

Development of Fatigue Detection System using Deep Learning Model

Abd Majid Darsono^{1*}, Nur Farah Izzati Ahmad Tarmizi¹, Abd Shukur Ja'afar¹, Anuar Jaafar¹, Haziezol Helmi Mohd Yusof¹, Mohamad Harris Misran¹, Nik Mohd Zarifie Hashim¹, Muhammad Imran Ahmad²

¹Faculty of Electronics and Computer Technology and Engineering, University Technical Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.

²Faculty of Electronic Engineering & Technology, University Malaysia Perlis, Pauh Putra Campus, 02600 Arau, Perlis, Malaysia

DOI: <https://dx.doi.org/10.47772/IJRISS.2025.908000658>

Received: 14 August 2025; Revised: 21 August 2025; Accepted: 24 August 2025; Published: 25 September 2025

ABSTRACT

Fatigue is a common issue that affects attention, cognitive performance, and overall well-being, particularly in educational settings. Detecting fatigue is essential where sustained focus and alertness are key to performance and safety, such as in educational, professional, and transportation environments. Traditional methods of detecting fatigue, such as educator observation, are often subjective and ineffective in identifying early signs of fatigue, which can lead to reduced student's engagement and academic performance. This project proposes a real-time fatigue detection system capable of identifying indicators such as drowsiness and sleepiness across multiple students in a classroom using the YOLOv8 deep learning model. YOLOv8 is a highly efficient object detection model that rapidly and accurately identifies and locates objects in images and videos. The project further evaluates the system's effectiveness in terms of accuracy and real-time processing within classroom environments. Experimental results demonstrate that the system achieves 92.8% mean average precision (mAP) and 91.4% testing accuracy, outperforming models such as YOLOv5 and Faster R-CNN. By enabling early and reliable detection of fatigue, this project has the potential to significantly enhance classroom engagement and improve learning outcomes.

Keywords— Fatigue Detection System, YOLO, Deep Learning Model, Convolution Neural Network (CNN)

INTRODUCTION

In the present quick-moving age, society confronts several obstacles that require efficient administration of time and energy. The requirements of everyday existence, be it at home, in the workplace, or within educational establishments, frequently cause individuals to feel intense fatigue. Fatigue has wide-ranging effects, impacting several sectors and presenting dangers beyond decreased productivity. In professional situations, for example, weariness is a main element in workplace mishaps and injuries. Workers who lack sufficient rest may have difficulty keeping focus, boosting the probability of mistakes that can result in hazardous situations. Reports in [1] emphasize that workplace fatigue not only elevates accident rates but also harms overall output. Acknowledging and tackling fatigue is thus vital to protect both safety and wellness. By equipping people with understanding and tactics to recognize and handle fatigue, society can foster a healthier balance between physical and mental wellness while lessening the hazards linked with exhaustion.

In the education sector [2], fatigue among students has become an increasingly pressing concern. It not only hampers academic performance but can also have long-term consequences on students' future development. In today's fast-paced world, students are often required to balance academic demands, extracurricular activities, and social responsibilities, which frequently leads to both physical and mental fatigue.

This project presents an innovative solution to address this issue by utilizing deep learning technology to detect signs of fatigue in classroom settings. The system analyzes data from multiple sources to identify fatigue

indicators such as eye closure, yawning, and sleep-related behaviors (e.g., resting the head on the desk). The primary objective is to assist teachers and educators in recognizing students who may be fatigued, allowing them to take timely and appropriate measures to provide support. By identifying fatigue early, teachers can adapt lesson plans, introduce breaks, or use more interactive methods to help students regain focus.

Beyond detection, the system also aims to raise awareness of the importance of fatigue management, encouraging students to achieve a healthier balance between academic responsibilities and mental well-being. By doing so, it is expected to reduce the negative effects of fatigue while enhancing student engagement and performance in the classroom. To achieve this, a real-time fatigue detection system was developed using the YOLOv8 deep learning model. The model detects student faces and analyzes fatigue indicators with high accuracy. The system was further evaluated under varying classroom conditions, including differences in lighting and camera angles, to ensure reliability and robustness in real-world environments.

Related Works

Fatigue is a condition that encompasses both mental and physical exhaustion, which can significantly impair an individual's ability to perform productive tasks. Among students, fatigue is a widespread issue that adversely affects concentration, engagement, and learning outcomes, and may even contribute to negative emotional states [3]. Extended periods of fatigue or reduced focus can diminish academic performance by weakening information retention and lowering motivation [4]. Conventional approaches to fatigue detection—such as relying on educators' observations—are largely subjective and prone to inconsistency, often resulting in poor accuracy and effectiveness in identifying fatigue among students [5]. Such observations are influenced by personal biases, teaching experience, attentiveness, and individual interpretations of fatigue symptoms [6]. Without a reliable, automated system for real-time fatigue detection, educators are deprived of actionable insights into students' alertness levels [7].

To address this challenge, [8] proposed a method that integrates Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), and Percentage of Eye Closure (PERCLOS) to evaluate a driver's alertness through facial expressions, focusing on the eyes and mouth. EAR measures eye openness, while MAR calculates mouth openness to identify yawning, both derived from vertical and horizontal distances between facial landmarks. PERCLOS, on the other hand, quantifies the proportion of time eyes remain closed within a given interval, where prolonged closures indicate drowsiness.

With the advent of deep learning, fatigue detection has evolved from manual feature extraction to automated neural network models capable of learning directly from raw data such as images and videos. Building on this, [9] conducted a study to assess online classroom engagement by analyzing students' concentration levels. Their system employed an enhanced Multitask Cascaded Convolutional Neural Network (MTCNN) to detect facial features, particularly the eyes and mouth. A refined AlexNet model then classified eye and mouth states as open or closed, with PERCLOS and Percentage of Mouth Open Time (PMOT) used to identify fatigue indicators such as extended eye closure or yawning. In addition, head pose estimation across pitch, yaw, and roll axes captured subtle movements like nodding or tilting, providing further indicators of drowsiness or disengagement.

Similarly, [10] proposed a driver drowsiness detection system using a deep Convolutional Neural Network (CNN). The architecture consisted of three convolutional layers with 3×3 kernels, where filter numbers doubled in successive layers, followed by two max-pooling layers (2×2 stride), three fully connected dense layers, and four dropout layers with a 0.25 drop rate to mitigate overfitting. Rectified Linear Unit (ReLU) served as the activation function for convolutional layers, while softmax was applied at the output for classification. The model, trained and tested on the Kaggle "Drowsiness dataset" of 1,448 images, classified states into "yawn" and "no-yawn" with an accuracy of 96%.

In the broader context, object detection provides a computer vision approach not only to classify images but also to locate objects within them using bounding boxes. Leveraging this, [11] introduced an eye detection model based on Faster RCNN (Region-based Convolutional Neural Network) for applications such as fatigue monitoring, activity recognition, liveness detection, and gaze estimation. Addressing challenges like pose and scale variation, face rotation, and occlusion, preprocessing steps such as contrast stretching and data

augmentation enhanced model robustness. The Faster RCNN, using ResNet-101 as its backbone, achieved accuracies of 98.32% and 98.11% on AR and GI4E datasets respectively, with computation times of 0.52 ms per image. These results demonstrated higher accuracy and efficiency compared to existing state-of-the-art models. The study further recommended exploring single-stage detectors like YOLO to optimize computation time for real-time applications.

Extending this line of research to education, [12] developed a real-time fatigue detection system for classroom settings using webcam-based monitoring. By adopting the YOLOv8 deep learning model, the system leveraged anchor-free methods, decoupled head architecture, and Distribution Focal Loss to improve detection accuracy. A dataset of 3,175 images (categorized as “awake” or “drowsy”), augmented through eight strategies and labeled with Roboflow, was used for training and testing on Google Colab. The system achieved a high mean average precision (mAP50) of 97% and overall accuracy of 85%, successfully identifying drowsy students in real time and providing visual alerts to assist teachers in sustaining engagement.

In another study, [13] focused on enhancing traffic safety by designing a non-intrusive, real-time drowsiness detection system for drivers based on facial features such as yawning and eye closure. Three benchmark datasets—NTHUDD, YawDD, and UTA-RLDD—were used for model training and testing. YOLOv5 and YOLOv8 produced outstanding results, achieving 100% precision and recall with a mAP@0.5 of 99.5% on UTA-RLDD. Additionally, the K-Nearest Neighbor (KNN) classifier achieved an accuracy of 98.89%, precision of 99.27%, and an F1 score of 98.86%.

Collectively, these studies highlight the transition from traditional observational methods to advanced deep learning-based approaches in fatigue detection. The findings demonstrate the effectiveness of neural network models in accurately identifying fatigue indicators, thereby enhancing safety and improving engagement across domains ranging from driver monitoring to classroom learning.

METHODOLOGY

The purpose of this study is to develop a real-time fatigue detection system for classroom environments using the YOLOv8 deep learning model for object detection. This section outlines the key stages of the model development process, which include data collection, preprocessing and augmentation, dataset splitting, model training, and performance testing.

Data Collection

In this phase, the dataset is compiled from a publicly available source, *Drowsiness Detection* on Roboflow, as well as a custom dataset obtained from the Internet. As shown in Fig. 1, dataset consists of facial images that display fatigue-related features, such as drowsiness (e.g., eye closure or yawning) and sleep (e.g., resting the head on a desk). These images are utilized for both training and testing the proposed models.



Fig. 1. Dataset collected from Roboflow and custom from the internet

Data Annotation, Preprocessing and Augmentation.

The next step involves processing the custom dataset through image annotation. For this project, the annotation process was carried out using the Roboflow platform. The collected images were uploaded to Roboflow, where each image was labeled and categorized into three classes: *Awake*, *Drowsy*, and *Sleep*. All images were resized to 640×640 pixels with a 'fit' adjustment (black padding on edges). To enhance dataset diversity and robustness, seven augmentation techniques were applied which were flip, rotation, brightness adjustment, blur, cropping, hue modification, and noise injection. This will be resulting in a total of 1,257 images across the three categories. The dataset was then divided into training (80%), validation (10%), and testing (10%) subsets. When combined with the publicly available dataset containing 5,166 images, the final dataset comprised 7,865 images.

The Architecture of YOLOv8

The proposed system is developed using the YOLOv8 deep learning model architecture. This architecture comprises several key components such as *backbone*, *neck*, and *head* layers which collectively demonstrate improvements over earlier YOLO versions, as illustrated in Fig. 2.

The *backbone* functions as the primary feature extractor from input images. YOLOv8 employs Cross Stage Partial Darknet (CSPDarknet), a highly efficient extractor that captures crucial visual patterns such as edges, shapes, and textures. As for *neck*, in order to enhance feature fusion across different scales, YOLOv8 integrates the Path Aggregation Network (PANet). This is particularly beneficial in classroom environments, where facial features may appear at varying sizes and positions. The *head* layer produces the final predictions, including bounding boxes, class labels, and confidence scores. Unlike previous versions, YOLOv8 employs an anchor-free mechanism, allowing it to directly predict bounding boxes and class probabilities with improved accuracy and efficiency.

Compared to earlier anchor-based versions such as YOLOv5, YOLOv8's anchor-free design reduces computational complexity, improves detection speed, and enhances performance for objects with diverse shapes and scales—making it more suitable for real-time fatigue detection in dynamic classroom settings.

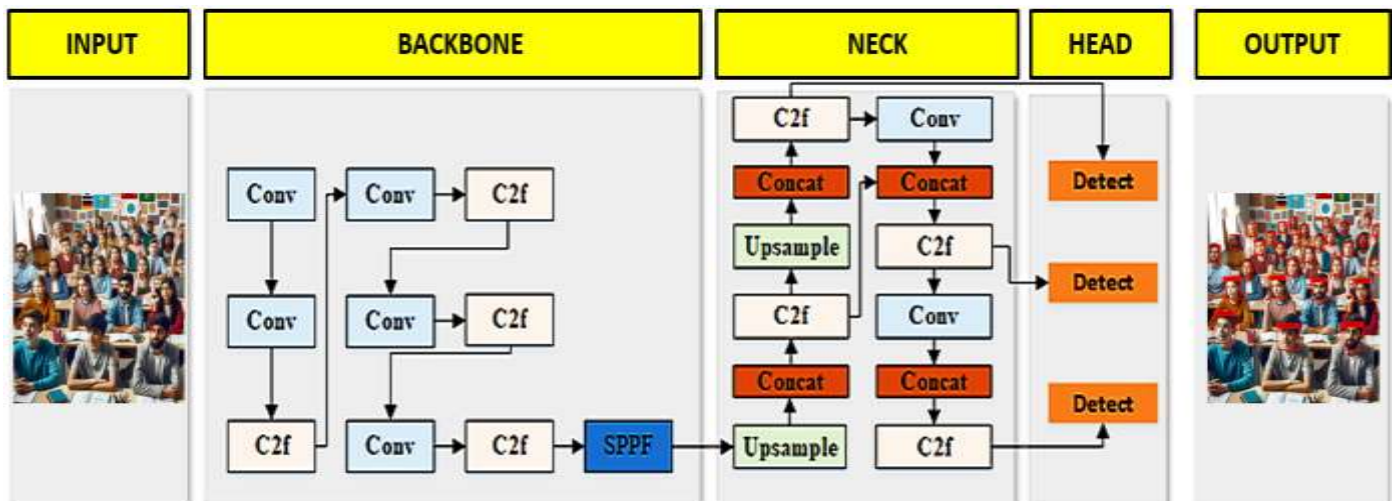


Fig. 2. YOLOv8 Model Architecture

Proposed fatigue detection system

The training and testing phases of the proposed system were implemented using Google Colab in conjunction with Google Drive. To optimize model performance and identify the best hyperparameters, cross-validation was employed. The dataset was divided into three folds, with training performed on two folds and testing on the remaining fold. This process was repeated until all three folds were used for training. The hyperparameter configuration that achieved the highest mean Average Precision (mAP) was then selected for the final training, which was conducted over 100 iterations on the entire dataset.

Using the YOLOv8 model, the system can detect and classify students into three categories: *Awake*, *Drowsy*, and *Sleep*. Model adjustments were made iteratively; if the performance metrics met the required thresholds, the process advanced to the next stage. Otherwise, the model was retrained with tuned hyperparameters. In the evaluation phase, the trained models were assessed using mAP, accuracy, precision, recall, and overall performance indicators.

The proposed system is designed for real-time monitoring and is integrated into a simple Graphical User Interface (GUI) called the Fatigue Detection System. This interface uses a live video stream, such as a webcam, to detect and classify student fatigue in small classroom settings. As illustrated in Fig. 3, bounding boxes of different colors are displayed to indicate the detected states: blue for *Awake*, green for *Drowsy*, and white for *Sleep*. Additionally, if more than 50% of students in the classroom are classified as drowsy or asleep, an alert notification is automatically triggered, allowing educators to quickly assess and respond to student engagement levels.

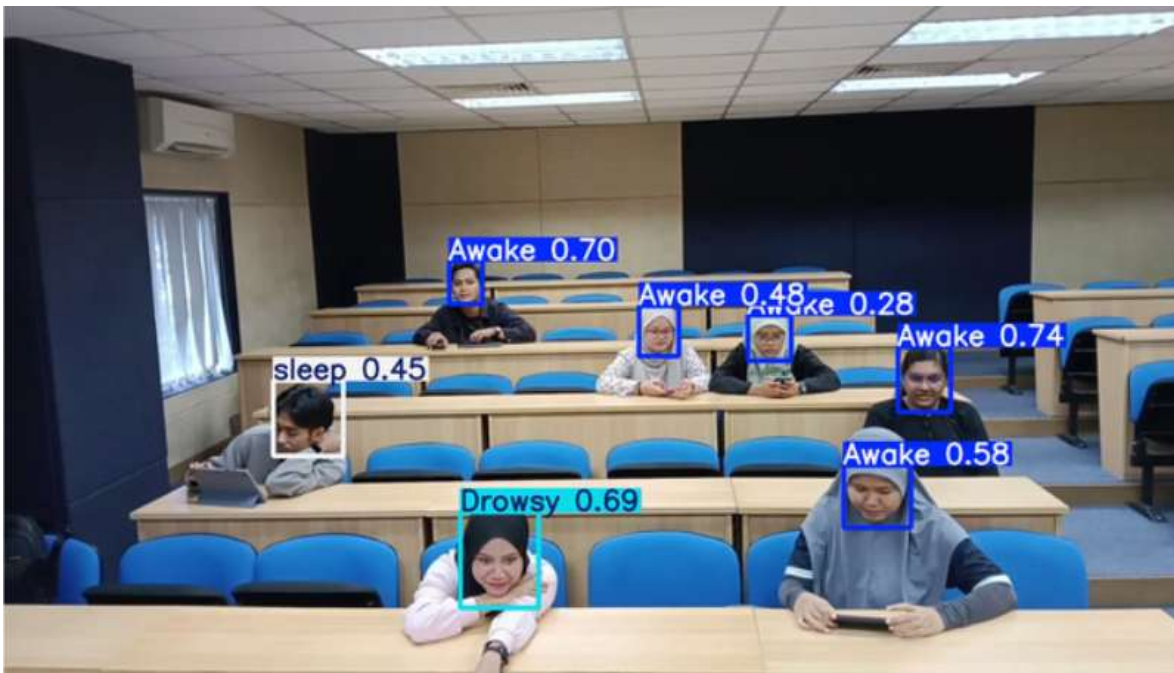


Fig. 1. Webcam page interface 1

Performance Evaluation

In the training phase of fatigue detection system, the model is set to 150 iterations (epochs) and evaluation metrics, such as mAP (mean Average Precision), precision, and recall are used to evaluate the performance of the models. These metrics are computed as follows:

$$mAP = \frac{1}{m} AP = \frac{1}{m} \int_0^1 P(R) dR \times 100\% \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

where:

True Positive (TP): The number of samples that were correctly predicted as positive.

False Positive (FP): The number of samples that were incorrectly predicted as positive.

False Negative (FN): The number of samples that were incorrectly predicted as negative.

RESULTS AND ANALYSIS

This section presents a detailed analysis, visualization, and evaluation of the simulation results and performance of the proposed fatigue detection system. The process begins with fine-tuning key training hyperparameters, such as the learning rate and batch size. Subsequently, the results are examined to assess the model's effectiveness in accurately detecting signs of fatigue specifically, drowsiness (e.g., eye closure and yawning) and sleep (e.g., resting the head on a desk) within real-time classroom environments.

Hyperparameter Tuning

Type of Hyperparameters	Hyperparameter Values
Batch size	8, 16, 32
Learning rate	0.1, 0.01, 0.001
Loss functions	Cross-Entropy (CE)
Weight decay	0.0001

Hyperparameter Tuning Results

The training of various hyperparameters was conducted using Google Colab. Two key hyperparameters were considered in this study: the learning rate and batch size. For each parameter, three different values were tested, as summarized in Table 1. The dataset was divided into K folds (with $K = 3$) for cross-validation. In each iteration, K-1 folds were used to train the model, while the remaining fold was used for validation. This process was repeated three times, ensuring that each fold served once as the validation set. Each split consisted of both training and validation subsets containing images and their corresponding labels. Finally, the performance metrics from all three iterations were averaged to provide a more reliable and stable estimate of the model's generalization ability.

Analyzing Various Batch Size and Learning Rate

Batch Size	Learning Rate	Precision	Recall	mAP50 (%)
8	0.1	88.3	84.63	92.1
	0.01	89.3	88.27	94.1
	0.001	86.5	85.53	91.4
16	0.1	89.73	85.0	92.3
	0.01	91.15	90.56	95.13
	0.001	86.53	87.77	92.38
32	0.1	88.1	86.73	92.8
	0.01	91.7	89.68	95.12
	0.001	87.47	87.36	92.7

Next, we examine the impact of different batch sizes on the performance of the proposed YOLOv8-nano model, as summarized in Table 2. Among the tested configurations, a batch size of 16 with a learning rate of 0.01 yielded the best results, achieving an overall mAP@50 of 95.13% and mAP@50–95 of 70.66%. This configuration also demonstrated high precision (91.15%) and the highest recall (90.56%) compared to all other parameter settings. The superior performance of batch size 16 suggests that it provides an effective balance between gradient stability and model generalization, enabling the network to learn more robust feature representations.

Figure 4 presents the validation results of the model's precision and mAP at an IoU threshold of 0.5 across

training epochs. Both metrics show a consistent upward trend before stabilizing within the range of 0.8 to 1.0. The training process was terminated early at Epoch 54, as no further improvement was observed over the last 10 epochs. Overall, the graph indicates that the model achieved strong performance, with high final scores in both precision and mAP@0.50, demonstrating that effective learning was successfully accomplished during training. The use of early stopping also helped to prevent overfitting, thereby enhancing the model's ability to generalize well to unseen data.

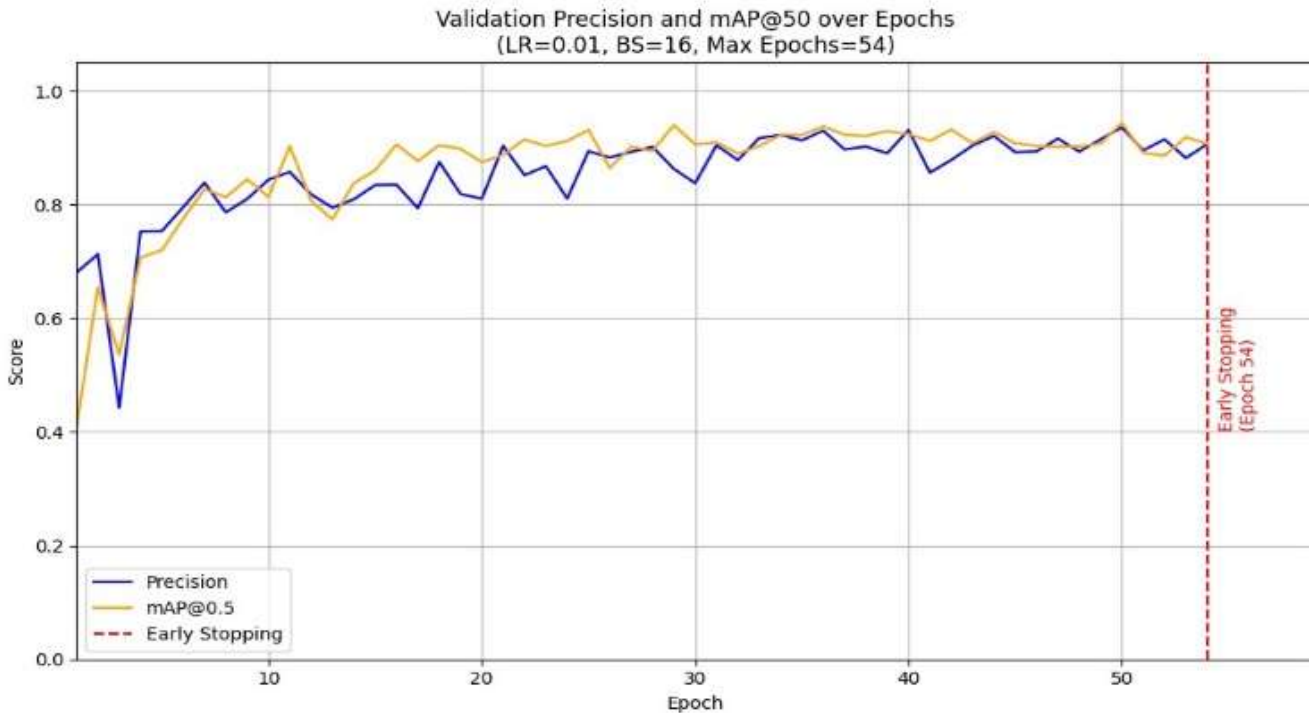


Fig. 2. Graph of Validation Precision and mAP@50 over Epochs

Analysis of Model Training Techniques

As shown in Table 3, the training progress of four different methods for fatigue detection is compared. These methods include the proposed YOLOv8-based model, YOLOv5, and Faster R-CNN, evaluated using standard performance metrics.

Comparison Between Proposed YOLOv8 Model with YOLOV5 and Faster R-CNN

Model	Parameter	Indicators Class		
		Awake	Drowsy	Sleep
YOLO-v8n	Precision	95.1	94.5	86.6
	Recall	95.5	93.4	81.2
	mAP50 (%)	92.8%		
YOLO-v5n	Precision	94.1	91.2	77.4
	Recall	92.9	88.4	81.2
	mAP50 (%)	91.3%		
Faster R-CNN	Precision	37.74		
	Recall	15.58		
	mAP50 (%)	38.06%		

Table 3 compares the performance of the proposed YOLOv8n model with YOLOv5n and Faster R-CNN using standard evaluation metrics. The YOLOv8n model demonstrates superior precision and recall across all categories, achieving 95.1% precision and 95.5% recall for *Awake*, 94.5% precision and 93.4% recall for *Drowsy*, and 86.6% precision and 81.2% recall for *Sleep*. This results in a mean Average Precision at 50% Intersection over Union (mAP@50) of 92.8%. In contrast, the YOLOv5n model shows slightly lower performance, with precision and recall of 94.1% and 92.9% for *Awake*, 91.2% and 88.4% for *Drowsy*, and 77.4% and 81.2% for *Sleep*, yielding an overall mAP@50 of 91.3%. Meanwhile, the Faster R-CNN model performs considerably worse, with only 37.74% precision, 15.58% recall, and a mAP@50 of 38.06%. These results highlight YOLOv8n as the most effective model for accurately detecting different states of fatigue, clearly surpassing YOLOv5n and Faster R-CNN, and confirming its suitability for real-time, high-accuracy object detection applications. The superior performance of YOLOv8n can be attributed to its anchor-free detection mechanism and enhanced feature extraction through CSPDarknet and PANet, which allow for more robust recognition of facial features across varying scales and positions.

Experimental Results in Real-time Situation

The real-time simulation experiments were conducted in small classroom environments using Visual Studio Code (VS Code). Examples of the real-time output are illustrated in Figure 6, where the system detects and classifies students into *Awake*, *Drowsy*, and *Sleep* states, along with the corresponding percentage distribution displayed in a bar chart within the GUI as shown in Fig. 5. The interface also provides a 'Use Live Webcam' button, which, when activated, automatically displays the live camera feed on the left side and the bar graph on the right side.

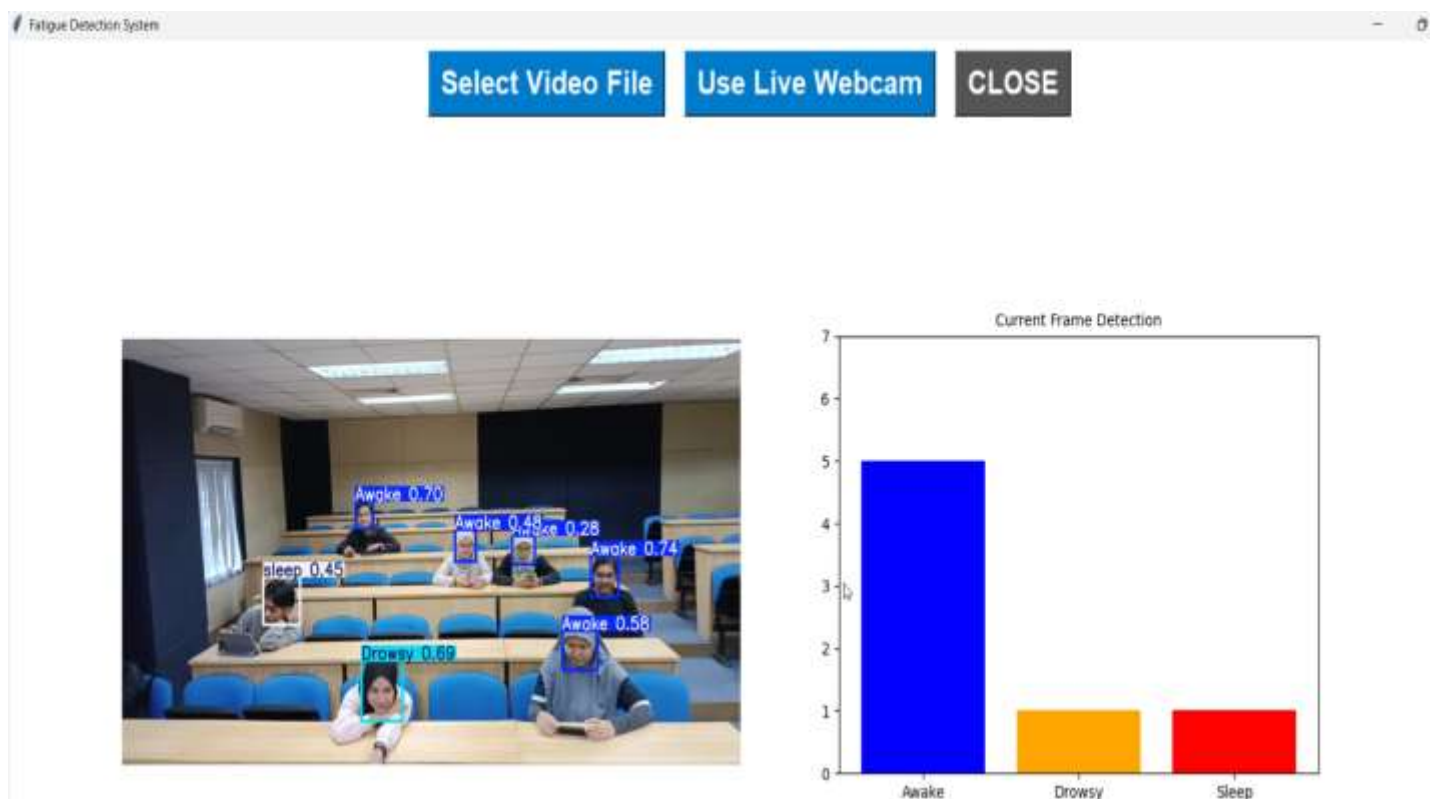


Fig. 5. Real-time Simulation Output 1

In the GUI, the y-axis represents the number of students in the class, while the x-axis corresponds to the detected fatigue states of students' faces. The bar chart is color-coded with blue representing *Awake*, orange representing *Drowsy*, and red representing *Sleep*. As illustrated in Fig. 6, the system also generates alert notifications when more than 50% of students are classified as *Drowsy* or *Sleep*. In such cases, an alert message '**ALERT: 50% of students are drowsy and sleeping!**' is displayed on the interface. This feature provides educators with an immediate visual cue, enabling them to respond proactively and re-engage students during teaching sessions. Ultimately, this mechanism supports the broader objective of enhancing classroom interaction, improving student attentiveness, and fostering more effective learning outcomes

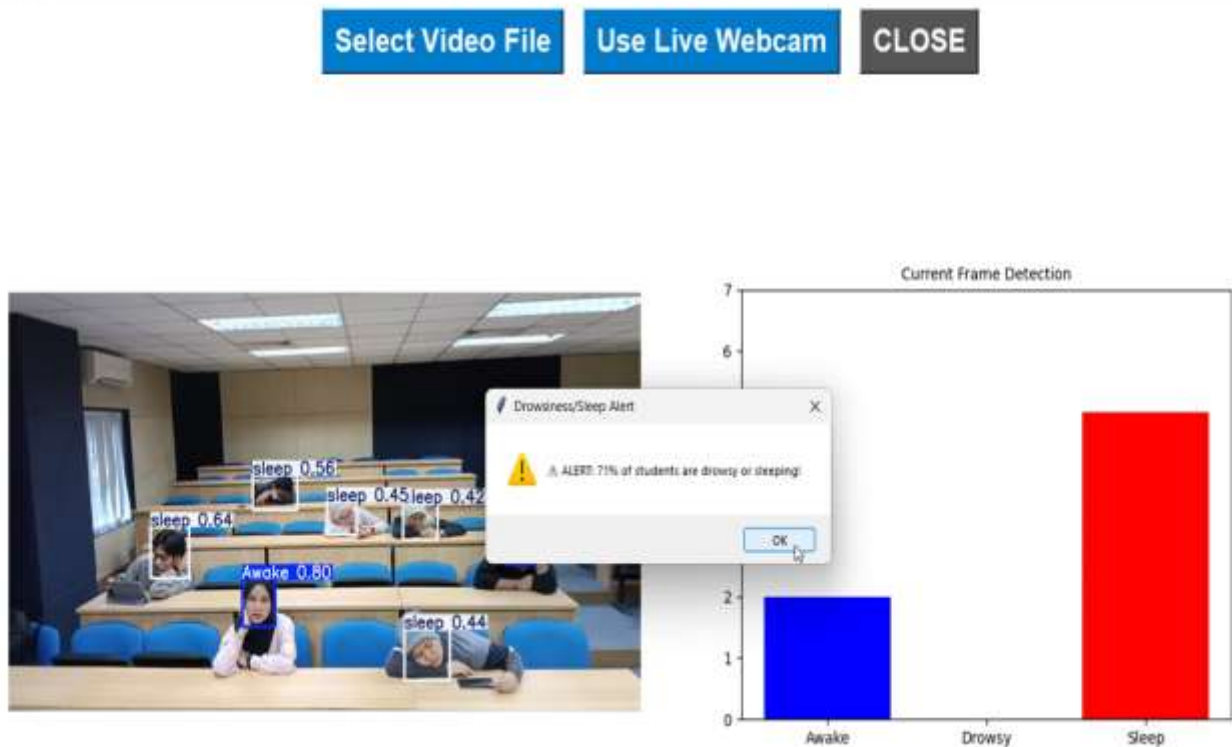


Fig. 6. Real-time Simulation Output 1

CONCLUSION

This project presents a real-time drowsiness detection system successfully implemented using the proposed YOLOv8 deep learning model to identify fatigue indicators such as drowsiness (e.g., eye closure and yawning) and sleep (e.g., resting the head on a desk) among students in a classroom. The system effectively detected multiple faces, analyzed the percentage of fatigue, and presented the results in graphical form, thereby providing valuable feedback to educators. The implementation proved to be both responsive and accurate, fulfilling the primary objectives established at the outset of this research. Overall, this project demonstrates the potential of deep learning for monitoring student behavior and enhancing academic engagement through the early identification of fatigue.

ACKNOWLEDGEMENT

The authors would like to express our gratitude to Faculty of Electronics and Computer Technology and Engineering (FTKEK) and Centre for Research and Innovation Management (CRIM) at University Technical Malaysia Melaka (UTeM) for their assistance in acquiring the essential information and resources for the successful completion of the research.

REFERENCES

1. T. R. Cunningham, R. J. Guerin, J. Ferguson, and J. Cavallari, "Work-related fatigue: A hazard for workers experiencing disproportionate occupational risks," *Am J Ind Med*, vol. 65, no. 11, pp. 913–925, Nov. 2022, doi: 10.1002/ajim.23325.
2. M. C. Pascoe, S. E. Hetrick, and A. G. Parker, "The impact of stress on students in secondary school and higher education," Jan. 02, 2020, Routledge. doi: 10.1080/02673843.2019.1596823.
3. J. Jeric et al., "Academic Performance, Fatigue, and Stress in Online Learning Among College Students: A Multivariate Analysis."
4. K. Okano, J. R. Kaczmarzyk, N. Dave, J. D. E. Gabrieli, and J. C. Grossman, "Sleep quality, duration, and consistency are associated with better academic performance in college students," *NPJ Sci Learn*, vol. 4, no. 1, Dec. 2019, doi: 10.1038/s41539-019-0055-z.

5. R. Rika Puspita, S. I. Banjaragung Rengel Tuban, J. Timur, and S. I. Karangtinoto Rengel Tuban, "Teacher Observation Assessment For Primary Education," vol. 11, no. 2, 2020, [Online]. Available: <http://journal.upgris.ac.id/index.php/eternal/index>
6. A. Greenhouse-Tucknott, J. B. Butterworth, J. G. Wrightson, N. A. Harrison, and J. Dekerle, "Effect of the subjective intensity of fatigue and interoception on perceptual regulation and performance during sustained physical activity," *PLoS One*, vol. 17, no. 1 January, Jan. 2022, doi: 10.1371/journal.pone.0262303.
7. K. M. Torres and A. Statti, "What Can Data Tell Us?," *International Journal of Curriculum Development and Learning Measurement*, vol. 4, no. 1, pp. 1–13, Mar. 2023, doi: 10.4018/ijcdlm.320219.
8. A. Chaurasiya, S. Sonsale, R. Daga, and A. Patankar, "Driver Drowsiness Detection System by Measuring EAR and MAR," *International Research Journal of Engineering and Technology*, 2021, [Online]. Available: www.irjet.net
9. F. Yuan and Y. Nie, "Online Classroom Teaching Quality Evaluation System Based on Facial Feature Recognition," *Sci Program*, vol. 2021, 2021, doi: 10.1155/2021/7374846.
10. L. Zhou, S. Li, and Y. Wang, "Fatigue Detection and Early Warning System for Drivers Based on Deep Learning," in *2023 IEEE 3rd International Conference on Data Science and Computer Application, ICDSCA 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 1348–1351. doi: 10.1109/ICDSCA59871.2023.10392792..
11. N. Ahmad, M. K. Anish, S. A. Barlaskar, K. S. Yadav, R. Laskar, and A. Hossain, "Eye detection using Faster-RCNN," in *2022 IEEE Region 10 Symposium, TENSYP 2022*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/TENSYP54529.2022.9864431.
12. N. H. H. M. Hamidi, N. A. Z. Abidin, R. Aminuddin, C. C. Sheng, K. A. F. A. Samah, and S. D. N. M. Nasir, "A Real-Time System for Monitoring Student Drowsiness in the Classroom Using the Deep Learning Model YOLOv8," in *2024 5th International Conference on Artificial Intelligence and Data Sciences, AiDAS 2024 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 105–110. doi: 10.1109/AiDAS63860.2024.10730451.
13. S. Essahraui et al., "Real-Time Driver Drowsiness Detection Using Facial Analysis and Machine Learning Techniques," *Sensors*, vol. 25, no. 3, Feb. 2025, doi: 10.3390/s25030812.