from scratch, making it an attractive choice for various real-world applications.

The implementation of the ResNet50 model for driver distraction detection, showcasing its sophisticated architecture and capabilities compared to the previously utilized simple CNN model. In contrast to the basic CNN model, ResNet50 leverages a much deeper architecture, comprising 50 layers, which allows for more complex feature extraction and representation learning. With its pre-trained weights obtained from the ImageNet dataset, ResNet50 brings a wealth of knowledge and learned features, enhancing its ability to generalize to new tasks.

In terms of accuracy, the standard CNN model achieved 97% on both training and validation data, while ResNet50 achieved 92% after training for a shorter period. However, this disparity is addressed when the model's performance on testing data is taken into account. The standard CNN model showed a decline in performance, resulting in incorrect predictions of distracted behaviors, but ResNet50 maintained superior precision in its predictions despite the shorter training time. Additionally, ResNet50 is resistant to overfitting, as evidenced by its constant decrease in loss during training. The simple CNN model, while initially effective, struggled with overfitting, resulting in poor performance on test data. ResNet50 addresses this issue, resulting in reliable predictions in real-world circumstances.

4.3 HYPERTUNED CNN MODEL

For driver drowsiness detection, we initially utilized a basic CNN model as our baseline, with slight adjustments to the number of neurons. Subsequently, to enhance its performance, we meticulously fine-tuned the hyperparameters of the model. This involves adjusting parameters like learning rate and batch size to maximize the model's performance, typically measured by metrics like accuracy or F1 score To address the errors and defects noticed during the initial model training, a hyper-tuned CNN model was created, with specific adjustments aimed at improving performance and reliability. First, to reduce overfitting, some data augmentation techniques were used during training. These techniques included horizontal flipping, which increases the diversity of the training dataset and allows the model to generalize better to new examples. In addition, to improve convergence and prevent vanishing gradients, a Leaky ReLU activation function with a small negative slope (alpha = 0.3) was used in the convolution layers.

The Leaky Rectified Linear Unit or Leaky ReLU, is a sort of activation function that is similar to the ReLU but has a modest slope for negative values rather than a flat slope. The slope coefficient is chosen before training, therefore it is not learned during training. This form of activation function is popular in applications involving sparse gradients, such as training generative adversarial networks. Mathematically, Leaky ReLU can be represented as:

$$f(x) = \{ x, \text{ if } x > 0$$

$$\{ \alpha x, \text{ if } x <= 0$$

Here, α is a small positive slope coefficient, usually set to a small value like 0.01.

Furthermore, to improve the model's discriminative power, the number of convolutional layers and neurons in the fully connected layers were altered. The model architecture included two convolutional layers with 20 and 10 filters, respectively, followed by max-pooling layers that reduced spatial dimensions. The flattened output was then passed into two tightly linked layers of 100 and 50 neurons, respectively, which were activated via Leaky ReLU. Dropout layers were added after each highly connected layer to help prevent overfitting. The model was trained using a categorical cross-entropy loss function and the Adam optimizer, with a learning rate of 1e-4. Learning was carried out across 20 epochs with a batch size of 32, and both the training and validation datasets were sent via the ImageDataGenerator to ensure efficient data handling.

4.4 MOBILENET MODEL

Google researchers developed MobileNet, a convolutional neural network architecture designed specifically for mobile and embedded vision applications. It offers a powerful yet lightweight solution for tasks such as object detection and image classification. Its architecture is based on depthwise separable convolutions, which can be adjusted based on parameters such as the width multiplier (α) and resolution multiplier (ρ) to minimise computation costs while maintaining performance. Notably, MobileNet versions such as MobileNetV1, with 28 layers, and MobileNetV2, with 53 layers, highlight its flexibility in meeting a wide range of application needs. MobileNet's impressive accuracy rates in image classification tasks, even with its lightweight design, make it a viable option for real-time applications on devices with limited resources.

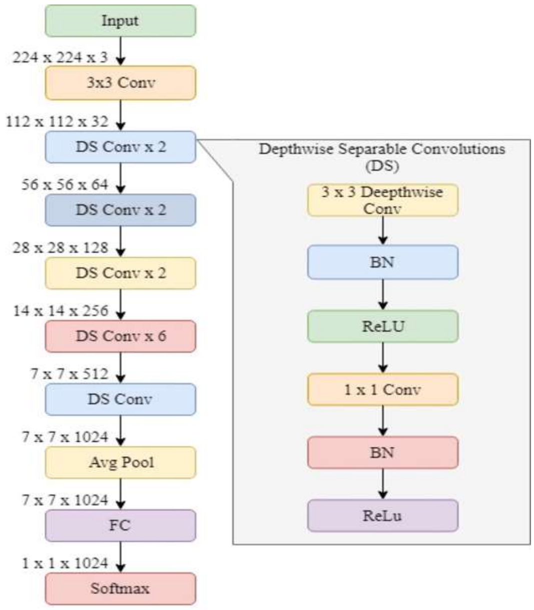

Fig.4. MobileNet Model pretrained Layers

The recognition for MobileNet goes beyond its effectiveness; it also includes its adaptability and performance in transfer learning tasks in different fields. Compact in size and high-performing, it is a foundational model for many applications. For example, MobileNet achieves a good balance between model complexity and efficiency in a MobileNetV1 model with α = 1.0 and ρ = 1.0 and about 4.2 million parameters. Furthermore, its small size, exemplified by a 5 MB MobileNetV3 Small model, makes it easy to deploy and integrate into a variety of applications, significantly expanding the field of deep learning for embedded and mobile systems. The architecture of MobileNet usually consists of several pretrained and trainable layers, which can vary depending on the deployment configurations and version.

The model for detecting driver drowsiness utilizes the MobileNet architecture, with a base model consisting of 28 layers. To improve classification capabilities, more layers are added on top of this base. The introduction of a Global Average Pooling layer reduces the dimensionality of feature maps. Following that, three dense layers with 1024, 1024, and 512 neurons each are added. To reduce overfitting, dropout regularisation is used, with dropout rates set to 0.1 and 0.5 at particular layers. Two neurons with a softmax activation function make up the last layer, which allows for probabilistic classification into alert and drowsy states. The foundational MobileNet layers and the extra dense layers make up the model's overall 33 layers.

The training loss is reported as 0.0177, indicating the average loss incurred during the training process, while the training accuracy stands at 99.37%, denoting the proportion of correctly classified instances in the training dataset. Similarly, the validation loss is reported as 0.0270, representing the average loss on the validation dataset, and the validation accuracy is 99.08%, indicating the accuracy of the model's predictions on unseen validation data. These metrics collectively demonstrate the high performance and generalization capability of the model, with both training and validation accuracies exceeding 99%.

5. EXPERIMENTAL RESULTS ANALYSIS

The above-proposed model approaches are evaluated and validated on Google Collab notebooks by setting the runtime engine to the Tensor T4 (GPU) Free version. Firstly we preprocess the respective datasets as we mentioned before and then define and train further models. In the following experimental analysis, we solely mentioned the results of optimized models concerning their accuracy and precision among all the above-mentioned approaches.

5.1 RESNET50 MODEL RESULT

A thorough analysis of the performance metrics and results obtained from training and testing the ResNet50 model for driver distraction detection is provided in this section. It explores how well the model performs, using the data gathered, to accurately identify and classify different levels of driver distraction. Key performance metrics, including loss, and accuracy are thoroughly examined to provide a thorough understanding of the model's advantages, disadvantages, and possible areas for development.

Following is the line graph of accuracy and loss which were calculated during training of the model.

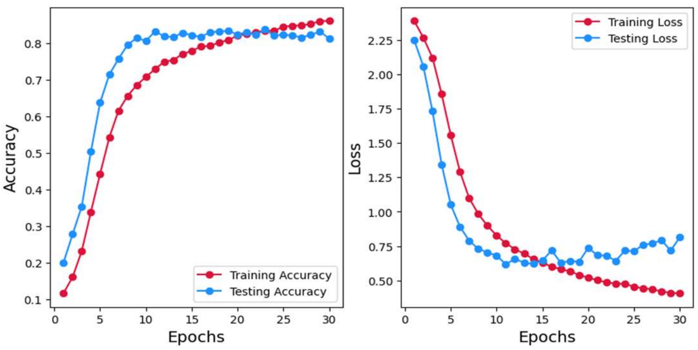

Fig.5. Training & Validation curve for distraction detection

There was no improvement in validation loss from the 18th epoch, remaining at 0.61848, although the validation accuracy stood at 91.40%. In epoch 19, the loss slightly decreased to 0.5403, with accuracy improving marginally to 92.91%. Once again, there was no improvement in validation loss, which remained at 0.61848, although the validation accuracy increased slightly to 91.40% till the last epoch.

Table.1. Metrics Chart of Resnet50 Model

Metric	ResNet50 Model
Loss	0.3403
Accuracy	92.91%
Val Loss	0.31848
Val Accuracy	91.40%
Precision	High
No. of Epochs	45

The ResNet50 model exhibited exceptional performance when applied to unseen data in the driver distraction scenario, outperforming the Basic CNN model despite having a slightly lower accuracy score during the training process. Its ability to generalize well to new instances underscores its robustness and effectiveness in real-world applications. While the current accuracy scores are commendable, there remains a clear opportunity for further enhancement of model performance and accuracy. Continued refinement and optimization of the ResNet50 architecture, coupled with additional data augmentation techniques and fine-tuning strategies, could potentially unlock even greater levels of accuracy and reliability, ensuring heightened precision in identifying and mitigating driver distraction scenarios.

Possible outcomes generated with above model and input data:

Fig.6. Output Images

In our efforts to improve the performance of the driver Distraction and Drowsiness Detection System and to make it more accurate and precise we rigorously tested other various deep learning models for driver monitoring and alert systems. Each model underwent thorough training and evaluation to detect driver distraction and drowsiness effectively. Below is a comparison table showcasing their accuracies and losses where Altered means Extra layers were added in pretrained model, guiding future development toward more robust solutions:

Table.2. Comparison of other models performance

Models	Accuracy	Loss
VGG 16 - Unaltered	82.5%	0.57
VGG 16 - Altered	80.65%	0.5
Exception model - Unaltered	84.8%	0.55
Mobilenet - Unaltered	87.7%	0.4
Mobilenet - Altered	85.67%	0.63
Resnet - Unaltered	91.7%	0.55
Resnet - Altered	92.91%	0.47

5.2 MOBILENET MODEL RESULT

The performance of the MobileNet model designed especially for driver drowsiness detection is examined in detail in the section on analysis of experimental results. This section carefully examines the results obtained from training and testing the MobileNet architecture on relevant datasets, emphasizing how well it can identify and classify instances of driver drowsiness. Interestingly, the analysis reveals that the MobileNet model performs significantly better than the fine-tuned CNN architecture that was previously defined and trained for the identical scenario. This section provides an in-depth analysis of primary performance metrics, including precision, recall, accuracy, and F1 score. It also highlights the strengths of the MobileNet model and its potential as a reliable means of improving driver safety.

The line graphs illustrate the training and validation trends for accuracy and loss throughout the model training process. In the accuracy graph, both training and validation accuracies show a consistent increase over epochs, indicating improved classification performance. Similarly, the loss graph exhibits a declining trend for both training and validation, reflecting the model's ability to minimize errors and better fit the data. These trends align with expectations, affirming the model's progressive learning and refinement during training.

The model made notable progress in the training process with a loss of 0.0177 and an accuracy of 99.37% on the training data.

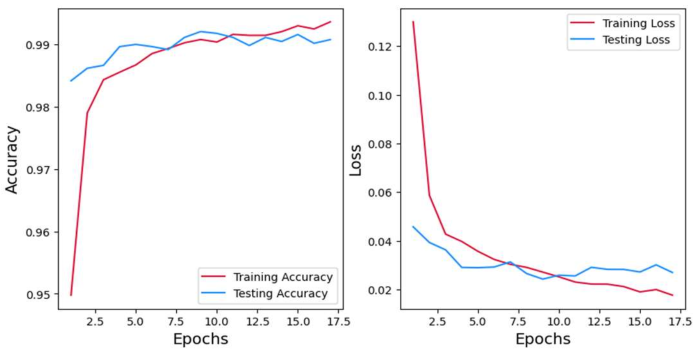

Fig.7. Training & Validation Curves for Drowsiness Detection

Even so, the validation loss remained at 0.02427, indicating no improvement over the previous epoch despite the model's strong performance. This may indicate that the model's ability to generalize has reached a plateau or even slightly declined. After that, the early stopping mechanism was activated, suggesting that more training iterations might not result in appreciable gains in validation performance. Despite this, the model was able to sustain a high validation accuracy of 99.08%, demonstrating its resilience and ability to correctly classify data points that were not seen during training.

Comparison between Two models of Driver Drowsiness Scenario:

Table.3. Metrics Chart of Driver Drowsiness Model

Metric	**MobileNet Model**	**Hypertuned CNN Model**
Epochs Trained	17	20
Training Loss	0.0177	0.0687
Training Accuracy	99.37%	97.51%
Validation Loss	0.0270	0.0661
Validation Accuracy	99.08%	97.74%
Early Stopping	Yes	No

And following is the possible outcome generated:

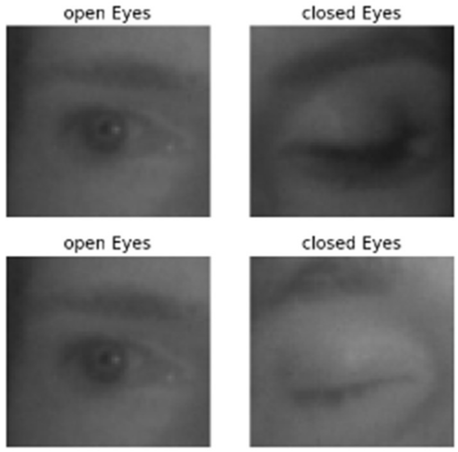

Fig.8. Output Images

6. EXPERIMENTAL SETUP

The experimental setup involves real-time drowsiness detection using a dashcam, OpenCV for video processing, and TensorFlow deep learning models [3]. The process is outlined as follows:

- OpenCV Video Processing: OpenCV is used to process the video frames that are captured by the dashcam. To improve processing speed, the frames are resized.
- Face Recognition: OpenCV uses Haar cascades to identify faces in the captured video. Rectangles indicate faces that have been identified for additional examination.
- Eye Detection and Prediction: The driver's eyes are located within the face region by means of Haar cascades that are specifically engineered for eye detection. For the purpose of detecting drowsiness, these eye regions are extracted and processed using a pre-trained deep learning model (MobileNet). Whether the eyes are open or closed is predicted by the model.
- Drowsiness Detection Algorithm: If both eyes are predicted to close for a certain amount of time, a drowsiness score is increased. If this score exceeds a predefined threshold, an alarm will sound to alert the driver. The alarm sounds until the driver's eyes are detected as open or until a predetermined time has passed [6].
- Visual and auditory feedback: The drowsiness score is indicated by overlaying text onto the video frame. When drowsiness is detected, audio feedback in the form of an alarm sound is generated, providing the driver with an additional alert mechanism.
- Dynamic Threshold and Alarm Control: Based on the driver's state, the algorithm adjusts the drowsiness detection threshold and controls the duration of the alarm. This ensures adaptability.

- Frame Skipping for Efficiency: To improve processing speed and reduce computational load, the script uses frame skipping. This technique skips a set number of frames (determined by the skip_frames variable) before proceeding to the next frame. By skipping frames, the system can maintain real-time performance without processing every single frame, resulting in increased efficiency.
- Impact on response time: The frame-skipping mechanism may cause a slight delay in the system's response time. This delay is primarily due to the lower frequency of frame processing. While frame skipping contributes to real-time performance, it can impair the system's ability to detect and respond quickly to changes in the driver's state, such as the onset of drowsiness.
- Trade-off Between Speed and Accuracy: The decision to skip frames requires a trade-off between processing speed and accuracy. While frame skipping speeds up processing, it may impair the system's ability to detect subtle changes in driver behaviour. Thus, developers must carefully balance the number of frames skipped to ensure an acceptable level of responsiveness while maintaining computational efficiency.
- Optimization parameters that can be adjusted include: To reduce the impact of frame skipping on response time, developers can change parameters like frame skip count (skip_frames) and drowsiness detection threshold (score_threshold). Fine-tuning these parameters enables optimization for specific hardware capabilities and performance requirements.
- Real-time Adaptability: Despite the inherent delay introduced by frame skipping, the system maintains real-time adaptability to changes in the driver's state. By dynamically adjusting parameters and monitoring drowsiness indicators, the system can promptly trigger alerts when necessary, thereby enhancing driver safety in real-world scenarios.

In summary, while frame skipping optimizes processing speed in the experimental setup, it may introduce a minor delay in response time. Developers must carefully balance frame-skipping parameters to achieve optimal performance while maintaining adequate responsiveness for timely detection of drowsiness and other safety-critical events.

7. CONCLUSION

The study on driver distraction and drowsiness detection using deep learning approaches reveals important findings about the efficacy of convolutional neural networks (CNNs) in addressing road safety concerns. The findings highlight the ability of deep learning techniques, such as ResNet50 and MobileNet, to accurately detect signs of driver distraction and drowsiness in real-time scenarios. The developed systems are efficient and reliable in real-world applications, providing proactive measures to reduce the risks associated with distracted and drowsy driving. By utilising advanced deep learning methodologies, these systems can analyse video feeds from in-car cameras, identify behavioural cues indicative of distraction or drowsiness, and send timely alerts to drivers, improving road safety.

Furthermore, the models' adaptability and generalization capabilities are critical for ensuring consistent performance under a variety of driving conditions and demographics. Despite the inherent trade-offs between processing speed and accuracy, the study emphasises the importance of optimization strategies for balancing these variables and achieving peak system performance. Overall, the study's findings demonstrate the promising potential of deep learning approaches for improving driver safety through real-time detection and intervention mechanisms. Continued research and optimisation efforts in this domain are critical to future advancements in proactive accident prevention and road safety enhancement.

REFERENCES

[1] A. Pinto, M. Bhasi, D. Bhalekar, P. Hegde and S. G. Koolagudi, "A Deep Learning Approach to Detect Drowsy Drivers in Real Time," 2019 IEEE 16th India Council International Conference (INDICON), Rajkot, India, 2019, pp. 1-4, doi: 10.1109/INDICON47234.2019.9030305.

[2] Dipu, Md.Tanvir & Hossain, Syeda & Arafat, Yeasir & Rafiq, Fatama. (2021). Real-time Driver Drowsiness Detection using Deep Learning. International Journal of Advanced Computer Science and Applications. 12. 844-850. 10.14569/IJACSA.2021.0120794.

[3] V B Navya Kiran, Raksha R, Anisoor Rahman, Varsha K N, Dr. Nagamani N P, 2020, Driver Drowsiness Detection, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NCAIT – 2020 (Volume 8 – Issue 15)

[4] 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

[5] Jabbar, Rateb & Shinoy, Mohammed & Kharbeche, Mohamed & Al-Khalifa, Khalifa & Krichen, Moez & Barkaoui, Kamel. (2020). Driver Drowsiness Detection Model Using Convolutional Neural Networks Techniques for Android Application.

[6] Perwej, Dr. Yusuf & Akhtar, Nikhat & Verma, Mrs & Chaturvedi, Shivam & Mishra, Shubham. (2022). An Intelligent Motion Detection Using OpenCV. International Journal of Scientific Research in Science, Engineering and Technology. Volume 9. Pages 51-63. 10.32628/IJSRSET22925.

[7] Hussain, Dostdar & Ismail, Muhammad & Hussain, Israr & Alroobaea, Roobaea & Ullah, Syed Sajid. (2022). Face Mask Detection Using Deep Convolutional Neural Network (DCNN) and MobileNetV2 Based Transfer Learning. Wireless Communications and Mobile Computing. 2022. 10.1155/2022/1536318.

[8] J. Yu, X. Hao and P. He, "Single-stage Face Detection under Extremely Low-light Conditions," 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, Canada, 2021, pp. 3516-3525, doi: 10.1109/ICCVW54120.2021.00392

[9] Alshalfan, Khalid & Zakariah, Mohammed. (2021). Detecting Driver Distraction Using Deep-Learning Approach. Computers, Materials & Continua. 680. 689-704. 10.32604/cmc.2021.015989.

[10] Orman, Zeynep & Battal, Abdulkadir & Kemer, Erdem. (2011). A Study On Face, Eye Detection And Gaze Estimation. International Journal of Computer Science & Engineering Survey. 2. 10.5121/ijcses.2011.2303.

[11] Vijayan, Vineetha & Elizabeth, Sherly. (2019). Real time detection system of driver drowsiness based on representation learning using deep neural networks. Journal of Intelligent & Fuzzy Systems. 36. 1-9. 10.3233/JIFS-169909.

[12] Torres, Renato & Ohashi, Orlando & Carvalho, Eduardo & Pessin, Gustavo. (2017). A Deep Learning Approach to Detect Distracted Drivers Using a Mobile Phone. 72-79. 10.1007/978-3-319-68612-7_9.