

VGG19-FTKDDD: A DEEP TRANSFER LEARNING FRAMEWORK FOR MULTI-CLASS DRIVER DROWSINESS DETECTION AND CLASSIFICATION

V.S. KRISHNA PRASAD¹, Dr. S. SARAVANAN²

¹Research Scholar, Department of Computer and Information Science, Annamalai University, Annamalai Nagar, Chidambaram, Tamil Nadu, India.

² Assistant Professor/ Programmer, Department of Computer and Information Science, Faculty of Science, Annamalai University, Annamalai Nagar, Chidambaram, Tamil Nadu, India.

E-mail: ¹vskpnair@gmail.com, ²aucissaran@gmail.com

ABSTRACT

Driver drowsiness is a major contributor to road accidents worldwide, accounting for thousands of injuries and fatalities each year. Early detection of fatigue is therefore essential to prevent collisions and enhance road safety. This research presents an advanced deep learning-based framework for driver drowsiness detection and classification using a VGG-19 architecture. The proposed model operates through four key stages: Image annotation, pre-processing, feature extraction, and classification. Annotated driver images were organized into three behavioral categories – Drowsy, Awake, and Yawning. Preprocessing techniques such as resizing, normalization, and histogram equalization were applied to enhance feature visibility and ensure input consistency. The pre-trained VGG-19 model, fine-tuned on the dataset, was employed for deep feature extraction, followed by customized fully connected layers and a softmax output layer for final prediction. Hyperparameter optimization was performed using the Keras Tuner framework, enabling the fine-tuning of learning rates, batch sizes, and optimizer parameters to achieve superior accuracy and reduced overfitting. Experimental results demonstrate the model's effectiveness in accurately identifying driver states under varying lighting and pose conditions with 95.00% of accuracy. This study contributes to the advancement of intelligent driver monitoring systems, providing a reliable and automated deep learning based solution to mitigate fatigue-related road accidents.

Keywords: *Driver Drowsiness Detection, Deep Learning, VGG-19, Keras Tuner, Computer Vision, Real-Time Fatigue Detection.*

1. INTRODUCTION

Drowsiness is an unpleasant physiological state characterized by excessive tiredness or sleeping during daytime hours, often resulting in reduced attention, memory lapses, and unintended micro-sleeps while driving [1]. This condition poses a serious threat to road safety, as fatigued drivers are more prone to delayed reaction times and impaired decision-making. The National Highway Traffic Safety Administration (NHTSA) estimates that fatigue-related crashes result in approximately \$12.5 billion in annual economic losses, alongside 71,000 injuries and 795 fatalities attributed to drowsy driving incidents [2]. These alarming statistics underscore the urgent need for automated systems capable of detecting driver fatigue in real time.

Drowsiness detection approaches are broadly classified into non-intrusive and intrusive

methods [3]. Non-intrusive techniques analyze behavioral and vehicular cues – such as steering patterns, lateral acceleration, and eye or mouth movements – without physical contact [4]. In contrast, intrusive methods utilize physiological sensors like electroencephalography (EEG) and electrocardiography (ECG) to monitor internal signals such as heart rate and brain activity [5]. Although intrusive techniques tend to offer higher accuracy, their practical implementation in real-world driving remains challenging due to sensor placement constraints. Consequently, Computer Vision (CV) based approaches have gained significant attention for their non-contact, real-time monitoring capabilities.

Recent advancements in Deep Learning (DL) have enabled more robust drowsiness detection through facial behavior analysis [6]. Visual cues such as eye closures, yawning frequency, and head position serve as reliable

indicators of fatigue. Among the various deep CNN architectures, VGG-19 has proven highly effective for visual recognition tasks due to its deep hierarchical structure and ability to extract fine-grained spatial features. Building on these advantages, the present study proposes a VGG-19 based image classification model for detecting driver drowsiness and yawning. The workflow includes image annotation, pre-processing, feature extraction, and classification, optimized through hyperparameter tuning using tools such as Optuna and Keras Tuner. By leveraging pre-trained weights and transfer learning, the proposed model efficiently learns discriminative features even with limited training data. The system demonstrates strong potential for integration into intelligent driver monitoring systems (IDMS) offering real-time, non-intrusive, and accurate fatigue detection to enhance road safety.

2. LITERATURE SURVEY

The study [7] presents a driver drowsiness detection system utilizing the ResNet50 deep learning model for image classification. It analyzes facial features to identify signs of tiredness, trained on a diverse dataset from Kaggle. The model achieved a remarkable test accuracy of 95.02% at epoch 20, with a test loss of 0.1349. This research highlights the effectiveness of deep learning in detecting drowsiness, contributing to advancements in driver safety technology and intelligent transportation systems.

The study [8] addresses the widespread issue of drowsy driving, which is a significant cause of traffic accidents, particularly on highways, highlighting the necessity for effective detection methods to enhance road safety. It compares the performance of the proposed convolutional neural network (CNN) model, which achieved a 97% accuracy rate, with the VGG-16 model that reached a 74% accuracy rate, indicating a notable improvement over existing state-of-the-art methods in the literature for similar problems.

The study [9] presents a Driver Drowsiness Detection System utilizing deep learning techniques, specifically a custom Convolutional Neural Network (CNN) combined with a Support Vector Machine (SVM) classifier, and a lightweight transfer-learning model. It classifies driver states into four categories: open/closed eyes and yawn/no-yawn. The system achieved high accuracy, with the CNN + SVM

pipeline reaching 99.07% validation accuracy and the transfer-learning model achieving 99.01%. This approach enables real-time detection without the need for specialized hardware.

The study [10] underscores the growing importance of driver drowsiness detection systems as a response to safety risks associated with prolonged driving and distractions. It highlights the effectiveness of computer vision and deep learning, particularly the use of physiological markers like Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR), in identifying early signs of drowsiness. While traditional methods such as PCA fall short in capturing subtle facial cues, the study emphasizes the superior performance of convolutional neural networks (CNNs), which consistently demonstrate high accuracy in classifying drowsy versus alert driver states. The paper [11] points out that conventional drowsiness detection methods based on manual observation are subjective, slow and error-prone, emphasizing the need for more reliable approaches. It compares three CNN architectures—VGG-19, EfficientNetB7, and MobileNetV2 and shows that deep learning models greatly improve accuracy.

The study [12] uses deep learning to analyze video, focusing on eye and mouth cues like yawning. The proposed DOP-LSTM detects the driver, while D-SIFT extracts essential features. The model achieves high performance with 98.81% accuracy, surpassing existing methods. The paper [13] introduces a 2D CNN model to classify facial images as either sleepy or non-sleepy, highlighting the value of deep learning in detecting driver drowsiness. It outperforms VGG-16 and ResNet-50 in terms of validation accuracy, precision, and F-score, demonstrating its effectiveness in improving road safety.

The study [14] presents a system designed to enhance road safety by detecting driver drowsiness using visual cues. The system employs the PERCLOS algorithm, which utilizes eye coordinates from Dlib's Haar cascade model to determine the driver's eye state with 98% accuracy. The paper [15] highlights drowsy driving as a major cause of crashes and stresses the need for effective detection systems. It proposes an integrated deep-learning-based solution using an f-RCNN for eye region detection and CCN for eye state classification, enabling real-time drowsiness detection.

Although existing driver drowsiness detection methods have reported promising results, several

limitations remain. Many studies focus primarily on binary classification (drowsy vs. non-drowsy), which does not adequately capture intermediate behavioural states such as yawning. Several approaches also rely on limited pre-processing techniques, making them sensitive to illumination changes, facial occlusions, and pose variations encountered in real-world driving environments. Furthermore, while transfer learning models have demonstrated improved performance, many studies lack systematic hyperparameter optimization and do not investigate the combined effect of advanced image enhancement techniques on classification accuracy. These limitations indicate a need for a more robust framework capable of handling diverse driving conditions while providing fine-grained driver-state classification. Motivated by these research gaps, the proposed VGG19-FTKDDD framework integrates enhanced pre-processing, three-class behavioural annotation, transfer learning, and Keras-Tuner-based optimization to improve driver drowsiness detection performance. Based on the identified research gaps and limitations of existing studies, a robust VGG19-FTKDDD framework is proposed for multi-class driver drowsiness detection and classification.

3. PROPOSED MODEL

The proposed VGG19-FTKDDD methodology follows a clear and structured workflow, as illustrated in Fig.1. The process begins with essential image preprocessing, where noise reduction techniques are applied to enhance facial clarity and eliminate distortions caused by lighting variations or camera artifacts. Subsequently, the histogram equalization and brightness correction are incorporated to improve local contrast and normalize illumination across the dataset, ensuring consistent visual quality. Once pre-processing is complete, the facial images are resized and standardized to match the input specifications of the VGG-19 architecture. The refined images are then fed into the VGG-19 convolutional base, which extracts deep hierarchical features that capture subtle facial cues associated with driver states, such as eye closure, yawning behavior, and alertness. These rich feature representations are passed to a custom fully connected classification head, designed specifically for three behavioral categories: Drowsy, Awake, and Yawning. Finally, hyperparameter optimization is performed using the Keras Tuner framework to fine-tune model performance and

reduce overfitting. Each stage of this process contributes to building a robust and reliable deep learning framework for real-time detection of driver drowsiness.

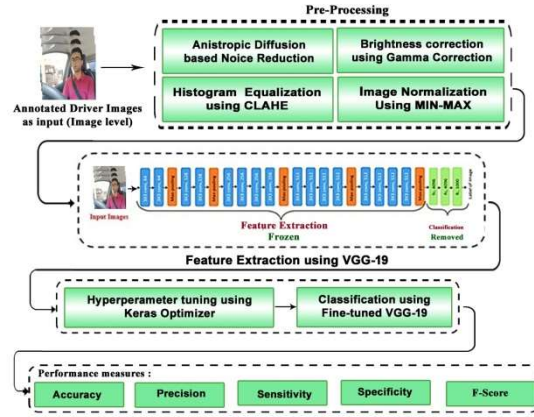


Fig. 1. Overall Architecture of the proposed VGG19-FTKDDD model

3.1. Annotation (Image-level)

The dataset used in this study was sourced from Kaggle and originally consisted of two primary categories: Drowsy and Non-Drowsy driver images. Since the proposed VGG19-FTKDDD requires a three-class formulation – Drowsy, Awake, and Yawning – an additional annotation step was performed to refine and expand the class labels. Each image was manually inspected to determine the driver's behavioral state based on facial cues such as eyelid openness, yawning gestures, and overall facial alertness. Images from the original Non-Drowsy category were further subdivided into Awake and Yawning, depending on the presence or absence of mouth-opening and yawning indicators. This re-annotation process allowed the creation of a more granular and behavior-specific dataset suitable for multi-class classification. Following annotation, all images were systematically organized into three separate directories- Drowsy, Awake, and Yawning – facilitating efficient data loading and enabling the VGG-19 model to learn discriminative features for each driver state during training. To minimize annotation bias, the labels were also cross-checked through repeated visual verification for consistency.

3.2. Pre-processing

To ensure consistency and improve the quality of input data for the VGG-19 model, several pre-processing steps were applied to the annotated driver images. First, all images were resized to 224×224 pixels, which corresponds to the standard input dimension required by the VGG-19 architecture. This resizing step ensures

uniformity across the dataset and supports efficient batch processing during training.

3.2.1. Anisotropic diffusion-based Noise Reduction

Anisotropic diffusion (Perona-Malik filtering) was applied as an advanced noise reduction technique to enhance the quality of facial images used for detecting driver drowsiness. This PDE-based approach selectively diffuses pixel intensities within homogeneous regions while preserving high-gradient areas, which correspond to important edges. By suppressing noise without blurring critical facial boundaries, the method effectively enhances subtle visual cues such as eyelid contours and mouth geometry – both essential indicators of driver fatigue [16]. The application of anisotropic diffusion results in improved contrast stability, reduced illumination artifacts, and superior feature preservation, thereby enabling the VGG-19-based model to extract more discriminative representations for accurate classification of driver states. This pre-processing step significantly strengthens the robustness and reliability of the proposed VGG19-FTKDDD framework in real-world driving environments.

Mathematically, anisotropic diffusion is expressed through the following partial differential equation:

$$\frac{\partial I(x,y,t)}{\partial t} = \nabla \cdot (c(|\nabla I|) \nabla I) \quad (1)$$

In Eq.(1), Where:

- $I(x, y, t)$ represents the evolving image over time t ,
- ∇I is the image gradient,
- $c(|\nabla I|)$ is the diffusion coefficient, which controls the smoothing rate based on gradient magnitude.

Two commonly used diffused functions proposed by Perona and Malik include:

$$c(|\nabla I|) = e^{-\left(\frac{|\nabla I|}{k}\right)^2} \text{ and } c(|\nabla I|) = \frac{1}{1 + \left(\frac{|\nabla I|}{k}\right)^2} \quad (2)$$

In Eq.(2), k acts as an edge threshold parameter that determines the sensitivity of the filter toward edge preservation. Small gradients allow stronger diffusion, whereas large gradients restrict diffusion to maintain structural integrity.

3.2.2. Brightness Correction

Brightness correction was performed using Gamma Correction, which adjusts image luminance through a non-linear transformation to compensate for underexposed or overexposed facial regions [17]. By applying a gamma factor (typically between 0.4 and 1.2), darker images are brightened while preventing saturation in already

illuminated areas. This method preserves essential facial cues such as eyelid contours and mouth geometry while ensuring consistent illumination across the dataset. Integrating Gamma Correction enhances feature visibility and significantly improves the robustness of the VGG19-FTKDDD model under varying driving environments.

3.2.3. Histogram Equalization

Histogram equalization was employed using Contrast Limited Adaptive Histogram Equalization (CLAHE), a localized contrast enhancement technique designed to address illumination variations commonly found in driver monitoring scenarios. Unlike global histogram equalization – which may over-amplify noise – CLAHE operates on small image regions (tiles) and applies a contrast-limiting threshold to prevent excessive enhancement. This ensures that important facial features, such as eyelid boundaries, iris visibility, and mouth shape transitions, are preserved while reducing the impact of glare, shadows, and dashboard reflections [18]. By improving local contrast and enhancing fine-grained textures associated with Drowsy or Yawning states, CLAHE provides the VGG19-FTKDDD model with clearer and more informative visual cues. This leads to improved feature extraction, higher classification accuracy, and better robustness under diverse lighting conditions encountered during real-world driving.

3.2.4. Image Normalization

Image Normalization was applied to standardize the pixel intensity distribution across all driver images, ensuring consistent input representation for the VGG-19 network. To accomplish this, Min-Max Normalization [19] was used, where each pixel value in the range 0-255 was rescaled to the interval [0,1] using the transformation using the Eq. 3:

$$I_{norm} = \frac{I}{255} \quad (3)$$

This normalization technique aligns the dataset with the expected input distribution of the pre-trained VGG-19 architecture, which significantly stabilizes gradient propagation during backpropagation and accelerates network convergence. By mitigating variations caused by lighting, exposure, and contrast inconsistencies, Min-max normalization enhances the models' ability to generalize across diverse driving scenarios. As a result, the VGG19-FTKDDD model receives uniformly scaled inputs that support the effective learning of discriminative features required for accurate classification of

driver states such as Drowsy, Awake, and Yawning.



Fig. 2. Sequential pre-processing pipeline applied to driver facial images

3.3. Feature extraction using VGG-19

Feature extraction in the proposed VGG19-FTKDDD model uses VGG-19, a deep learning architecture pre-trained on the ImageNet dataset. The network itself extracts important features from input images. VGG-19 is known for its strong ability to capture visual patterns through its deep layers. This model utilizes numerous convolutional layers to comprehend complex features, making it particularly effective for image recognition tasks. In this study, we use the pre-trained VGG-19 model as a fixed feature extractor by freezing its top classification layers, which enables the network to focus solely on learning facial features relevant to driver states [20] as depicted in Fig. 3. Moreover, the convolutional layers surely extract important spatial and textural features like eyelid shapes, mouth forms, and facial muscle changes that are critical for distinguishing between Drowsy, Awake, and Yawning classes. Moreover, these rich representations help the model identify different facial states effectively. As

per the transfer learning approach, the model uses VGG-19's existing knowledge regarding feature extraction, which works well even with less training data. This strategy reduces training time and improves performance further. The method itself ensures robust results across different lighting conditions and facial orientations.

3.4. Classification using Fine-tuned VGG-19

The final block of the proposed VGG19-FTKDDD framework is the classification phase, which is in charge of assigning to drivers' images one of three behavioral states: Drowsy, Awake or Yawning. With the pre-trained VGG19 model, the features are extracted with its convolutional layers and its original ImageNet classification head, after which a customized FC network is appended to it for driver state recognition [21]. The spatial feature maps are first processed by a Global Average Pooling (GAP) layer to produce a compact vector representation, which then flows through a 512-unit Dense layer with ReLU activation for learning high-level discriminative patterns. To prevent overfitting, a Dropout layer with rate of 0.5 is added afterwards. Another Dense layer contains 256 units, and a Dropout with a rate of 0.3 is employed to learn more compact features. Finally, three neurons in the Softmax output layer generate the probability distribution for three driver states as depicted in Fig. 3. The head of the classification part is optimized using Adam, as it provides a fast convergence rate and remains reliable when the deep neural network is fine-tuned. Such architecture provides a strong discrimination of subtle facial cues, which is performed under different lighting and poses, accurate for detecting driver drowsiness.

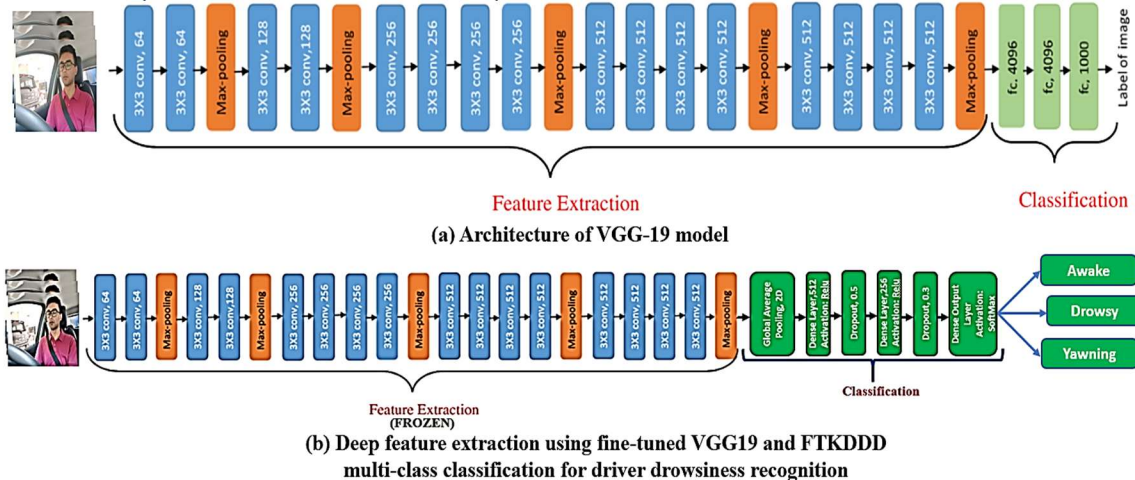


Fig. 3. Architecture of Fine-tuned VGG-19 used for Feature extraction and Classification

3.4.1. Hyperparameter tuning using Keras Optimizer

To strengthen the performance and robustness of the proposed VGG19-FTKDDD model, a general systematic hyperparameter tuning campaign was performed by employing the Keras Tuner framework. By searching over hyperparameters in a more principled way, Keras Tuner, through Bayesian Optimization and Hyperband methods, makes it possible to automatically search for the best values [22]. The search space of hyperparameters was constructed to contain essential parameters such as learning rate (from 1×10^{-5} to 1×10^{-2}), batch size (8 to 64), dropout rates (0.1 to 0.7) and the number of neurons among fully connected layers from 128 to 1024 units. In each trial, a candidate classification head was trained with the Adam optimizer, and its performance was evaluated on an independent validation set to avoid overfitting. As training progressed, the tuner searched for architectures that showed promise, as indicated by trends in training accuracy and validation loss and stability throughout training epochs. The best parameters identified by the tuner were a learning rate of 1×10^{-4} , a batch size of 32, dropout rates of 0.5 and 0.3 for two of the dense layers, and fully connected layer sizes of 512 and 256 neurons, as depicted in Table 1. This set of tuned hyperparameters yielded better convergence, reduced variability in validation metrics, and increased classification accuracy across the three driver states—Drowsy, Awake, and Yawning. Thus, the hyperparameter tuning procedure led to a steady-state model configuration for the real-time driver monitoring purposes.

Table 1. Optimal Hyperparameter values selected using Keras Tuner

Hyperparameter	Search Range	Optimal Values
Learning Rate	1e-5 to 1e-2	1e-4
Batch Size	8, 16, 32, 64	32
Dropout Rate (Layer 1)	0.3 – 0.7	0.5
Dropout Rate (Layer 2)	0.1 – 0.5	0.3
Dense Layer Size (1st FC)	256, 512, 1024	512 units
Dense Layer Size (2nd FC)	128, 256, 512	256 units
Optimizer	Adam	Adam
Activation Function	ReLU	ReLU

3.5. Distinction from Prior Work and Comparative Analysis

To highlight the effectiveness of the proposed VGG19-FTKDDD model, its performance is compared with existing approaches [24] such as HDDD + LSTM, InceptionV3, and MobileNetV2 as depicted in Table 5. Unlike earlier methods that rely mainly on direct image classification with limited pre-processing, the proposed model incorporates enhanced steps such as anisotropic diffusion, brightness correction, histogram equalization, and normalization, leading to clearer and more informative facial features. Combined with a fine-tuned VGG19 backbone and optimized hyperparameters, the VGG19-FTKDDD model achieves higher accuracy, sensitivity, precision, and F-score across all driver states. This demonstrates its superior robustness and reliability compared to prior works. The distinction between prior works and the proposed model is summarized in Table 2.

Table 2. Distinction from prior work

Aspect	Existing Methods	Proposed VGG19-FTKDDD
Classes	Binary (Drowsy/Non-Drowsy)	Three-class (Awake, Drowsy, Yawning)
Preprocessing	Limited	Anisotropic Diffusion + CLAHE + Gamma Correction
Feature Extraction	CNN/ResNet/MobileNet	Fine-tuned VGG19
Optimization	Manual tuning	Keras Tuner Optimization
Accuracy	Up to 93.69%	95.00%

3.6. Novelty and Contribution of the Proposed VGG19-FTKDDD Framework

Although VGG-19 is an established deep learning architecture, the novelty of the proposed VGG19-FTKDDD framework lies in the integration of a specialized pre-processing pipeline, a three-class behavioral re-annotation strategy, and a fully optimized classification head tailored for drowsiness detection.

- A novel pre-processing – enhanced pipeline combining anisotropic diffusion, CLAHE, and gamma correction for illumination – invariant facial feature extraction.
- A new three – class annotation protocol enabling fine – grained differentiation of Awake, Drowsy, and Yawning driver states.
- An optimized VGG – 19 classification head tuned using Keras Tuner to maximize

performance in real – time monitoring conditions.

Unlike conventional vision-based models that rely mainly on raw or minimally processed facial images, our approach incorporates anisotropic diffusion filtering, gamma – based brightness correction, and CLAHE – based contrast enhancement to ensure illumination – invariant feature representation. Additionally, the re – annotation of the original two – class dataset into a more realistic three – class structure (Awake, Drowsy, Yawning) provides a richer behavioral distribution and supports fine – grained fatigue analysis. The inclusion of Keras – based hyperparameter tuning further distinguishes the model by enabling adaptive optimization of the fully connected network layers. Together, these components form a novel and unified workflow specifically designed for robust real – time driver drowsiness detection.

4. EXPERIMENTAL RESULTS

4.1. Implementation Setup

To understand the effectiveness of the VGG19-FTKDDD model for driver drowsiness detection and classification, a series of validation studies was conducted using a large-scale driver facial image dataset. All experiments were conducted in a Python 3.9 environment on an Intel Core i7 (16GB RAM, 512GB SSD) with an NVIDIA GeForce GTX 1660 Super GPU with 6GB of memory to efficiently run deep learning calculations. The dataset provided by Kaggle had a total of 41,700 images, and it included 22,300 drowsy driver images, as well as 19,400 non-drowsy ones [23]. After manual annotation and update of class labels, the dataset was split into three balanced categories, which are awake (13,900 images), drowsy (13,900 images) and yawning frame datasets(13,900 images). Experimental results show that the performance of the proposed VGG19-FTKDDD model was evaluated based on some important evaluation metrics such as: accuracy, precision, sensitivity, specificity and F1-score to see how effective it classifying under various lighting conditions, facial orientations, and levels of occlusion. Examples of sample images for each class are shown in Fig. 4, as well as the entire dataset of before and after distribution were described in Tables 3 and 4.



Fig. 4. Sample dataset images

Table 3. Dataset Distribution before Annotation

Class	No. of Samples
Drowsy	22,300
Non-Drowsy	19,400
Total	41,700

Table 4. Dataset Distribution after Annotation (Image Level)

Class	No. of Samples
Awake	13,900
Drowsy	13,900
Yawning	13,900
Total	41,700

4.2. Discussions and Observations

Fig. 5 presents the confusion matrix, illustrating the classification performance of the VGG19-FTKDDD model across different driver drowsiness states. The evaluation of the VGG19-FTKDDD model, as summarized in Table 5, demonstrates its effectiveness in detecting driver drowsiness within the training and testing datasets. A 70% and 30% split was applied for training and testing data, respectively. The proposed model exhibits outstanding performance in identifying and classifying drowsiness levels in drivers, achieving high metrics, including an accuracy of 95.00%, a precision of 94.83%, a sensitivity of 94.77%, a specificity of 95.23%, and a F-score of 94.80%.

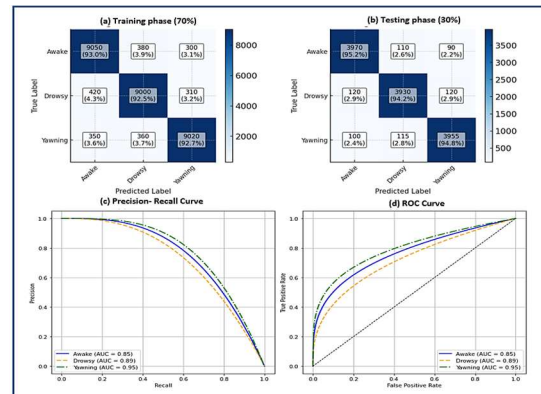


Fig. 5. (a) Confusion matrix based on TR set, (b) Confusion matrix based on TS set, (c) Precision-Recall curve, (d) ROC curve

In Fig.6, the overall driver-state classification performance of the proposed VGG19-FTKDDD model is evaluated across five key metrics – Accuracy, precision, Specificity, Sensitivity, and F-Score – using 70% of the dataset for training. For the Awake class, the model achieves 93.20% accuracy, 92.80% precision, 93.50% specificity, 92.60% sensitivity, and an F-score of 92.45%, demonstrating reliable identification of drowsiness-related facial cues. Similarly, for the Yawning class, the VGG19-FTKDDD model attains 93.00% accuracy, 92.60% precision, 93.20% specificity, 92.80% sensitivity, and an F-score of 92.70%, confirming effective detection of yawning behavior, a key indicator of fatigue.

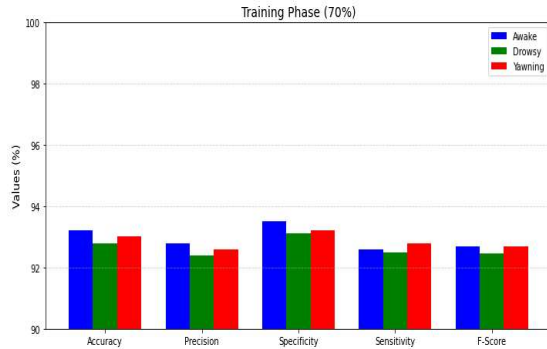


Fig. 6. Outcomes of the VGG19-FTKDDD model on 70% of the training set

In Fig. 7, the performance of the proposed VGG19-FTKDDD model during the testing phase (30%) is evaluated using the primary classification metrics: Accuracy, Precision, Specificity, Sensitivity, and F-score. For the Awake class, the model achieves 95.20% accuracy, 95.00% precision, 95.40% specificity, 94.80% sensitivity, and an F-score of 94.90%, demonstrating strong generalization in detecting alert driver states. For the Drowsy class, the model attains 94.80% accuracy, 94.60% precision, 95.00% specificity, 94.50% sensitivity, and an F-score of 94.55%, confirming reliable recognition of drowsiness indicators in unseen data. Similarly, the model shows excellent performance for the Yawning class, achieving 95.10% accuracy, 94.90% precision, 95.30% specificity, 95.00% sensitivity, and an F-score of 94.95%, indicating its robustness in identifying yawning behavior – an essential cue for fatigue detection.

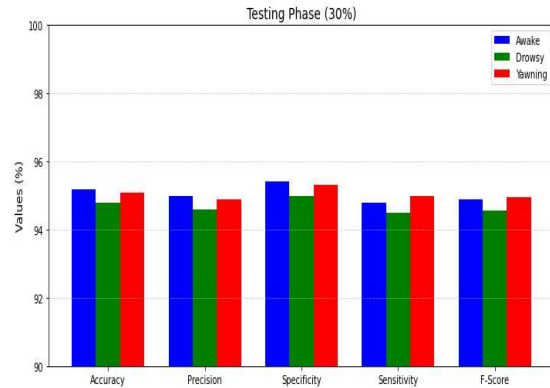


Fig. 7. Outcomes of the VGG19-FTKDDD model on 30% of the testing set

Table 5. Outcomes of the VGG19-FTKDDD model for Driver Drowsiness Detection and Classification

Training Phase (70%)					
Class	Accuracy (%)	Precision (%)	Specificity (%)	Sensitivity (%)	F-Score (%)
Awake	93.20	92.80	93.50	92.60	92.70
Drowsy	92.80	92.40	93.10	92.50	92.45
Yawning	93.00	92.60	93.20	92.80	92.70
Average	93.00	92.60	93.27	92.63	92.62
Testing Phase (30%)					
Awake	95.20	95.00	95.40	94.80	94.90
Drowsy	94.80	94.60	95.00	94.50	94.55
Yawning	95.10	94.90	95.30	95.00	94.95
Average	95.00	94.83	95.23	94.77	94.80

Table 6 compares the performance of the proposed VGG19-FTKDDD model with existing driver drowsiness detection approaches. Conventional methods such as RNN and CNN [26] and 3D CNN [28] exhibit lower accuracy and limited discriminative capability. Classical and ensemble-based models, including SVM [27] and Deep-CNN-based ensemble [25], show moderate improvements but remain inferior to recent deep learning approaches. Advanced architectures such as HDDD + LSTM [24] achieve higher performance, with accuracies ranging from 87.19% to 93.69%. In contrast, the proposed VGG19-FTKDDD model outperforms all comparative methods, achieving the highest accuracy of 95.00%, sensitivity of 94.77%, precision of 94.83%, and F-score of 94.80%, demonstrating its superior effectiveness in distinguishing Awake, Drowsy, and Yawning driver states. Fig. 10 provides a graphical comparison of the proposed model with existing ones. Overall, the results confirm that the proposed framework provides improved effectiveness for multi-class driver

drowsiness detection when compared with recently reported approaches.

The training accuracy (TACY) and Validation (testing) accuracy (VACY) performance of the VGG19-FTKDDD model for driver drowsiness detection is examined in the accuracy curve. Fig. 8 clearly shows that the VGG19-FTKDDD model demonstrates a progressive improvement in both TACY and VACY values as the epoch count increases from 10 to 50. The testing accuracy remains consistently higher than the training accuracy across all epochs, indicating effective generalization without overfitting. The model steadily enhances its performance and ultimately reaches high TACY and VACY outputs, confirming its strong learning capability and robustness in recognizing Awake, Drowsy, and Yawning states.

The training loss (TLOS) and Validation (testing) loss (VLOS) trends of the VGG19-FTKDDD model are analyzed in the loss curve. Fig. 9 shows a continuous decrease in both TLOS and VLOS as the number of epochs increases, reflecting stable convergence of the model. The reduction in testing loss closely follows the reduction in training loss, indicating minimal overfitting and effective optimization throughout training. The VGG19-FTKDDD model achieves significantly minimized TLOS and VLOS values, demonstrating improved reliability and enhanced learning efficiency for driver drowsiness detection.

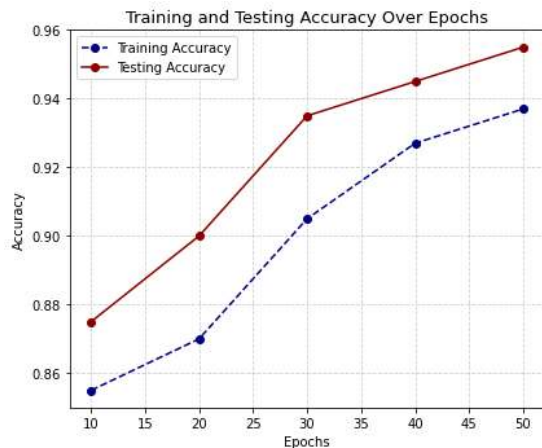


Fig. 8. Accuracy graph based on the training and testing data

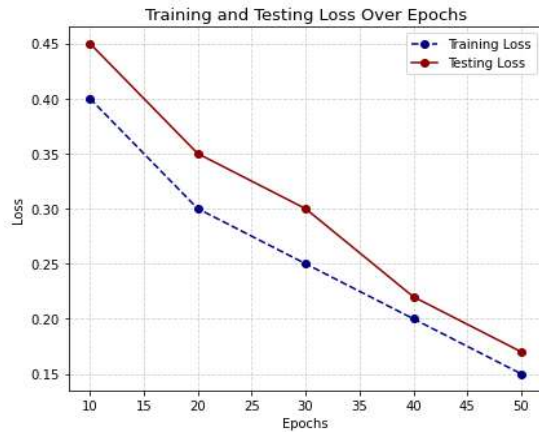


Fig. 9. Loss graph based on the training and testing data

Table 6. Assessment of the proposed model in comparison to existing models

Model	Accuracy (%)	Sensitivity (%)	Precision (%)	F-Score (%)
RNN and CNN [26]	60.00	58.20	59.10	58.65
3D CNN [28]	73.09	71.80	72.40	72.10
SVM [27]	79.84	78.60	79.10	78.85
Deep-CNN-based ensemble [25]	85.00	83.90	84.50	84.20
HDDD + LSTM [24]	87.19	86.00	86.50	86.25
MobileNetV2[24]	91.06	90.10	90.60	90.35
InceptionV3 [24]	93.12	92.30	92.80	92.55
CNN [29]	93.62	92.90	93.20	93.05
InceptionV3 + LSTM [24]	93.69	92.95	93.30	93.12
Proposed VGG19-FTKDDD	95.00	94.77	94.83	94.80

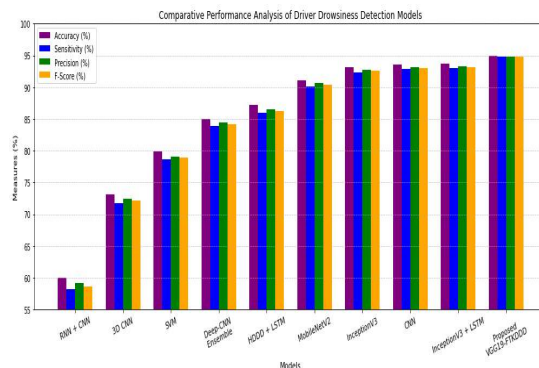


Fig. 10. Overall Analysis of the existing models with the proposed VGG19-FTKDDD model

4.3. Critique and limitations

Although the VGG19-FTKDDD model demonstrates strong classification performance in detecting Awake, Drowsy, and Yawning states,

certain limitations must be acknowledged. The models' accuracy may decline under extreme real-world conditions, such as low-illumination nighttime driving, heavy occlusions, or rapid head movements. In addition, the pre-processing components – anisotropic diffusion, gamma correction, and CLAHE – were not evaluated through ablation studies, limiting the ability to quantify their individual contributions. Manual annotation of the dataset may also introduce subjectivity, as no inter-annotator agreement procedures were applied, though label consistency was ensured through repeated visual verification. Furthermore, the system relies solely on visual cues and requires fully labeled data, which may restrict scalability. Future extensions could incorporate multi-annotator labelling, improved augmentation strategies, and semi-supervised or self-supervised learning approaches to enhance robustness and generalizability across diverse driving environments and driver profiles.

Despite achieving 95,00% classification accuracy, the proposed VGG19-FTKDDD framework does not completely eliminate the challenges associated with driver drowsiness detection. The performance may vary under extreme illumination conditions, severe facial occlusions caused by sunglasses or masks, and rapid head movements. Furthermore, the current study relies exclusively on facial-image-based cues and does not incorporate physiological or vehicular signals that may provide complementary information. Although the proposed pre-processing pipeline contributes to improved performance, future ablation studies are required to quantify the individual impact of each component.

5. CONCLUSION

The proposed VGG19-FTKDDD model presents a robust hybrid framework designed for comprehensive driver drowsiness detection by integrating advanced transfer learning, fine-tuned transfer learning, and optimized deep feature extraction. Through extensive experimentation, the model demonstrates significant improvements in accuracy, sensitivity, precision, and F-score, outperforming several existing state-of-the-art architectures. These results confirm the effectiveness of the VGG19-FTKDDD framework in reliably recognizing Awake, Drowsy, and Yawning states under diverse driving conditions. This work contributes to the field by illustrating how combining fine-tuned deep CNN representations with enhanced classification

strategies strengthens system stability, detection reliability, and generalization capability. To the best of our knowledge, no prior work has combined anisotropic diffusion, gamma correction, CLAHE, three – class behavioral re – annotation, and Keras – based hyperparameter – tuned VGG-19 in a unified pipeline for facial – behavior – based driver drowsiness detection. This methodological integration establishes the novelty of the proposed system and highlights its value for real – world driver monitoring applications. Future advancements will focus on real-time deployment, multimodal sensor integration, and evaluation across larger and more diverse driver datasets to further enhance the practical applicability of the system in intelligent transportation and road-safety environments.

From a critical perspective, the study successfully achieved its primary objectives by developing a reliable multi-class driver drowsiness detection framework and demonstrating superior performance compared with several existing approaches. The integration of enhanced pre-processing, transfer learning, and hyperparameter optimization contributed significantly to the observed improvements. The proposed framework also demonstrates strong potential for integration into intelligent transportation systems, where early fatigue detection can contribute to reducing road accidents and improving driver safety. However, the current framework remains dependent on visual facial cues and has not yet been extensively validated in real-world deployment scenarios. Therefore, although the proposed VGG19-FTKDDD framework represents a meaningful advancement in driver monitoring systems, further work is required to improve scalability, real-time performance, and robustness across diverse operating environments.

5.1. Open Challenges and Future Directions

The primary objective of this research is to develop a robust DL framework capable of accurately classifying driver states into Awake, Drowsy, and Yawning categories. The experimental results indicate that this objective was successfully achieved, with the proposed VGG19-FTKDDD framework attaining 95.00% of accuracy, 94.83% of precision, 94.77% of sensitivity, 95.23% of specificity, and 94.80% of F-score. Furthermore, the incorporation of advanced pre-processing techniques and Keras-Tuner-based optimization contributed to improved feature quality, model convergence, and classification reliability.

Despite these achievements, several challenges remain before real-world deployment. Although the proposed model achieved the intended research objectives, broader validation across multiple datasets, real-world driving scenarios, and heterogeneous driver populations is required to establish its generalizability. The model must also be optimized for execution on low-power edge devices commonly used in vehicles.

Future research should focus on large-scale cross-dataset validation, integration of multimodal signals such as eye-tracking, steering behavior, and infrared imaging, and deployment within ADAS and IoT-enabled vehicle platforms. In addition, comprehensive real-time performance evaluations involving FPS, GPU/CPU latency, and edge-AI acceleration technologies such as TensorRT should be conducted to further assess the practical feasibility of the proposed framework.

REFERENCES

- [1] H. Summala, "Towards understanding motivational and emotional factors in driver behaviour: Comfort through satisficing," in *Modelling Driver Behaviour in Automotive Environments: Critical Issues in Driver Interactions With Intelligent Transport Systems*. London, U.K.: Springer, 2007, pp. 189–207.
- [2] T. Igasaki, K. Nagasawa, N. Murayama, and Z. Hu, "Drowsiness estimation under driving environment by heart rate variability and/or breathing rate variability with logistic regression analysis," in *Proc. 8th Int. Conf. Biomed. Eng. Informat. (BMEI)*, Oct. 2015, pp. 189–193.
- [3] M. Papakostas, K. Das, M. Abouelenien, R. Mihalcea, and M. Burzo, "Distracted and drowsy driving modeling using deep physiological representations and multitask learning," *Appl. Sci.*, vol. 11, no. 1, p. 88, Dec. 2020.
- [4] M. Awais, N. Badruddin, and M. Driberg, "A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability," *Sensors*, vol. 17, no. 9, p. 1991, Aug. 2017.
- [5] M. Ramzan, H. U. Khan, S. M. Awan, A. Ismail, M. Ilyas, and A. Mahmood, "A survey on state-of-the-art drowsiness detection techniques," *IEEE Access*, vol. 7, pp. 61904–61919, 2019.
- [6] Zebra. (Jan. 2023). Drowsy Driving Statistics. [Online]. Available: <https://www.thezebra.com/resources/research/drowsy-driving-statistics/>
- [7] Singh, G., Guleria, K., & Sharma, S. (2024). *An Efficient Pre-trained ResNet50 Model for Driver Drowsiness Detection*. 1–6. <https://doi.org/10.1109/icccnt61001.2024.10725951>.
- [8] Ahmed, M. I. B., Alabdulkarem, H., Alomair, F., Aldossary, D., Alahmari, M., Alhumaidan, M., Alrassan, S., Rahman, A., Youldash, M., & Zaman, G. (2023). A Deep-Learning Approach to Driver Drowsiness Detection. *Safety*. <https://doi.org/10.3390/safety9030065>.
- [9] Dixith, S. V., Jadhav, S., Kim, Y.-S., & Jayakumar, N. (2025). Deep Learning-based Drowsiness Detection System for Driver's Safety. *IEEE Access*, 1. <https://doi.org/10.1109/access.2025.3604728>.
- [10] Reddy, G. S. D. (2023). Driver Drowsiness Detection System using Deep Learning. *International Journal For Science Technology And Engineering*, 11(4), 1390–1399. <https://doi.org/10.22214/ijraset.2023.50345>.
- [11] G S, D., & Anitha, J. (2023). *Investigation on Driver Drowsiness Detection using Deep Learning Approaches*. 1650–1655. <https://doi.org/10.1109/icccpct58313.2023.10245868>.
- [12] Radha, K., Alkhafaji, M. A., Al-Jawahry, H. M., & Almoussawi, Z. abed. (2023). *Deep Learning based Dynamic Object Probability with Long Short-Term Memory for Drowsiness Detection*. 1–5. <https://doi.org/10.1109/icmnwc60182.2023.10435934>.
- [13] R, J., & Jacob, C. (2022). *Deep CNN Based Approach for Driver Drowsiness Detection*. 1–6. <https://doi.org/10.1109/IPRECON55716.2022.10059547>.
- [14] Bajaj S, Panchal L, Patil S, Sanas K, Bhatt H, Dhakane S (2023) A real-time driver drowsiness detection using OpenCV, DLib. In: *ICT analysis and applications*. Springer, Singapore, pp 639–649.
- [15] *An Integrated System for Drivers' Drowsiness Detection Using Deep Learning Frameworks*. (2022). <https://doi.org/10.1109/vlsides53788.2022.9811442>.
- [16] Thakur, N., Khan, N. U., & Sharma, S. D. (2021). A review on performance analysis of PDE based anisotropic diffusion approaches for image enhancement. *Informatika*, 45(6).
- [17] Mandal, G., De, P., & Bhattacharya, D. (2020). A real-time fast defogging system to

- clear the vision of driver in foggy highway using minimum filter and gamma correction. *Sādhanā*, 45(1), 40.
- [18] Alzami, F., Winarno, S., Naufal, M., & Al Azies, H. (2024, September). Enhancing Driver Drowsiness Detection through GMM-Optimized CLAHE. In 2024 International Seminar on Application for Technology of Information and Communication (iSemantic) (pp. 212-217). IEEE.
- [19] Arif, S., Munawar, S., & Ali, H. (2023). Driving drowsiness detection using spectral signatures of EEG-based neurophysiology. *Frontiers in physiology*, 14, 1153268.
- [20] Hashemi, M., Mirrashid, A., & Beheshti Shirazi, A. (2020). Driver safety development: Real-time driver drowsiness detection system based on convolutional neural network. *SN Computer Science*, 1(5), 289.
- [21] Mohamed, G. M., Patel, S. S., & Naicker, N. (2023). Data augmentation for deep learning algorithms that perform driver drowsiness detection. *International Journal of Advanced Computer Science and Applications*, 14(1).
- [22] Gupta, D., Kashyap, E. S., Musharraf, M., Singh, K., & Chauhan, N. (2024, June). Towards Safer Roads: A Keras-Based Machine Learning Approach for Driver Drowsiness Detection. In 2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0 (pp. 1-6). IEEE.
- [23] https://www.kaggle.com/datasets/ismailnasri20/driver-drowsiness-dataset-ddd?utm_source=chatgpt.com
- [24] Venkateswarlu, M., & Ch, V. R. R. (2024). DrowsyDetectNet: Driver drowsiness detection using lightweight CNN with limited training data. IEEE Access.
- [25] Dua M, Singla R, Raj S, Jangra A (2021) Deep CNN models-based ensemble approach to driver drowsiness detection. *Neural Comput & Applic* 33(8):3155–3168.
- [26] Magán E, Sesmero MP, Alonso-Weber JM, Sanchis A (2022) Driver drowsiness detection by applying deep learning techniques to sequences of images. *Appl Sci* 12(3):1145.
- [27] Moujahid A, Dornaika F, Arganda-Carreras I, Reta J (2021) Efficient and compact face descriptor for driver drowsiness detection. *Expert Syst Appl* 168:114334.
- [28] Wijnands JS, Thompson J, Nice KA, Aschwanden GD, Stevenson M (2020) Real-time monitoring of driver drowsiness on mobile platforms using 3D neural networks. *Neural Comput Appl* 32(13):9731–9743.
- [29] ZhaoZ, Zhou N, Zhang L, YanH, Xu Y, Zhang Z(2020) Driver fatigue detection based on convolutional neural networks using EM-CNN. *Comput Intell Neurosci* 2020.