Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)
Vol. 8, No. 3, November 2025, pp. 800 – 810
p-ISSN: 2614-3372 | e-ISSN: 2614-6150

Intelligent Alert System With Yolo V8 Algorithm for Early Detection of Microsleep In Vehicle Drivers

[1*]Ria Citra Desiany Putri, [2]Kusumanto, [3]Dewi Permata Sari
[1,2,3]Electrical Engineering D4 Study Program, Department of Electrical Engineering,
Sriwijaya State Polytechnic, Indonesia
Email: [1]riacitradesianyputri@gmail.com, [2]manto_6611@yahoo.co.id, [3]dewipematasari@polsri.ac.id

Article Info

Article history:
Received Jul 10th, 2025
Revised Oct 12th, 2025
Accepted Nov 31th, 2025

Keyword:
Computer Vision
Microsleep
Object Detection
Raspberry Pi
YOLO

ABSTRACT

Microsleep is a brief state of sleep that occurs suddenly without the person being aware of it and poses a serious risk to drivers, especially on long journeys. This study developed an intelligent alert system based on the YOLOv8 algorithm for the early detection of microsleep in drivers in real time by analyzing the state of the eyes and the position of the head. Using 3,458 annotated facial images as training data, the model was implemented on the Raspberry Pi platform for local processing without cloud dependency. The system activates a buzzer and warning light when it detects signs of drowsiness. Test results show the effectiveness of this method in the early detection of microsleep with 90.3% precision, 91.3% accuracy, 96.8% recall, and an F1 score of 93.9%. It has been shown to function optimally in a variety of lighting conditions to improve road safety.

Corresponding Author:
Ria Citra Desiany Putri,
Electrical Engineering, Electrical Engineering D4 Progran Study,
Sriwijaya State Polytechnic,
Jl. Srijaya Negara, Bukit Lama, Bukit Besar, Palembang City, South Sumatera, 30139.
riacitradesianyputri@gmail.com

DOI: http://dx.doi.org/10.24014/ijaidm.v8i3.37766

1. INTRODUCTION

Microsleep is a brief state of sleep that occurs unconsciously in a person and usually lasts from a few seconds to several tens of seconds. In the context of driving, this state is very dangerous because it can lead to loss of consciousness and a slowing of the driver's reaction time, even if only for a very short period of time. Microsleep often occurs in drivers who are tired, drowsy, or driving for long periods of time, especially at night or on unstimulating roads. The World Health Organization (WHO) reports that fatigue and drowsiness while driving are one of the leading causes of road accidents in many countries, with thousands of victims each year due to reduced driver alertness [1]. Therefore, early detection of microsleep is a very important aspect of efforts to improve road safety and minimize the risk of road accidents.

With technological advances, various approaches have been developed to detect drowsiness and microsleep in drivers, ranging from methods based on physiological sensors to approaches based on computer vision. The computer vision approach is considered more practical and non-invasive, as it uses only the image of the driver's face to analyze visual indicators such as eye state, facial expressions, and head position in real time. In recent years, deep learning-based object detection algorithms, particularly the You Only Look Once (YOLO) family, have been widely used due to their ability to quickly and accurately detect objects in a single processing step. The YOLO model is considered particularly suitable for real-time applications because it can classify and locate objects simultaneously in a single pass, making it more efficient than conventional two-step methods [2]. Thanks to these advantages, YOLO is a promising solution for driver monitoring systems that require rapid response and high accuracy in detecting the early signs of microsleep.

Iwan Virgiawan, et al, 2024[3], used the YOLO V3 method for real-time object detection and tracking based on computer vision. The results of this study show that the system can detect a range of up to 100 cm when the object is close to the camera and up to 150 cd/m2 when the object is in the most crowded room. For now, this research has not been able to detect or analyze the profile of the person being detected. However, this research can be further developed by improving the analysis of the person's profile and also by offering assistance when an object is detected so that it is easier to analyze it.

Dw Ayu Agung Indra, et al., 2023[4], Implementation of the Yolo V5 Detector Algorithm on Motor Vehicle Drivers in Realtime, showed that the model trained using the Yolo V5 algorithm with 247 image data and split data with a ratio of 80:20 managed to detect drowsiness in motor vehicle drivers correctly and reliably. However, this is done so that future research can provide more complete and varied datasets so that the model can learn more features and data. Mihuandayani,et al, 2025[5], Implementation of YOLOv8 for Human Object Detection and Counting with Convolutional Neural Networks, Test results show that the system is capable of correctly detecting human objects with an 85% success rate. However, some obstacles remain, such as the difficulty in detecting people walking close to each other and those carrying large bags.

The latest development in the YOLO algorithm family is marked by the introduction of YOLOv8, which is designed to improve detection accuracy while maintaining computational efficiency in real-time applications. YOLOv8 adopts a lighter and more optimized architecture than previous versions, and supports an anchor-free detection approach that can improve object localization accuracy, especially for small and complex objects. Research by Chen et al. shows that YOLOv8 is capable of integrating various facial features simultaneously to detect driver fatigue in real-time with a high degree of accuracy, even in less than ideal lighting conditions [6]. This advantage makes YOLOv8 particularly relevant for microsleep detection applications, where small changes in eye condition and facial expression must be detected quickly and accurately.

In terms of performance, various studies show that YOLOv8 has significant advantages over conventional object detection models and previous versions of YOLO. Golfantara reports that YOLOv8 exhibits more stable and accurate real-time detection performance than YOLOv5 and YOLOv7, especially in scenarios with a wide variety of objects and dynamic environments [7]. Statistical analysis also shows that YOLOv8 offers a better balance between precision and recall, reducing detection errors (false positives) and detection failures (false negatives). This advantage is particularly important in microsleep early warning systems, as detection errors can lead to unnecessary alerts or, conversely, failures to detect dangerous conditions in the driver.

In addition to its advantages in terms of accuracy and speed, YOLOv8 also offers great flexibility in the training and implementation process of the system. The Ultralytics framework used in YOLOv8 supports various data augmentation and transfer learning techniques, making it easy to adapt the model to limited datasets and varied environmental conditions. Jonathan and Hermanto point out that YOLOv8's ability to adjust input resolution, number of epochs, and training conFigureuration makes it particularly suitable for implementation on resource-constrained devices such as the Raspberry Pi [8]. This flexibility enables the development of a microsleep detection system that is not only accurate, but also efficient and autonomous, without relying on cloud computing, making it more applicable for real-world use in vehicles.

2. RESEARCH METHOD

This research takes an experimental approach to develop a microsleep detection system based on the YOLOv8 algorithm, with the goal of enabling real-time detection through facial image analysis. The research process includes several integrated steps, including data collection, data labeling, model training, and system evaluation on the target device. The proposed system combines digital image processing, deep learning-based object detection, and embedded system implementation. Facial images are captured using a camera, annotated to represent open, closed, and irrelevant eye conditions, and then trained using the YOLOv8 architecture. The system is deployed on a Raspberry Pi as the main processing unit, supported by a USB camera for visual input and output devices such as buzzers and LEDs to provide real-time alerts when signs of microsleep are detected.

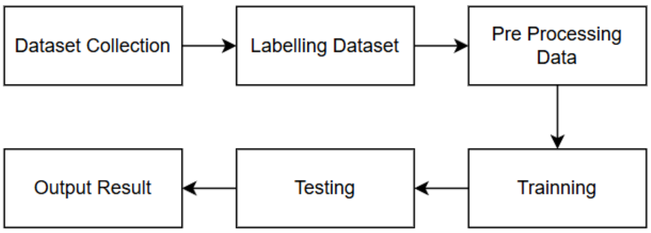

Figure 1. YOLO Blok Diagram

A diagram of the data collection process in a deep learning-based detection system, such as YOLOv8, is shown in Figure 1. In simple terms, the following is an explanation. The process begins with data input, which is the collection of data in the form of images or videos. Next, the data is presented through the Data

Labeling stage, where objects in the image, such as mats or faces, are annotated. Once labeled, the data proceeds to the Pre-Processing stage, which includes normalization, image size modification, and format conversion to fit the model. After data collection, the two main steps are training and testing. In the training stage, the model learns from the data to analyze the target. On the other hand, testing is used to assess the performance of the model in relation to unexamined data. Finally, the system produces output results that are either real-time object detection or classification predictions.

2.1. Literature Review

The selection of YOLO v8 in this study was based on previous research on microsleep detection. This study comprehensively reviewed various academic literature and previous research discussing the concept of microsleep, object detection algorithms, and the application of computer vision in driving safety systems, as described in Table 1.

Table 1. Academic literature and previous research discussing the concept of microsleep

Author	Title of the Study	Research Issue	Research Method	Results
Edmund Ucok Armin, Anggun Purnama Edra, Fakhri Ikhwanul Alifin, Ikhwanussafa Sadidan, Indri Purwita Sary, Ulinnuha Latifa, 2023 [9]	Performance of the YOLO v8 Model for Detecting Drowsiness in Car Drivers	a drowsiness detection system that is less accurate without direct contact with the driver.	YOLO v8 Algorithm	Based on the study's findings, the YOLOv8 model was able to use 2996 photos from the secondary dataset to train, achieving a mAP value of 0.96092. Using the same dataset they had gathered, the researchers further assessed the YOLOv8 model's performance and contrasted it with other object identification models, such as SSD-Resnet, YOLOv3, and YOLOv5. According to the evaluation results, YOLOv8 exhibits superior object identification accuracy even though its frame rate is 50 FPS lower than SSD-Resnet's.
Rhadis Steffani Saputri1, Aulia Apriliani, Rizky Syahrul Amar, Lola Yorita Astri, 2025 [10]	Early Detection of Microsleep in Four-Wheeled Vehicle Drivers Through Eye Images Using the YOLOv8 Algorithm	The system has difficulty detecting when the driver turns their head or tilts their head.	YOLO v8 Algorithm	Previous research successfully developed a microsleep early detection system based on the lightweight and efficient YOLO v8n algorithm, enabling high-precision real-time recognition of a driver's eye state. Evaluation results demonstrated outstanding performance with mAP@0.597.9%, precision 97.2%, recall 96.7%, and an F1 score of 0.97, proving extremely low detection error under various lighting conditions
Saepudin, Nana Sujana, Muhamad Malik Mutoffar, Alfian Azzam Haryanto, 2024 [11]	YOLO v8 Performance Analysis Roboflow Optimization for Emotional Facial Expression Detection with Machine Learning	Face detection is often inaccurate due to poor data quality.	YOLO v8 Algorithm	The research results show that optimizing RoboFlow significantly improves YOLOv8's ability to recognize emotional facial expressions, with higher accuracy than models trained without optimization. In addition, there is also an increase in processing time efficiency, making the model suitable for real-time applications. This study demonstrates that combining YOLOv8 and RoboFlow results in a more reliable and efficient model for recognizing emotional facial expressions.
P.Vijay Bhaskar Reddy, S.Sai Rachana, 2023 [12]	Driver Drowsiness Monitoring System Using Yolov8	The results are quite good, but remain weak in terms of performance evaluation and error analysis.	YOLO v8 Algorithm	In this project, we used the YOLO (You Only Look Once) V8 object recognition system to train the data set. This model successfully recognized both classes (asleep and awake) as specified in the data set and successfully recognized and monitored the driver's condition. Our accuracy reached 91.1%. This model is faster and more accurate than other models based on the KNN algorithm because it recognizes objects in one step, not in two stages.

2.2. Microsleep

Microsleep is a very brief phase of sleep that usually lasts only a few hours, although it can also result in prolonged periods of inactivity. This phenomenon often occurs without anyone realizing it, and can happen when someone is sitting in a focused state, especially when they are driving, especially when they are feeling very sad or sleepy. Microsleep often occurs in situations that require a high level of alertness, such as when you feel sleepy, which can help you and others around you. Because the brain needs to be active to get information from the surrounding environment in a timely manner, when microsleep occurs, drivers can lose control of their vehicles. And this is one of the main causes of traffic accidents, especially when traveling long distances or at night [13].

2.3. Artificial Intelligence (AI)

The creation of robots or systems that can mimic human intellect and reasoning particularly in terms of decision-making, problem-solving, and even pattern recognition is the focus of the computer science discipline known as artificial intelligence (AI). AI is designed to make it possible for computers to do specific tasks reliably and automatically, possibly eliminating the necessity for peaceful human contact. One of AI's main characteristics is its ability to learn from data and function in a chaotic environment. With this skill, AI can not only execute statistical orders but also adapt to its environment and operating context. Modern AI has developed largely due to large-scale data analysis and automated learning techniques like machine learning and deep learning, which form the basis of many AI applications today [14].

2.4. Machine learning

Machine learning (ML), a branch of artificial intelligence, develops methods for systems to learn from data. To put it briefly, machine learning algorithms make predictions or judgments about new data by using historical data. In the domain of microsleep detection, machine learning allows computers to analyze the aspects of drowsiness based on visual characteristics identified from photographs of people or floors. The ability of machine learning to generalize from historical data makes it an essential tool for developing reliable systems. YOLO (You Only Look Once), a machine learning technique utilized in photo applications, is one type of deep learning-based object recognition technique used in this study to identify conditions in real time [15].

2.5. Deep learning

A subfield of machine learning called “deep learning” uses the design of a deep neural network to analyze and interpret data. This approach is inspired by the way people decipher patterns and draw inferences from their experiences. Deep learning is particularly useful in dealing with high-dimensional data, such as text, writing, and images, as it can automatically extract features from raw data. In image processing, deep learning - specifically, the Convolutional Neural Network (CNN) architecture - is used for accurate image classification, object evaluation, and localization. One of the main benefits of deep learning is its ability to make complex representations of features without the need to create human features, making it ideal for applications such as camera-based partial sleep detection [16].

2.6. Roboflow

Roboflow is a web platform that offers features related to dataset management. Roboflow is a computer vision development framework that enables better data collection, preprocessing, and model training techniques. Roboflow allows users to share datasets while processing them, annotate or label objects to be detected using selection frames, and preprocess datasets, for example by converting them to grayscale and augmenting them using Roboflow. To test object detection, a dataset is required at the data acquisition stage (data collection), which poses a challenge for researchers who need to collect a dataset of high-quality vehicle images for object detection [17].

2.7. You Only Look Once (YOLO)

YOLO, which stands for “You Only Look Once”, is one of the deep learning-based object detection algorithms designed to identify and locate objects in images or videos in real time. It describes the traditional method of working in object detection, which is often done through two stages: classification and region proposal (the study of regions that may include objects). In contrast to these approaches, YOLO reduces the process to a single-stage detector, where the model performs detection and classification in one forward pass on each image[18]. The YOLO V8 architecture is shown in Figure 2.

YOLO works quickly because all predictions are made simultaneously in one pass, which makes it ideal for real-time applications. In the learning process, the model teaches the relationship between visual characteristics and object locations based on pre-labeled datasets. Each bounding box has dimensions (width and height), location information (x, y), and probability values associated with a particular object class. To display prediction results and predict overlap, YOLO also uses the Non-Maximum Suppression (NMS) technique, which only strengthens the predicted boxes with the highest scores[20].

2.8. Object Detection

A computer vision and image analysis approach called object detection is used to locate and identify one or more objects in an image or video. Object detection provides information in the form of bounding boxes and labels for the identified objects, unlike biased image classification, which only provides labels for the entire image. This method makes it easier to achieve the two main goals of object localization and object classification. One of the most widely used and popular object detection methods is the YOLO (Look Only

Once) method. This method is highly effective and ideal for real-time applications such as drowsiness detection systems as it can identify multiple objects in a single operation [21].

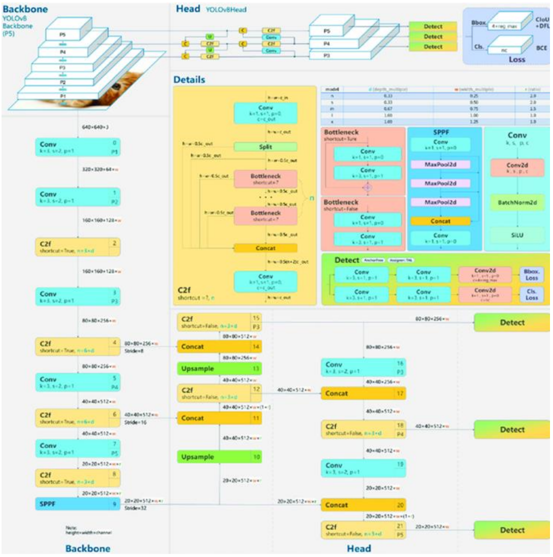

Figure 2. YOLO V8 Architecture [19]

2.9. Confusion Matrix

One type of tabular representation used to evaluate the performance of classification models is the confusion matrix. This matrix compares the labels derived from the data (actual classes) with the predictions made by the model (predicted classes). The three main components of this matrix are: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) in the basic form of binary classification [22].

This matrix can be used to generate different evaluation metrics, such as F1 score, precision, recall, and accuracy. These metrics are useful for evaluating the strengths and weaknesses of the model, especially in cases where the distribution of students is not uniform. Predictions with IoU > 0.5 are considered as true positives, while predictions with IoU < 0.5 are categorized as false positives or false negatives. The confusion matrix also displays the detection results based on the IoU associated with object detection [23].

3. RESULTS AND ANALYSIS

A number of performance measures and visual diagnostics were used to assess the YOLOv8 microsleep detection system. The model was evaluated using validation data that represented three predetermined focal point and microsleep classes after it had been trained on around 3,458 annotated pictures. Training curves, model accuracy measures, and matrix analysis were used to evaluate performance. The model is also captured by a camera in this study, and the data is analyzed by a Raspberry Pi before being sent to a You Only Look Once (YOLO) V8 for processing to an LED and buzzer.

3.1. Collection Dataset

The system needs a collection of pictures or videos of the user in different states, such as motion and focus, in order to proceed with the initial stage, data collecting (Figure 3).

Figure 3. Collection Dataset Microsleep

A dataset of driver face photos in several conditions, including awake, eyes open, and non-drowsy, is displayed in Figure 3. The driver's eyes are half-closed or closed when they are drowsy, and they yawn.

3.2. Labelling Dataset

The labeling process is an important step in deep learning-based microsleep detection systems such as YOLOv8. Its purpose is to provide the model with structured data on the location and classification of the items to be identified. This procedure is performed after the collection of facial image data from several sources, including videos, live cameras, and external databases. A bounding box is used to annotate each collected image, indicating where the driver's face is located. Next, a class is assigned to each object in the image, such as focus for normal conditions or micro-sleep for drowsy conditions [24]. A dataset labelling example can be seen in Figure 4.

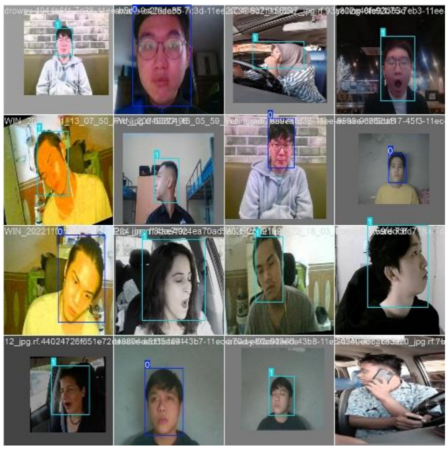

Figure 4. Dataset labelling example

The model must distinguish between the driver's eyes being fully open (focused), half-closed, and completely closed (microsleep) in the context of microsleep detection in Figure 4. The model's ability to identify these important states when the driver is operating the vehicle increases with annotation accuracy and consistency. Inaccurate class labels, missing faces, or redundant bounding boxes are examples of small labeling errors that can drastically decrease the model's prediction accuracy. Therefore, monitoring the quality of labeling is an important step in the data preparation process [24].

3.3. Pre-Processing Data

The initial action performed before using the data for training is called preprocessing. During this phase, a number of transformation procedures are used to prepare the raw data, such as images taken by a camera or external datasets. These procedures involve shrinking the image to fit the input model (640 x 640 pixels), normalizing pixels, changing the color format, and adding data (e.g., flipping, rotation, or brightness modification) to increase the variety of data. Preprocessing is mostly done to ensure that all input data represents the actual situation that the model will face and has a consistent structure and format [25].

Figure 5. Pre-Processing Data

Figure 5 displays a collection of photos of the driver's face that have been analyzed and captioned. Each image displays the outcome of the face area selection box using the class labels from the prior classification step, which were partial sleep or attention.

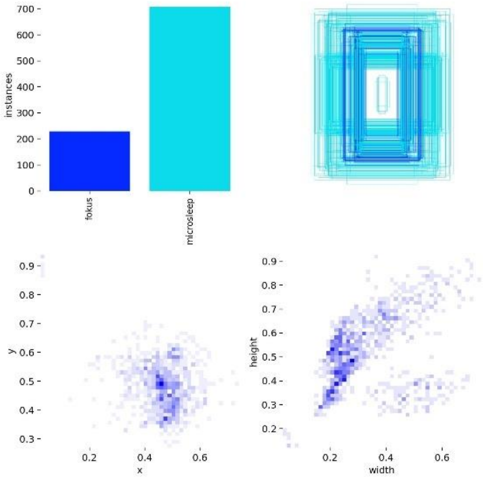

Figure 6. Exploration Visualization Dataset

The results of the initial exploration of the dataset are visualized in Figure 6 after the annotation process but before model training. Since the image shows the class and bounding box distributions, as well as differences in object size and location within the dataset, this visualization is an important component of the data preparation step. This visualization can be used to determine whether the dataset is representative and suitable for training with a detection model such as YOLOv8.

There is an imbalance in the data, as evidenced by the remaining class data exceeding the focal class data in the bar graph on the top left, which displays the number of examples of each class. By increasing the data in the minority class or changing the weights during training, this imbalance can prevent bias during training. The facial objects are regularly distributed in the center of the image, as shown in the bounding box overlay visualization on the top right, indicating a systematic and purposeful data collection procedure. The coordinate distribution of the bounding box center, which is mostly centered around the center of the image, is shown in the heat map on the bottom left. This shows that the faces of the objects are often centered in the middle. In contrast, the size distribution of the bounding box in the bottom right falls within a predefined range in terms of length and width, indicating consistent element dimensions that support model learning. Although some tactics are required to overcome class imbalance, annotated data is often of high quality and structure for use in training detection models.

3.4. Training

Training image processing, identifying patterns in each class, and evaluating model performance using metrics such as loss, precision, recall, and MAP are all part of the training process.

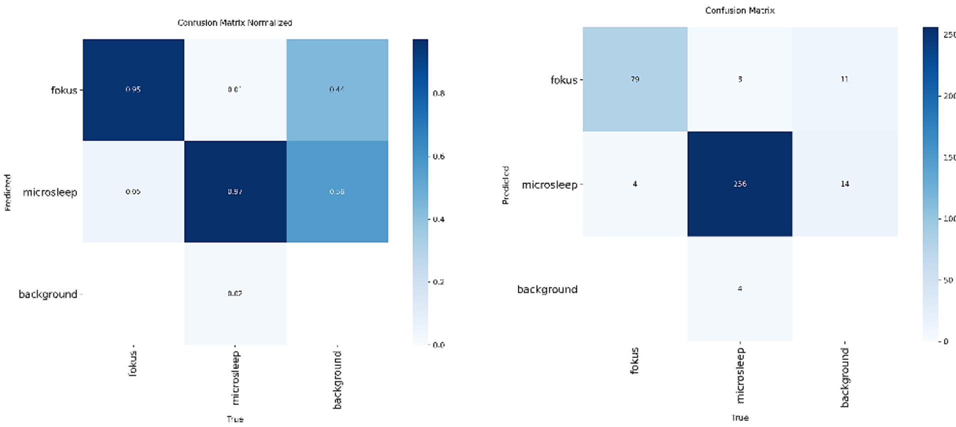

Figure 7. Confusion Matrix

The first matrix in Figure 7 is the confusion matrix, which shows the percentage of predictions for each category. For example, about 44% of the actual background data was not identified as focal, indicating a large degree of uncertainty in the background category, while about 95% of the actual focal data was correctly identified. The absolute version of the confusion matrix, which shows the number of correct model predictions, comes in second. While there were still some prediction errors in other categories, including focus and background, it is clear that the model performed very well in identifying partially silent data (256 accurate data out of 274).

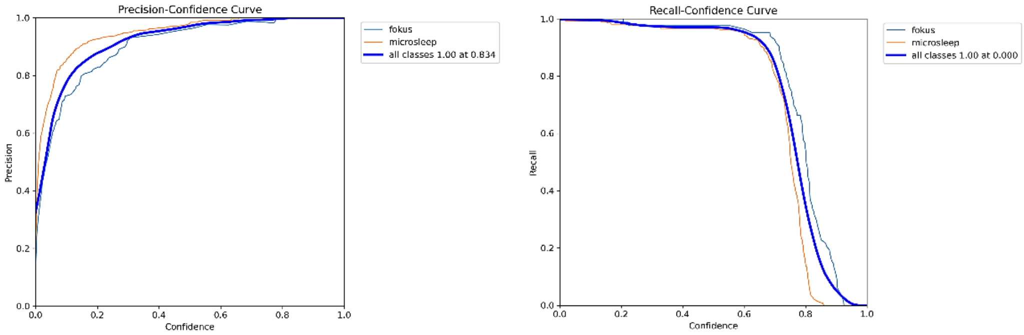

Figure 8. Confusion Matrix PR

The recall-confidence curve, which shows the relationship between recall value and confidence level, is first shown in Figure 7. Moving to the right (high confidence) causes the recall value to decrease dramatically. This means that the number of correctly identified cases decreases when the model only makes predictions at a high confidence level. This suggests that some data may not actually be observed despite the model having high confidence in some predictions, especially in the partial sleep category which seems to decline faster than the attention category. The confidence-accuracy curve, the second graph, shows the relationship between confidence and accuracy. The accuracy value increases as the confidence level increases. Predictions with higher confidence levels tend to be more accurate, as shown in this graph. In most confidence ranges, the accuracy of the partial sleep category is greater than the accuracy of the attention category, indicating that the model tends to be more accurate when predicting partial sleep with high confidence.

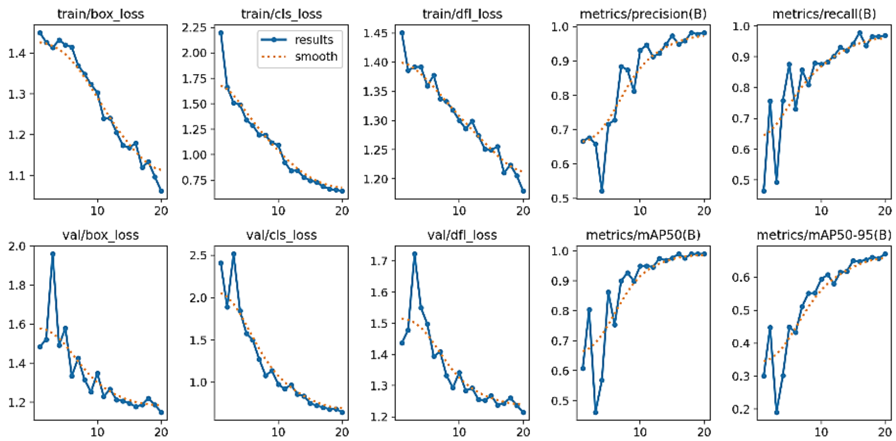

Figure 9. Performance Graph YOLO V8

Figure 8 shows the accuracy, or the proportion of the model's predictions that are actually relevant. For example, the accuracy value will decrease if the model incorrectly predicts too many “partial sleep” images. The graph shows that the accuracy of the “Partial Sleep” category is somewhat greater than the accuracy of the “Concentration” category, indicating that the model is very accurate and confident in identifying partial sleep states. Recall refers to how accurately the model represents each scenario in the real world. A low recall rate indicates that many cases were missed. The graph shows that the model did not miss too many “Focused” or “Partial Sleep” events as the recall for both categories is very high and balanced. F1-Score provides a balanced view of precision and recall by aligning the two. A high F1-Score indicates that the model is complete (high recall) and also accurate (high precision). The graph shows that the F1-Score for both categories is almost the same, which indicates that the model has a fairly consistent performance in both categories.

3.5. Testing

In the testing phase, new data is used to evaluate the ability of the trained YOLO model to determine the driver's state, especially in distinguishing between attention and partial sleep. The purpose of this testing is to see if the model can apply the information it has acquired during training to real-world scenarios that have never been observed before.

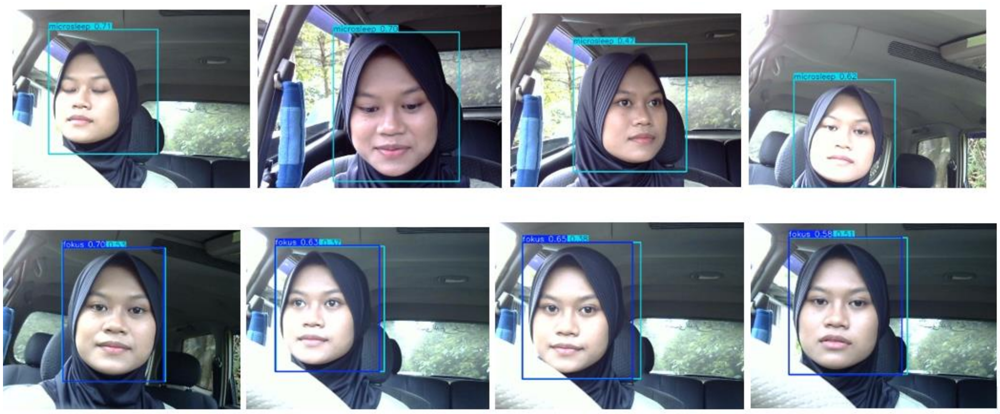

Figure 10. Training Result

According to the test results in Figure 10, the model can recognize drivers' faces and use their eye expressions to categorize their condition. The model classifies the eyes as “focused” with a fairly high confidence level of 0.63 or 0.65 when the eyes are open. On the other hand, with confidence values between 0.47 and 0.71, the model categorized the eyes as “partially asleep” when the eyes were closed or microsleep. The precision was 90.3%, accuracy was 91.3%, recall was 96.8%, and the F1 score was 93.9%, indicating that the model consistently recognized concentration and microsleep conditions across various scenarios.

When compared to previous studies that used older versions of YOLO, there was a significant improvement in performance. Research by Virgiawan et al. using YOLOv3 showed that the system was capable of real-time object detection but was still limited to object-level analysis and had not been specifically optimized for driver facial conditions or microsleep [3]. Meanwhile, research by Indra et al. that applied YOLOv5 for driver drowsiness detection obtained fairly good results, but used a relatively small dataset (247 images), thus potentially limiting the model's generalization ability in various lighting and face angle conditions [4]. Compared to the two studies, the use of YOLOv8 in this study showed more stable and consistent performance, especially when implemented on devices with limited resources such as Raspberry Pi. This is in line with the findings of Armin et al., who reported that YOLOv8 has higher detection accuracy than YOLOv3 and YOLOv5, despite a slightly lower frame rate [9]. In addition, Saputri et al.'s research also reported YOLOv8's very high performance in microsleep detection with precision and recall values above 96% in varying lighting conditions [10].

The main difference that is the advantage of this study is the success in maintaining high performance on the Raspberry Pi platform, which has computational limitations compared to PCs or high-end GPUs. Reddy and Rachana's research reported an accuracy of 91.1% using YOLOv8, but the evaluation of hardware performance and error analysis was still limited [12]. In contrast, this study shows that with dataset optimization, training conFigureuration, and parameter adjustment, YOLOv8 is still able to work optimally on embedded systems without relying on cloud computing.

4. CONCLUSION

Testing results on 3,458 datasets show that the YOLOv8 model is capable of performing quite well. Overall, the system achieves an accuracy of 90.3%, with an average accuracy of 91.3% and a recall of 96.8% for the two main classes, namely concentration and microsleep. The F1 score of 93.9% shows a good balance between the model's ability to identify drowsiness states and avoid detection errors. This result confirms that the system is capable of consistently detecting signs of microsleep in most test scenarios.

The developed system shows good detection capability even in less than ideal lighting conditions, such as when there is backlighting behind the driver. Under these conditions, the face can still be detected and classified fairly consistently, whether the eyes are open or beginning to close. This shows that the YOLOv8 model is capable of adapting to variations in light intensity and changes in the direction of the light source. However, test results also show that under certain conditions, particularly when the driver is focused, the microsleep label often appears at the same time as the “focus” label, resulting in double detection (double labeling). This shortcoming can lead to unnecessary alerts. It is therefore necessary, in future research, to improve the dataset, increase the quality of annotations, and adjust the confidence threshold, particularly in extreme lighting conditions and with more varied face angles, as well as improve the consistency of annotations so that the model can better distinguish between concentration and microsleep.

REFERENCES

[1] World Health Organization, "Road Traffic Injuries," WHO. 2020 [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.

[2] A. Bochkovskiy, C-Y. et al., "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv:2004.10934v1 [cs.CV], pp. 1–17, 2020. Available: https://arxiv.org/abs/2004.10934

[3] I. Virgiawan, "Deteksi dan tracking objek secara real-time berbasis computer vision menggunakan metode YOLO V3," Humantech Jurnal Ilmu Multidisiplin Indonesia., vol. 3 No. 3, 2024, doi: 10.32670/ht.v3i3.4348.

[4] D. A. A. Indra, "Implementasi Algoritma YOLO V5 Detektor pada Pengemudi Kendaraan Bermotor Secara Real-Time," J. Ilm. KOMPUTASI, vol. 22, pp. 575–582, 2023, doi: 10.32409/jikstik.22.4.3488.

[5] Mihuandayani et al., "Implementasi YOLOv8 untuk Deteksi dan Hitung Objek Manusia dengan Convolutional Neural Network," Jurnal Buffer Informatika, vol. 11 no. 1, pp 27-35, 2025, doi: 10.25134/buffer.v11i1.356

[6] C. Chen, X. Liu, M. Zhou, Z. Li, Z. Du, and Y. Lin, "Lightweight and Real-Time Driver Fatigue Detection Based on MG-YOLOv8 with Facial Multi-Feature Fusion," J. Imaging, no. 1, pp. 1–25, 2025.

[7] M. F. Golfantara, "Penggunaan Algoritma YOLOv8 untuk Identifikasi Rempah-Rempah", Jurnal JITET, vol. 12 no. 3, pp. 3867-3873, 2024, doi: 10.23960/jitet.v12i3S1.5221.

[8] J. Jonathan and D. Hermanto, "Penentuan Epochs Hasil Model Terbaik : Studi Kasus Algoritma YOLOv8," Digital Transformation Technology, vol. 4, no. 2, pp. 792–798, 2024, doi: 10.47709/digitech.v4i2.4640

[9] E. U. Armin, A. P. Edra, F. I. Alifin, I. Sadidan, I. P. Sary, and U. Latifa, "Performa Model YOLOv8 untuk Deteksi Kondisi Mengantuk pada pengendara mobil," BRAHMANA, vol. 5, no. 1, pp. 67–76, 2023. DOI:10.30645/brahmana.v5i1.279

[10] R. S. Saputri, et al., "Deteksi Dini Microsleep pada Pengemudi berkendara Roda Empat Melalui Citra Mata Menggunakan Algoritma YOLOv8", Jurnal Processor, vol. 20 no. 2, pp. 165-179, 2025, doi: 10.33998/processor.2025.20.2.2529

[11] Saepudin, et al., "Analisis Kinerja YOLOv8 Optimalisasi, Roboflow untuk Deteksi Ekspresi Wajah Emosional dengan Machine Learning", Naratif, Vol. 6 no. 2, 2024, doi: 10.53580/naratif.v6i2.292

[12] P. V. B. Reddy and S. S. Rachana, "DRIVER DROWSINESS MONITORING," IJCRT, vol. 11, no. 6, pp. 930–934, 2023.

[13] A. I. Moorjani, "Hubungan antara rawan bosan dan kemudahan tertidur dengan microsleep saat mengemudi," Jurnal Mitra Teknik Sipil, vol. 4, no. 3, pp. 729–736, 2021. Available: https://journal.untar.ac.id/index.php/jmts/article/view/12642

[14] A. Samuel, "Two machine-learning procedures have been investigated in some detail using the game of checkers. Enough work has been done to verify the fact that a computer can be programmed so that it will learn to play a better game of checkers than can be played by ," IBM J. Res. Dev., vol. 3, no. 3, pp. 210–229, 1959, doi: 10.1147/rd.33.0210.

[15] R. Stuard, Artificial Intelligence: A Modern Approach (4th ed.). Pearson, 2020. [Online]. Available: https://aima.cs.berkeley.edu/

[16] I. Goodfellow, Deep Learning. MIT Press, 2016. [Online]. Available: https://www.deeplearningbook.org/

[17] N. J. Hayati, "Object Tracking Menggunakan Algoritma You Only Look Once (YOLO)v8 untuk Menghitung Kendaraan", Jurnal Komputa, vol. 12 no. 2, pp 91-99, 2023, doi: 10.34010/komputa.v12i2.10654

[18] J. Zophie and H. H. Triharminto, "Implemetasi Algoritma You Only Look Once (YOLO) menggunakan Web Camera untuk Mendeteksi Objek Statis dan Dinamis Implementation of You Only Look Once (YOLO) Algorithm using Web Camera for Static dan Dinamic Object Detection," Jurnal TNI Angkatan Udara, vol. 1, no. 1, pp. 98–109, 2022, [Online]. Available: https://e-jurnal.tni-au.mil.id/index.php/jpb/article/view/50/44

[19] O. Khare, "YOLOv8-Based Visual Detection of Road Hazards : Potholes , Sewer Covers , and Manholes," arXiv:2311.00073v1, 2022. [Online]. Available: https://arxiv.org/abs/2311.00073

[20] S. N. Mohamad et al., "Development of Real Time Cat Auto Feeder Dispenser Using Arduino," International Journal of Synergy Engineering Technologyl., vol. 3, no. 1, pp. 52–57, 2022. [Online]. Available: https://tatiuc.edu.my/ijset/index.php/ijset/article/download/117/99

[21] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once : Unified , Real-Time Object Detection," CVPR, 2016, doi: 10.1109/CVPR.2016.91.

[22] M. Telçeken and D. Akgun, "An Evaluation of Image Slicing and YOLO Architectures for Object Detection in UAV Images," MDPI, vol. 14, no. 23, pp. 1–16, 2024, doi: https://doi.org/10.3390/app142311293.

[23] Y. Akbar et al., "Deteksi microsleep menggunakan convolutional neural network dengan sensor esp32 pada pengemudi microsleep detection using convolutional neural network with esp32 sensor on the driver," Intecoms, vol. 7, no. 5, pp. 1668–1674, 2024, doi: 10.31539/intecoms.v7i5.12142.

[24] Ultralytics, "Dataset Preparation Documentation," Ultralytics.com. 2023. [Online]. Available: https://docs.ultralytics.com/datasets/

[25] E. Syahrudin, E. Utami, and A. D. Hartanto, "Augmentation for Accuracy Improvement of YOLOv8 in Blind Navigation," Jurnal RESTI, vol. 5, no. 158, pp. 579–588, 2026. doi:10.29207/resti.v8i4.5931

BIBLIOGRAPHY OF AUTHORS

Ria Citra Desiany Putri is an active student in the Electrical Engineering Program at Sriwijaya State Polytechnic. She is interested in embedded systems and computer vision, particularly the application of technology to improve driving safety. During her studies, she has been actively involved in various research and development projects related to microcontrollers, image processing, and sensor integration. The author continues to strive to develop her technical and research skills through academic activities and practical training in smart technology.

Dr. Ing. R.D. Kusumanto, S.T., M.M. is a Senior Lecturer in Electronics Engineering at Sriwijaya State Polytechnic, with expertise in the field of renewable energy. He has a strong commitment to the campus community, including serving as director of Sriwijaya State Polytechnic (2012–2016) and actively conducting research and community service to this day.

Dewi Permata Sari is a lecturer in the Department of Electronics Engineering at the Polytechnic of Sriwijaya, Palembang, Indonesia. She obtained her bachelor's degree in Electrical Engineering from Universitas Sriwijaya in 2001, and later completed her master's degree in High Voltage Electrical Engineering at Universitas Bina Darma in 2012. Her teaching areas include SCADA systems, automation, LabVIEW, microcontrollers, and industrial safety standards. Her research interests focus on embedded systems, IoT-based automation, and robotics for education and public service.