Article

Enhancing Road Safety: Deep Learning-Based Intelligent Driver Drowsiness Detection for Advanced Driver-Assistance Systems

Eunmok Yang [ID] and Okyeon Yi *

Department of Financial Information Security, Kookmin University, Seoul 02707, Republic of Korea; emyang@kookmin.ac.kr
* Correspondence: oyyi@kookmin.ac.kr

Abstract: Driver drowsiness detection is a significant element of Advanced Driver-Assistance Systems (ADASs), which utilize deep learning (DL) methods to improve road safety. A driver drowsiness detection system can trigger timely alerts like auditory or visual warnings, thereby stimulating drivers to take corrective measures and ultimately avoiding possible accidents caused by impaired driving. This study presents a Deep Learning-based Intelligent Driver Drowsiness Detection for Advanced Driver-Assistance Systems (DLID3-ADAS) technique. The DLID3-ADAS technique aims to enhance road safety via the detection of drowsiness among drivers. Using the DLID3-ADAS technique, complex features from images are derived through the use of the ShuffleNet approach. Moreover, the Northern Goshawk Optimization (NGO) algorithm is exploited for the selection of optimum hyperparameters for the ShuffleNet model. Lastly, an extreme learning machine (ELM) model is used to properly detect and classify the drowsiness states of drivers. The extensive set of experiments conducted based on the Yawdd driver database showed that the DLID3-ADAS technique achieves a higher performance compared to existing models, with a maximum accuracy of 97.05% and minimum computational time of 0.60 s.

Keywords: driver drowsiness detection; deep learning; northern goshawk optimization; road safety; ShuffleNet

1. Introduction

Transportation has been instrumental in human life, and a major portion of South Korea's economy arises from the transportation industry [1]. Although driving can be a means for quick and safe travel, fatigue, drowsiness, and a shortage of driver vigilance can result in accidents, including damages and fatalities. Driver drowsiness is accountable for a huge number of accidents all over the world [2]. Driver distraction signifies reduced focus on one's actions, which can be serious for protective driving without challenging activities. There are various factors that distract drivers' attention; however, in practice, only two major types have been studied: (1) fatigue and (2) distraction [3]. The term fatigue describes the integration of signs that diminish performance and a subjective sense of tiredness. Despite extensive research, the term fatigue still does not have a commonly recognized definition. The European Transport Safety Council (ETSC) refers to fatigue as tiredness in relation to an incapability or a reluctance to endure activities, commonly because such activities have been occurring for an extended period [4]. In addition, tiredness begins as an exterior representation of fatigue, which is predominant during driving. According to the aim of a study, the words fatigue and drowsiness are either utilized interchangeably or specified [5]. Consequently, a few research studies performed by the ETSC described four drowsiness stages according to user behavior, such as moderately awake, completely awake, strictly sleepy, and drowsy [6]. In this study, we categorized two stages, namely drowsy and alert, to minimize complexity and attain generalizable outcomes by employing a binary classification. Fatigue during driving has been shown to significantly increase

the risk of suffering from an accident. In particular, driving when feeling fatigue increases the possibility of a crash by 4–6 times compared to alert driving. Indeed, a major study draws the conclusion that fatigue is featured in 15% to 20% of accidents [7]. Therefore, the expansion of techniques efficient in monitoring driver fatigue conditions in real time, as well as having the ability to warn drivers shortly prior to falling asleep, is essential for avoiding collapses.

Driver drowsiness detection is dependent upon having a technique that starts recording a driver's steering behavior as soon as the activities of a trip begin [8]. Common indicators of waning concentration include signs showing that the driver is hardly steering, along with small, yet fast and sudden steering activities for maintaining the car on its path. Remarkably, the incorporation of deep learning (DL) methods has considerably improved signal and image processing functions to address real-time complexity. The advance of DL methods is evidenced in the significant developments in the domain of object recognition [9]. These improvements have played a crucial role in different industries, such as the field of security systems, the medical field, and the field of autonomous vehicles. Recently, the application of DL has led to a revolution in the identification of drowsiness indicators of every type, including vehicle motion, behavioral, and biological indicators [10]. The significant impact of DL in identifying these important features of drowsiness is transformational.

This study presents a Deep Learning-based Intelligent Driver Drowsiness Detection for Advanced Driver-Assistance Systems (DLID3-ADAS) technique. To accomplish this, the DLID3-ADAS technique initially pre-processes the input frames using a median filtering (MF) approach. Additionally, the complex features of the images are derived through the use of the ShuffleNet approach. Moreover, the Northern Goshawk Optimization (NGO) methodology is exploited to arrive at an optimum hyperparameter choice for the ShuffleNet algorithm. Lastly, an extreme learning machine (ELM) model is proposed to properly detect and classify the drowsiness states of drivers. This comprehensive comparison study shows that the DLID3-ADAS technique achieves enhanced performance over other approaches on various aspects. The key contributions are as follows:

- This study presents a new DLID3-ADAS method that follows a multi-stage approach to ensure that subsequent steps operate on refined and improved data, potentially leading to more accurate drowsiness detection.
- This study employs the ShuffleNet approach for feature extraction from input images. ShuffleNet, which is known for its performance and lower computational requirements, permits the extraction of complex features from image frames, allowing the model to capture intricate patterns connected with driver drowsiness.
- The hyperparameter optimization of the ShuffleNet model using the NGO algorithm with cross-validation helps boost the predictive outcomes of the DLID3-ADAS technique for unobserved data. This ensures the efficient tuning of model parameters, thereby improving the overall performance of the drowsiness detection technique.
- This study employs an ELM model for the final stage of driver drowsiness detection and classification. ELM, which is known for its fast training speed and simplicity, is effectively utilized to make timely and accurate predictions.
- The proposed model offers accurate driver drowsiness detection with maximum classification performance and minimum computational complexity.

2. Relevant Works in the Literature

In [11], a bio-sensing probe that coupled LEDs in the near-infrared (NiR) spectrum with a photodetector named PhotoPlethysmoGraphy (PPG) was developed. The PPG signal generation was controlled by modifying non-oxygenated and oxygenated hemoglobin concentrations in a monitored subject's bloodstream, which could be promptly associated with cardiac activities as measured in the Autonomic Nervous System (ANS) to describe the subject's drowsiness stage. Phan et al. [12] introduced two effective techniques with three conditions for a doze alert model. One technique utilizes previously implemented

facial landmarks for identifying yawns and blinks, which are dependent upon the correct thresholds for all drivers. Then, it employs DL methods with two modified deep neural networks (DNNs) based on ResNet-50V2 and MobileNet-V2. The other technique examines videos and recognizes the activities of the driver in all frames to automatically learn each feature. In [13], the applicability of an Advanced Driver-Assistance System (ADAS), a DL-based driver tiredness identification technique, is presented. Initially, the face area of drivers could be recognized by employing the SSD MobileNet object detection algorithm. The identified head, mouth, and eye positions were monitored and verified over time. Lastly, these models can be combined with Convolutional Neural Networks (CNNs).

Rundo et al. [14] designed an ADAS based upon the deployment of an ad hoc devised bio-inspired sensing platform that samples drivers' PPG signal, which is correctly inter-related with the corresponding stage of subjective attention. A novel downstream deep model is implemented to sufficiently process the driver's PPG signal by reconstructing the corresponding attention stage. Sinha et al. [15] implemented various frameworks for analyzing the effectiveness of drowsiness detection based on facial areas. An innovative identification technique was developed by employing DL methods. To evaluate drivers' condition, facial areas related to the whole face could be utilized. Various approaches were implemented for facial recognition, such as Yolo V3, Viola Jones, and DLib. The CNN method implemented for drowsiness identification was adapted from LeNet for classification. Kumar et al. [16] proposed a non-invasive technique for detecting drowsiness in drivers. Facial features were utilized to detect the drowsiness of drivers. The eye and mouth regions were extracted and analyzed using a hybrid DL algorithm. The hybrid DL technique was developed by integrating both a long short-term memory (LSTM) technique and an adapted InceptionV3. InceptionV3 was adapted by including a global average pooling layer for dropout and spatial robustness methods.

In [17], a solution was established that addresses all the difficulties of the Raspberry Pi camera but remains quite portable and proficient. By employing the OpenCV technique and tasks, the Raspberry Pi camera converts the video stream of a driver to grayscale and frame images. The eye sets are taken into consideration, and the Eye Aspect Ratio and face landmarks are obtained through Euclidean distances. Patel et al. [18] developed a driver tiredness identification method that employs eye blinking, mouth closing, and mouth opening amounts to recognize drowsiness. An alert sound is produced once the driver's eyes are closed for a long period. The output of the model is established based on the DL algorithm of Dlib, which implements a CNN as its baseline method for accurate recognition, OpenCV, and the Raspberry Pi platform with an attached camera.

Jeong and Ko [19] developed a fast FER method to monitor drivers' reactions, which is capable of working in lower-measurement devices installed in vehicles. For this purpose, a hierarchical weighted random forest (WRF) algorithm, which is trained based upon the similarity of data to increase its accurateness, is exploited. Zhao et al. [20] introduced an innovative research study using a dynamic FER, while employing near-infrared (NIR) video sequences and LBP-TOP (local binary patterns from three orthogonal planes) feature descriptors. NIR images integrated with LBP-TOP features offer an illumination-invariant description of facial video frames. Chen et al. [21] developed a multi-modal fusion-based FER system that is efficient in precisely identifying facial expressions irrespective of lighting cases and head positions, utilizing a structured-light imaging camera that gives three modalities of images—Depth Maps, near-infrared (NIR), and RGB. Majeed et al. [22] designed a deep neural network (DNN) framework for driver drowsiness identification by exploiting an CNN.

The existing landscape of driver drowsiness detection through the use of ADASs that employ DL techniques has witnessed important developments. However, a major research gap exists, highlighting the vital requirement for systematic exploration and optimization of hyperparameters. While there have been significant developments in DL approaches in this field, the literature has only studied the impact of hyperparameter tuning on the overall effectiveness of driver drowsiness detection models. Overcoming this research

gap is essential to fully utilize DL techniques, confirm their robustness and adaptability in real-time driving conditions, and finally enrich the reliability and efficiency of ADASs.

3. The Proposed Method

In this study, we established a novel DLID3-ADAS technique. The DLID3-ADAS technique aims to enhance road safety via the detection of drowsiness among drivers. It comprises many processes, namely MF preprocessing, ShuffleNet-based feature extraction, NGO-based parameter tuning, and ELM classification. Figure 1 describes the overall procedure of the presented DLID3-ADAS technique. In this study, the selection of methods, including ShuffleNet and MF, is based on their proven efficiency in different fields and their possible collaboration to address the particular challenge of driver drowsiness detection. While each individual method has been commonly used, their incorporation within the DLID3-ADAS architecture is new and strategic. The reason behind using MF lies in its capability to alleviate noise in an input frame, which provides a clean foundation for the extraction of succeeding features. ShuffleNet, renowned for its efficacy in extracting complicated features while maintaining computation efficacy, is chosen to capture complex patterns that are critical for identifying signs of driver drowsiness. The combination of the NGO technique for tuning the hyperparameters in ShuffleNet is a deliberate attempt to optimize the accuracy of the model. The incorporation of these techniques within the DLID3-ADAS method is a thoughtful synthesis and a simple combination, which leverages the strength of all the methods to create an advanced and cohesive technique for detecting driver drowsiness. The contribution lies in the strategic incorporation and optimization of these algorithms within a unified framework that is specifically tailored for enhancing road safety while building upon existing methods.
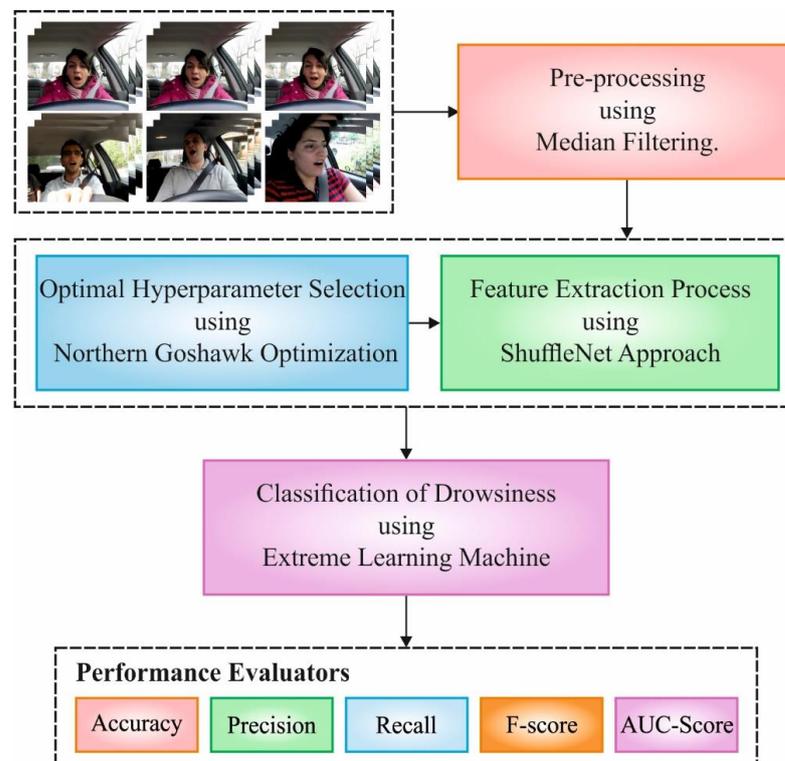

Figure 1. The overall procedure of the DLID3-ADAS system.

3.1. Preprocessing

The DLID3-ADAS technique initially pre-processes the input frames using the MF approach to improve the quality and robustness of the input frames [23]. This method enrolls the help of a MF technique to pre-process raw image frames taken by in-car cameras.

The MF technique efficiently decreases noise and artifacts in the images and provides enhanced feature extraction in the following analysis. By changing all pixel intensity values to the median value and its local neighborhood, MF effectively mitigates the effect of outliers and enriches the precision of the input frames. This pre-processing stage not only enhances the visual input but also ensures the preservation of reliability and accuracy in subsequent stages with the use of other DL methods. This contributes to achieving a more effective and robust driver drowsiness identification method in the ADAS.

3.2. ShuffleNet-Based Feature Extraction

In this work, the complex features of images are derived through the use of the ShuffleNet approach. ShuffleNet significantly decreases the computational rate and attains exceptional effectiveness while maintaining computational accuracy [24]. Essentially, combined convolution has been employed in the AlexNet network model, and a few effective NN architectures, namely MobileNetv2 and Xception, present depthwise separable convolution depending on the group convolution. While the capability of the method as well as the quantity of computation are coordinated, the addition of point-by-point convolution takes a huge portion; thus, pixel-level group convolution is presented in the ShuffleNet architecture to decrease the 1×1 convolutional function. However, the convolutional process could be limited for all group-by-group convolutions, which decreases the model's computational complexity. However, while numerous group convolutions are weighted, the output channel can arise from a smaller section of the input network to be positioned. The output is only compared to the inputs, and other groups' data could be attained. The data among the groups are not compatible with each other, which hinders the data flow among the networks within the groups [25]. The input and output networks of ShuffleNet could be fixed to a similar amount to decrease memory consumption. In the case that the feature map size is $h \times w$, the convolution of height and width is 1×1 as a model, and input and output counts are C_1 and C_2, respectively. Based on the multiply–accumulate operations (MACs) and Float Operations (FLOPs), the computational equation is as follows:

$$B = h \cdot w(C_1 \times (1 \times 1) \times C_2) = h \cdot w \cdot C_1 \cdot C_2, \tag{1}$$

$$MAC = h \cdot w \cdot C_1 + h \cdot w \cdot C_2 + (1 \times 1) \times C_1 \times C_2 = h \cdot w \cdot (C_1 + C_2) + C_1 \times C_2 \tag{2}$$

By using inequality,

$$(C_1 + C_2) \geq 2\sqrt{C_1 C_2}, \tag{3}$$

when

$$MAC \geq 2\sqrt{h \cdot w \cdot B} + \frac{B}{h \cdot w} \tag{4}$$

In Equation (4), w and h are the mapping feature's width and height, B represents the FLOPs, and MAC denotes the network layer memory accessibility as well as the write and read consumption rates. Thus, if $C_1 = C_2$, once the input channel corresponds to the output channel, the memory consumption will be minimum.

3.3. Parameter Tuning Using the NGO Model

At this stage, the NGO method is exploited to attain a better hyperparameter choice for the ShuffleNet algorithm. The NGO approach shows its performance in parameter tuning the ShuffleNet method. By leveraging the model's ability to effectively explore the performance space and utilize potential areas, the NGO approach systematically fine-tunes the parameters, namely the learning rates, batch sizes, and architectural parameters of ShuffleNet. This optimization method is proposed to improve the model's efficiency by modifying its configuration, thereby leading to enhanced accuracy, decreased overfitting, and improved generalization abilities. The NGO approach's adaptability and versatility generate the potential for efficiently directing the complex hyperparameter landscape of deep neural networks (DNNs), thus contributing to the optimization and modification of ShuffleNet for various tasks and databases. The NGO algorithm is a new swarm intelligent

optimization technique that stimulates the hunting behaviors of northern goshawks [26]. The NGO algorithm has been developed to enhance this selection process and demonstrates high accuracy and stability, as well as exceptional optimization performance. The fundamental steps of the NGO algorithm are given below:

Step 1: Population initialization

A population matrix X is first generated, and population members are initialized randomly within the search range as follows:

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{N \times m} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,j} & \cdots & x_{1,m} \\ \vdots & & \vdots & & \vdots \\ x_{i,1} & \cdots & x_{i,j} & \cdots & x_{i,m} \\ \vdots & & \vdots & & \vdots \\ x_{N,1} & \cdots & x_{N,j} & \cdots & x_{N,m} \end{bmatrix}_{N \times m} \tag{5}$$

where the location of the ith northern goshawk is X_i; N denotes the number of population members and set the maximum dimension m of the solution; $x_{i,j}$ represents the ith northern goshawk in the jth dimension; and F and F_i are the objective function values of the ith population and their formula is expressed as follows:

$$F(X) = \begin{bmatrix} F_1 = F(X_1) \\ F_i = F(X_i) \\ F_1 = F(X_N) \end{bmatrix}_{N \times 1} \tag{6}$$

Step 2: Recognize prey and attack

The initial phase of hunting a target when emulating a northern goshawk is to randomly choose the target and perform a rapid hunt. The aim is to detect a better area in the search range, and a global search is performed. The ith prey is fixed to $P_{i,j}$ as the newest location in the jth dimension; r and I are randomly generated integers during the iteration process; and $x_{i,j}$ and $P_{i,j}$ are the locations of the goshawk and the prey.

$$\begin{cases} P_j = X_k \\ i \in 1, 2, \cdots, N \\ k \in 1, 2, \cdots, i-1, i+1, \cdots, N \end{cases} \tag{7}$$

$$x_{i,j}^{new,P_1} = \begin{cases} x_{i,j} + r\left(p_{i,j} - Ix_{i,j}\right), & F_{P_i} < F_i \\ x_{i,j} + r\left(x_{i,j} - p_{i,j}\right), & F_{P_j} \geq F_i \end{cases} \tag{8}$$

The updated prey location P_1, the updated northern goshawk location x_i^{new, P_1}, and the updated objective function value F_i^{new, P_1} are obtained after the first stage of the attack:

$$X_i = \begin{cases} X_i^{new,P_1}, & F_i^{new,P_1} < F_j \\ X_{i,} & F_i^{new,P_1} \geq F_j \end{cases} \tag{9}$$

Step 3: Pursuit and Escape

Once the northern goshawk has attacked the prey, it then attempts to get out and continues hunting for another target. Based on its speed, the northern goshawk can capture a target in almost any situation [27]. These simulation behaviors increase the model's ability to conduct a local search within the search range. It is assumed that the northern goshawk has the current iteration number t, a hunting range of about R, and the maximum iteration

counts T. Then, the location $x_{i,j}^{new, P_2}$ in the jth dimension of the ith goshawk is attained as follows:

$$R = 0.02\left(1 - \frac{t}{T}\right) \tag{10}$$

$$x_{i,j}^{new,P_2} = x_{i,j} + R(2r - 1)x_{i,j} \tag{11}$$

$$X_i = \begin{cases} X_i^{new,P_2}, & F_i^{new,P_2} < F_j \\ X_{i,} & F_i^{new,P_2} \geq F_j \end{cases} \tag{12}$$

where X_i^{new, P_2} and F_i^{new, P_2} are the position and objective function values of the ith northern goshawk after the iterative and second-stage upgrade.

Fitness selection is a crucial factor in the GWO approach. Solution encoding is exploited for assessing the aptitude (goodness) of the candidate solution. Here, the accuracy value is the main condition utilized for designing a fitness function:

$$\text{Fitness} = \max(P) \tag{13}$$

$$P = \frac{TP}{TP + FP} \tag{14}$$

Based on the above the mathematical formulae, FP and TP are the false-positive and true-positive values.

3.4. ELM Classification

Finally, the ELM model detects and classifies the drowsiness of drivers. ELM is a method based on single-layer feedforward networks (FFNs) whose main goal is to interlink a natural learning device to the neural networks [28]. Owing to its superior structure, which depends on random hidden neuron devices that do not need to be adjusted to be parallel to conventional ANNs, it can deliver precise outcomes with a low computational cost. Furthermore, ELM provides other benefits like comfort of execution, superior simplification capacity, and minimum human involvement. Because of these advantages, ELM is selected as the central machine learning (ML) technique. The main scientific method underlying ELM is defined below.

A single hidden layer (HL) exists in the system of ELM. The input weights of the input layers and HL are set once and do not need to be trained. Iterative testing is mainly utilized to fix the outcome weights of the outcome layers and HL. The training and development period of an ELM model is faster when compared to single-layer FFNs, so the input weights endure in their early state and only the output weights are trained [29]. Figure 2 portrays the architecture of the ELM model. In 2006, Huang et al. proposed an ELM model with a simple 3-layer structure procedure to find the faults of conventional soft calculating methods. The bias values and input weight are produced arbitrarily in the structure of the ELM model. It employs a simple inverse process of the HL output matrix to scale an output weighted matrix among the output layers and HL logically. In addition, it is a great time-sequence forecast technique due to its interpolation and estimation abilities. The ELM model is signified arithmetically as a function with training data (N) and hidden nodes (L) as follows:

$$\sum_{i=1}^{L} w_i g\left(W_{in(i)}, b_i, x_j\right) = \sum_{i=1}^{L} w_i g\left(W_{in(i)}, x_j + b_i\right) = Y_j, j = 1, \dots, N. \tag{15}$$

where w_i implies the weighted matrix; x_j signifies the input vector; $W_{in(i)}$ represents the weighted vector; $W_{in(i)}.x_j$ denotes the inner product of $W_{in(i)}$; x_j and b_i symbolize the bias of ith hidden nodes; $g(\bullet)$ represents the sigmoid function; and y_j stands for the output. An input weight and bias are chosen arbitrarily at the beginning of the ELM procedure.
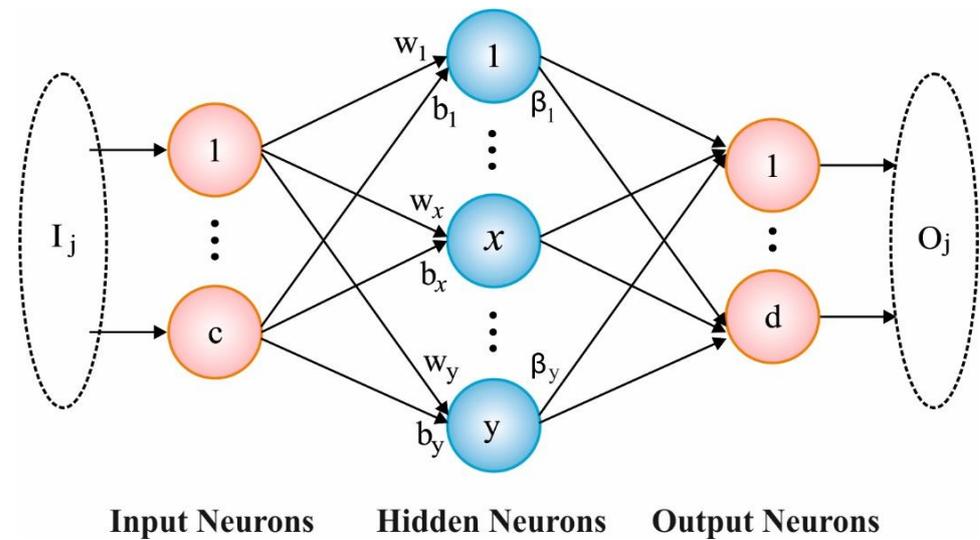
Figure 2. ELM structure.

4. Result Analysis and Discussion

The proposed model was simulated using the Python 3.8.5 tool. The proposed model was experimented on a PC with i5-8600k, GeForce 1050Ti 4 GB, 16 GB RAM, 250 GB SSD, and 1 TB HDD. In this study, drowsiness detection among drivers was investigated by using the Yawdd driver database from the Kaggle repository [30]. The dataset comprises 500 instances divided into two classes, as shown in Table 1. Figure 3 displays sample drowsiness and non-drowsiness images.

Table 1. Details of dataset.

Classes	No. of Samples
Drowsiness	250
Non-drowsiness	250
Total Samples	500

Figure 4 shows the classification analysis of the DLID3-ADAS technique at a ratio of 80:20 for TRPH/TSPH. Figure 4a,b display the confusion matrices accomplished by the DLID3-ADAS technique. This simulation value shows that the DLID3-ADAS method can precisely categorize and identify images into two class labels. Figure 4c shows the PR analysis of the DLID3-ADAS method. This figure shows that the DLID3-ADAS technique attains exceptional PR analysis under each class. To conclude, Figure 4d displays the ROC analysis of the DLID3-ADAS technique. The outcome shows that the DLID3-ADAS technique provides efficient findings with increased ROC values for diverse class labels.

The drowsiness detection ability of the DLID3-ADAS technique was examined at a ratio of 80:20 for TRPH/TSPH, and the results are shown in Table 2 and Figure 5. The obtained findings show that the DLID3-ADAS technique can properly differentiate the drowsiness and non-drowsiness classes. According to an 80% of TRPH, the DLID3-ADAS technique offers an enhanced $accu_y$ of 96.23%, a $prec_n$ of 96.29%, a $reca_l$ of 96.23%, an F_{score} of 96.25%, and an AUC_{score} of 96.23%. Additionally, based on a 20% of TSPH, the DLID3-ADAS technique boosts the values to an $accu_y$ of 97.05%, a $prec_n$ of 96.96%, a $reca_l$ of 97.05%, an F_{score} of 96.99%, and an AUC_{score} of 97.05%.

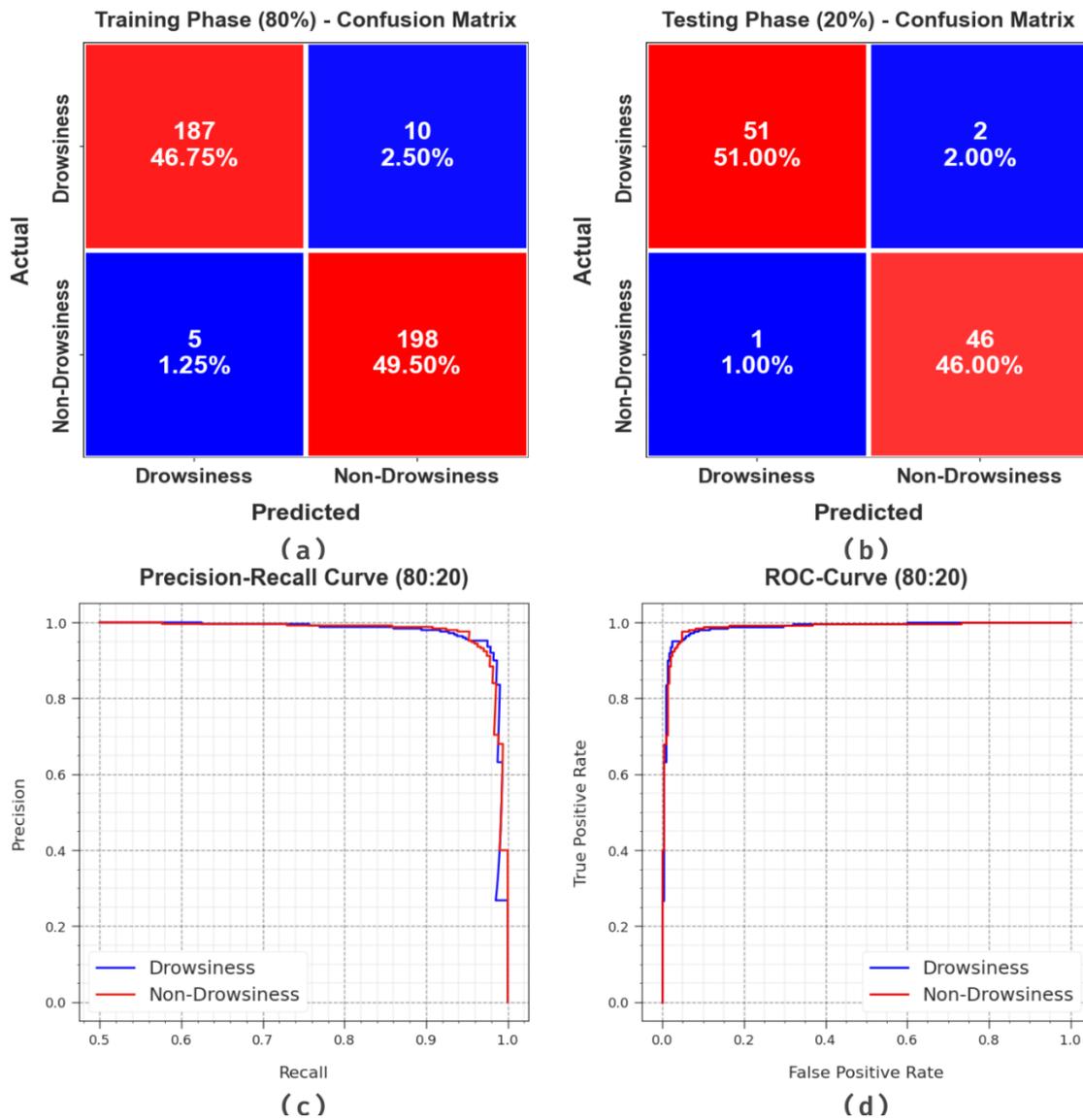**Figure 3.** (**a**) Drowsiness and (**b**) non-drowsiness images.

Figure 4. (**a**,**b**) Confusion matrices based on a ratio of 80:20 for TRPH/TSPH; (**c**) PR curve based on a ratio of 80:20 for TRPH/TSPH; and (**d**) ROC curve based on a ratio of 80:20 for TRPH/TSPH.

Table 2. Drowsiness detection of the DLID3-ADAS model at a ratio of 80:20 for TRPH/TSPH.

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	AUC_{score}
TRPH (80%)					
Drowsiness	94.92	97.40	94.92	96.14	96.23
Non-drowsiness	97.54	95.19	97.54	96.35	96.23
Average	96.23	96.29	96.23	96.25	96.23
TSPH (20%)					
Drowsiness	96.23	98.08	96.23	97.14	97.05
Non-drowsiness	97.87	95.83	97.87	96.84	97.05
Average	97.05	96.96	97.05	96.99	97.05
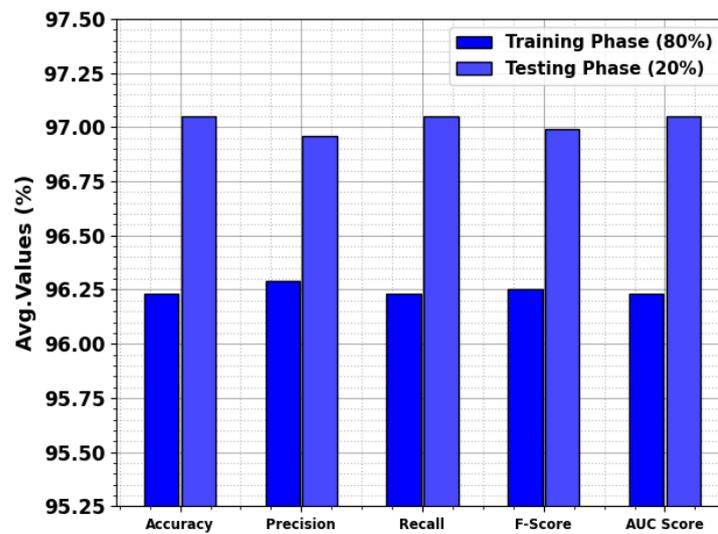
Figure 5. The average outcome of the DLID3-ADAS system under a ratio of 80:20 for TRPH/TSPH.

The $accu_y$ curves for training (TR) and validation (VL) portrayed in Figure 6 for the DLID3-ADAS method under a ratio of 80:20 for TRPH/TSPH offer valuable insights into its effectiveness on varied epochs. Primarily, they show a constant upgrading in both TR and TS $accu_y$ with higher epochs, demonstrating the effectiveness of the model in learning and recognizing the patterns of both data of TR and TS. The upgrading tendency in the TS $accu_y$ highlights the adaptability of the model for the TR dataset and its ability to generate correct predictions when using unobserved data, underscoring its capability in generating results with robust generalization.
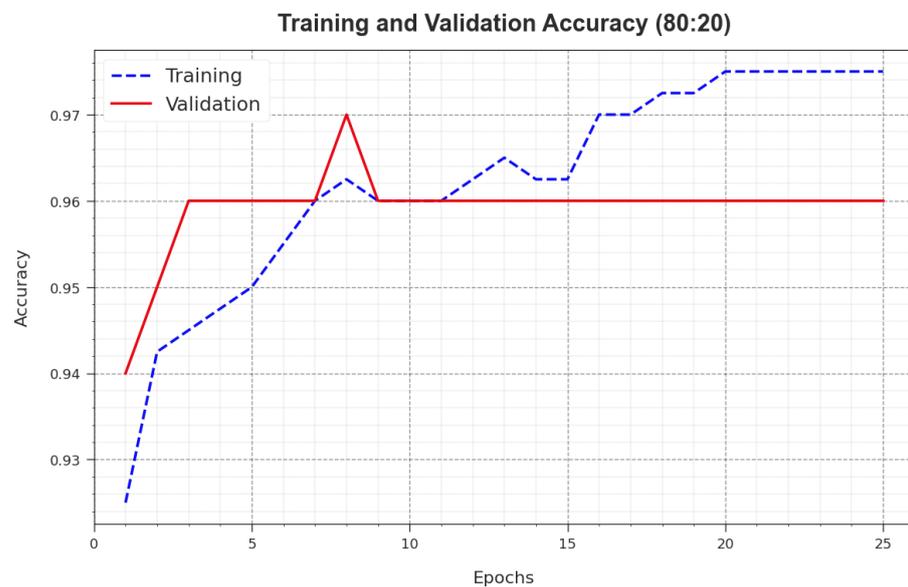
Figure 6. $Accu_y$ curves of the DLID3-ADAS model under a ratio of 80:20 for TRPH/TSPH.

Figure 7 illustrates a wide-ranging overview of the TR and TS loss values obtained by the DLID3-ADAS technique under a ratio of 80:20 for TRPH/TSPH across diverse epochs. The TR loss reliably decreases as the model upgrades the weights to minimize the classification errors in these datasets. The loss curves represent the model's alignment with the TR data, underscoring its ability to capture patterns. What is significant is the incessant improvement in parameters when using the DLID3-ADAS technique, which is developed to diminish the differences between predictions and actual TR labels.
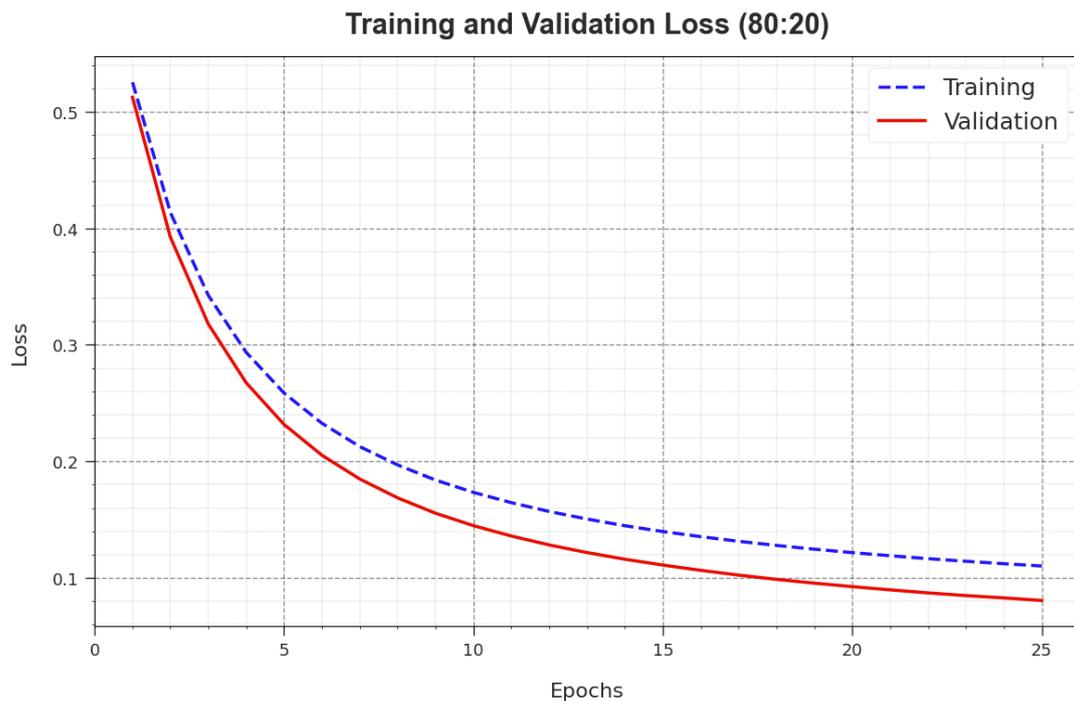
Training and Validation Loss (80:20)

Figure 7. Loss curve of the DLID3-ADAS technique under a ratio of 80:20 for TRPH/TSPH.

Figure 8 shows the classification analysis of the DLID3-ADAS technique under a ratio of 70:30 for TRPH/TSPH. Figure 8a,b illustrate the confusion matrices acquired by the DLID3-ADAS algorithm. This outcome reveals that the DLID3-ADAS technique can correctly categorize and classify the images into two class labels. Next, Figure 8c presents the PR curve of the DLID3-ADAS technique. This result shows that the DLID3-ADAS technique attains exceptional PR analysis under each class. In conclusion, Figure 8d demonstrates the ROC curve of the DLID3-ADAS technique. The simulation value shows that the DLID3-ADAS technique offers efficient findings with boosted ROC values for diverse classes.

The drowsiness detection ability of the DLID3-ADAS method was examined under a ratio of 70:30 for TRPH/TSPH, and the results are shown in Table 3 and Figure 9. The experimental outcome shows that the DLID3-ADAS method appropriately distinguishes the non-drowsiness and drowsiness classes. Based on a 70% of TRPH, the DLID3-ADAS method achieves an increased $accu_y$ of 95.03%, a $prec_n$ of 95.40%, a $reca_l$ of 95.03%, an F_{score} of 95.12%, and an AUC_{score} of 95.03%. In addition, with a 30% of TSPH, the DLID3-ADAS method attains an increased $accu_y$ of 95.68%, a $prec_n$ of 95.39%, a $reca_l$ of 95.68%, an F_{score} of 95.33%, and an AUC_{score} of 95.68%.

The $accu_y$ curves for TR and VL shown in Figure 10 for the DLID3-ADAS technique under a ratio of 70:30 for TRPH/TSPH offers valuable insights into its effectiveness across various epochs. Essentially, they show a continuous upgrading in both TR and TS $accu_y$ with increased epochs, demonstrating the effectiveness of the model in learning and recognizing the patterns in both data of TR and TS. The increasing trend in TS $accu_y$ highlights the adaptability of the model for the dataset of TR and its ability to make precise predictions on unobserved data, underscoring its capability in generating results with robust generalization.
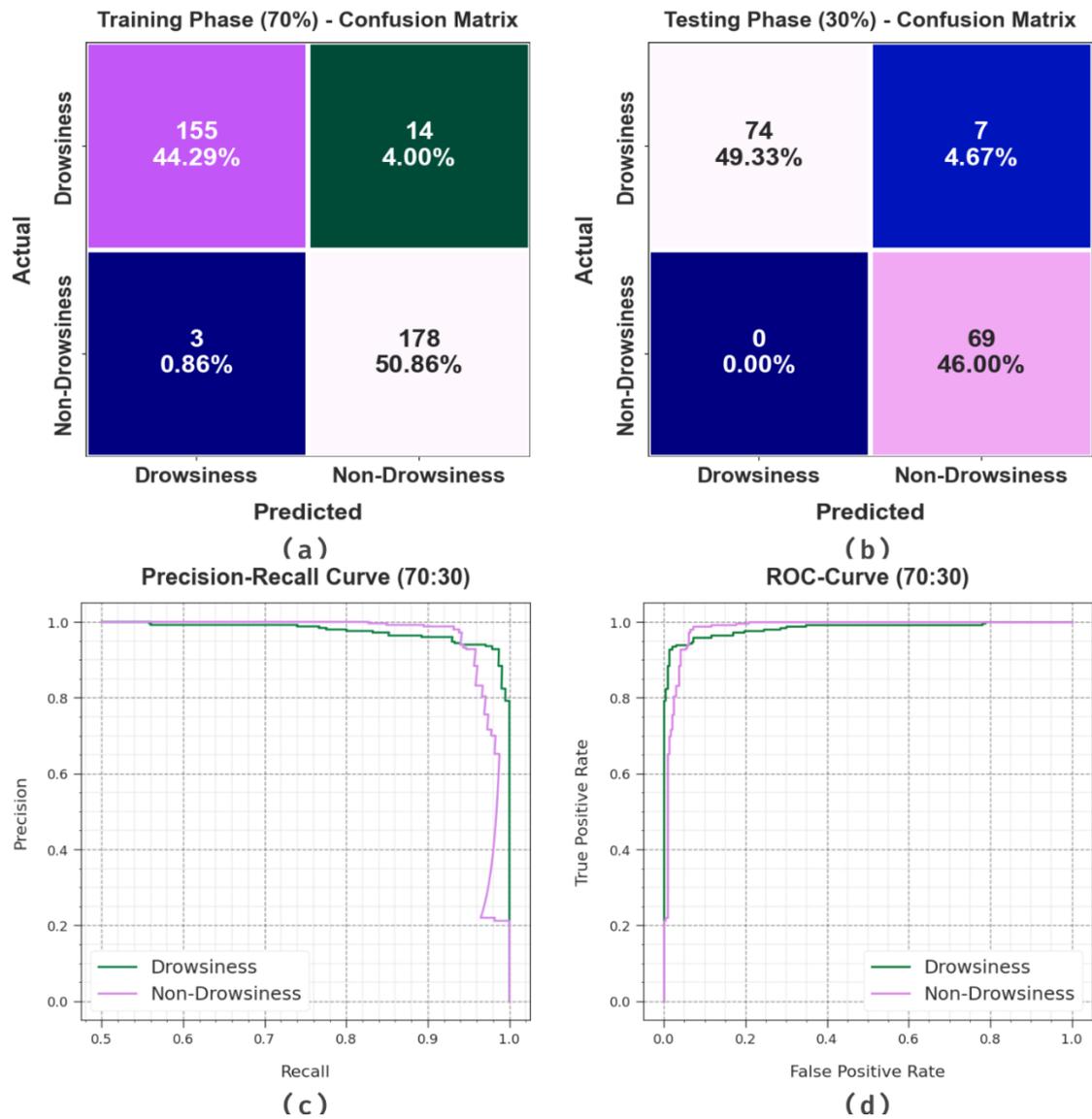
Figure 8. (**a**,**b**) Confusion matrices with a ratio of 70:30 for TRPH/TSPH; (**c**) PR curve under a ratio of 70:30 for TRPH/TSPH; and (**d**) ROC curve under a ratio of 70:30 for TRPH/TSPH.

Table 3. Drowsiness detection of the DLID3-ADAS system under a ratio of 70:30 for TRPH/TSPH.

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	AUC_{score}
TRPH (70%)					
Drowsiness	91.72	98.10	91.72	94.80	95.03
Non-drowsiness	98.34	92.71	98.34	95.44	95.03
Average	95.03	95.40	95.03	95.12	95.03
TSPH (30%)					
Drowsiness	91.36	100.00	91.36	95.48	95.68
Non-drowsiness	100.00	90.79	100.00	95.17	95.68
Average	95.68	95.39	95.68	95.33	95.68
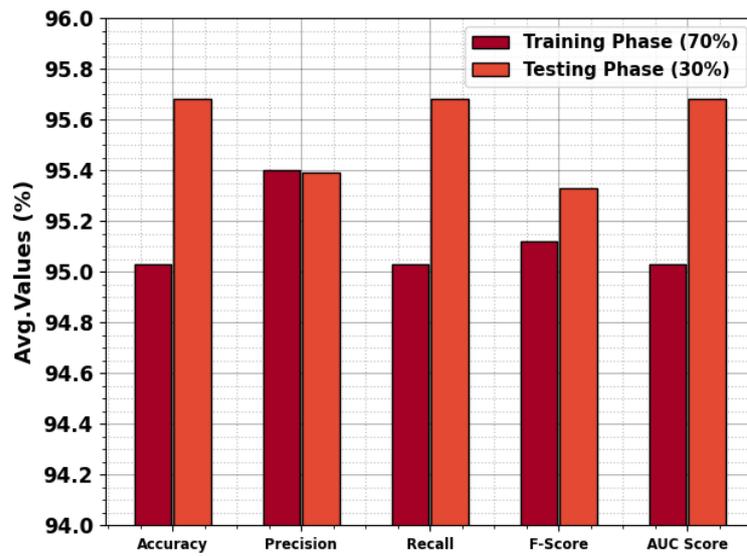
Figure 9. The average outcome of the DLID3-ADAS algorithm under a ratio of 70:30 for TRPH/TSPH.
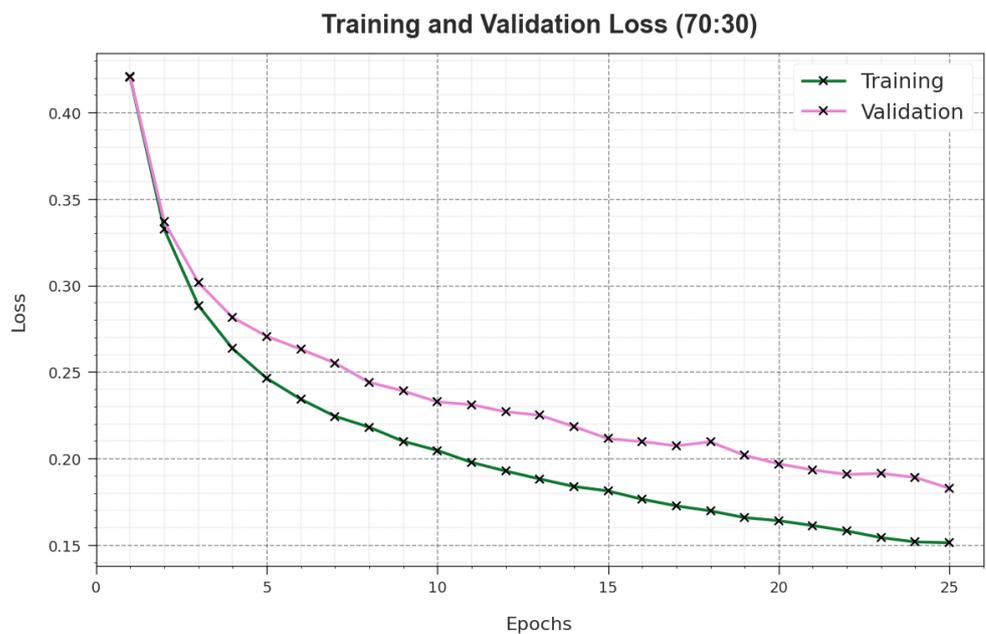
Figure 10. $Accu_y$ curves of the DLID3-ADAS model under a ratio of 70:30 for TRPH/TSPH.

Figure 11 displays a comprehensive overview of the TR and TS loss values for the DLID3-ADAS technique under a ratio of 70:30 for TRPH/TSPH across diverse epochs. The TR loss dependably diminishes as the model upgrades the weights to decrease the classification errors in both datasets. These loss curves noticeably reveal the model's alignment with the TR dataset, underscoring its abilities to effectively capture patterns. What is significant is the continuous enhancement in parameters when using the DLID3-ADAS technique, which is aimed at minimizing the discrepancies between predictions and actual TR labels.

Training and Validation Accuracy (70:30)
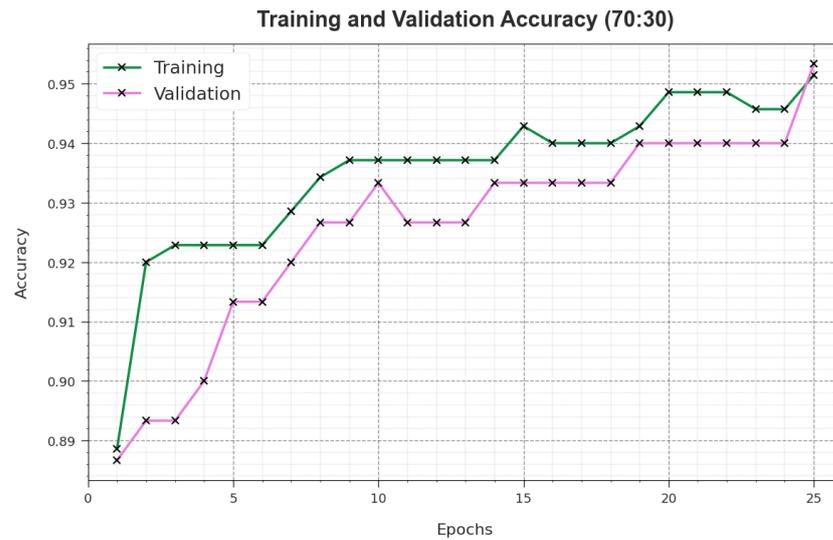
Figure 11. Loss curves of the DLID3-ADAS method under a ratio of 70:30 for TRPH/TSPH.

To confirm the improved performance of the DLID3-ADAS method, a brief comparison with other methods was conducted, and the results are shown in Table 4 and Figure 12 [22]. The comparative findings demonstrate that the DLID3-ADAS technique achieves exceptional performance over other models. In terms of $accu_y$, the DLID3-ADAS technique achieves an increased $accu_y$ of 97.05% while the YOLOv3-tiny CNN, SVM, LSTM NN, Dlib + linear SVM, 2s-STGCN, Dlib + 15-layer CNN, and 3D Deep CNN models attain smaller $accu_y$ values of 94.32%, 89.00%, 88.00%, 92.50%, 93.40%, 96.69%, and 96.80%, respectively.

Table 4. Comparison analysis of the DLID3-ADAS technique with other models.

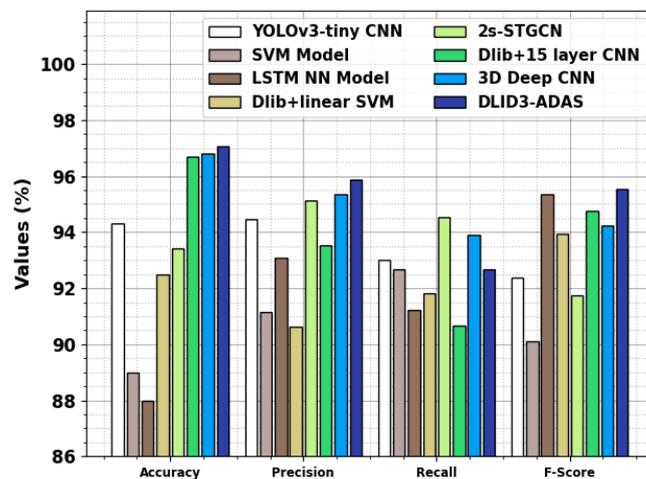| Methodology	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}
YOLOv3-tiny CNN	94.32	94.46	93.01	92.37
SVM Model	89.00	91.16	92.67	90.10
LSTM NN Model	88.00	93.08	91.22	95.35
Dlib + linear SVM	92.50	90.61	91.80	93.92
2s-STGCN	93.40	95.13	94.55	91.74
Dlib + 15-layer CNN	96.69	93.51	90.67	94.76
3D Deep CNN	96.80	95.36	93.90	94.22
DLID3-ADAS	97.05	95.87	92.69	95.53

Figure 12. Comparison analysis of the DLID3-ADAS technique with other models.

The results of a computational time (CT) analysis of the DLID3-ADAS technique when compared to other models are shown in Table 5 and Figure 13. These findings show that the DLID3-ADAS technique obtains a higher performance compared to other algorithms. In terms of *CT*, the DLID3-ADAS method achieves a minimum *CT* of 0.60 s, whereas the YOLOv3-tiny CNN, SVM, LSTM NN, Dlib + linear SVM, 2s-STGCN, Dlib + 15-layer CNN, and 3D Deep CNN methods attain larger *CT* values of 2.16 s, 1.60 s, 2.26 s, 1.67 s, 1.24 s, 1.52 s, and 2.16 s, respectively.

Table 5. CT analysis of the DLID3-ADAS technique compared to other models.

Methodology	Computational Time (s)
YOLOv3-tiny CNN	2.16
SVM Model	1.60
LSTM NN Model	2.26
Dlib + linear SVM	1.67
2s-STGCN	1.24
Dlib + 15-layer CNN	1.52
3D Deep CNN	2.16
DLID3-ADAS	0.60
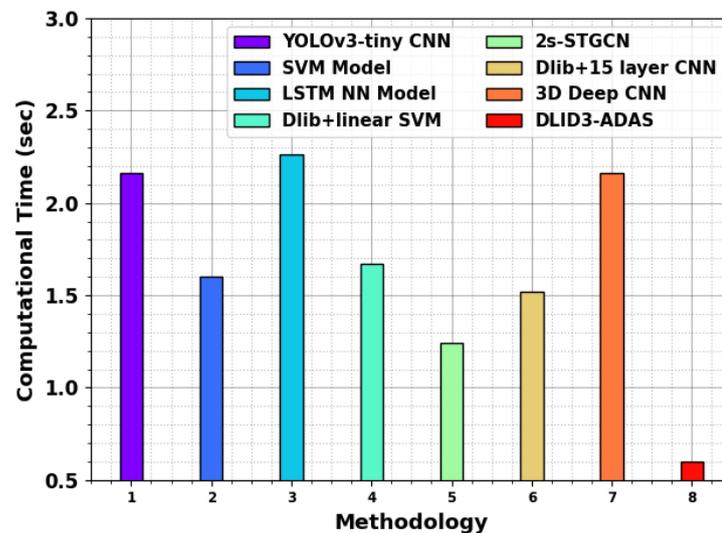
Figure 13. CT analysis of the DLID3-ADAS technique compared to other models.

5. Conclusions

In this study, we presented an innovative DLID3-ADAS technique. The DLID3-ADAS technique aims to enhance road safety via the detection of drowsiness among drivers. It comprises many processes, namely MF preprocessing, ShuffleNet-based feature extraction, NGO-based parameter tuning, and ELM classification. To accomplish drowsiness detection, the DLID3-ADAS technique initially pre-processes the input frames using the MF approach. Additionally, the complex features of the images are derived through the use of the ShuffleNet approach. Moreover, the NGO algorithm is exploited to select the optimum hyperparameter choice for the ShuffleNet approach. Finally, the ELM technique detects and classifies the drowsiness states of drivers. To examine the performance of the DLID3-ADAS technique, a sequence of simulations was conducted. The comprehensive comparative analysis shows that the DLID3-ADAS technique achieves superior performance when compared to other methods, with a maximum accuracy of 97.05% and a minimum computational time of 0.60 s.

The DLID3-ADAS architecture has certain limitations despite the promising outcomes. The binary classification of non-drowsy and drowsy states, while practical, may oversimplify the states of driver fatigue, potentially missing slight variation in drowsiness level.

Furthermore, the model's efficiency depends on the availability of comprehensive and diverse datasets, and the model's generalization to diverse driving conditions requires further exploration. In future work, an objective is to resolve this limitation by integrating nuanced drowsiness levels, leveraging large and different datasets, and expanding the applicability of the model to various driving scenarios. Furthermore, exploring adaptive learning mechanisms and refining the algorithm's real-time abilities will be vital to enhance the robustness and real-world deployment of the DLID3-ADAS architecture. An avenue for future research is to improve the accuracy and sophistication of the proposed model in detecting driver drowsiness, ultimately contributing to enhanced road safety outcomes.

Data Availability Statement: The data presented in this study are openly available in [Kaggle. YawDD driver dataset, reference number [30].

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ashraf, I.; Hur, S.; Shafiq, M.; Park, Y. Catastrophic factors involved in road accidents: Underlying causes and descriptive analysis. *PLoS ONE* **2019**, *14*, e0223473. [CrossRef]
2. Rosen, H.E.; Bari, I.; Paichadze, N.; Peden, M.; Khayesi, M.; Monclús, J.; Hyder, A.A. Global road safety 2010–2018: An analysis of global status reports. *Injury* **2022**. [CrossRef]
3. Saleem, A.A.; Siddiqui, H.U.R.; Raza, M.A.; Rustam, F.; Dudley, S.; Ashraf, I. A systematic review of physiological signals based driver drowsiness detection systems. *Cogn. Neurodyn.* **2023**, *17*, 1229–1259. [CrossRef]
4. Albadawi, Y.; Takruri, M.; Awad, M. A review of recent developments in driver drowsiness detection systems. *Sensors* **2022**, *22*, 2069. [CrossRef]
5. Forsman, P.M.; Vila, B.J.; Short, R.A.; Mott, C.G.; Van Dongen, H.P. Efficient driver drowsiness detection at moderate levels of drowsiness. *Accid. Anal. Prev.* **2013**, *50*, 341–350. [CrossRef]
6. Caldwell, J.A.; Caldwell, J.L.; Thompson, L.A.; Lieberman, H.R. Fatigue and its management in the workplace. *Neurosci. Biobehav. Rev.* **2019**, *96*, 272–289. [CrossRef] [PubMed]
7. Kanwal, K.; Rustam, F.; Chaganti, R.; Jurcut, A.D.; Ashraf, I. Smartphone Inertial Measurement Unit Data Features for Analyzing Driver Driving Behavior. *IEEE Sens. J.* **2023**, *23*, 11308–11323. [CrossRef]
8. Cui, Y.; Xu, Y.; Wu, D. EEG-based driver drowsiness estimation using feature weighted episodic training. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2019**, *27*, 2263–2273. [CrossRef] [PubMed]
9. Tsuzuki, Y.; Mizusako, M.; Yasushi, M.; Hashimoto, H. Sleepiness detection system based on facial expressions. In Proceedings of the IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society, Lisbon, Portugal, 14–17 October 2019; Volume 1, pp. 6934–6939. [CrossRef]
10. Maior, C.B.S.; das Chagas Moura, M.J.; Santana, J.M.M.; Lins, I.D. Real-time classification for autonomous drowsiness detection using eye aspect ratio. *Expert Syst. Appl.* **2020**, *158*, 113505. [CrossRef]
11. Rundo, F. Deep Learning Systems for Advanced Driving Assistance. *arXiv* **2023**, arXiv:2304.06041.
12. Phan, A.C.; Nguyen, N.H.Q.; Trieu, T.N.; Phan, T.C. An efficient approach for detecting driver drowsiness based on deep learning. *Appl. Sci.* **2021**, *11*, 8441. [CrossRef]
13. Kumral, F.; Küçükmanisa, A.Y.H.A.N. Temporal Analysis Based Driver Drowsiness Detection System Using Deep Learning Approaches. *Sak. Univ. J. Sci.* **2022**, *26*, 710–719. [CrossRef]
14. Rundo, F.; Pino, C.; Castagnolo, G.; Spampinato, C. Advanced Intelligent deep learning-based system for Robust Driving Assistance. In Proceedings of the 2023 AEIT International Conference on Electrical and Electronic Technologies for Automotive (AEIT AUTOMOTIVE), Modena, Italy, 17–19 July 2023; pp. 1–5. [CrossRef]

15. Sinha, A.; Aneesh, R.P.; Gopal, S.K. Drowsiness detection system using deep learning. In Proceedings of the 2021 Seventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, 25–27 March 2021; pp. 1–6. [CrossRef]

16. Kumar, V.; Sharma, S.; Ranjeet. Driver drowsiness detection using modified deep learning architecture. In *Evolutionary Intelligence*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 1–10. [CrossRef]

17. Himaswi, N.; Sathvika, K.; Pathima, S. Deep Learning-Based Drowsiness Detection System Using IoT. In Proceedings of the 2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), Bengaluru, India, 27–28 January 2023; pp. 961–966. [CrossRef]

18. Patel, P.P.; Pavesha, C.L.; Sabat, S.S.; More, S.S. Deep Learning based Driver Drowsiness Detection. In Proceedings of the 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 9–11 May 2022; pp. 380–386. [CrossRef]

19. Jeong, M.; Ko, B.C. Driver's facial expression recognition in real-time for safe driving. *Sensors* **2018**, *12*, 4270. [CrossRef]

20. Zhao, G.; Huang, X.; Taini, M.; Li, S.Z.; PietikäInen, M. Facial expression recognition from near-infrared videos. *Image Vis. Comput.* **2011**, *29*, 607–619. [CrossRef]

21. Chen, J.; Dey, S.; Wang, L.; Bi, N.; Liu, P. Multi-modal fusion enhanced model for driver's facial expression recognition. In Proceedings of the 2021 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Shenzhen, China, 5–9 July 2021; pp. 1–4. [CrossRef]

22. Majeed, F.; Shafique, U.; Safran, M.; Alfarhood, S.; Ashraf, I. Detection of drowsiness among drivers using novel deep convolutional neural network model. *Sensors* **2023**, *23*, 8741. [CrossRef]

23. Kheradmandi, N.; Mehranfar, V. A critical review and comparative study on image segmentation-based techniques for pavement crack detection. *Constr. Build. Mater.* **2022**, *321*, 126162. [CrossRef]

24. Lin, S.L. Research on tire crack detection using image deep learning method. *Sci. Rep.* **2023**, *13*, 8027. [CrossRef] [PubMed]

25. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 6848–6856.

26. Sun, L.; Lin, Y.; Pan, N.; Fu, Q.; Chen, L.; Yang, J. Demand-Side Electricity Load Forecasting Based on Time-Series Decomposition Combined with Kernel Extreme Learning Machine Improved by Sparrow Algorithm. *Energies* **2023**, *16*, 7714. [CrossRef]

27. Dehghani, M.; Hubálovský, Š.; Trojovský, P. Northern goshawk optimization: A new swarm-based algorithm for solving optimization problems. *IEEE Access* **2021**, *9*, 162059–162080. [CrossRef]

28. Adnan, R.M.; Dai, H.L.; Mostafa, R.R.; Islam, A.R.M.T.; Kisi, O.; Elbeltagi, A.; Zounemat-Kermani, M. Application of novel binary optimized machine learning models for monthly streamflow prediction. *Appl. Water Sci.* **2023**, *13*, 110. [CrossRef]

29. Ding, S.; Xu, X.; Nie, R. Extreme learning machine and its applications. *Neural Comput. Appl.* **2014**, *25*, 549–556. [CrossRef]

30. Available online: https://www.kaggle.com/datasets/aamir2000/yawdd-driver (accessed on 16 October 2023).