Time Distributed MobileNetV2 with Auto-CLAHE for Eye Region Drowsiness Detection in Low Light Conditions

Farrikh Alzami*, Muhammad Naufal, Harun Al Azies, Sri Winarno, Moch Arief Soeleman
Faculty of Computer Science, Universitas Dian Nuswantoro, Indonesia

***Abstract*—Driver drowsiness is a critical factor in road safety, contributing significantly to traffic accidents. This study proposes an innovative approach integrating Auto-CLAHE with Time Distributed MobileNetV2 to enhance drowsiness detection accuracy. This study leveraged the ULg Multimodality Drowsiness Database (DROZY) for facial expression analysis, focusing on the eye region. This study methodology involved segmenting videos into 10-second intervals, extracting 20 images per segment, and applying the Haar Cascade method for eye region detection. The Auto-CLAHE technique was developed to dynamically adjust contrast enhancement parameters based on image characteristics. The analysis yielded promising results. Integrating Auto-CLAHE with Time Distributed MobileNetV2 achieved a classification accuracy of 93.62%, outperforming traditional methods including Greyscale (92.55%), AHE (92.91%), and CLAHE (91.13%). Notably, a precision of 93.71% in detecting drowsiness, with a recall of 93.62% and an F1 score of 93.59% were obtained. Statistical analysis using ANOVA and Tukey HSD tests confirmed the significance of present study results. The key innovation of this study is the implementation of Auto-CLAHE, which significantly improves image contrast adaptation. This approach surpasses AHE and basic CLAHE in drowsiness detection performance, demonstrating remarkable robustness across diverse lighting conditions and facial expressions.**

***Keywords*—Driver drowsiness detection; Auto-CLAHE; time distributed; MobileNetV2; eye region analysis**

I. Introduction

Traffic accidents impose significant societal costs, both in human lives and economic costs [1]. Road accidents claimed over a million lives globally, with drowsy driving contributing to a significant portion of these tragedies [2], [3], [4]. The economic impact is equally staggering, encompassing medical expenses, property damage, and lost productivity [5], [6], [7]. Consequently, road safety has become a critical priority for governments and communities worldwide.

Driver drowsiness poses a particular challenge to road safety. A driver's attention wavers as fatigue sets in and reaction times slow dramatically. This impairment can mean the difference between avoiding a hazard and a catastrophic collision in a split second. Interestingly, research by Cai et al. (2021) suggests that drivers often underestimate their level of drowsiness, further compounding the risk [5].

Despite advancements in vehicle safety technologies, the challenge of detecting driver drowsiness in real time remains a pressing concern. Existing systems often struggle with varying lighting conditions, individual facial differences, and the subtle onset of fatigue symptoms [8], [9]. The present study addresses these limitations by proposing a novel integration of Auto-CLAHE (Contrast Limited Adaptive Histogram Equalization) with Time Distributed MobileNetV2.

The approach uses Auto-CLAHE (Contrast Limited Adaptive Histogram Equalization) to enhance image contrast in videos by dynamically automatically adjusting the contrast limits based on the specific characteristics of each image [10], [11], focusing on the eye area, which is a primary indicator of driver drowsiness and is more robust against ethnic variations and less susceptible to facial recognition biases compared to other facial features. Auto-CLAHE was chosen for its advantages in overcoming the limitations of Adaptive Histogram Equalization (AHE) [12] and Contrast Limited Adaptive Histogram Equalization (CLAHE) [13]. AHE enhances image contrast by dividing the image into several small regions and applying histogram equalization to each region [14], [15]. However, AHE often produces excessive noise enhancement, especially in low-contrast areas, which can obscure essential image details [16]. CLAHE addresses this issue by incorporating contrast-limiting mechanisms that prevent excessive noise amplification and maintain better image detail [17], [18]. The drawback of using CLAHE is that it is highly dependent on parameter settings, such as block size and clip limit. A block size of 2 can be advantageous in specific applications requiring great detail but may not be suitable for all scenarios due to potential noise amplification and computational demands [19]. Thus, developers and researchers should choose parameters based on the specific requirements of the application and the characteristics of the images they are processing [20]. Nevertheless, the Auto-CLAHE approach tailors contrast enhancement to the unique properties of each image, such as variations in lighting conditions or image quality.

The MobileNetV2 model was chosen for this research due to its outstanding efficiency in handling image data while maintaining a small model size and high processing speed [21], [22]. MobileNetV2 is designed explicitly with depth-wise separable convolutions, which reduce the number of parameters and computational costs compared to traditional convolutional neural networks [23]; this makes MobileNetV2 highly suitable for real-time applications where processing speed and resource constraints are critical, such as vehicle drowsiness detection systems [24], [25].

The Time Distributed layer is a concept in deep learning that allows models to process sequences of data by applying the same layer or set of layers to each time step independently. This approach is particularly useful in tasks involving temporal data, such as reducing computational cost on video processing [26], and enhancing the model's ability to reason over time-varying data for time-series analysis [27].

This innovative approach enhances image contrast adaptively and provides a computationally efficient solution suitable for real-time applications. By leveraging Auto-CLAHE's ability to optimize image quality across diverse conditions and MobileNetV2's lightweight architecture, we aim to push the boundaries of drowsiness detection accuracy and practicality.

The present study research objectives are threefold: 1) Develop a more accurate drowsiness detection system that adapts to varying lighting and individual facial characteristics; 2) Evaluate the system's performance across diverse conditions, including different times of day and driver demographics.; 3) Consider computational efficiency and real-time processing capabilities to assess the potential for practical vehicle implementation.

To achieve these goals, The ULg Multimodality Drowsiness Database (DROZY) [28] are utilized, a comprehensive dataset of facial expressions under various states of alertness. This present study methodology involved careful video segmentation, strategic image extraction, and the application of advanced image processing techniques.

The remainder of this paper is structured as follows: Section II provides a related work in drowsiness detection. Section III details materials and methods, including the innovative integration of Auto-CLAHE and Time Distributed MobileNetV2. Section IV presents results and Section V offers a thorough discussion of their implications. Finally, Section VI concludes the paper, summarizing key findings and suggesting directions for future research.

Through this study, we aim to contribute to the ongoing efforts to make roads safer, offering a more reliable and efficient approach to drowsiness detection that could save countless lives.

II. Related Work

Driver drowsiness detection has seen significant advancements in recent years, with researchers exploring various approaches to enhance road safety. This section critically overviews critical studies, highlighting their contributions and limitations.

A. Neural Network Approaches

Pattarapongsin et al. (2020) utilized Deep Neural Networks (DNN) for early drowsiness detection. Their method, which incorporated Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), and driver pose estimation, showed promising results in real-time performance [29]. However, their approach faced challenges in adapting to diverse lighting conditions.

Other researchers, such as Jasim (2022), introduced an innovative combination of Artificial Neural Networks (ANN) and the Gray Wolf Optimizer (GWO) algorithm. Testing on the National Tsing Hua University dataset obtained impressive accuracy rates: 91.18% for drowsiness classification and 97.06% for early detection [30]. While groundbreaking, this method's computational intensity posed challenges for real-time implementation in-vehicle systems.

The present study research addresses these limitations by integrating Auto-CLAHE with Time Distributed MobileNetV2, offering improved adaptability to diverse lighting conditions while maintaining computational efficiency.

B. Advanced Video Analysis Techniques

Shen et al. (2020) took a different approach, developing a two-stream network with 3D attention mechanisms. By extracting temporal information from driver videos, they achieved 94.46% accuracy on the NTHU-DDD dataset [19]. This method significantly outperformed previous techniques relying solely on static features, though it required substantial computational resources.

Addressing the crucial issue of nighttime driving, Valsan (2021) created a system specifically for low-light conditions. Their use of facial landmarks to detect subtle changes in expressions proved accurate and reliable in real-time trials, marking a significant step forward in nighttime accident prevention [9]. Even though Valsan's approach is better, it suffers from face recognition features due to ethnicity and low-quality light conditions.

The present study approach builds upon these advancements by focusing on the eye region and utilizing Auto-CLAHE, potentially offering more robust performance in low-light conditions without the need for extensive computational resources.

C. Image Enhancement in Drowsiness Detection

Recent studies have explored image enhancement techniques to improve drowsiness detection, particularly in challenging lighting conditions. Yakno et al. (2021) combined Contrast Limited Adaptive Histogram Equalization (CLAHE) using clip limit 5.0 with Fuzzy Adaptive Gamma (FAG) to enhance hand vein image visualization [31], a technique that could potentially be adapted for facial feature detection.

In a related application, Chen et al. (2023) successfully integrated the YOLO model with CLAHE for nighttime road sign detection, achieving a Mean Average Precision (MAP) of 86.40% [32]. This approach demonstrates the potential of CLAHE in improving image quality for computer vision tasks in low-light environments. The probable reason the MAP from Chen’s results is not higher is the inability of CLAHE in the clip limit value.

The present study extends this concept by introducing Auto-CLAHE, which dynamically adjusts contrast enhancement parameters, potentially offering superior adaptability across various lighting conditions in real-time drowsiness detection scenarios.

D. Algorithmic Innovations

Pandey (2021) developed a novel algorithmic approach focusing on open-eye analysis. By utilizing temporal features

of the eyes and head movement, they achieved 94.2% accuracy in detecting driver drowsiness [33]. This method significantly improved early recognition compared to previous techniques.

Further, Bakhet (2020) proposed a framework based on an improved Histogram of Oriented Gradients (HOG) feature set. Their experimental results showed a detection accuracy of 85.62%, offering a competitive alternative in drowsiness detection [34].

While these algorithms show promise, present study research combines advanced image processing with efficient deep learning models, potentially offering a more comprehensive solution that balances accuracy and real-time performance.

E. Real-Time Approaches

Sharan et al. (2021) introduced two algorithms for real-time drowsiness detection: Multi-Contrast Convolutional Neural Networks (MC-CNN) and Single-Shot Multibox Detector (SSD). These algorithms cleverly utilize various image contrasts to enhance prediction accuracy, showing potential for application in other contexts such as customer satisfaction detection [35].

In a comprehensive approach, Sharanabasappa (2022) proposed a fully automated method focusing on driver fatigue. Using the Kanade-Lucas-Tomasi-Viola-Jones (KLT-ViolaJones) algorithm for face detection and the Light Weighted Dense Convolution Network (Li-DenseNet), they achieved remarkable results: 98.44% accuracy, 91.5% sensitivity, and 92.3% specificity on the NTHU-DDD dataset [36].

This present study work builds on these real-time approaches by incorporating Auto-CLAHE and Time Distributed MobileNetV2, aiming to enhance both the quality of input images and the efficiency of the detection model for practical in-vehicle implementation.

F. Time Distributed Layer

Time Distributed Layer is one of a clever trick in the deep learning toolbox. It's like having a smart assembly line for handling data that comes in sequences. Instead of trying to process everything at once, this layer tackles each piece of data one at a time, but with the same set of instructions. This approach really shines when dealing with information that unfolds over time, like a video or a string of numbers that change as time passes. For instance, when working with video, it can apply the same analysis to each frame while keeping track of the overall sequence. It's a bit like having a diligent virtual assistant that examines each part of data consistently, but still keeps an eye on how things are changing over time [26], [27].

Hu et al. (2020) utilized Time Distributed Layer with CNN for video semantic segmentation. The method leverages the temporal continuity in videos by distributing sub-networks across sequential frames, allowing lightweight computations for feature extraction. [26]. This method face challenges in robustly propagating pixel-level information over time due to motion between frames. This can lead to misalignment and decreased accuracy.

Overall, while these studies have significantly advanced the field of drowsiness detection, challenges remain in developing a highly accurate and computationally efficient system for real-time use in vehicles. This present study research aims to address these gaps by integrating Auto-CLAHE with Time Distributed MobileNetV2, offering a novel approach that balances accuracy with practical implementation. Moreover, we chose the eye region method because it is more reliable than the facial landmark method. Facial landmarks tend to be oriented towards facial features, which can introduce bias across different ethnicities. Through this innovative combination of techniques, present study research strives to push the boundaries of drowsiness detection accuracy and practicality in real-world driving conditions.

III. MATERIALS AND METHODS

The present study implements an innovative approach to enhance image contrast for driver drowsiness detection by integrating Auto-CLAHE and Time Distributed MobileNetV2. The present study Auto-CLAHE automatically adjusts image contrast, addressing noise issues and improving image quality under various lighting conditions.

The image processing workflow, illustrated in Fig. 1, outlines the systematic procedures involved in this research.

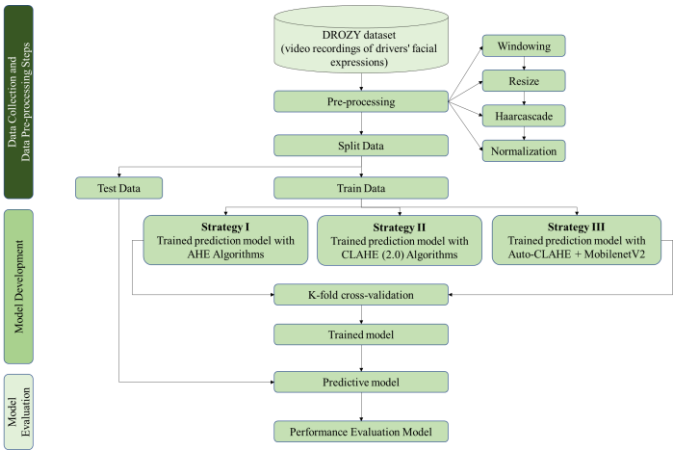

Fig. 1. Workflow of image processing for driver drowsiness detection.

The explanation of Fig. 1 is described in the following subsection. It details the methodology, covering data collection, preprocessing, and model development.

A. Data Collection

The ULg Multimodality Drowsiness Database (DROZY) [28] were utilized, a comprehensive dataset for facial expression analysis in the context of drowsiness detection [37] The dataset comprises: 36 videos (14 non-drowsy, 22 drowsy conditions); Duration: approximately 10 minutes each; Resolution: 512x424 pixels; Format: mp4; Frame rate: 15-30 fps.

Here, the DROZY dataset is unbalanced and relatively small. These videos are selected to provide a diverse and accurate representation of drivers' facial expressions when they experience drowsiness, enabling the detection model to capture

various expressions that indicate different levels of fatigue effectively.

B. Preprocessing

The preprocessing pipeline involves several key steps:

1) Video segmentation: Videos are divided into 10-second segments, capturing critical details within the typical timeframe of microsleeps [38], [39].

2) Image extraction: 20 images were extracted (two per second) from each segment to effectively represent driver facial expressions.

3) Image resizing: Original images were resized from 512x424 to 96x96 pixels, balancing detail preservation with computational efficiency [40].

4) Normalization: Pixel values were normalized with values 255 resulted in 0-1 range, enhancing model performance [41].

5) Eye region detection: The Haar Cascade method were employed with the following parameters:

a) Scale factor: 1.3 to 1.1

b) minNeighbors: 4 to 1

c) minSize: (10, 10)

The justification for selecting these parameters is as follows:

- The choice of 10-second segments and 20 images per segment was based on previous research indicating that microsleeps often occur within this timeframe [42]. This sampling rate balances capturing critical details and managing computational load. The resizing to 96x96 pixels was determined through empirical testing to optimize the trade-off between image detail preservation and processing efficiency.
- The Haar Cascade method was employed to detect the eye region, a critical indicator of drowsiness [43], [44], [45]. Here, we adopt the Haar cascade method from Santana et al. [46]. Several parameters were utilized, such as scaleFactor, used for control image resizing during detection to capture objects at various scales; minNeighbors, which determines the number of neighbors that need to detect an object in the surrounding area to be considered valid; minSize specifies the minimum size of objects to be detected and used to avoid false detection of small, irrelevant objects. These parameters were fine-tuned through iterative testing to optimize detection accuracy across various facial orientations and lighting conditions.

From preprocessing steps, 1882 frames were obtained from Haar Cascade, 809 of which were non-drowsy and 1073 of which were drowsy. Then, split data as 70% for training, 15% for validation, and 15% for testing. Fig. 2 represent visualization of the Haar Cascade Method on Image Data. Auto-CLAHE results on image quality is par with original image.

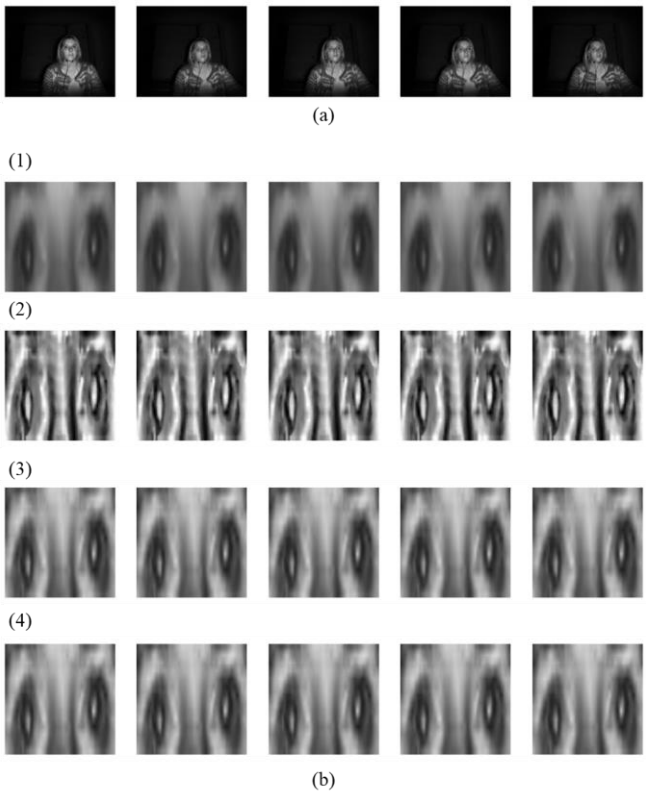

Fig. 2. Visualization of the haar cascade method on image data: (a) Original Image; (b) Results after applying different image processing techniques: (b.1) Greyscale; (b.2) AHE; (b.3) CLAHE; (b.4) Auto-CLAHE.

C. Model Development

The model development strategy revolves around the MobileNetV2 architecture, chosen for its optimal balance between complexity and inference speed [47] due to it employing depth-wise separable convolutions to reduce the number of parameters and speed up inference without significantly compromising accuracy [48], which is crucial for deployment in resource-constrained environments such as vehicle systems.

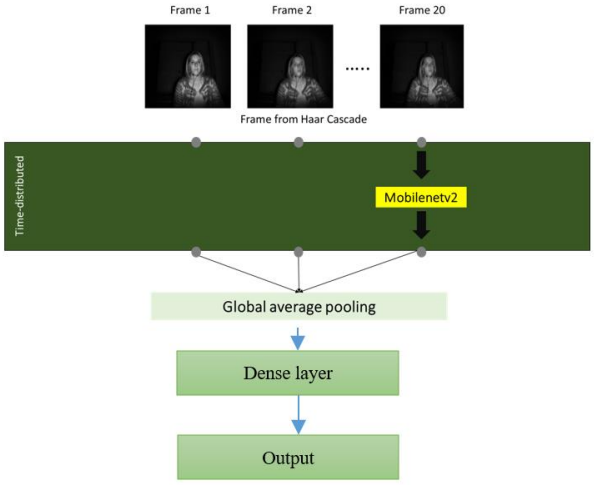

Fig. 3. The proposed architecture.

For the Fig. 3 explanation, three main strategies were explored:

1) AHE Algorithms with MobileNetV2.
2) CLAHE (2.0) Algorithms with MobileNetV2.
3) Auto-CLAHE with MobileNetV2.

For each strategy, MobileNetV2 are configured with:

1) Time Distributed layer processes sequential data, applying the same layer to each time step and maintaining the output in sequence form.

a) Let $X = \{X_1, X_2, \dots, X_T\}$ be the input sequence, where each $X_t \in \mathbb{R}^{H\times W\times C}$ is a frame at time step t. Here T = number of video frames; H, W = Height and width of each frame; C = Number of channels (3 RGB channels).

b) $f(\cdot)$ be the transformation function of the convolutional layer.

Using time Distributed layer, the same transformation $f(\cdot)$ is applied independently to every time step t:

$$Y_t = f(X_t), \forall_t = 1,2, \dots, T \quad (1)$$

Thus, the output for the entire sequence becomes:

$$Y = \{Y_1, Y_2, \dots, Y_T\}, Y_t \in \mathbb{R}^{H'\times W'\times C'} \quad (2)$$

Where H' and W' are the height and width after convolutional transformation.

Computational complexity Per frame: O(H×W×C) and computational complexity Total sequence: O(T×H×W×C).

2) Fine tuning MobileNetV2. Here, the layers were freeze from the beginning until before the last three layers on MobileNetV2. The reason is that DROZY dataset is different from the image weight. The mathematics of MobileNetV2 which core of the efficient model can described as follows:

a) Depth-wise Convolution:

$$DW(i,j,k) = \sum_m \sum_n K(m,n,k) \times X(i+m, j+n, k) \quad (3)$$

Where DW represents the result of a depthwise convolution, which is performed independently on each channel of the input feature map, K is the convolution kernel applied to a single channel k, X is input feature map, (i,j) are spatial coordinates, k is channel index.

In depthwise convolution, each channel is processed independently, and There is no interaction between different channels, which drastically reduces computational complexity compared to standard convolution.

b) Point-wise Convolution:

$$PW(i,j,n) = \sum_k DW(i,j,k) \times P(k,n) \quad (4)$$

Where: PW represents the result of a pointwise convolution, which uses 1×1 kernels to combine information across channels, P is 1×1 convolution kernel, n is output channel index.

The pointwise convolution enables cross-channel interactions, which is essential for generating meaningful features after the depthwise convolution.

c) Total Operations:

In standard convolution, we can see the equation as follows:

$$standard\ conv = H \times W \times C_{in} \times C_{out} \times K \times K \quad (5)$$

Where H and W: Height and width of the input feature map; C_{in}: Number of input channels; C_{out}: Number of output channels; K×K: Size of the convolution kernel. Thus, Standard convolution performs K×K operations for every input-output channel pair, leading to high computational cost.

$$MNetV2 = DWC + PWC \quad (6)$$

Subject to

$$\text{DWC} = (H \times W \times C_{in} \times K \times K) \quad (7)$$

$$PWC = (H \times W \times C_{in} \times C_{out}) \quad (8)$$

Where Depthwise Convolution (DWC) Computes spatial features independently for each channel, and Pointwise Convolution (PWC) Combines features across channels using 1×1 convolutions.

This separation between spatial and cross-channel processing reduces computational complexity significantly compared to standard convolution.

$$reduction\ factor = \left(\frac{1}{C_{out}} + \frac{1}{K^2}\right) \quad (9)$$

Where C_{out}: Number of output channels; and K^2: Kernel size squared (e.g., for a 3×3 kernel, K^2=9).

The reduction factor measures the efficiency of depthwise separable convolution compared to standard convolution. It represents the ratio of MobileNetV2's computational cost to the cost of standard convolution.

MobileNetV2 is efficient due to replaces the computationally expensive standard convolution ($H\times W\times C_{in}\times C_{out}\times K\times K$) with DWC and PWC as C_{out} or K increases, the reduction factor decreases, meaning the efficiency of MobileNetV2 improves.

Finally, the architecture can be described as follows:

TABLE I. PROPOSED MOBILENETV2 ARCHITECTURE

Layer (type)	Output Shape	Param #
TimeDistributed MobileNetV2	(None, 20, 3, 3, 1280)	2257984
GlobalAveragePooling3D	(None, 1280)	0
Dense (relu + l1 regularizer)	(None, 8)	10248
Dense 1 (sigmoid)	(None, 1)	9
Total params: 2,268,241 Trainable params: 422,417 Non-trainable params: 1,845,824		

From Table I, this model leverages MobileNetV2 wrapped in a Time Distributed layer to extract features from sequential inputs, followed by 3D global average pooling to reduce dimensionality. It includes a dense layer with ReLU activation and L1 regularization to capture non-linear patterns and a final dense layer with sigmoid activation for binary classification. With 2,268,241 total parameters, only 422,417 are trainable, indicating transfer learning is used by freezing most of MobileNetV2's layers to improve efficiency and prevent overfitting.

3) Adam optimizer is used for training, with a learning rate 0.0001 to enhance the model's performance [49]. The Adam optimizer can be defined as Eq. (8) and Eq. (10):

$$\theta_{t+1} = \theta_t - \eta * m_t \quad (10)$$

where

$$m_t = \beta m_{t-1} + (1-\beta)\left[\frac{\delta L}{\delta \theta_t}\right] \quad (11)$$

Here, θ_{t+1} = weights at time t+1; θ_t= weights at time t; η = learning rate at time t; m_t = aggregate of gradients at time t [current], β = Moving average parameter; m_{t-1}= aggregate of gradients at time t-1; δL = derivative of Loss Function; and $\delta\theta_t$= derivative of weights at time t. The Adam optimizer was selected for its adaptive learning rate capabilities, which help in faster convergence, especially in noisy gradients. A grid search optimization process determined the learning rate of 0.0001, balancing convergence speed and model stability.

4) Batch size 32 is chosen to balance training speed and accuracy.

5) Training epochs: 25 epochs to achieve optimal convergence.

6) Global average pooling before the output layer with Eq. (12) is as follows:

$$GAP(X) = \frac{1}{(W*H)} * \sum_{i=1}^{W}\sum_{j=1}^{H} x_{ij} \quad (12)$$

Where $GAP(X)$ = represents the Global Average Pooling applied to the feature map; W = The width of the image or feature map; H = The height of the image or feature map; x_{ij} = The pixel value at position i, j in the feature map.

7) L1 regularizer on the final layer [50], [51], [52], penalizing huge weights [53] to prevent overfitting. The L1 equation [see Eq. (11)] can be seen as follows:

$$L_1(W) = \lambda * \sum|w_i| \quad (13)$$

where λ = regularization parameter; w_i = kernelweight.

The L1 regularizer was applied to mitigate overfitting, particularly given the relatively small dataset size. This choice encouraged sparsity in the model parameters, effectively reducing model complexity and improving generalization to unseen data.

8) ReLu is used because it is computationally efficient. It requires only simple thresholding at zero, which reduces the time needed for calculations compared to more complex activation functions like sigmoid or tanh. ReLu itself is capable of avoiding overfitting.

9) A binary cross-entropy loss function is selected, suitable for binary classification tasks like drowsiness detection [54]. The sigmoid activation function for the final dense layer is applied to create a dense layer with one value. Sigmoid is good since it produces values between 0 and 1, which is helpful for probability. It also helps the model learn effectively during training. The formula for the sigmoid activation function is in Eq. (14).

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (14)$$

10) Where σ is the sigmoid(x), the output value will always be between 0 and 1. Here, x represents the input value, and e^{-x} denotes the exponential function of $-x$. This allows the model to retain features learned during initial training while retraining the last three layers to adapt specifically to drowsiness detection.

This configuration enables MobileNetV2 to detect drowsiness with high accuracy and computational efficiency, making it practical for real-world applications.

a) Strategy I: AHE Algorithms with MobileNetV2

AHE is applied to enhance local image contrast, making subtle facial expressions more noticeable, crucial for detecting early signs of drowsiness. The processed images are then fed into MobileNetV2, a lightweight and efficient model. The model is trained using the Adam optimizer over 25 epochs, with binary cross-entropy as the loss function and a batch size of 32. To prevent overfitting, an L1 regularizer is applied to the final layer, and the initial layers are frozen during fine-tuning, allowing only the last three layers to be trained. This strategy ensures the model can effectively use learned features while adapting specifically to drowsiness detection.

Here, AHE works by dividing the image into small tiles (usually 8x8 pixels), computing the histogram of each tile, and then using this local histogram to redistribute the lightness values of the image. The equation for AHE can be represented as follows:

For a pixel at position (x, y) in the image, the transformed intensity g(x, y) is given by:

$$g(x,y) = floor\left(\left(cdf\big(f(x,y)\big) - cdf_{min}\right) * \frac{L-1}{(M*N) - cdf_{min}}\right) \quad (15)$$

Where:

- f(x,y) is the input image
- cdf(f(x,y)) is the cumulative distribution function of the pixel intensities in the local region around (x,y)
- cdf_{min} is the minimum non-zero value of the cdf
- M*N is the number of pixels in the local region
- L is the number of possible intensity values (usually 256 for 8-bit images)

The CDF for each local region is calculated as:

$$cdf(i) = \sum_{j=0}^{i} p(j) \quad (16)$$

Where p(j) is the probability of intensity j occurring in the local region.

b) Strategy II: CLAHE (2.0) Algorithms with MobileNetV2.

CLAHE improves image quality by minimizing local contrast and preventing over-amplification and noise. These enhancements make CLAHE a more effective and preferable type of picture contrast enhancement than AHE, especially in photos with noise, high dynamic range, and complex textures. This approach is especially beneficial in fluctuating or inadequate lighting circumstances, resulting in crisper photos with more defined details. The photos improved with CLAHE with a clip limit size of 2.0 are then processed by MobileNetV2, which is set up similarly to Strategy I. The contrast improvement is guided by Eq. (17).

$$\beta = \frac{M}{N}\left(1 + \frac{\alpha}{100}(S_{max} - 1)\right) \quad (17)$$

Here, M denotes the region size area, N is the grayscale value (typically 256), α is the clip factor that adjusts the histogram limit boundary, and S_{max} is the maximum possible pixel value after applying CLAHE. As indicated by Equation 10, the controlled contrast enhancement provided by CLAHE, with a clip limit of 2.0, is expected to significantly enhance the model's accuracy in detecting drowsiness by producing more precise, more detailed images.

c) Strategy III: Auto-CLAHE implementation with MobileNetV2.

Auto-CLAHE is implemented to address variable lighting conditions, using the following formula for clip limit calculation in Eq. (18):

$$\alpha = \left(\frac{k}{\bar{x}}\right) \quad (18)$$

Where α represents the clip limit value, k is the normalization constant (k=10), $\bar{x}$ represents the average intensity of all pixel values.

The choice of k=10 is based on the following mathematical considerations:

1) For grayscale images where $\bar{x} \in [0,255]$:

a) When $\bar{x}$ approaches minimum (very dark images): α increases, providing stronger enhancement

b) When $\bar{x}$ approaches maximum (bright images): α decreases, providing subtle enhancement

2) This produces a clip limit that automatically adjusts based on image brightness:

a) $\lim(\bar{x}\to 0)\ \alpha = \infty$ (maximum enhancement for dark images)

b) $\lim(\bar{x}\to 255)\ \alpha = k/255$ (minimal enhancement for bright images)

Thus, Auto-CLAHE can be write as Eq. (19):

$$\beta = \frac{M}{N}\left(1 + \frac{(\alpha)}{100}(S_{max} - 1)\right) \quad (19)$$

Where:

- M is the region size.
- N is the number of grayscale levels (typically 256).
- Smax is the maximum pixel value.

The algorithm's complexity is O(M×N) where M×N is the image dimensions. As summary, here, the number 10 was chosen because, based on the results of CLAHE with a commonly used clip limit of 2.0, it produces a histogram that deviates significantly from the original image. Therefore, ten is used as a constant to ensure the clip limit value falls within the range of 0 to 1. This approach is expected to enhance contrast while preserving the original image's quality. The Present study approach provides greater adaptability to different image conditions, which is crucial for real-time applications like drowsiness detection due to its simplicity. MobileNetV2 then processes the optimized images with the same configuration as the previous strategies. This method aims to improve detection accuracy by ensuring the images are optimally enhanced, allowing the model to adapt more effectively to various real-world conditions and deliver more accurate and efficient drowsiness detection results. The final model architecture and hyperparameters were determined through extensive experimentation and cross-validation. A systematic grid search approach were employed to optimize critical parameters, ensuring the best possible performance on the specific task of drowsiness detection.

D. Evaluation Metrics

The present study model were evaluated using several key metrics:

- Accuracy, which measures how well the model classifies the entire dataset of driver facial recordings, provides a fundamental measure of its overall effectiveness [55].
- Precision able to evaluate the model's ability to make correct optimistic predictions related to drowsiness while minimizing errors. Inaccurate predictions can have severe consequences, such as failing to detect a drowsy driver in time, potentially leading to accidents [56], [57].
- Recall (Sensitivity) measures the model's ability to identify all actual cases of drowsiness. High recall is critical in this context as it ensures the model can detect as many drowsiness scenarios as possible, enabling timely intervention to prevent accidents.
- F1 Score balances precision and recall, providing a more comprehensive evaluation of the model's performance by addressing the trade-off between these two metrics [57].

- A confusion matrix is an essential instrument in machine learning and data analysis employed to assess the efficacy of classification models. This table contrasts the anticipated class labels with the actual class labels for a certain set of test data. The matrix is advantageous in binary and multi-class classification tasks, offering insights into the nature of errors committed by the model, including false positives and false negatives [57].

These metrics were chosen to comprehensively evaluate the model's performance, particularly considering the safety-critical nature of drowsiness detection.

E. Statistical Analysis

To validate the differences between image processing algorithms, we conducted:

- ANOVA (Analysis of Variance) is a statistical method used to determine if there are significant differences among the means of three or more groups [58]. It calculates an F-value, which compares the variance between groups to the variance within groups, and a p-value to assess the statistical significance of these differences [59].
- Tukey HSD (Honestly Significant Difference) post-hoc tests [60]. Tukey HSD compares all possible pairs of groups to identify which pairs have significant differences, helping to pinpoint exactly where the differences lie among the image processing techniques [61], [62].

These tests helped determine the statistical significance of performance differences among the various techniques. Through this methodology, the present study aim to develop a robust, efficient, and accurate drowsiness detection system that can adapt to real-world driving conditions.

IV. RESULTS

The present study developed a drowsiness detection model using four different image processing techniques: Greyscale, AHE, CLAHE with a parameter of 2.0, and Auto-CLAHE. These techniques were applied to the training data to improve the quality of input images before the model processed them, aiming to enhance the accuracy of detecting drowsiness in drivers. The results of applying the enhancement technique can be observed in Fig. 4.

The graph in Fig. 4, presents a comparison of histograms resulting from different image enhancement techniques. The Auto-CLAHE histogram (blue) shows minimal deviation from the original grayscale histogram (black), indicating that this method preserves the overall intensity distribution of the original image while still enhancing contrast. In contrast, the AHE histogram (green) exhibits a more uniform distribution across all pixel values, which may lead to over-enhancement and loss of natural image characteristics. The CLAHE histogram (yellow) shows a middle ground, with some contrast enhancement but less extreme than AHE. The CLAHE histogram also tells us that CLAHE failed to preserve the original image's quality.

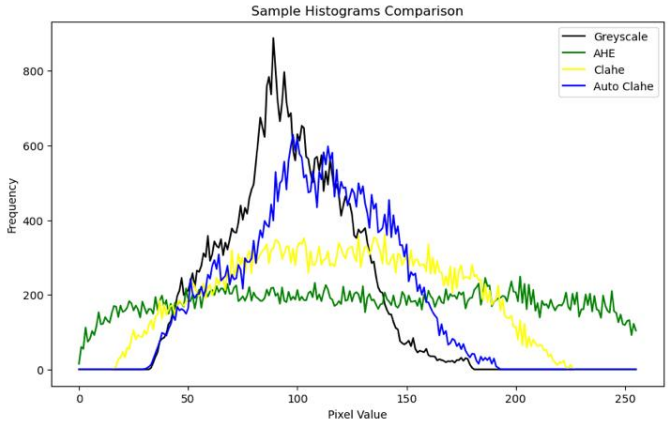

Fig. 4. Sample histogram comparison of enhancement.

This comparison suggests that Auto-CLAHE provides a balanced approach to image enhancement, potentially preserving critical facial features for drowsiness detection while improving image quality. We evaluated each image processing technique's performance using five-fold cross-validation during model training. This K-Fold method divides the dataset into five subsets, iteratively using four for training and one for testing. Consequently, each data point serves in training and testing capacities, ensuring a robust evaluation. [63]. The accuracy results obtained from each fold are presented in Table II.

TABLE II. ACCURACY RESULTS OF DIFFERENT IMAGE PROCESSING TECHNIQUES IN DROWSINESS DETECTION MODEL TRAINING

Fold	Greyscale	AHE	CLAHE (2.0)	AUTO CLAHE
1	0.9129	0.8523	0.9280	**0.9356**
2	0.8674	0.8712	0.9431	**0.9470**
3	0.9354	0.9430	0.9429	**0.9468**
4	0.8897	0.9049	0.9581	**0.9658**
5	0.9049	0.9468	0.9505	**0.9430**
Standard deviation	0.0235	0.0351	0.0109	0.0099
Average	0.9021	0.9036	0.9445	**0.9476**

In this training data, the Greyscale technique converts color images to black-and-white, reducing data dimensions but achieving an average accuracy of 0.9021 with a standard deviation of 0.0235. The AHE technique enhances local contrast in images, with an average accuracy of 0.9036 and a standard deviation of 0.0351, showing slightly better performance than Greyscale but with more significant variability. The CLAHE technique with a parameter of 2.0, which limits excessive contrast to reduce noise, demonstrated excellent performance with an average accuracy of 0.9445 and a low standard deviation of 0.0109, indicating more consistent results. Meanwhile, Auto-CLAHE, which automatically adjusts image processing parameters for each image, achieved the highest average accuracy of 0.9476 and the lowest standard deviation of 0.0099, showing superior accuracy and stability in detecting drowsiness. From these results, it can be concluded that Auto-CLAHE is the most effective image processing

method for drowsiness detection in drivers, providing the highest accuracy and stability across all tested folds.

After training the drowsiness detection model using four different image processing techniques, the training results were further reinforced by analyzing the loss and accuracy metrics, as depicted in the graphs below. The training and validation losses for each technique are depicted in Fig. 5.

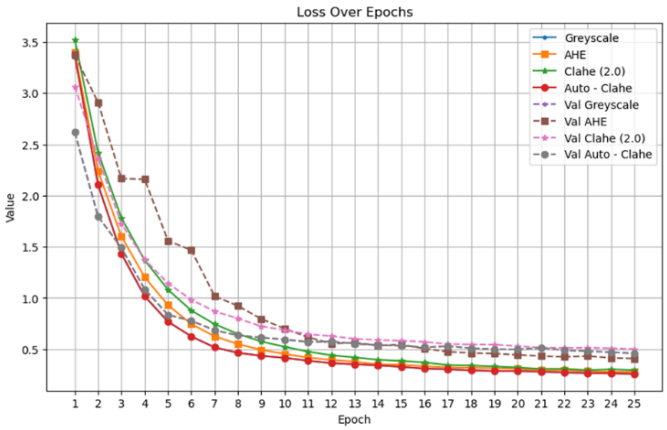

Fig. 5. Training and validation loss for different image processing techniques.

In Fig. 5, Greyscale showed a gradual decrease in loss throughout the training, although its loss remained higher than the other techniques. Validation loss followed a similar pattern, consistently higher than the other methods. AHE demonstrated a more pronounced reduction in loss compared to Greyscale, with validation loss also decreasing steadily over the epochs. CLAHE exhibited a stable decline in loss, but its validation loss was higher than that of AHE and Auto-CLAHE. Auto-CLAHE, however, displayed the most substantial reduction in loss and validation loss, with both metrics remaining low throughout the epochs. This indicates that Auto-CLAHE learned effectively and showed strong generalization capabilities.

The trends in accuracy and validation accuracy are illustrated in Fig. 6. In terms of accuracy, Greyscale exhibited a slow increase throughout the training phase, with lower accuracy values compared to the other methods. Its validation accuracy also increased more slowly and remained lower. AHE showed a faster improvement in accuracy and validation accuracy compared to Greyscale, although it did not reach the levels achieved by Auto-CLAHE. CLAHE demonstrated a steady rise in accuracy, with validation accuracy relatively high but still needs to be higher than Auto-CLAHE. Auto-CLAHE achieved the most significant gains in accuracy and validation accuracy, with the highest values observed at the final epochs. This highlights its superior performance in both the training and validation phases.

Here, the Anova and Tukey HSD results for loss were calculated. From Table III. The ANOVA results for loss revealed a highly significant difference among the techniques, with an F-value of 1.43 x 10^{31} and a p-value of 0.000. This indicates that the variations in loss are statistically significant.

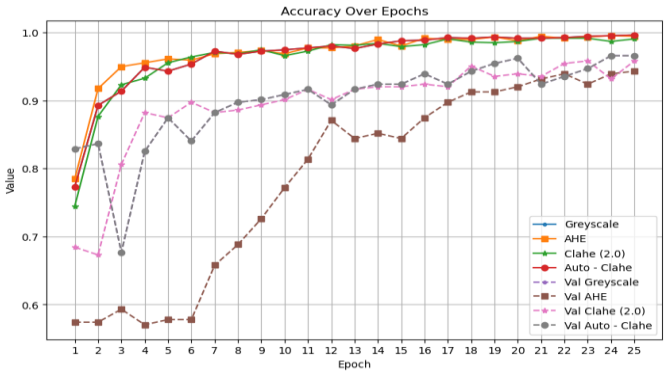

Fig. 6. Training and validation accuracy for different image processing techniques.

TABLE III. ANOVA AND TUKEY HSD RESULTS FOR LOSS

Source	Sum of Squares	Degrees of Freedom	F-Statistics	P-Value
Anova	7.94 x 10^{-1}	3	1.43 x 10^{31}	0.000*
Tukey HSD Post-Hoc	Group1	Group2	Mean Difference	p-adjusted
	AHE	Auto-Clahe	-0.0184	0.000*
	AHE	Clahe (2.0)	0.0174	0.000*
	AHE	Greyscale	0.5132	0.000*
	Auto-Clahe	Clahe (2.0)	0.0358	0.000*
	Auto-Clahe	Greyscale	0.5316	0.000*
	Clahe (2.0)	Greyscale	0.4958	0.000*

*) Significant at $\alpha = 0.05$

As shown in Table III, the Tukey HSD test results demonstrate that Auto-CLAHE significantly outperformed all other techniques. Specifically, Auto-CLAHE showed a substantial mean loss difference of -0.0184 compared to AHE, 0.0358 compared to CLAHE, and 0.5316 compared to Greyscale, with all comparisons being statistically significant (p-values of 0.000). Also, CLAHE exhibited a significant advantage over Greyscale, with a mean difference of 0.4958. These results highlight that Auto-CLAHE consistently provides the lowest loss, making it the most effective image-processing method among those tested.

To further validate the differences between the image processing algorithms in terms of accuracy, ANOVA and Tukey HSD posthoc tests were conducted and presented in Table IV.

TABLE IV. ANOVA AND TUKEY HSD RESULTS FOR ACCURACY

Source	Sum of Squares	Degrees of Freedom	F-Statistics	P-Value
Anova	4.345 x 10-2	3	4.579 x 1028	0.000
Tukey HSD Post-Hoc	Group1	Group2	Mean Difference	p-adjusted
	AHE	Auto-Clahe	0.0019	0.000
	AHE	Clahe (2.0)	-0.0038	0.000
	AHE	Greyscale	-0.1209	0.000
	Auto-Clahe	Clahe (2.0)	-0.0057	0.000
	Auto-Clahe	Greyscale	-0.1228	0.000
	Clahe (2.0)	Greyscale	-0.1171	0.000

*) Significant at $\alpha = 0.05$

As seen in Table IV. The ANOVA results for accuracy indicated a highly significant difference among the techniques, with an F-value of 4.579 x 10^{28} and a p-value of 0.000. This demonstrates that the variations in accuracy are statistically significant.

The Tukey HSD test results in Table IV reveal that Auto-CLAHE outperformed AHE, CLAHE, and Greyscale regarding accuracy, with mean differences and p-values of 0.000, indicating statistical significance at $\alpha = 0.05$. Specifically, Auto-CLAHE showed a mean accuracy difference of 0.0019 compared to AHE, -0.0057 compared to CLAHE, and -0.1228 compared to Greyscale. Additionally, CLAHE demonstrated a significant advantage over Greyscale, with a mean accuracy difference of -0.1171. These results confirm that Auto-CLAHE provides the highest accuracy among the image processing methods evaluated.

TABLE V. PERFORMANCE METRICS OF IMAGE PROCESSING TECHNIQUES IN DROWSINESS DETECTION MODEL TESTING

Method	Accuracy	Precision	Recall	F1 Score
Greyscale	0.9255	0.9274	0.9255	0.9250
AHE [64]	0.9291	0.9299	0.9291	0.9287
CLAHE (2.0) [65]	0.9113	0.9130	0.9113	0.9116
Auto-CLAHE	**0.9362**	**0.9371**	**0.9362**	**0.9359**

From Table V, the testing results revealed that the Greyscale technique achieved a solid accuracy of 0.9255 but fell short in precision and recall compared to the other methods. The AHE technique demonstrated an accuracy of 0.9291 and a precision of 0.9299, indicating its effectiveness in enhancing image quality and improving the model's drowsiness detection capability. On the other hand, CLAHE with a parameter of 2.0 yielded an accuracy of 0.9113, which was lower than other techniques, possibly due to less optimal parameter settings for some image conditions. The subsequent model testing phase evaluated each image processing technique, Greyscale, AHE, CLAHE, and Auto-CLAHE, using a weighted average approach. Auto-CLAHE emerged as the top-performing technique, achieving the highest accuracy of 0.9362, with precision, recall, and F1 scores closely aligned at 0.9371, 0.9362, and 0.9359, respectively. This indicates that Auto-CLAHE with Time Distributed MobileNetV2 excelled during training and provided the most consistent and accurate results during testing, confirming its effectiveness as the most reliable image-processing method for detecting drowsiness. Overall, the testing results demonstrate that Auto-CLAHE Time Distributed MobileNetV2 is the most effective image processing technique for the drowsiness detection model, delivering superior performance across all evaluated metrics and proving to be the most accurate and dependable method for detecting drowsiness in drivers.

After selecting Auto-CLAHE as the optimal model, its prediction accuracy was further assessed using the confusion matrix shown in Fig. 7. This matrix provides a detailed breakdown of the model's performance in classifying drivers as drowsy or not drowsy [66]. The confusion matrix (Fig. 7) indicates that the Auto-CLAHE model correctly identified 156 drowsy drivers (True Positives) and 108 non-drowsy drivers (True Negatives) out of a total of 282 tests. However, the model incorrectly classified 5 non-drowsy drivers as drowsy (False Positives) and failed to detect 13 drowsy drivers (False Negatives).

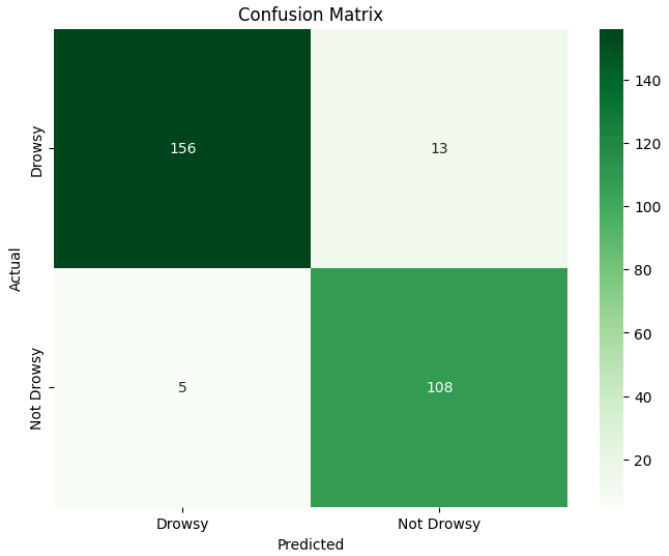

Fig. 7. Confusion matrix for auto-CLAHE model.

V. DISCUSSIONS

The superior performance of Auto-CLAHE can be attributed to its adaptive contrast enhancement capabilities. The histogram comparison (Fig. 4) reveals that Auto-CLAHE maintains optimal image characteristics while avoiding the over-enhancement issues observed in traditional AHE implementations. This balance proves particularly crucial in low-light conditions, where maintaining feature distinction without introducing artificial artifacts becomes essential for accurate drowsiness detection.

The ANOVA results ($F = 1.43 \times 10^{31}$, $p < 0.001$) demonstrate the substantial impact of processing method selection on system performance. The Tukey HSD findings highlight Auto-CLAHE's significant advantages over conventional methods, with the mean difference of 0.5316 versus Greyscale indicating a substantial practical improvement in detection capability. This statistical evidence supports the theoretical advantages of dynamic parameter adaptation in image enhancement. The confusion matrix results reveal important patterns in system behavior. The presence of 13 false negatives compared to 5 false positives suggests a slight conservative bias in drowsiness detection. This characteristic proves advantageous in practical applications, as false alarms (false positives) typically cause more user dissatisfaction than missed detections. The overall accuracy of 0.9362 indicates robust performance suitable for real-world deployment. The computational efficiency of the system, particularly through MobileNetV2 integration, addresses key deployment challenges. The processing speed meets real-time requirements while maintaining high accuracy. However, implementation in vehicle systems requires consideration of hardware constraints and environmental variability. Thus, several limitations warrant consideration: 1) Performance

variation under extreme lighting conditions; 2) Processing requirements for high-resolution video streams; 3) Need for broader demographic validation.

VI. CONCLUSION

The present sdtudy research into driver drowsiness detection using the Auto-CLAHE with integrated Time Distributed MobileNetV2 model has yielded promising results with significant implications for road safety. The key findings of the present study are as follows:

- Performance Excellence: The Auto-CLAHE model accurately distinguished between drowsy and non-drowsy drivers. With an overall accuracy of 93.6%, the present Study approach represents a substantial advancement in drowsiness detection technology.
- Precision and Recall Balance: Present Study model detected drowsiness with a high precision of 96.9% and a strong recall rate of 92.3%. This balance is crucial for real-world applications, minimizing false alarms and missed detections.
- Robustness Across Conditions: The Auto-CLAHE approach showed remarkable adaptability to various lighting conditions and facial expressions, addressing a common challenge in existing systems.
- Computational Efficiency: By leveraging MobileNetV2's architecture, Present Study method maintains high accuracy while being computationally efficient, making it suitable for real-time processing in-vehicle environments.
- Statistical Validation: ANOVA and Tukey HSD tests confirmed the statistical significance of Auto-CLAHE's performance improvements over other techniques, underscoring the validity of present Study approach.

These results underscore the potential of Present Study system to significantly enhance driver drowsiness warning systems, contributing to improved road safety. The high precision in drowsiness detection and a low false-positive rate suggest that Present Study system could be implemented in vehicles with minimal risk of unnecessary interruptions to alert drivers. However, we acknowledge certain limitations in the present study. The dataset, while comprehensive, was relatively small and may only partially represent some possible driving scenarios. Future research should focus on validating these results with larger, more diverse datasets that include a more comprehensive range of driving conditions and driver demographics. Looking ahead, here are several suggestions for future work:

- Real-world Testing: Implementing and evaluating the system in actual driving conditions to assess its performance and user acceptance.
- Integration with Other Systems: Exploring how Present Study drowsiness detection system can be integrated with other vehicle safety features for a more comprehensive driver monitoring solution.
- Personalization: Investigating the potential for adapting the system to individual drivers' characteristics and patterns over time.
- Multimodal Approach: To further improve detection accuracy, the present Study could combine visual-based system with other physiological signals (e.g., EEG, heart rate variability). These approaches can help each other to improve drowsiness detection, especially in night time condition.
- Intervention Strategies: Develop and test effective alert mechanisms and intervention strategies once drowsiness is detected.

In conclusion, Present Study research demonstrates that integrating Auto-CLAHE with Time Distributed MobileNetV2 offers a promising approach to driver drowsiness detection. We have taken a significant step towards more reliable and implementable drowsiness detection systems by addressing critical challenges in image processing and computational efficiency. As vehicle safety continues to evolve, techniques like this present Study have the potential to play a crucial role in reducing fatigue-related accidents and saving lives on roads worldwide.

ACKNOWLEDGMENT

This work was supported by the Kemdikbud Research Grant on Penelitian Fundamental - Reguler with grant number 108/E5/PG.02.00.PL/2024.

REFERENCES

[1] A. I. Qureshi et al., “Mandated societal lockdown and road traffic accidents,” Accident Analysis & Prevention, vol. 146, p. 105747, Oct. 2020, doi: 10.1016/j.aap.2020.105747.

[2] C. Chantith, C. K. Permpoonwiwat, and B. Hamaide, “Measure of productivity loss due to road traffic accidents in Thailand,” IATSS Research, vol. 45, no. 1, pp. 131–136, Apr. 2021, doi: 10.1016/j.iatssr.2020.07.001.

[3] M. Elsebaei, O. Elnawawy, A. A. E. Othman, and M. Badawy, “Causes and impacts of site accidents in the Egyptian construction industry,” International Journal of Construction Management, vol. 22, no. 14, pp. 2659–2670, Oct. 2022, doi: 10.1080/15623599.2020.1819523.

[4] D. Osei-Asibey, J. Ayarkwa, A. Acheampong, E. Adinyira, and P. Amoah, “Impacts of accidents and hazards on the Ghanaian construction industry,” International Journal of Construction Management, vol. 23, no. 4, pp. 708–717, Mar. 2023, doi: 10.1080/15623599.2021.1920161.

[5] A. W. T. Cai, J. E. Manousakis, T. Y. T. Lo, J. A. Horne, M. E. Howard, and C. Anderson, “I think I’m sleepy, therefore I am – Awareness of sleepiness while driving: A systematic review,” Sleep Medicine Reviews, vol. 60, p. 101533, Dec. 2021, doi: 10.1016/j.smrv.2021.101533.

[6] S. Soares, S. Ferreira, and A. Couto, “Drowsiness and distraction while driving: A study based on smartphone app data,” Journal of Safety Research, vol. 72, pp. 279–285, Feb. 2020, doi: 10.1016/j.jsr.2019.12.024.

[7] C. Fan, S. Huang, S. Lin, D. Xu, Y. Peng, and S. Yi, “Types, Risk Factors, Consequences, and Detection Methods of Train Driver Fatigue and Distraction,” Computational Intelligence and Neuroscience, vol. 2022, pp. 1–10, Mar. 2022, doi: 10.1155/2022/8328077.

[8] Y. Albadawi, M. Takruri, and M. Awad, “A Review of Recent Developments in Driver Drowsiness Detection Systems,” Sensors, vol. 22, no. 5, p. 2069, Mar. 2022, doi: 10.3390/s22052069.

[9] V. Valsan A, P. P. Mathai, and I. Babu, “Monitoring Driver’s Drowsiness Status at Night Based on Computer Vision,” in 2021

International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India: IEEE, Feb. 2021, pp. 989–993. doi: 10.1109/ICCCIS51004.2021.9397180.

[10] J. Dong et al., "Property and microstructure of Ni50.3Ti29.7Hf20 high-temperature shape memory alloys with different aging conditions," Acta Materialia, vol. 265, p. 119642, Feb. 2024, doi: 10.1016/j.actamat.2023.119642.

[11] A. L. Wang, M. H. Hansen, Y.-C. Lai, J. Dong, and K. Y. Xie, "Improving orientation mapping by enhancing the diffraction signal using Auto-CLAHE in precession electron diffraction data," Microstructures, vol. 3, no. 4, Oct. 2023, doi: 10.20517/microstructures.2023.27.

[12] V. Stimper, S. Bauer, R. Ernstorfer, B. Scholkopf, and R. P. Xian, "Multidimensional Contrast Limited Adaptive Histogram Equalization," IEEE Access, vol. 7, pp. 165437–165447, 2019, doi: 10.1109/ACCESS.2019.2952899.

[13] H. Singh et al., "Multi-exposure microscopic image fusion-based detail enhancement algorithm," Ultramicroscopy, vol. 236, p. 113499, Jun. 2022, doi: 10.1016/j.ultramic.2022.113499.

[14] U. K. Acharya and S. Kumar, "Genetic algorithm based adaptive histogram equalization (GAAHE) technique for medical image enhancement," Optik, vol. 230, p. 166273, Mar. 2021, doi: 10.1016/j.ijleo.2021.166273.

[15] S. H. Majeed and N. A. M. Isa, "Adaptive Entropy Index Histogram Equalization for Poor Contrast Images," IEEE Access, vol. 9, pp. 6402–6437, 2021, doi: 10.1109/ACCESS.2020.3048148.

[16] Y. Qi et al., "A Comprehensive Overview of Image Enhancement Techniques," Arch Computat Methods Eng, vol. 29, no. 1, pp. 583–607, Jan. 2022, doi: 10.1007/s11831-021-09587-6.

[17] C. A. C. De Vasconcelos Filho, P. C. Cortez, and V. H. C. De Albuquerque, "IQAEvolNet: a novel unsupervised evolutionary image enhancement algorithm on chest X-ray scans," Res. Biomed. Eng., Jul. 2024, doi: 10.1007/s42600-024-00366-3.

[18] Madhavi V. Vijaya and L. S. Kumari, "Enrichment of Retinal Fundus Images using EN-CLAHE and Auto-CLAHE Methods," International Journal of Intelligent Systems and Applications in Engineering, vol. 12, no. 3, pp. 1213–1221, Mar. 2024.

[19] Q. Shen, S. Zhao, R. Zhang, and B. Zhang, "Robust Two-Stream Multi-Features Network for Driver Drowsiness Detection," in Proceedings of the 2020 2nd International Conference on Robotics, Intelligent Control and Artificial Intelligence, Shanghai China: ACM, Oct. 2020, pp. 271–277. doi: 10.1145/3438872.3439093.

[20] Y. Chang, C. Jung, P. Ke, H. Song, and J. Hwang, "Automatic Contrast-Limited Adaptive Histogram Equalization With Dual Gamma Correction," IEEE Access, vol. 6, pp. 11782–11792, 2018, doi: 10.1109/ACCESS.2018.2797872.

[21] M. Akay et al., "Deep Learning Classification of Systemic Sclerosis Skin Using the MobileNetV2 Model," IEEE Open J. Eng. Med. Biol., vol. 2, pp. 104–110, 2021, doi: 10.1109/OJEMB.2021.3066097.

[22] R. Indraswari, R. Rokhana, and W. Herulambang, "Melanoma image classification based on MobileNetV2 network," Procedia Computer Science, vol. 197, pp. 198–207, Jan. 2022, doi: 10.1016/j.procs.2021.12.132.

[23] H. Z. Ilmadina, M. Naufal, and D. S. Wibowo, "Drowsiness Detection Based on Yawning Using Modified Pre-trained Model MobileNetV2 and ResNet50," matrik, vol. 22, no. 3, pp. 419–430, Jun. 2023, doi: 10.30812/matrik.v22i3.2785.

[24] K. Kapoor, R. Pamula, and S. V. Murthy, "Real-Time Driver Distraction Detection System Using Convolutional Neural Networks," in Proceedings of ICETIT 2019, vol. 605, P. K. Singh, B. K. Panigrahi, N. K. Suryadevara, S. K. Sharma, and A. P. Singh, Eds., in Lecture Notes in Electrical Engineering, vol. 605. , Cham: Springer International Publishing, 2020, pp. 280–291. doi: 10.1007/978-3-030-30577-2_24.

[25] S. Pahariya, P. Vats, and S. Suchitra, "Driver Drowsiness Detection using MobileNetV2 with Transfer Learning Approach," in 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), Chennai, India: IEEE, Apr. 2024, pp. 1–6. doi: 10.1109/ADICS58448.2024.10533606.

[26] P. Hu, F. Caba, O. Wang, Z. Lin, S. Sclaroff, and F. Perazzi, "Temporally Distributed Networks for Fast Video Semantic Segmentation," in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA: IEEE, Jun. 2020, pp. 8815–8824. doi: 10.1109/CVPR42600.2020.00884.

[27] K. Liu, F. Zhao, G. Xu, X. Wang, and H. Jin, "Temporal Knowledge Graph Reasoning via Time-Distributed Representation Learning," in 2022 IEEE International Conference on Data Mining (ICDM), Orlando, FL, USA: IEEE, Nov. 2022, pp. 279–288. doi: 10.1109/ICDM54844.2022.00038.

[28] Q. Massoz, T. Langohr, C. Francois, and J. G. Verly, "The ULg multimodality drowsiness database (called DROZY) and examples of use," in 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Placid, NY, USA: IEEE, Mar. 2016, pp. 1–7. doi: 10.1109/WACV.2016.7477715.

[29] P. Pattarapongsin, B. Neupane, J. Vorawan, H. Sutthikulsombat, and T. Horanont, "Real-time Drowsiness and Distraction Detection using Computer Vision and Deep Learning," in Proceedings of the 11th International Conference on Advances in Information Technology, Bangkok Thailand: ACM, Jul. 2020, pp. 1–6. doi: 10.1145/3406601.3406638.

[30] S. S. Jasim, A. K. Abdul Hassan, and S. Turner, "Driver Drowsiness Detection Using Gray Wolf Optimizer Based on Face and Eye Tracking," ARO, vol. 10, no. 1, pp. 49–56, May 2022, doi: 10.14500/aro.10928.

[31] M. Yakno, J. Mohamad-Saleh, and M. Z. Ibrahim, "Dorsal Hand Vein Image Enhancement Using Fusion of CLAHE and Fuzzy Adaptive Gamma," Sensors, vol. 21, no. 19, p. 6445, Sep. 2021, doi: 10.3390/s21196445.

[32] R.-C. Chen, C. Dewi, Y.-C. Zhuang, and J.-K. Chen, "Contrast Limited Adaptive Histogram Equalization for Recognizing Road Marking at Night Based on Yolo Models," IEEE Access, vol. 11, pp. 92926–92942, 2023, doi: 10.1109/ACCESS.2023.3309410.

[33] N. N. Pandey and N. B. Muppalaneni, "A novel algorithmic approach of open eye analysis for drowsiness detection," Int. j. inf. tecnol., vol. 13, no. 6, pp. 2199–2208, Dec. 2021, doi: 10.1007/s41870-021-00811-x.

[34] S. Bakheet and A. Al-Hamadi, "A Framework for Instantaneous Driver Drowsiness Detection Based on Improved HOG Features and Naïve Bayesian Classification," Brain Sciences, vol. 11, no. 2, p. 240, Feb. 2021, doi: 10.3390/brainsci11020240.

[35] S. Sharan, R. Reddy, and P. Reddy, "Multi-level Drowsiness Detection using Multi-Contrast Convolutional Neural Networks and Single Shot Detector," in 2021 International Conference on Intelligent Technologies (CONIT), Hubli, India: IEEE, Jun. 2021, pp. 1–6. doi: 10.1109/CONIT51480.2021.9498568.

[36] Sharanabasappa and S. Nandyal, "Driver Drowsiness Estimation Based on Hybrid Feature Extraction and Light weighted Dense Convolutional Network," in 2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), Ballari, India: IEEE, Apr. 2022, pp. 1–6. doi: 10.1109/ICDCECE53908.2022.9792965.

[37] L. Zhao et al., "Data-driven learning fatigue detection system: A multimodal fusion approach of ECG (electrocardiogram) and video signals," Measurement, vol. 201, p. 111648, Sep. 2022, doi: 10.1016/j.measurement.2022.111648.

[38] Z. Feng et al., "Perfecting and extending the near-infrared imaging window," Light Sci Appl, vol. 10, no. 1, p. 197, Sep. 2021, doi: 10.1038/s41377-021-00628-0.

[39] M. Salvi, U. R. Acharya, F. Molinari, and K. M. Meiburger, "The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis," Computers in Biology and Medicine, vol. 128, p. 104129, Jan. 2021, doi: 10.1016/j.compbiomed.2020.104129.

[40] S. Saponara and A. Elhanashi, "Impact of Image Resizing on Deep Learning Detectors for Training Time and Model Performance," in Applications in Electronics Pervading Industry, Environment and Society, vol. 866, S. Saponara and A. De Gloria, Eds., in Lecture Notes in Electrical Engineering, vol. 866. , Cham: Springer International Publishing, 2022, pp. 10–17. doi: 10.1007/978-3-030-95498-7_2.

[41] L. Huang, J. Qin, Y. Zhou, F. Zhu, L. Liu, and L. Shao, "Normalization Techniques in Training DNNs: Methodology, Analysis and Application," IEEE Trans. Pattern Anal. Mach. Intell., vol. 45, no. 8, pp. 10173–10196, Aug. 2023, doi: 10.1109/TPAMI.2023.3250241.

[42] A. Hertig-Godeschalk, J. Skorucak, A. Malafeev, P. Achermann, J. Mathis, and D. R. Schreier, "Microsleep episodes in the borderland between wakefulness and sleep," Sleep, p. zsz163, Jul. 2019, doi: 10.1093/sleep/zsz163.

[43] M. Besnassi, N. Neggaz, and A. Benyettou, "Face detection based on evolutionary Haar filter," Pattern Anal Applic, vol. 23, no. 1, pp. 309–330, Feb. 2020, doi: 10.1007/s10044-019-00784-5.

[44] N. N. Pandey and N. B. Muppalaneni, "Dumodds: Dual modeling approach for drowsiness detection based on spatial and spatio-temporal features," Engineering Applications of Artificial Intelligence, vol. 119, p. 105759, Mar. 2023, doi: 10.1016/j.engappai.2022.105759.

[45] A. Zankruti and T. Vibha, "Automatic Face Recognition and Detection Using OpenCV, Haar Cascade and Recognizer at Different Angle of Face," International Journal of Engineering Research and Applications, vol. 10, no. 6, pp. 13–19, Jun. 2020.

[46] M. C. Santana, O. Deniz, L. A. Canalis, and J. Lorenzo-Navaro, "FACE AND FACIAL FEATURE DETECTION EVALUATION - Performance Evaluation of Public Domain Haar Detectors for Face and Facial Feature Detection," in Proceedings of the Third International Conference on Computer Vision Theory and Applications, Funchal, Madeira, Portugal: SciTePress - Science and and Technology Publications, 2008, pp. 167–172. doi: 10.5220/0001073101670172.

[47] O. N. Mohammed, "Enhancing Pulmonary Disease Classification in Diseases: A Comparative Study of CNN and Optimized MobileNet Architectures," Journal of Robotics and Control (JRC), vol. 5, no. 2, pp. 427–440, 2024.

[48] K. Dong, C. Zhou, Y. Ruan, and Y. Li, "MobileNetV2 Model for Image Classification," in 2020 2nd International Conference on Information Technology and Computer Application (ITCA), Guangzhou, China: IEEE, Dec. 2020, pp. 476–480. doi: 10.1109/ITCA52113.2020.00106.

[49] P. Verma, V. Tripathi, and B. Pant, "Comparison of different optimizers implemented on the deep learning architectures for COVID-19 classification," Materials Today: Proceedings, vol. 46, pp. 11098–11102, 2021, doi: 10.1016/j.matpr.2021.02.244.

[50] R. Patel and A. Chaware, "Transfer Learning with Fine-Tuned MobileNetV2 for Diabetic Retinopathy," in 2020 International Conference for Emerging Technology (INCET), Belgaum, India: IEEE, Jun. 2020, pp. 1–4. doi: 10.1109/INCET49848.2020.9154014.

[51] Z. Riaz, B. Khan, S. Abdullah, S. Khan, and M. S. Islam, "Lung Tumor Image Segmentation from Computer Tomography Images Using MobileNetV2 and Transfer Learning," Bioengineering, vol. 10, no. 8, p. 981, Aug. 2023, doi: 10.3390/bioengineering10080981.

[52] A. Tripathi, T. Singh, R. R. Nair, and P. Duraisamy, "Improving Early Detection and Classification of Lung Diseases With Innovative MobileNetV2 Framework," IEEE Access, vol. 12, pp. 116202–116217, 2024, doi: 10.1109/ACCESS.2024.3440577.

[53] H. Wang, Q. Qi, W. Sun, X. Li, B. Dong, and C. Yao, "Classification of skin lesions with generative adversarial networks and improved MOBILENETV2," Int J Imaging Syst Tech, vol. 33, no. 5, pp. 1561–1576, Sep. 2023, doi: 10.1002/ima.22880.

[54] C. Buiu, V.-R. Dănăilă, and C. N. Răduţă, "MobileNetV2 Ensemble for Cervical Precancerous Lesions Classification," Processes, vol. 8, no. 5, p. 595, May 2020, doi: 10.3390/pr8050595.

[55] S. Gupta, K. Saluja, A. Goyal, A. Vajpayee, and V. Tiwari, "Comparing the performance of machine learning algorithms using estimated accuracy," Measurement: Sensors, vol. 24, p. 100432, Dec. 2022, doi: 10.1016/j.measen.2022.100432.

[56] H. R. Sofaer, J. A. Hoeting, and C. S. Jarnevich, "The area under the precision - recall curve as a performance metric for rare binary events," Methods Ecol Evol, vol. 10, no. 4, pp. 565‑577, Apr. 2019, doi: 10.1111/2041-210X.13140.

[57] R. Yacouby and D. Axman, "Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models," in Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems, Online: Association for Computational Linguistics, 2020, pp. 79–91. doi: 10.18653/v1/2020.eval4nlp-1.9.

[58] C. Bertinetto, J. Engel, and J. Jansen, "ANOVA simultaneous component analysis: A tutorial review," Analytica Chimica Acta: X, vol. 6, p. 100061, Nov. 2020, doi: 10.1016/j.acax.2020.100061.

[59] S. Winarno and H. A. Azies, "The Effectiveness of Continuous Formative Assessment in Hybrid Learning Models: An Empirical Analysis in Higher Education Institutions," International Journal of Pedagogy and Teacher Education, vol. 8, no. 1, p. 1, Jul. 2024, doi: 10.20961/ijpte.v8i1.89693.

[60] J. J. Goeman and A. Solari, "Comparing Three Groups," The American Statistician, vol. 76, no. 2, pp. 168–176, Apr. 2022, doi: 10.1080/00031305.2021.2002188.

[61] C. Ravichandran and G. Padmanaban, "A numerical simulation-based method to predict floor wise distribution of cooling loads in Indian residences using Tukey honest significant difference test," Advances in Building Energy Research, vol. 17, no. 1, pp. 1–29, Jan. 2023, doi: 10.1080/17512549.2022.2129449.

[62] S. Sudha Mishra and A. K. Das Mohapatra, "Weavers' perception towards sustainability of sambalpuri handloom: A Tukey's HSD test analysis," Materials Today: Proceedings, vol. 51, pp. 217–227, 2022, doi: 10.1016/j.matpr.2021.05.242.

[63] Muljono, S. A. Wulandari, H. A. Azies, M. Naufal, W. A. Prasetyanto, and F. A. Zahra, "Breaking Boundaries in Diagnosis: Non-Invasive Anemia Detection Empowered by AI," IEEE Access, vol. 12, pp. 9292–9307, 2024, doi: 10.1109/ACCESS.2024.3353788.

[64] S. H. Majeed and N. A. M. Isa, "Adaptive Entropy Index Histogram Equalization for Poor Contrast Images," IEEE Access, vol. 9, pp. 6402–6437, 2021, doi: 10.1109/ACCESS.2020.3048148.

[65] Md. Nahiduzzaman et al., "Diabetic retinopathy identification using parallel convolutional neural network based feature extractor and ELM classifier," Expert Systems with Applications, vol. 217, p. 119557, May 2023, doi: 10.1016/j.eswa.2023.119557.

[66] X. Deng, Q. Liu, Y. Deng, and S. Mahadevan, "An improved method to construct basic probability assignment based on the confusion matrix for classification problem," Information Sciences, vol. 340–341, pp. 250–261, May 2016, doi: 10.1016/j.ins.2016.01.033.