

Ceng Giap Yo

IJTDI Establishing SSD-MobileNetV2 - Paper Rivised (1).docx

 Subclass 4
 18-25 Mei 2026
 Universitas Dian Nuswantoro

Document Details

Submission ID**trn:oid:::1:3575426299****Submission Date****May 21, 2026, 12:38 PM GMT+7****Download Date****May 21, 2026, 12:40 PM GMT+7****File Name****IJTDI_Establishing_SSD-MobileNetV2_-_Paper_Rivised_1_.docx****File Size****2.7 MB****13 Pages****6,995 Words****43,098 Characters**

15% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography

Match Groups

-  **82 Not Cited or Quoted 14%**
Matches with neither in-text citation nor quotation marks
-  **10 Missing Quotations 1%**
Matches that are still very similar to source material
-  **0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
-  **0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 10%  Internet sources
- 13%  Publications
- 3%  Submitted works (Student Papers)

Match Groups

- 82 Not Cited or Quoted 14%**
Matches with neither in-text citation nor quotation marks
- 10 Missing Quotations 1%**
Matches that are still very similar to source material
- 0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 10% Internet sources
- 13% Publications
- 3% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	library.acadlore.com	1%
2	Internet	www.mdpi.com	1%
3	Internet	thesai.org	<1%
4	Student papers	Westford School of Management	<1%
5	Internet	download.bibis.ir	<1%
6	Publication	Thangaprakash Sengodan, Sanjay Misra, M Murugappan. "Advances in Electrical ...	<1%
7	Publication	K. V. Sambasivarao, Anasuya Sesha Roopa Devi Bhima. "Artificial Intelligence, Co...	<1%
8	Publication	Hui Chen, Jianyi Yang. "Defect Classification Dataset and Algorithm for Magnetic ...	<1%
9	Publication	Marissa Utami, Ermatita Ermatita, Abdiansah Abdiansah. "Improving the transfer...	<1%
10	Publication	Sama Hussein Al-Gburi, Kanar Alaa Al-Sammak, Ion Marghescu, Claudia Cristina ...	<1%

11	Internet	www.acadlore.com	<1%
12	Internet	www.researchgate.net	<1%
13	Internet	e-journal.unair.ac.id	<1%
14	Internet	pmc.ncbi.nlm.nih.gov	<1%
15	Publication	V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challeng...	<1%
16	Student papers	Liverpool John Moores University	<1%
17	Publication	Xinqi Hao, Jianzhao Cao, Changtao Wang, Nan Chen. "MACS-Pose: Topological-Co...	<1%
18	Publication	Bodaghi, Morteza. "Adaptive Attention-Based LSTM Framework for Multimodal D...	<1%
19	Internet	repository.nusamandiri.ac.id	<1%
20	Internet	ebin.pub	<1%
21	Student papers	CSU, Fullerton	<1%
22	Student papers	Manchester Metropolitan University	<1%
23	Internet	theses.lib.polyu.edu.hk	<1%
24	Publication	Mohan Arava, Divya Meena Sundaram. "Integrating lightweight YOLOv5s and fac...	<1%

25	Publication	Shalli Rani, Ayush Dogra. "Emerging Trends in Smart Computing and Sustainable ...	<1%
26	Internet	arxiv.org	<1%
27	Internet	journal.50sea.com	<1%
28	Internet	rodin.uca.es	<1%
29	Publication	Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dharendra Kumar Shukla. "Intelli...	<1%
30	Publication	Govindarajan Subburaman, Mary Vennila Selvadurai. "Unified voting-based ense...	<1%
31	Internet	hcisj.com	<1%
32	Internet	ijeecs.iaescore.com	<1%
33	Internet	rctce2026.tce.edu	<1%
34	Internet	section.iaesonline.com	<1%
35	Internet	stax.strath.ac.uk	<1%
36	Internet	www.scik.org	<1%
37	Publication	"Next-Generation Networks and Deployable Artificial Intelligence", Springer Scie...	<1%
38	Publication	Boubiche Djallel Eddine, Sedat Akleylek. "Generative AI for Cybersecurity", CRC Pr...	<1%

39	Publication	Hasanah Putri, Rendy Munadi, Sofia Naning Hertiana, Alfin Hikmaturokhman. "To...	<1%
40	Publication	Samy Abd El-Nabi, Khalil F. Ramadan, El-Sayed M. El-Rabaie, Ahmed Emam, Walid ...	<1%
41	Publication	Shatakshi Saxena, Angel, Aaryaveer Gill, Shailendra Tiwari. "Chapter 27 A Real-Ti...	<1%
42	Publication	Tahesin Samira Delwar, Mangal Singh, Sayak Mukhopadhyay, Akshay Kumar et al...	<1%
43	Publication	Zidi Zhao, Hong Zhang, Liqiang Wang, Haijun Huang. "A Multi-model Edge Compu...	<1%
44	Publication	Zizheng Guo, Guanglei Wang, Ming Zhou, Guofa Li. "Monitoring and Detection of ...	<1%
45	Internet	bright-journal.org	<1%
46	Internet	dirros.openscience.si	<1%
47	Internet	discovery.ucl.ac.uk	<1%
48	Internet	ijai.iaescore.com	<1%
49	Internet	jurnal.ubhinus.ac.id	<1%
50	Internet	oaepublish.com	<1%
51	Internet	www.bright-journal.org	<1%
52	Internet	www.ijjsae.org	<1%

53	Publication	"Proceedings of International Conference on Advanced Computing and Systems",...	<1%
54	Publication	Jichi Chen, Shengkai Jin, Yuguo Cui, Chunfeng Wei, Kemal Polat, Fayadh Alenezi. "...	<1%
55	Publication	Oliveira, Daniel Sousa. "Combination of Physiological Signals and Image Processi...	<1%
56	Publication	Paolo Ferro, Harinadh Vemanaboina, Chander Prakash. "Computational Techniqu...	<1%
57	Publication	Salma Anber, Wafaa Alsaggaf, Wafaa Shalash. "A Hybrid Driver Fatigue and Distra...	<1%
58	Publication	Yo Ceng Giap, Muljono -, Affandy -, Ruri Suko Basuki, Harun Al Azies, R. Rizal Isna...	<1%
59	Publication	Pushpa Choudhary, Sambit Satpathy, Arvind Dagur, Dhirendra Kumar Shukla. "Re...	<1%



Establishing SSD-MobileNetV2 as a Robust Baseline for Driver Drowsiness Detection Toward IoT-Ready In-Driving Safety Systems

Yo Ceng Giap^{1,2}, Muljono Muljono^{1*}, Affandy¹, Ruri Suko Basuki¹, Harun Al Azies¹,
Deshinta Arrova Dewi³

¹ Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang, Indonesia

² Faculty of Science and Technology, Universitas Buddhi Dharma, Banten, Indonesia

³ Faculty of Data Science and Information Technology, INTI International University, Malaysia

*Correspondence: Muljono Muljono (muljono@dsn.dinus.ac.id)

Received: MM-DD-YYYY

Revised: MM-DD-YYYY

Accepted: MM-DD-YYYY

Citation: A. A. Almazroi, M. A. Alqarni, N. Aslam, and R. A. Shah, "Real-Time CNN-Based Driver Distraction & Drowsiness Detection System," *Intelligent Automation and Soft Computing*, vol. 37, no. 2, pp. 2153–2174, 2023, doi: 10.32604/iasc.2023.039732



© 2025 by the authors. Published by Acadlore Publishing Services Limited, Hong Kong. This article is available for free download and can be reused and cited, provided that the original published version is credited, under the CC BY 4.0 license.

Abstract: Driver drowsiness is one of the major reasons behind road accidents, emphasizing the need for accurate and efficient fatigue detection systems which can help monitor practical in-vehicle environments. While significant progress has been made in visual fatigue detection based on deep learning, many previous studies have been performed using a single dataset for training or controlled environments for testing. In this paper, we examine the reliability of lightweight driver-monitoring architectures for vision-based driver drowsiness detection based on three heterogeneous public datasets (YawDD, DDD, and NTHU-DDD) which cover different lighting conditions, facial characteristics, and head poses as encountered in driving scenarios. Among the considered architectures, SSD-MobileNetV2 was the most consistent, yielding an accuracy of 92%, precision of 93%, recall of 92%, and F1-score of 92% while also being computationally lighter than the other considered architectures. Reliability of the proposed architecture was statistically validated using the McNemar Test and 95% confidence intervals. Our results show that SSD-MobileNetV2 could be a promising baseline for future lightweight driver-monitoring systems for heterogeneous driving environments.

Keywords: Driver drowsiness detection; Edge-based inference; Embedded driver monitoring; IoT-ready safety systems; Lightweight deep learning; Intelligent transportation systems; In-vehicle monitoring; Vision-based driver monitoring; SSD-MobileNetV2.

1. Introduction

Transportation safety remains a major global concern due to the high number of traffic accidents and fatalities reported annually. As reported by the World Health Organization (WHO), the road traffic accidents are among the top causes of mortality in the age group from 5 to 29 years old. The main reason for road traffic accidents is driver fatigue. It was estimated that in 2023, about 1.19 million people worldwide died due to road traffic accidents [1]. Driver fatigue leads to decreased alertness and slowed reactions, which result in higher risk of accidents and poor performance especially at night and while driving for a long period of time [2], [3]. Although traditional methods of addressing driver fatigue including rest stops, safety campaigns [4], and technology-enabled learning systems were proposed, there is a great need for effective systems of fatigue detection [5].

Deep learning methods recently developed for the visual detection of driver drowsiness have been successful by analyzing eye movement, facial expressions, and head pose [6]. Yet, the performance of these models depends heavily on dataset quality, environment variation, and architectural complexity [6], [7], [8]. Some of the most common issues include changes in lighting conditions, facial features, camera view angles, head motion, and resolution, which may cause instability of models during driving operations. Thus, assessing the stability of such models under different driving conditions is an important requirement for ITS applications.

In this research work, three different but related publicly available datasets, namely, YawDD, DDD, and NTHU-DDD, are used. YawDD provides different head poses and eye-state information, whereas DDD gives fatigue-

related information [9]. In addition, NTHU-DDD dataset consists of different fatigue states and also covers night time driving conditions [10]. Thus, by combining these datasets, we can have a heterogenous evaluation environment.

For this investigation, three representative deep learning models are considered in order to establish a reliable benchmark model that can detect drowsiness in drivers in real time. The first one is YOLO11n, known for its real-time inference capabilities [6], the second one is Resnet50, famous for its excellent feature extraction; and the third one is SSD-MobileNetV2, known for its tradeoff between computational efficiency and detection accuracy [11].

However, even after considerable advancements in the use of deep learning techniques for detecting driver drowsiness, some issues related to robustness in various heterogeneous driving environments and implementation on constrained automotive platforms still remain important concerns for intelligent transportation systems. In terms of intelligent transportation systems, detection of driver drowsiness should not be seen only as a problem of computer vision but rather an integral part of intelligent transportation safety and accident prevention systems [12]. Lightweight and reliable designs are crucial for real-time implementation on embedded and edge computing automotive platforms.

The study stands out in that it critically analyzes SSD-MobileNetV2 as a possible baseline model for real-time drowsiness detection of drivers through an evaluation of three different public datasets (YawDD, DDD, and NTHU-DDD). Additionally, McNemar's Test and 95% Confidence Interval have been used to enhance the reliability analysis of the tested models. The results of this study will form a solid foundation for future research on IoT-enabled driver monitoring systems.

2. Related Work

2.1. Driver Drowsiness Detection in Intelligent Transportation System

The driver drowsiness detection is now part of the intelligent transportation safety system because of its significance in lowering risks of accidents and facilitating real time monitoring of drivers. The choice of appropriate datasets is crucial in deep learning approach to driver drowsiness detection. Several public datasets have been widely used for fatigue detection research, including YawDD [6], [13], [14], [15], NITYMED [15], [16], Drozy [17], DDD [7], NTHU-DDD [10], [18], [19]. Each dataset provides different visual characteristics, driver fatigue behaviors, and environmental conditions. However, these datasets also present limitations that may affect model performance in real-time deployment scenarios.

Researchers have evaluated various deep learning approaches using these datasets. Vijaypriya and Uma utilized the YawDD dataset with a Multi-Scale Convolutional Neural Network (MCNN) and achieved an accuracy of 98.38% [18]. Zhang et al. evaluated a multi-granularity deep framework using the NTHU-DDD dataset and reported an accuracy of 90.05% with approximately 37 FPS, demonstrating suitability for real-time fatigue detection [19]. Devi et al. utilized the DDD dataset and achieved approximately 99% accuracy using a COOT-optimized deep learning framework for real-time driver drowsiness detection [9].

2.2. Vision-Based Driver Monitoring

Vision-based driver monitoring systems commonly utilize facial cues such as eye closure, yawning behavior, facial expressions, and head movement to identify fatigue conditions. In intelligent transportation systems, vision-based driver monitoring plays an important role in improving road safety by enabling real-time observation of driver alertness, fatigue symptoms, and behavioral conditions during vehicle operation. Various vision-based monitoring approaches have been proposed to support real-time driver fatigue detection, including Convolutional Neural Networks (CNN) [20], [21], [22], Yolo [6], Resnet50 [23], [24], SSD [11], [25], Long Short-Term Memory (LSTM) [26], [27], [28] architectures.

In his research, Jahan developed a CNN-based deep learning model to detect drowsiness based on eye conditions. This model showed increased accuracy in detecting fatigue, reaching 97.53% [8]. These approaches demonstrate that visual monitoring methods can effectively capture fatigue-related facial features and driver behavior patterns in real-time driving environments.

Table 1. Comparative performance of vision-based driver-monitoring approaches for driver drowsiness detection

Model	Dataset	Dataset Size (images/videos/samples)	Optimization	Accuracy	Precision	Recall	F1 Score	Ref
COOT	DDD	-	√	99	98	-	98	[9]
Resnet	DDD	41.790 images	-	85.4	86.1	84.3	85.2	[29]
Multi-layer perceptron	NTHU	200 videos	-	81	-	-	-	[30]
KNN	NTHU	-	-	72.5	-	-	-	[31]

MCNN	NTHU	-	-	90.05	-	-	-	[19]
KNN	NTHU	-	-	89.86	96.35	89.18	92.63	[18]
Naïve Bayes	NTHU	-	-	87.92	95.82	86.89	91.14	[18]
MCNN + FSA	NTHU	-	√	98.26	99.45	98.10	98.77	[18]
SVM	YawDD	-	-	92.5	-	-	-	[31]
CNN	YawDD	-	-	93.83	-	-	-	[32]
SDM algorithm + CNN	YawDD	-	√	89.55	-	-	-	[33]
KNN	YawDD	-	-	87.69	77.27	86.82	81.77	[18]
Naïve Bayes	YawDD	-	-	83.42	70.41	82.52	75.99	[18]
MCNN + FSA	YawDD	-	√	98.38	97.06	97.84	97.45	[18]
MTCNN + DLIB	YawDD	100 videos	√	79	-	-	-	[34]
DLIB + LSTM	YawDD	100 videos	√	74	-	-	-	[34]
ConNN	YawDD+NTHU	96.817 images	√	98.81	-	-	-	[35]

Note: Dataset size values reported in previous studies vary depending on whether the datasets are represented as image frames, videos, or extracted samples, following the original experimental settings of each referenced study.

Based on the researchers' model tests, using the DDD, NTHU-DDD, and YawDD datasets, or a combination of two datasets, yielded accuracy values ranging from 72.5% to 99%. However, the models achieving accuracy above 90% have mostly undergone improvements, resulting in high accuracy. Only the CNN model trained on the YawDD dataset did not undergo model improvements, yet still achieved an accuracy of 93.83%. Overall, the results of this test indicate that models using combination techniques, such as MCNN + FSA and ConNN, tend to achieve better performance, particularly in terms of accuracy and the balance between precision and recall. Because this study evaluates lightweight architectures across three heterogeneous datasets without model-specific optimizations, the 92% accuracy represents a realistic, practically meaningful performance level for IoT-ready, real-time driver-monitoring systems. These comparative results indicate that lightweight driver-monitoring architectures can provide sufficiently stable performance for practical transportation safety applications, particularly under heterogeneous driving conditions.

2.3. Lightweight Deep Learning for Edge Deployment

Lightweight deep learning architectures are increasingly important for embedded automotive systems and edge-based driver monitoring platforms due to latency and computational constraints. When choosing a deep learning model to detect driver fatigue, it is necessary to balance inference speed, computational efficiency, and detection accuracy.

In real-time transportation monitoring systems, computationally efficient architecture such as YOLO11n are advantageous because they enable fast inference under limited computational resources. However, it does not perform as effectively in detecting small or overlapping objects, which can make it harder to identify subtle facial features that indicate fatigue [6]. Resnet50 achieves high accuracy and is effective across a wide range of data, although it requires greater computational resources [23], studies on image authenticity detection demonstrated that Resnet achieves consistent and stable performance across heterogeneous datasets. SSD architectures offer a balance between speed and accuracy; however, they are less effective at detecting small objects [11]. Therefore, the selection of the appropriate lightweight architecture must be tailored to the specific requirements of real-time driver monitoring and available hardware resources.

2.4. Cross-Dataset Generalization Challenges

One of the main challenges of deep learning driver drowsiness detection is the limited cross-dataset generalization. Previous work has often evaluated their models on a single dataset collected in a controlled environment, resulting in models that perform well on particular benchmarks but with less robustness to different visual environments. Different illumination, facial features, camera angles, head poses and fatigue expressions of different datasets usually lead to the performance degradation of the models when they are confronted with unseen driving conditions.

This limitation becomes more and more important for real-world in-vehicle monitoring systems where the visual conditions are dynamic and unpredictable. Hence, it is necessary to evaluate lightweight architectures on heterogeneous datasets, such as DDD, NTHU-DDD, and YawDD, to verify the model stability, robustness, and practicality for intelligent transportation safety systems.

2.5. Research Gap and Contribution

Even though many deep learning approaches have been suggested for detecting drowsy drivers, almost all of them were evaluated using only one data set in experimental settings. Therefore, there is a lack of research on how robust these lightweight models are to variations in real-world driving scenarios. Also, the importance of efficient

computation and reliable performance under real-world conditions is not emphasized enough in current literature.

From previous comparative evaluations and optimizations of lightweight machine learning and deep learning models, it has been shown that there is a need to balance between performance and computational efficiency in intelligent systems [36], [37]. Past comparative studies of computer vision applications have shown that lightweight and hybrid models can still perform strongly in terms of classification ability while lowering computation costs in embedded systems applications. There is however a lack of such studies that focus on heterogeneous driver monitoring and transportation safety edge applications.

Thus, this paper analyzes the effectiveness of three selected lightweight models namely YOLO11n, ResNet50, and SSD-MobileNetV2 under three heterogeneous publicly available datasets such as DDD, NTHU-DDD, and YawDD that cover a wide variety of illumination, facial features, and pose of the head. Moreover, a statistical evaluation method based on McNemar's Test and 95% confidence interval estimation is performed to improve the credibility of the analyzed models. The results are supposed to contribute significantly to the design of future driver monitoring systems.

3. Methodology

3.1. Research Framework

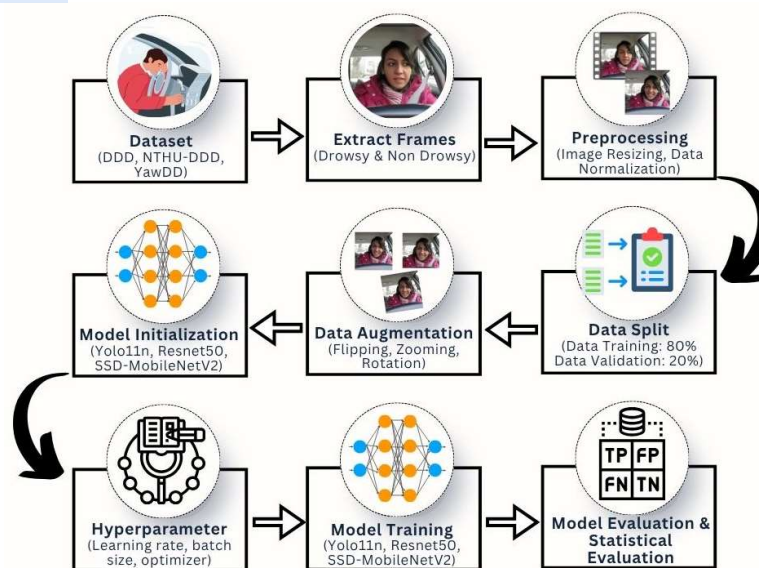


Figure 1. Methodological Pipeline

Figure 1 shows the entire framework that was employed in this study to detect driver drowsiness using lightweight deep learning frameworks. The process started with acquiring data from different public sources including DDD, NTHU-DDD, and YawDD, followed by frame extraction to get the images of drowsy and non-drowsy drivers. This was followed by pre-processing, which involved resizing and normalizing of images.

In order to ensure model robustness and avoid overfitting, augmentation methods such as flipping, zooming, and rotation were used. Three light models, namely YOLO11n, Resnet50, and SSD-MobileNetV2, were subsequently trained based on optimized values of hyperparameters. Lastly, model performance was evaluated via confusion matrix analysis for robustness and reliability testing under heterogeneous driving scenarios.

3.2. Dataset Description and Preparation

The initial step is to collect a dataset, which is specifically created for detecting the state of drowsiness in drivers. The dataset used in the current study includes a variety of images, which can vary in terms of various factors like driving condition, fatigue level, illumination, and facial pose. In this research, three heterogeneous datasets have been used, which have been collected from Kaggle, including Driver Drowsiness Dataset (DDD) [38], NTHU-DDD [39], and YawDD [40]. These datasets include images and videos of drowsy and non-drowsy drivers. This dataset will be split into two distinct subsets, including the training set and the validation set. The training set will be used to train deep learning algorithms, such as YOLO11n, Resnet50, and SSD-MobileNetV2.

The result of the integration of the three datasets led to an increase in the variability of the combined dataset. Thus, data augmentation was done to ensure that there is still good quality and variability of the data and avoid overfitting. Below is the number of datasets created through data augmentation for testing the three models:

Table 2. Distribution of image samples across drowsy and nondrowsy classes

Dataset	Drowsy	Nondrowsy	Total
DDD	83	83	166
NTHU-DDD	84	84	168
YAWDD	83	83	166
Total dataset			500

3.3. Lightweight Deep Learning Architectures

In this research work, three lightweight deep learning models such as YOLO11n, Resnet50, and SSD-MobileNetV2 are considered in order to develop a reliable baseline model for drowsiness detection among drivers in real-time scenarios by implementing intelligent transportation safety systems. These three lightweight models have been selected on account of their varying design aspects that are important for embedded vision applications for monitoring of drivers. All three models are pre-trained with the help of initial weights before training.

The selection of YOLO11n is based on its light-weighted architecture and real-time object detection feature, making it feasible to deploy in an embedded monitoring system where computation power is limited. With its ability to do fast inference, it can be deployed in a transportation safety system that requires instant driver state detection. But previous research revealed that lightweight YOLO-based architecture faces difficulties in detecting facial fatigue due to visual challenges [6].

The Resnet50 network was used as a representative deep residual learning architecture, which can extract high-level visual features and maintain stable performance on different datasets [23]. The residual learning algorithm provides deep feature representation, which is advantageous for identifying fatigue characteristics of faces. However, the architecture consumes more computational resources than lightweight embedded systems.

SSD-MobileNetV2 integrates the Single Shot Detector (SSD) structure and the efficient MobileNetV2 model, which is the combination of computational efficiency and high detection accuracy [11]. This architecture is suitable for real-time deployment scenarios, and at the same time, it ensures the stable detection performance despite the different conditions of driving. Previous research indicated that SSD-based lightweight structures work well in real-time driver monitoring systems [25].

The comparison of these architectures enables evaluation of the trade-off between accuracy, computational efficiency, and deployment feasibility for intelligent transportation safety systems and embedded driver-monitoring environments.

3.4. Model Training Configuration

The tested architectures, which include YOLO11n, Resnet50, and SSD-MobileNetV2, were initialized and trained under similar experimental conditions to ensure a fair comparison between them. The experiments were performed using Google Colab with NVIDIA Tesla T4 GPU support.

Prior to training, all image samples were resized to 640×640 pixels and normalized for consistent inputs to the models. The dataset from DDD, NTHU-DDD, and YawDD was split into train and validation sets in an 80:20 proportion. In order to make the models more robust to noise and minimize overfitting, data augmentation including flips, zooms, and rotations was used while training the models.

Training parameters of all models were set up using the Adam optimizer with learning rate 0.01, batch size 16, and 50 epochs. For consistency, same training parameters were used in YOLO11n, Resnet50, and SSD-MobileNetV2 so that performance difference between them could be due to architectural differences only. These parameters were chosen considering practicality as well as efficiency in terms of computational speed and reliability.

3.5. Performance Evaluation

Following the training phase, the model's performance was analyzed based on confusion matrix criteria such as accuracy, precision, recall, and F1 score. These measures were adopted to analyze the performance of the evaluated architectures in discriminating between drowsy and non-drowsy drivers under various driving conditions. Recall and false negatives were particularly considered, because a failed identification of drowsiness could directly impact transportation safety systems.

Besides the evaluation measures, statistical validation was performed using the McNemar Test and 95% Confidence Interval (CI). The McNemar Test was used to determine the statistical significance of differences in the performance of the examined systems, whereas the 95% confidence interval was utilized to measure the reliability and consistency of the achieved results.

4. Results

In this section, we will explore in depth the results of tests performed to evaluate the performance of models for detecting drowsiness. In order to establish a reliable baseline, three representative models were tested: YOLO11n, Resnet50, and SSD-MobileNetV2. The results obtained from each of these models give us valuable information

on how well they perform.

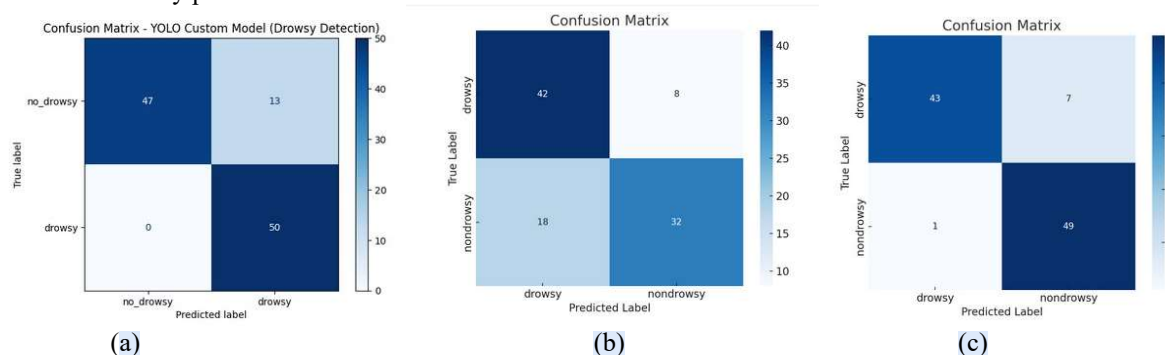


Figure 2. Confusion matrix (a) YOLO11n, (b) Resnet50, and (c) SSD-MobileNetV2

As shown in the confusion matrix in Figure 2, the YOLO11n algorithm was able to detect all drowsiness cases correctly (50/50), giving a perfect recall rate for the drowsy class. This indicates that the YOLO11n model is less prone to missing drowsiness, an aspect that is very important especially in critical areas where safety is paramount. However, 13 samples of non-drowsy were falsely categorized as drowsy, indicating over-detection.

The performance of the Resnet50 model was found to be moderate when differentiating between the two states of drivers. Nevertheless, this model is prone to making false positive and negative predictions, which implies poor stability in classifying drivers under heterogeneous driving conditions. The training and validation losses illustrate that even though the former gradually reduces throughout the training process, the latter remains relatively unstable.

On the other hand, SSD-MobileNetV2 showed better generalization ability compared to Resnet50. This is because SSD-MobileNetV2 made predictions with higher accuracy on both the drowsy and non-drowsy classes while committing fewer classification errors. Furthermore, the training and validation loss kept reducing throughout the training period, suggesting better generalization ability. It can be observed that SSD-MobileNetV2 performs better under heterogeneous visual scenarios provided by DDD, NTHU-DDD, and YawDD datasets.

From the results, it is clear that SSD-MobileNetV2 is the best in terms of generalization and stability whereas YOLO11n performs well in terms of precision. Despite its strengths, Resnet50 suffers from overfitting. Both architectures demonstrate practical applicability in terms of drowsiness detection but SSD-MobileNetV2 provides a better balance between stability and generalization.

Table 3. Train and Validation Loss

Epoch	YOLO11n		Resnet50		SSD-MobileNetV2	
	Train Loss	Val Loss	Train Loss	Val Loss	Train Loss	Val Loss
1	1.6731	1.6633	0.9616	0.8708	1.5757	0.8071
10	0.6028	1.0908	0.6894	0.6907	0.3438	0.2521
20	0.4335	0.7331	0.6806	0.6636	0.1780	0.1344
30	0.3190	0.4516	0.6212	0.6324	0.1364	0.0861
40	0.2898	0.3930	0.6121	0.6009	0.0992	0.0585
50	0.2004	0.7858	0.5823	0.5711	0.0822	0.0438

Table 3 shows how the three architectures performed in the same experiments by comparing their training loss and validation loss values at different epochs (1, 10, 20, 30, 40, and 50). As the number of epochs increases, the training loss decreases for all three models. This shows that all three have learned the training data well. But each model is better or worse at generalizing to validation data.

The YOLO11n model's training loss steadily dropped from 1.6731 at epoch 1 to 0.2004 at epoch 50. But its validation loss did not decrease as much. The validation loss decreased from 1.6633 (epoch 1) to 0.3930 (epoch 40), then increased to 0.7858 (epoch 50). This indicates that the model is overfitting: it adapts too closely to the training data and performs worse on the validation data.

Meanwhile, Resnet50 showed a steady downward trend. Train loss decreased from 0.9616 (epoch 1) to 0.5823 (epoch 50), and validation loss also reduced from 0.8708 to 0.5711 at the same epoch. This indicates that the training process is proceeding smoothly and the model has sufficient generalisation capabilities, although the model still exhibits limited generalization capability compared to SSD-MobileNetV2.

On the other hand, the SSD-MobileNetV2 model shows the best performance. Train loss decreases sharply from 1.5757 at epoch 1 to just 0.0822 at epoch 50. The validation loss also followed a positive trend, dropping dramatically from 0.8071 to 0.0438. This shows that SSD-MobileNetV2 not only learns well from the training data but also generalises well to the validation data, without any signs of overfitting. Based on these results, SSD-MobileNetV2 demonstrated the most balanced and stable overall performance in this training scenario.

All reported evaluation metrics were calculated based on binary classification performance across the drowsy

and non-drowsy classes.

Table 4. Results of testing three models

Model	Accuracy	Precision	Recall	F1 Score
YOLO11n	88.18	90.62	88.18	88.14
Resnet50	74	75	74	74
SSD-MobileNetV2	92	93	92	92

The test results for the three models in Table 4 show significant differences across several important metrics, including accuracy, precision, recall, F1 score, and processing time. YOLO11n demonstrated strong precision performance, with a precision of 90.62%. The high precision value indicates that YOLO11n produces relatively few false-positive predictions. But YOLO11n achieved only 88.18% accuracy, lower than SSD-MobileNetV2. The recall rate for YOLO11n was 88.18%, indicating that the model can detect existing objects, but it is still slightly behind SSD-MobileNetV2. The YOLO11n F1 score of 88.14% indicates a good balance between precision and recall. YOLO11n demonstrated relatively stable detection capability under heterogeneous driving conditions, although its overall performance remained below SSD-MobileNetV2. However, its overall stability and processing efficiency remained lower than SSD-MobileNetV2 for practical embedded deployment scenarios.

Resnet50 is more unstable in comparison to others regarding different performance measures, meaning that it is probably not the most appropriate architecture for real-time applications. Resnet50 can only give 74% of correct predictions, meaning that this network has a very high percentage of errors. Also, in terms of precision, recall, and F1 score, this network performs at 74%, implying that this neural network performs worse in object detection compared to other networks. While having relatively reasonable processing time, Resnet50 cannot be used in real-time applications because of poor accuracy and stability.

SSD-MobileNetV2, on the other hand, was the best-performing network across almost all measures. SSD-MobileNetV2 had the highest accuracy of 92%. The SSD-MobileNetV2 model also performed extremely well when it comes to precision and recall, achieving values of 93% and 92% respectively. These numbers demonstrate that the SSD-MobileNetV2 is capable of object detection and classification with high accuracy.

Figure 3 shows that SSD-MobileNetV2 achieved the shortest overall processing duration among the evaluated architectures, requiring approximately 0.0058 hours (about 21 seconds), compared to YOLO11n (0.068 hours) and Resnet50 (0.0233 hours). The processing time is measured based on the same experimental setup using Google Colab with NVIDIA Tesla T4 GPU support. The numbers presented above show the total processing time during the experimental evaluation period and not the inference latency time for individual images. The results show that SSD-MobileNetV2 has less computational overhead and higher computational efficiency compared to the other models tested. The proposed architecture demonstrates potential for future embedded transportation safety applications.

Given the same experimental environment, the reported values should be interpreted as relative indicators of the computational workload rather than real-time inference latency measurements. The actual deployment performance in real transportation systems may be dependent on embedded hardware specifications, camera resolution and streaming conditions. Future studies should evaluate the inference latency at frame level, FPS performance and energy efficiency on embedded automotive platforms such as NVIDIA Jetson or edge-based IoT devices.

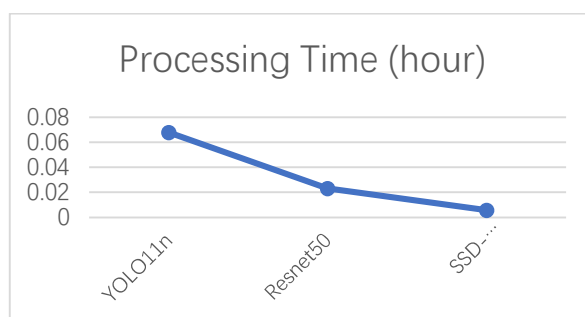


Figure 3. Total experimental processing duration under identical GPU-based evaluation settings

In order to further prove the statistical significance of the differences in performances between YOLO11n, Resnet50, and SSD-MobileNetV2, two statistical tests were performed; these include the McNemar significance test and the calculation of 95% confidence intervals of accuracy. These tests evaluate not only the numerical superiority of each model but their statistical significance as well.

The McNemar test is a non-parametric test commonly used to evaluate whether two classifiers differ significantly in their error distributions when applied to the same dataset [41]. Several deep learning evaluation

studies recommend it because it directly compares disagreement cases rather than aggregate accuracy.

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c} \quad (1)$$

Where:

b = cases misclassified by model A but correctly classified by model B

c = cases misclassified by model B but correctly classified by model A

H_0 = There is no significant difference between the two models

Significance threshold: $p < 0.05$

Table 5. McNemar Test

Model	χ^2	p-value	Significant?
SSD vs Resnet50	4.00	0.045	Yes
SSD vs YOLO11n	1.25	0.26	No
YOLO11n vs Resnet50	0.516	0.47	No

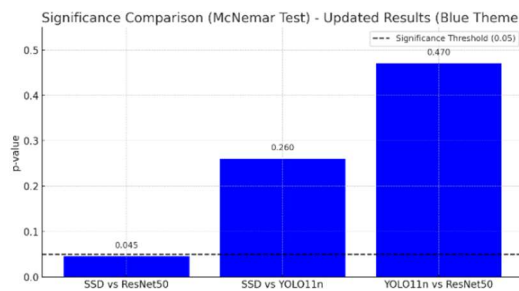


Figure 4. Significance Comparison

To quantify the statistical certainty of accuracy values, 95% confidence intervals (CI) were computed using the binomial proportion formula [42].

$$CI = p \pm 1.96 \sqrt{\frac{p(1-p)}{n}} \quad (2)$$

Where:

p = accuracy

n = total test samples

1.96 = z-score for 95% confidence

Table 6. Confidence Interval (CI 95%)

Model	Accuracy	CI 95%
SSD-MobileNetV2	92%	86.7% - 97.3%
YOLO11n	88.18%	82.1% - 94.2%
Resnet50	74%	65.8% - 82.2%

The statistical analysis reveals that SSD-MobileNetV2 is significantly better than Resnet50 ($p = 0.045$), whereas the differences in the accuracy rates of SSD-MobileNetV2 compared to YOLO11n ($p = 0.26$) and YOLO11n versus Resnet50 ($p = 0.47$) are not statistically significant. Moreover, based on the results of 95% confidence interval analysis, it can be stated that SSD-MobileNetV2 demonstrates the narrowest interval (86.7% - 97.3%), which is evidence of higher stability and reliability of the algorithm.

Generally, SSD-MobileNetV2 emerges as the best-performing among all three in terms of overall performance metrics including accuracy, precision, recall, F1 Score and time. Performance of SSD-MobileNetV2 was equal to that of various previously reported light-weight driver monitoring systems [32]. On the other hand, YOLO11n shows strengths in specific situations where precision is of utmost importance, but it is less stable in comparison with SSD-MobileNetV2 as a baseline. In case when efficiency and high performance are the primary goals, then SSD-MobileNetV2 is the model of choice. As mentioned above, application of hybrid methods like COOT, MCNN + FSA, and ConNN proved to be more efficient in obtaining higher accuracy (above 90%) and achieving the required balance between precision and recall [9], [18], [35]. Thus, it could be stated that more advanced combination techniques result in greater gains in comparison with traditional models like YOLO11n, Resnet50, and SSD-MobileNetV2.

The superior performance of SSD-MobileNetV2 in the present work can be attributed to its lightweight structure, which allows for efficient feature extraction along with a low computational complexity and stable detection capacity in different driving environments. The architecture showed higher consistency in performance on the integrated DDD, NTHU-DDD, and YawDD datasets, thereby suggesting that it had superior generalization capabilities under different conditions of illumination, facial appearance, and head orientations often witnessed in practical driving situations. Such an architectural advantage makes SSD-MobileNetV2 a suitable baseline framework for embedded driver monitoring systems.

5. Discussion

The results of this study highlight the importance of computationally efficient architecture for driver monitoring systems in achieving stable performance under heterogeneous driving conditions. This aspect becomes increasingly important as modern vehicles are progressively integrated with intelligent transportation safety systems utilizing IoT and edge-computing technologies. Among the evaluated architectures, SSD-MobileNetV2 demonstrated the most stable generalization performance across the YawDD, DDD, and NTHU-DDD datasets under varying lighting conditions, facial appearances, and head poses.

These findings are consistent with previous studies showing that lightweight deep learning architectures can maintain stable classification performance while reducing computational overhead [37]. Previous comparative studies in computer vision have also demonstrated that lightweight and hybrid CNN-based architectures are capable of achieving scalable performance suitable for practical deployment environments [36]. In intelligent transportation systems, balancing computational efficiency and detection reliability is particularly important because embedded driver monitoring systems often operate under constrained hardware resources and dynamically changing visual conditions. Therefore, the lightweight MobileNetV2 architecture combined with a single-shot detection framework provides a promising baseline for future studies involving sequence modeling, sensor fusion, and edge-based transportation safety applications.



Figure 5. Proposed practical deployment workflow of the SSD-MobileNetV2-based driver monitoring system for intelligent transportation safety applications.

Figure 5 illustrates the proposed practical deployment workflow of the SSD-MobileNetV2-based driver monitoring system. An in-vehicle camera continuously captures the driver's facial condition and transmits the visual data to an embedded AI processing unit for real-time analysis. The SSD-MobileNetV2 architecture analyzes fatigue-related features such as eye behavior, yawning activity, facial expressions, and head posture to estimate the driver's alertness state. When drowsiness symptoms are detected, the system activates warning mechanisms including visual alerts and audible alarms to improve driver awareness and reduce accident risk. Due to its lightweight computational design, the proposed architecture is suitable for deployment on embedded and edge-based transportation platforms with limited hardware resources. In fleet safety environments, the system may also support centralized driver monitoring and transportation safety management through IoT-enabled monitoring infrastructure.

6. Conclusions

This study evaluated the reliability of lightweight driver monitoring architectures for vision-based driver drowsiness detection, with three different heterogeneous public datasets used to reflect various driving conditions. From the experimental results, it can be seen that SSD-MobileNetV2 performed the best with respect to accuracy, precision, recall, and F1 score with less computational cost compared with other architectures. The statistical analysis of the experiment with the help of McNemar's test and 95% confidence intervals also shows that the

model is reliable and stable. It appears that SSD-MobileNetV2 can be considered a good lightweight baseline architecture for driver monitoring system in intelligent transportation safety environments.

Nevertheless, despite these promising results, there are certain limitations to the study such as use of open-source datasets, lack of temporal modeling in frame-level analysis, and evaluation on only three representative lightweight networks. However, the study can serve as a significant stepping stone for future studies related to intelligent transportation safety. In addition, although three heterogeneous public datasets were utilized, the overall experimental scale is still relatively limited when considering real-world transportation environments. In real-world deployment scenarios, model generalization can be affected by changes in traffic conditions, driver behavior, camera placement, weather conditions, and uncontrolled illumination. Therefore, additional validation in real-world driving scenarios and more complex transportation conditions is needed to evaluate the robustness and scalability of the proposed architecture.

7. Future Work

The future research directions would involve enhancing the SSD-MobileNetV2 architecture to enable practical implementation within embedded and IoT-enabled transportation platforms. Future research may involve comparisons with more advanced lightweight networks like MobileNetV3, EfficientNet, TinyML, temporal deep learning models, among others, to assess their computational efficiency and resilience when facing varying driving conditions. Furthermore, hybrid methods that utilize physiological parameters, vehicle telemetry, and/or driver behavior may be considered to enhance the accuracy of the algorithm in challenging driving situations. Real-world driving settings and automotive platforms would be essential in validating deployment feasibility and practicality of intelligent transportation systems, including frame-level inference latency, frames per second (FPS) measurements, memory consumption, and energy efficiency analysis on embedded and edge automotive platforms to better assess the performance of real-time implementations.

Author Contributions

The contributions of all the authors were equally significant towards this work. Yo Ceng Giap contributed by developing the idea for research, designing the methodology, improving the model and preparing the first draft of the manuscript. Muljono coordinated the research, provided theoretical support and made substantial revisions on the manuscript. Affandy was responsible for data collection, data pre-processing, experimental analysis, and evaluation of model performance. Ruri Suko Basuki assisted in literature survey, preparation of related work, and preparation of figures and charts. Harun Al Azies performed statistical validation through CI95 and Two-Proportion Z-Test analysis and interpretation of results. Deshinta Arrova Dewi contributed towards software implementation, embedded system optimization, and manuscript editing. This manuscript has been read and approved by all the authors.

Data Availability

The datasets used in this study are publicly available through Kaggle and include the Driver Drowsiness Dataset (DDD), NTHU-DDD, and YawDD datasets cited in references [38], [39], [40]. All datasets were accessed and used solely for academic research purposes in accordance with their respective public usage policies. Processed data and experimental configurations used in this study are available from the corresponding author upon reasonable request.

Acknowledgements

The Ministry of Education, Culture, Research, and Technology (Kemdiktisaintek) expresses its most profound appreciation for this study. The Ministry financed the research that led to this article as part of the BIMA Fundamental Research 2026 initiative, with grant contract number 133/C3/DT.05.00/ PL-MULTITAHUN LANJUTAN/2026. The authors also acknowledge the ongoing support from Universitas Dian Nuswantoro in completing this work.

Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] World Health Organization, *Global status report on road safety 2023*. Switzerland, 2023.
- [2] E. H. Rad, M. Hosseinnia, N. Mousavi, A. Shekari, L. Kouchakinejad-Eramsadati, and N. Khodadadi-Hassankiadeh, "Fatigue in taxi drivers and its relationship with traffic accident history and experiences: a cross-sectional study in the north of Iran," *BMC Public Health*, vol. 24, no. 1, Dec. 2024, doi: 10.1186/s12889-024-18044-5.

- [3] D. A. L. S. F. J. S. T. K. C. J. L. C. Ann Williamson, "The link between fatigue and safety," *Accid. Anal. Prev.*, vol. 43, no. 2, pp. 498–515, Mar. 2011.
- [4] D. F. D. Julian Lim, "A Meta-Analysis of the Impact of Short-Term Sleep Deprivation on Cognitive Variables," *Psychol. Bull.*, 2010, doi: 10.1037/a0018883.supp.
- [5] L. Chen and W. Zheng, "Research on Railway Dispatcher Fatigue Detection Method Based on Deep Learning with Multi-Feature Fusion," *Electronics (Switzerland)*, vol. 12, no. 10, 2023, doi: 10.3390/electronics12102303.
- [6] S. Liu, Y. Wang, Q. Yu, J. Zhan, H. Liu, and J. Liu, "A Driver Fatigue Detection Algorithm Based on Dynamic Tracking of Small Facial Targets Using YOLOv7," *IEICE Trans. Inf. Syst.*, vol. E106.D, no. 11, pp. 1881–1890, 2023, doi: 10.1587/transinf.2023EDP7093.
- [7] A. Sohail, A. A. Shah, S. Ilyas, and N. Alshammry, "A CNN-based Deep Learning Framework for Driver's Drowsiness Detection," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 3, pp. 169–178, 2024, doi: 10.14569/IJACSA.2024.0150317.
- [8] I. Jahan *et al.*, "4D: A Real-Time Driver Drowsiness Detector Using Deep Learning," *Electronics (Switzerland)*, vol. 12, no. 1, 2023, doi: 10.3390/electronics12010235.
- [9] G. R. Devi *et al.*, "COOT-Optimized Real-Time Drowsiness Detection using GRU and Enhanced Deep Belief Networks for Advanced Driver Safety," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 4, pp. 804–814, 2024, doi: 10.14569/IJACSA.2024.0150483.
- [10] S. Anber, W. Alsaggaf, and W. Shalash, "A Hybrid Driver Fatigue and Distraction Detection Model Using AlexNet Based on Facial Features," *Electronics (Switzerland)*, vol. 11, no. 2, 2022, doi: 10.3390/electronics11020285.
- [11] A. A. Almazroi, M. A. Alqarni, N. Aslam, and R. A. Shah, "Real-Time CNN-Based Driver Distraction & Drowsiness Detection System," *Intelligent Automation and Soft Computing*, vol. 37, no. 2, pp. 2153–2174, 2023, doi: 10.32604/iasc.2023.039732.
- [12] A. A. Sheikh and I. Z. Khan, "Enhancing Road Safety: Real-Time Detection of Driver Distraction through Convolutional Neural Networks," May 2024, [Online]. Available: <http://arxiv.org/abs/2405.17788>
- [13] H. Zheng, Y. Wang, and X. Liu, "Adaptive Driver Face Feature Fatigue Detection Algorithm Research," *Applied Sciences (Switzerland)*, vol. 13, no. 8, 2023, doi: 10.3390/app13085074.
- [14] F. Majeed, U. Shafique, M. Safran, S. Alfarhood, and I. Ashraf, "Detection of Drowsiness among Drivers Using Novel Deep Convolutional Neural Network Model," *Sensors (Basel)*, vol. 23, no. 21, 2023, doi: 10.3390/s23218741.
- [15] R. Florez, F. Palomino-Quispe, R. J. Coaquira-Castillo, J. C. Herrera-Levano, T. Paixão, and A. B. Alvarez, "A CNN-Based Approach for Driver Drowsiness Detection by Real-Time Eye State Identification," *Applied Sciences (Switzerland)*, vol. 13, no. 13, 2023, doi: 10.3390/app13137849.
- [16] P. Christakos, N. Petrellis, P. Mousoulitis, G. Keramidas, C. P. Antonopoulos, and N. Voros, "A High Performance and Robust FPGA Implementation of a Driver State Monitoring Application," *Sensors*, vol. 23, no. 14, 2023, doi: 10.3390/s23146344.
- [17] A. Sedik, M. Marey, and H. Mostafa, "An Adaptive Fatigue Detection System Based on 3D CNNs and Ensemble Models," *Symmetry (Basel)*, vol. 15, no. 6, 2023, doi: 10.3390/sym15061274.
- [18] V. Vijaypriya and M. Uma, "Facial Feature-Based Drowsiness Detection with Multi-Scale Convolutional Neural Network," *IEEE Access*, vol. 11, pp. 63417–63429, 2023, doi: 10.1109/ACCESS.2023.3288008.
- [19] H. Zhang, T. Liu, J. Lyu, D. Chen, and Z. Yuan, "Integrate memory mechanism in multi-granularity deep framework for driver drowsiness detection," *Intelligence and Robotics*, vol. 3, no. 4, pp. 614–631, 2023, doi: 10.20517/ir.2023.34.
- [20] N. N. Alajlan and D. M. Ibrahim, "DDD TinyML: A TinyML-Based Driver Drowsiness Detection Model

- Using Deep Learning,” *Sensors*, vol. 23, no. 12, 2023, doi: 10.3390/s23125696.
- [21] M. Peivandi, S. Z. Ardabili, S. Sheykhivand, and S. Danishvar, “Deep Learning for Detecting Multi-Level Driver Fatigue Using Physiological Signals: A Comprehensive Approach,” *Sensors*, vol. 23, no. 19, 2023, doi: 10.3390/s23198171.
- [22] W. Sun, Y. Wang, B. Hu, and Q. Wang, “Exploration of Eye Fatigue Detection Features and Algorithm Based on Eye-Tracking Signal †,” *Electronics (Switzerland)*, vol. 13, no. 10, 2024, doi: 10.3390/electronics13101798.
- [23] B.-E. B. Lim, K.-W. Ng, and S.-L. Ng, “Drowsiness Detection System through Eye and Mouth Analysis,” *International Journal on Informatics Visualization*, vol. 7, no. 4, pp. 2489–2497, 2023, doi: 10.30630/joiv.7.4.2288.
- [24] S. Das, S. Pratihari, B. Pradhan, R. H. Jhaveri, and F. Benedetto, “IoT-Assisted Automatic Driver Drowsiness Detection through Facial Movement Analysis Using Deep Learning and a U-Net-Based Architecture,” *Information (Switzerland)*, vol. 15, no. 1, 2024, doi: 10.3390/info15010030.
- [25] W. Gao, L. Li, and T. Yang, “Driving State Discrimination Algorithm Based on Lightweight Network and Contrast Learning,” *Intern. J. Pattern Recognit. Artif. Intell.*, vol. 36, no. 11, 2022, doi: 10.1142/S0218001422520255.
- [26] Z. Huang, W. Tang, Q. Tian, T. Huang, and J. Li, “Air Traffic Controller Fatigue Detection Based on Facial and Vocal Features Using Long Short-Term Memory,” *IEEE Access*, vol. 12, pp. 56663–56682, 2024, doi: 10.1109/ACCESS.2024.3391427.
- [27] S. A. Alameen and A. M. Alhothali, “A Lightweight Driver Drowsiness Detection System Using 3DCNN With LSTM,” *Computer Systems Science and Engineering*, vol. 44, no. 1, pp. 895–912, 2022, doi: 10.32604/csse.2023.024643.
- [28] B. Akrouf and S. Fakhfakh, “How to Prevent Drivers before Their Sleepiness Using Deep Learning-Based Approach,” *Electronics (Switzerland)*, vol. 12, no. 4, 2023, doi: 10.3390/electronics12040965.
- [29] Z. Huang, “Integrating Attention Mechanisms and ResNet-50 For Enhanced Driver Sleepiness Detection,” *Informatica (Slovenia)*, vol. 49, no. 15, pp. 135–144, 2025, doi: 10.31449/inf.v49i15.7977.
- [30] R. Jabbar, K. Al-Khalifa, M. Kharbeche, W. Alhajyaseen, M. Jafari, and S. Jiang, “Real-time Driver Drowsiness Detection for Android Application Using Deep Neural Networks Techniques,” in *Procedia Computer Science*, Elsevier B.V., 2018, pp. 400–407. doi: 10.1016/j.procs.2018.04.060.
- [31] V. Uma Maheswari, R. Aluvalu, M. V. V. Prasad Kantipudi, K. K. Chennam, K. Kotecha, and J. R. Saini, “Driver Drowsiness Prediction Based on Multiple Aspects Using Image Processing Techniques,” *IEEE Access*, vol. 10, pp. 54980–54990, 2022, doi: 10.1109/ACCESS.2022.3176451.
- [32] H. He *et al.*, “A Real-time Driver Fatigue Detection Method Based on Two-Stage Convolutional Neural Network,” in *IFAC-PapersOnLine*, Elsevier B.V., 2020, pp. 15374–15379. doi: 10.1016/j.ifacol.2020.12.2357.
- [33] X. Li, J. Xia, L. Cao, G. Zhang, and X. Feng, “Driver fatigue detection based on convolutional neural network and face alignment for edge computing device,” *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 235, no. 10–11, pp. 2699–2711, Sep. 2021, doi: 10.1177/0954407021999485.
- [34] L. Chen, G. Xin, Y. Liu, and J. Huang, “Driver Fatigue Detection Based on Facial Key Points and LSTM,” *Security and Communication Networks*, vol. 2021, pp. 1–9, Jun. 2021, doi: 10.1155/2021/5383573.
- [35] B. K. Savaş and Y. Becerikli, “Real Time Driver Fatigue Detection System Based on Multi-Task ConNN,” *IEEE Access*, vol. 8, pp. 12491–12498, 2020, doi: 10.1109/ACCESS.2020.2963960.
- [36] S. Makwana, D. Vaghela, C. K. Chan, and N. Naik, “Novel hybrid approach of random forest and stacking ensemble to improve fresh water yield prediction in mobile wick solar still,” *Desalination Water Treat.*, vol.

- 324, Oct. 2025, doi: 10.1016/j.dwt.2025.101490.
- [37] L. D, P. C, M. Batumalay, and Manoj R Karthick, “Comparative Study of CNN-Based Architectures for Early Brain Tumor Diagnosis,” *Journal of Applied Data Sciences*, vol. 7, no. 1, pp. 529–540, Jan. 2026, doi: 10.47738/jads.v7i1.920.
- [38] Ismail Nasri, “Driver Drowsiness Dataset (DDD),” <https://www.kaggle.com/datasets/ismailnasri20/driver-drowsiness-dataset-ddd>.
- [39] Ikhlas El Hamly, “NTHU-DDD,” <https://www.kaggle.com/datasets/ikhlaselhamly/nthu-ddd>.
- [40] Enider, “Yawdd Dataset,” <https://www.kaggle.com/datasets/enider/yawdd-dataset>.
- [41] T. G. Dietterich, “Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms,” *Neural Comput.*, vol. 10, no. 7, pp. 1895–1923, Oct. 1998, doi: 10.1162/089976698300017197.
- [42] R. G. Newcombe, “Two-sided confidence intervals for the single proportion: comparison of seven methods,” *Stat. Med.*, vol. 17, no. 8, pp. 857–872, Apr. 1998, doi: 10.1002/(SICI)1097-0258(19980430)17:8<857::AID-SIM777>3.0.CO;2-E.