Drowsy Driver Detection System Using Deep Learning

Amal Al-Shahrani, Amjad Saeed Al-Ghamdi, Areej Saud Al-Gurashi, Raghad Saeed Al-Zahrani, Seham Awadh Al-Malki, and Nuha Idris Imam

Umm Al-Qura University, College Computing, Computer Science and Artificial Intelligence Department, KSA

Abstract
The research paper provides an overview of various approaches and methodologies for detecting driver drowsiness using computer vision and machine learning techniques. The primary focus is on detecting drowsiness indicators, such as eye closure and yawning, in challenging low-light conditions. Several studies have been conducted, employing different algorithms and models, to achieve accurate drowsiness detection. In this context, our contribution to the field involves incorporating YOLOv8, YOLOv5, and VGG16 as integral components of the methodology. By leveraging these advanced technologies, we aim to enhance the accuracy and effectiveness of driver drowsiness detection systems. This, in turn, has the potential to improve road safety and prevent accidents caused by driver fatigue.

Keywords:
Computer vision, CNN, YOLO, VGG

1. Introduction

Driver drowsiness is a significant factor contributing to road accidents, posing serious risks to both individual safety and public health. According to various traffic safety reports, a substantial proportion of vehicular accidents are attributed to driver fatigue, which impairs reaction times, decision-making abilities, and overall awareness. As a result, the development of reliable and effective driver drowsiness detection systems has become an imperative goal in the field of automotive safety [1] [2].

However, the effectiveness of these computer vision systems is often challenged by real-world conditions; addressing these challenges requires robust algorithms and models capable of maintaining high performance [3]. This research paper provides a comprehensive overview of the methodologies employed for driver drowsiness detection. Our primary contribution lies in the comparison between two advanced technologies: YOLOv8 and YOLOv5 (You Only Look Once, version 8), and VGG16 (Visual Geometry Group 16-layer network).

Our proposed system aims to enhance the detection accuracy of drowsiness indicators such as eye closure and yawning, even in suboptimal lighting scenarios. These models are anticipated to significantly improve the reliability of drowsiness detection systems, thereby contributing to road safety and the prevention of accidents caused by driver fatigue.

In the following sections, we will review related work in the field, detail our methodology, present experimental results, and discuss the implications of our findings. Through this research, we hope to provide valuable insights and advancements that will aid in the development of more effective driver drowsiness detection systems.

2. Literature Review

In 2019, Wisaroot Tipprasert et al. [4] discussed the method for detecting driver drowsiness in low-light conditions by utilizing an infrared camera. The authors focus on the detection of driver's eye closure and yawning as indicators of drowsiness. The proposed methodology consists of four steps: face detection, eye detection, mouth detection, and eye closure and yawning detection. Through experiments conducted on a dataset of 3,761 images, impressive accuracy rates were achieved, with 99.47% for face detection, 94.33% for eye detection, 99.80% for mouth detection, and 92.5% for eye closure and yawning detection. These results highlight the effectiveness of the proposed method in accurately identifying eye closure and yawning, even in challenging low-light situations. For future work, the authors aim to enhance the detection of other drowsiness symptoms and optimize the implementation of the method by exploring different camera capture angles.

Also, in 2020, Dogiwal et al. [5] employed a Support Vector Machine (SVM) approach for the detection of drowsiness. They conducted image segmentation and emotion detection, specifically tracking facial expressions such as eye and mouth movements, using their own private dataset. The model demonstrated robustness to changes in illumination, enabling it to effectively perform in varying lighting conditions with an accuracy of 93%. The researchers also expressed their intention to enhance the model's adaptability to different environmental conditions as part of their future work to optimize its performance.

Manuscript received June 5, 2025
Manuscript revised June 20, 2025
https://doi.org/10.22937/IJCSNS.2025.25.6.7

While, in 2020, N.Divya J et al. [6] used Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) as part of the algorithm to detect drowsy drivers, CNNs are used for image processing and extracting information from video clips captured by the camera. RNNs are used to handle sequential data such as physiological signals or facial movements. The proposed model relies on image processing techniques to locate the face and detect major facial structures like the eyes and mouth. The Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR) are then calculated to determine whether the eyes are closed or if mouth opening is detected, as these are primary indicators of drowsiness. CNNs and RNNs are used to analyze this data and classify the driver's condition based on the extracted observations; the level of accuracy was estimated at 93%. Future enhancements for the model include incorporating parameters, such as blink rate, yawning, and car state to improve accuracy. Plans involve adding a heart rate sensor to prevent accidents from sudden heart attacks on drivers.

Additionally, in 2020, Kiran et al. [7] proposed research about driver drowsiness detection. This paper provides a comprehensive survey of recent studies focusing on driver drowsiness detection and alert systems. Also, it examines various machine learning techniques employed in this domain, including the PERCLOS algorithm, the HAAR-based cascade classifier, and the utilization of OpenCV. In the future, the frequency of yawning can be considered as a potential parameter for detecting drowsiness. The software underwent partial testing and demonstrated effectiveness. However, there is ample room for further enhancements. The proposed system identifies drowsiness when the eyes remain closed for at least four frames and effectively distinguishes normal eye blinks from drowsiness. It is a noninvasive system that could be enhanced by incorporating different types of sensors.

Furthermore, in 2021, Chand et al. [8] proposed an emotion analysis-driven CNN system for detecting driver drowsiness. This research paper introduces an innovative approach utilizing Convolutional Neural Networks (CNN) to detect driver drowsiness through a multi-level distribution model. The future work in this paper, it is not explicitly stated. The experimental results show that the proposed two-level Convolutional Neural Network (CNN) model achieves a high accuracy level of 93% in detecting driver behavior and emotions. It outperforms conventional classifiers like KNN and SVM. Increasing the processing duration improves accuracy, but a minimum duration is necessary for accurate distraction detection. The model effectively identifies the drowsy state of the driver, and using more than two levels may lead to overfitting and reduced accuracy. Overall, this model holds promise for reducing road accidents and improving road safety.

Moreover, Hashemi et al. [9], discussed a study on driver safety development, specifically addressing the issue of real-time driver drowsiness detection. The authors propose a system that employs Convolutional Neural Networks (CNN) for real-time eye closure detection, a prominent indicator of drowsiness. Additionally, they collected a dataset specifically designed for driver drowsiness detection, encompassing a novel state of the eye. Notably, the expanded dataset incorporates oblique views, enabling the system to function effectively in diverse scenarios. Comparative analysis of three networks, including a Fully Designed Neural Network (FD-NN) and two networks utilizing transfer learning in VGG16 and VGG19 architectures with additional custom layers (TL-VGG), demonstrates the system's high accuracy and low computational complexity in estimating eye closure. The FDNN network achieves an impressive accuracy of 98.15% and a 99.8% AUC. In future work, the authors plan to focus on yawning analysis for fatigue detection using landmark points of the mouth.

On the other hand, in 2021, Bakheet et al. [10] developed a vision-based system for driver drowsiness detection. Their approach involved utilizing the histogram of oriented gradient (HOG) technique for feature extraction and the Naïve Bayes (NB) algorithm for classification. They trained and evaluated their model using a dataset called NTHU-DDD, which comprised 376 videos. The proposed model achieved an accuracy of 85.62%. To improve the model's generalization capability, the authors plan to incorporate different datasets into their future research endeavors.

Additionally, in 2021, Faraji et al. [11] presented an approach to drowsiness detection using a combination of trained YOLOv3 (You Only Look Once) CNN and Long Short-Term Memory (LSTM) neural networks. The study also highlighted the creation of a custom dataset for training CNN. The programming language Python was used for the creation of the dataset by extracting 1042 images from videos. and for the LSTM dataset, it is a series of eye closure and yawning duration times. The study reported impressive results, achieving an accuracy of 91.7% in drowsiness detection.

Although, in 2021, Adithya Sajikumar et al. [12] investigated the Drowsy Driver Detection system, the system incorporates the Viola-Jones algorithm, leveraging HAAR features and AdaBoost technology, for efficient face detection. Additionally, the CAMSHIFT algorithm, which relies on mean shift for target tracking, is employed. Notably, eye movement analysis is utilized to identify signs of fatigue, prompting the activation of an alert system. By monitoring if the driver's eyes remain closed for a minimum of four frames, the system effectively detects drowsiness. This system proves highly effective in mitigating accidents caused by driver drowsiness in the context of automotive settings. Looking ahead, future work in this domain could concentrate on enhancing the real-time drowsiness detection system, with the aim of further refining its capabilities in accurately identifying and alerting drowsy drivers.

In conclusion, in 2022, Mohd Azlan et al. [13] focused on developing a fatigue and drowsiness detection system for car drivers using artificial intelligence techniques. The study sequentially examined two prominent models: the Haar cascade classifier and the Histogram of Oriented Gradient (HOG) + linear Support Vector Machine (SVM). The Haar cascade classifier showcased impressive performance, achieving an average frames per second (FPS) rate of 13.5 and a response time of 1.12 seconds. In contrast, the HOG + SVM model achieved a slightly lower average FPS rate of 6.2 but had a longer response time of 6.95 seconds. The primary objective of the research was to strike a balance between fast processing speed and accurate drowsiness detection. Moreover, the study highlighted the paramount importance of optimizing the system to minimize the occurrence of false positives and false negatives. Looking ahead, future research in this domain aims to explore the integration of advanced face detection technologies in order to enhance the overall effectiveness of the drowsiness detection system.

3. Dataset

The Drowsiness dataset, which consists of images of Ritesh Kanjee from Augmented Startups simulating both "drowsy," where there are 525 images, and "awake," where there are 705 images of facial expressions, is a valuable resource for developing a computer vision model focused on driver alertness or driver safety. This dataset can serve as a benchmark for training and evaluating machine learning algorithms aimed at detecting drowsiness in drivers [14].

Additionally, we will document the practical use of YOLOv8 and YOLOv5 by utilizing a different dataset than the one used in previous studies, (specifically data from RoboFlow), and we will compare the performance of YOLOv8 and YOLOv5 to the previous versions in our study. This comparative analysis contributes to a comprehensive evaluation of the ability of YOLOv8 and YOLOv5 to enhance the efficiency of drowsiness detection systems for drivers.

4. Research methods

A. YOLOv8

YOLO (You Only Look Once) is an advanced algorithm used in the field of computer vision for rapid detection and classification of objects in images and videos. The algorithm consists of multiple layers and multi-resolution learning modules that work together to improve the accuracy of detection and classification. The process of detection and classification in YOLO is an integrated single process, where the image is divided into a grid of cells and potential boxes containing objects are identified, and then those boxes are classified to determine the object type and its location. YOLOv8 is a significant development in the successful YOLO model series, and one of the key changes in YOLOv8 is the use of AnchorFree boxes. Anchor-Free is typically used in object detection models to assist in predicting the location and size of objects in the image. This feature helps to speed up the non-maximum suppression (NMS) process, which is a preprocessing step that disregards incorrect predictions. Another important change in YOLOv8 is the use of new loss functions for bounding box loss and classification loss. These functions work to improve performance, especially when dealing with small objects [15].

B. YOLOv5

YOLOv5 belongs to the YOLO (You Only Look Once) series, renowned for its object detection capabilities in images and videos. It's designed to enhance both performance and speed in identifying objects. This model partitions the image into a network of backbone networks and employs deep neural network techniques to efficiently recognize and categorize. objects within the image A distinctive feature of YOLOv5 is its single-stage architecture, which enables direct object detection from the image without intermediate steps. Through deep learning on extensive labeled datasets, YOLOv5 learns to identify various objects such as people, cars, animals, and more Known for its high performance and rapid inference speed, YOLOv5 is well-suited for real-time applications [16].

C. VGG16

The VGG-16 model, developed by the Visual Geometry Group (VGG) at the University of Oxford, is a convolutional neural network (CNN) architecture known for its simplicity and effectiveness. It consists of 16 layers, including 13 convolutional layers and 3 fully connected layers, making it a deep model. VGG-16 excels in various computer vision tasks, such as image classification and object recognition. Its architecture comprises a series of convolutional layers followed by max-pooling layers, progressively increasing in depth. This design allows the model to learn complex hierarchical representations of visual features, resulting in reliable and accurate predictions. Despite newer architectures, VGG-16 remains popular in deep learning applications due to its versatility and impressive performance.

E. Training Methodology for YOLO

We utilized the YOLOv8 and YOLOv5 models to train our dataset, achieving high accuracy. The dataset consists of two distinct classes and a total of 1,230 images. These images are divided into 1,056 training samples, 103 validation samples, and 71 test samples. During model training, we specified hyperparameters with a learning rate of 0.01 and a batch size of 16. The training process spanned 70 epochs, as summarized in Table 1.

TABLE 1. THE HYPERPARAMETERS SET DURING MODEL TRAINING

Hyperparameter	**YOLOv8,YOLOv5**
Input image size	640
Epochs	70
Batch size	16
Optimizer	SGD
learning rate	0.01

F. Training Methodology for VGG16

We used the VGG16 model to train a dataset containing two distinct classes and a total of 1230 images. The dataset was split into 1056 training samples, each containing two classes (awake, Drowsy) and 71 test samples, also containing two classes (awake and Drowsy). During the model training process, we specified hyperparameters, set the batch size to 100, and trained it over 30 epoch, as shown in Table 2.

TABLE 2: THE HYPERPARAMETERS SET DURING MODEL TRAINING

Hyperparameter	**VGG16**
Input image size	224
Epochs	30
Batch size	100
Optimizer	Adam

5. Results

A. First, we need to define some concepts related to YOLO

Mean Average Precision (mAP): mAP extends the concept of AP by calculating the average AP values across multiple object classes. This is useful in multi-class object detection scenarios to provide a comprehensive evaluation of the model's performance.
P (precision): quantifies the proportion of true positives among all positive predictions, assessing the model's capability to avoid false positives. On the other hand, R (recall) calculates the proportion of true positives among all actual positives, measuring the model's ability to detect all instances of a class [17].

B. YOLOv8 training results using 70 epochs and 16 batch size

The results Figure 1, 2, 3, and 4 obtained from training and evaluating the YOLOv8 model are undeniably impressive, highlighting its remarkable accuracy across all classes. The model achieves an outstanding accuracy rate of 98.8% for all classes, with a precision of 0.924 and a recall of 0.987. These results demonstrate the model's robustness in effectively identifying and classifying objects within the test dataset. It is particularly noteworthy that the model consistently performs well, especially in distinguishing between the “awake” and “drowsy” classes, with accuracy rates of 99.2% and 98.4%, respectively. These findings emphasize the model's reliability and efficacy in object detection tasks.

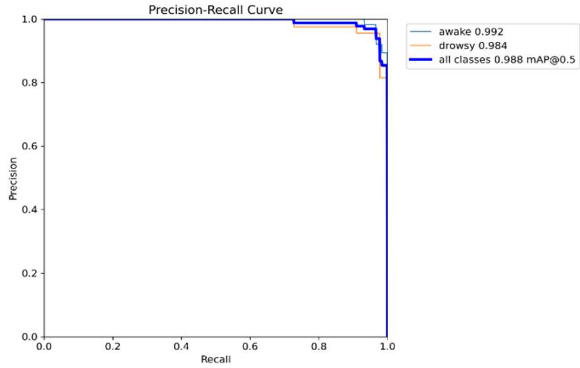

Figure 1: Training PR_Curve and mAP0.5 for YOLOv8

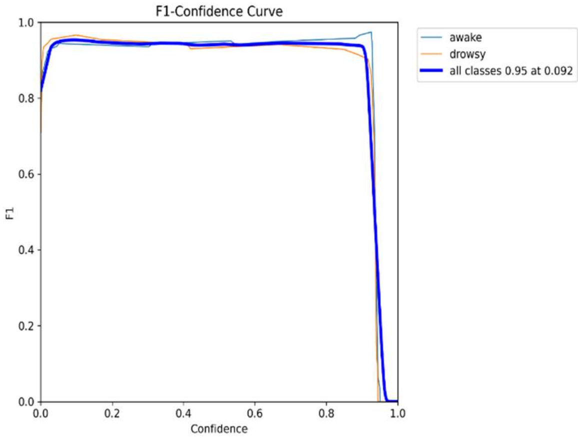

Figure 2: Training F1_Curve for YOLOv8

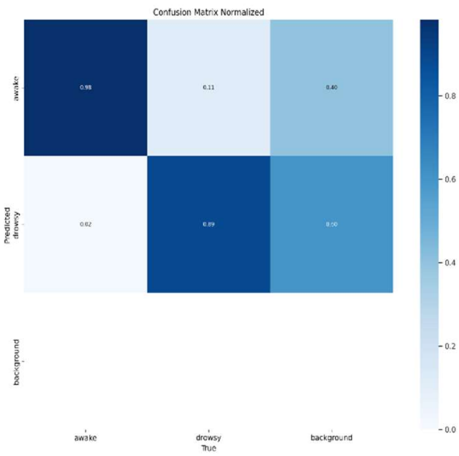

Figure 3: Training Confusion matrix normalized for YOLOv8

Figure 4: Examples of Validation Results for YOLOv8

C. YOLOv5 training results using 70 epochs and 16 batch size.

Our training model (Figure 5, 6, 7, and 8) has yielded impressive results for YOLOv5. The model achieved an overall accuracy of 96.3% with a precision of 0.952 and a recall of 0.949. When focusing on classifying the "awake" category, YOLOv5 demonstrated an accuracy of 96.7%. Similarly, for the "drowsy" class, it achieved an accuracy rate of 95.9%. These findings underscore the strong performance and effectiveness of YOLOv5 in accurately identifying and categorizing objects, particularly in distinguishing between the "awake" and "drowsy" classes.

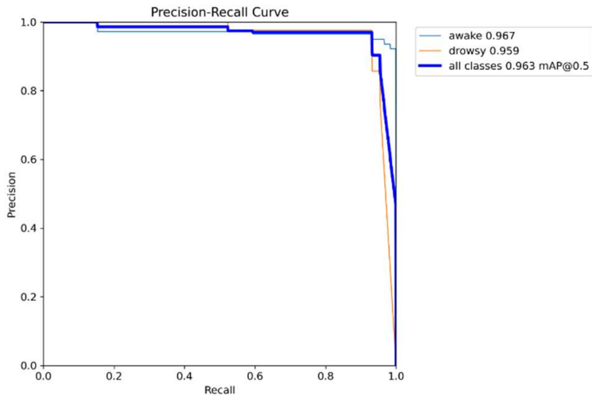

Figure 5: Training PR_Curve and mAP0.5 for YOLOv5

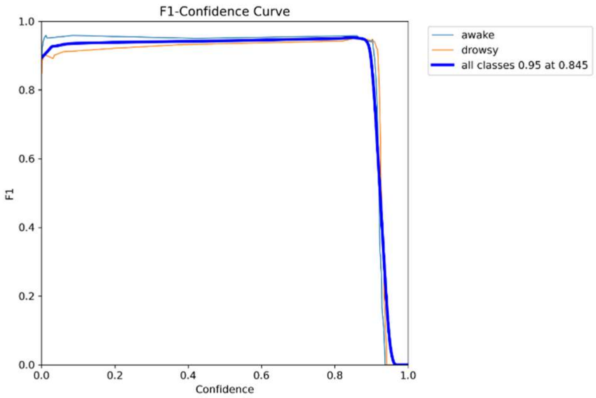

Figure 6: Training F1_Curve for YOLOv5

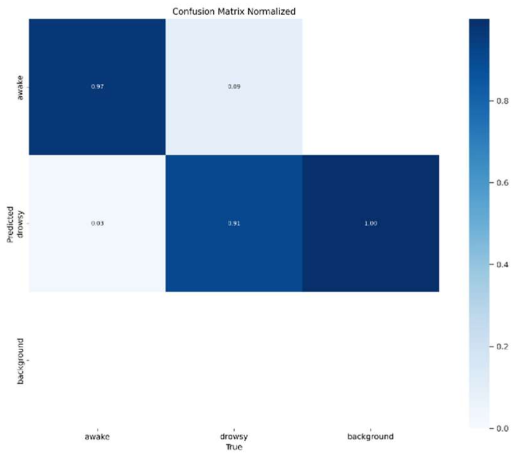

Figure 7: Training Confusion matrix normalized for YOLOv5

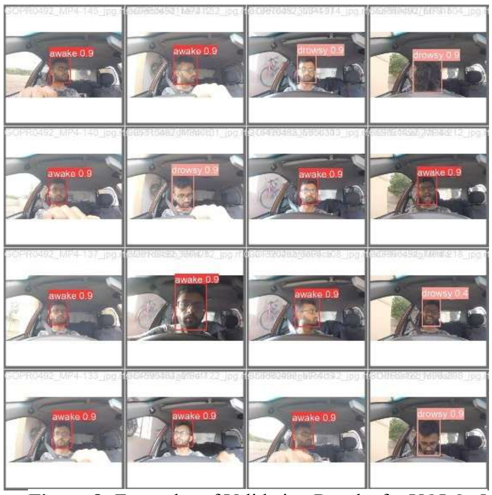

Figure 8: Examples of Validation Results for YOLOv5

TABLE.3: COMPRESSION BETWEEN YOLOv5 ANA YOLOv8 RESULTS

	"Drowsy" class	"Awake" class	Overall Accuracy
YOLOv5	95.9%	96.7%	96.3%
YOLOv8	98.4%	99.2%	98.8%

D. VGG16 training results using 7 epochs 2 verbose.

In the initial trial using the VGG16 model, we conducted training for 7 epochs with 2 levels of verbosity, resulting in an achieved accuracy of 57.39%.

E. VGG16 training results using 30 epochs 2 verbose.

By increasing the number of epochs to 30 while maintaining two levels of verbosity, a notable improvement in accuracy was observed. The model achieved an impressive accuracy of 97.73% in this subsequent attempt.

TABLE.4: COMPRESSION BETWEEN YOLOv5, YOLOv8 and VGG16 IN ACCURACY RESULTS

Model	Number of epochs	Accuracy results
YOLOv5	70	96.3%
YOLOv8	70	98.8%
VGG16	30	97.73%

Finally, the accuracy results obtained from the experiments demonstrate varying performance among the three models: YOLOv5, YOLOv8, and VGG16. YOLOv8 emerged as the top-performing model, achieving the highest accuracy of 98.8% after 70 epochs. Following closely was VGG16, with an accuracy of 97.73% after 30 epochs. YOLOv5 also demonstrated a respectable accuracy of 96.3% after 70 epochs.

Analyzing the class-wise accuracy, YOLOv8 showcased superior performance, achieving 99.2% accuracy for the "Awake" class and 98.4% accuracy for the "Drowsy" class. YOLOv5 exhibited slightly lower class-wise accuracies of 96.7% for the "Awake" class and 95.9% for the "Drowsy" class.

Considering these results, it can be concluded that YOLOv8 is the best-performing model, offering the highest overall accuracy and superior performance in class-wise accuracy compared to YOLOv5 and VGG16. However, the choice of the optimal model should also consider specific requirements, computational resources, and the desired balance between accuracy and speed. A comprehensive evaluation of these factors is crucial in selecting the most suitable model for a given task.

6. Discussion

YOLOv8 is an advanced algorithm used for object detection and classification in computer vision. It employs a multi-layered approach and multi-resolution learning modules to improve detection accuracy. One notable improvement in YOLOv8 is the use of anchor-free boxes, which aid in predicting object locations and sizes. This feature enhances the non-maximum suppression process, leading to more accurate predictions, especially for small objects. Additionally, YOLOv8 utilizes new loss functions for bounding box and classification losses, further improving performance. YOLOv5, another model in the YOLO series, is designed to enhance performance and speed in object detection. It partitions the image into a network of backbone networks and employs deep learning techniques for efficient object recognition. YOLOv5's single-stage architecture allows for direct object detection without intermediate steps, making it well-suited for real-time applications. It has gained recognition for its high performance and rapid inference speed. VGG16, developed by the Visual Geometry Group at the University of Oxford, is a convolutional neural network architecture known for its simplicity and effectiveness. With 16 layers, including convolutional and fully connected layers, VGG16 excels in image classification and object recognition tasks. Its hierarchical design enables it to learn complex visual features, resulting in reliable and accurate predictions. Despite newer architectures, VGG16 remains popular due to its versatility and impressive performance. In terms of training methodology, both YOLOv8 and YOLOv5 were trained on a dataset comprising 1,230 images with two distinct classes. The dataset was divided into training, validation, and test samples. The hyperparameters used for training included a learning rate of 0.01 and a batch size of 16. YOLOv8 was trained for 70 epochs, while YOLOv5 was trained for an unspecified number of epochs. On the other hand, VGG16 was trained on a dataset consisting of 1,230 images with two distinct classes. The dataset was split into training and test samples, with a batch size of 100. VGG16 was trained for 30 epochs. Overall, the YOLOv8, YOLOv5, and VGG16 models have different architectures and training methodologies, but they all excel in various computer vision tasks. YOLOv8 and YOLOv5 offer advanced object detection capabilities, with YOLOv8 incorporating anchor-free boxes and new loss functions, while VGG16 is known for its simplicity and effectiveness in image classification and object recognition. The chosen hyperparameters and training processes aim to achieve high accuracy and performance for the respective models and datasets used.

7. CONCLUSION AND FUTURE WORK

This paper presents a comprehensive exploration of the training methodologies and performance characteristics exhibited by the YOLOv8, YOLOv5, and VGG16 models in the domain of computer vision tasks. Through meticulous experimentation and analysis, we have endeavored to shed light on the inherent strengths and capabilities of each model architecture as they pertain to the critical tasks of object detection and classification within a dataset containing two distinct classes. The primary objective of our study has been to develop an advanced deep learning model specifically tailored to address the issue of driver drowsiness, with an unwavering focus on achieving a remarkable level of accuracy. The outcomes of our exhaustive experiments have yielded highly encouraging results, showcasing exceptional accuracy rates of 98.8% and 96.3% for the YOLOv8 and YOLOv5 models, respectively. Additionally, the VGG16 model has demonstrated commendable accuracy, attaining an impressive rate of 97.73%. These remarkable accuracy rates not only serve as a testament to the efficacy of the employed methodologies but also underscore their potential to significantly enhance road safety by effectively alerting drowsy drivers and proactively preventing potential accidents.

For future work, the highly anticipated YOLOv9 holds the promise of pushing the boundaries of accuracy and performance even further. Moreover, expanding the dataset's size and diversity would enable a more comprehensive evaluation, ensuring the robustness and generalizability of the developed deep learning model. By capitalizing on these opportunities for refinement and enhancement, we can continue to make substantial strides toward improving driver safety and mitigating the risks associated with drowsy driving.

References

[1] Fossati, A., Dimitrijevic, M., Lepetit, V., Fua, P. and Pollefeys, M., 2008. "Using Projective Depth to

Compute Stereo Disparity". In: R. Koch and F. Huang, eds. 2008. *Computer Vision Systems*. Lecture Notes in Computer Science, vol 5008. Berlin, Heidelberg: Springer, pp.23-34. Available at: https://link.springer.com/chapter/10.1007/978-3-540-75773-3_2 [Accessed 20 May 2024].

[2] Tommasi, M., Thiebaut, N., Derkenne, S. and Verhulst, B., 2022. "Neurocognitive and mental health impacts of the COVID-19 pandemic on children and adolescents: A systematic review". *Frontiers in Psychiatry*, [online] 13, Article 891489. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8914892/ [Accessed 20 May 2024].

[3] Fossati, A., Dimitrijevic, M., Lepetit, V., Fua, P. and Pollefeys, M., 2008. "Using Projective Depth to Compute Stereo Disparity". In: R. Koch and F. Huang, eds. *Computer Vision Systems*. Lecture Notes in Computer Science, vol. 5008. Berlin, Heidelberg: Springer, pp.23-34. Available at: https://link.springer.com/chapter/10.1007/978-3-540-75773-3_2 [Accessed 20 May 2024].

[4] Tipprasert, W., Chianrabutra, C., Charoenpong, T., & Sukjamsri, C. (2019). A Method of Driver's Eyes Closure and Yawning Detection for Drowsiness Analysis by Infrared Camera. Engineering Research Unit for Enhancing the Quality of Life, Srinakarinwirot University, Thailand, Available at: https://ieeexplore.ieee.org/abstract/document/8646001

[5] Dogiwal, S.R. and Sharma, V. (2020) 'Driver Fatigue Detection Analysis Based on Image Segmentation & Feature Extraction Using SVM', SKIT Research Journal, 10, pp. 1–5. Available at: https://doi.org/10.47904/IJSKIT.10.1.2020.1-5.

[6] Jeevan, P., & Divya, N. (2020). Drowsy Driver Detection with CNN and RNN. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 11(3), 2222-2226. Available at https://turcomat.org/index.php/turkbilmat/article/view/123

[7] V B Navya Kiran, Raksha R, Anisoor Rahman, Varsha K N, Dr. Nagamani N P, 2020, Driver Drowsiness Detection, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NCAIT – 2020, Available: https://www.ijert.org/driver-drowsiness-detection

[8] H. Varun Chand and J. Karthikeyan, "CNN Based Driver Drowsiness Detection System Using Emotion Analysis," 2021, Available: https://www.techscience.com/iasc/v31n2/44533/pdf

[9] Hashemi, M., Mirrashid, A., Beheshti Shirazi, A. (2021). "Driver Safety Development: Real-Time Driver Drowsiness Detection System Based on Convolutional Neural Network" Available at: https://link.springer.com/article/10.1007/s42979-020-00306-9

[10] Bakheet, S. and Al-Hamadi, A. (2021) 'A framework for instantaneous driver drowsiness detection based on improved HOG features and naïve Bayesian classification', Brain Sciences, 11, p. 240. Available at: https://doi.org/10.3390/brainsci11020240

[11] Faraji, F. Lotfi, F., Khorramdel, J., Najafi, A., & Ghaffari, A. (2021). "Drowsiness Detection Based On Driver Temporal Behavior Using a New Developed Dataset." arXiv preprint arXiv:2104.00125 Available at https://arxiv.org/abs/2104.00125

[12] Sajikumar, A, Aji, A, Baby, J. Prasad, A., Stanly, A. K., & John, N. M. (2021). Save The Drowsy Driver Drowsy Driver Detection. International Journal of Engineering Research & Technology (IJERT), Volume 9(7), 53. Available at https://www.ijert.org/save-the-drowsy-driver-drowsy-driver-detection

[13] Abu, M. A., Ishak, I. D., Basarudin, H., Ramli, A. F., & Shapiai, M. I. (2022). Fatigue and Drowsiness Detection System Using Artificial Intelligence Technique for Car Drivers. In A. Ismail et al. (Eds.), Design in Maritime Engineering (pp. 421-424). Springer Nature Switzerland AG. Available at https://link.springer.com/chapter/10.1007/978-3-030-89988-2_31

[14] Benjumea, A., Teeti, I., & Cuzzolin, F. (2022). YOLO-Z: Improving small object detection in YOLOv5 for autonomous vehicles.

[15] Benjumea, A., Teeti, I., & Cuzzolin, F. (2022). YOLO-Z: Improving small object detection in YOLOv5 for autonomous vehicles

[16] [Roboflow. (n.d.). Roboflow. Retrieved from https://roboflow.com/

[17] Ultralytics. (2024). YOLO Performance Metrics - Ultralytics YOLOv8 Docs. Retrieved from https://docs.ultralytics.com/guides/yolo-performance-metrics/