RAJALAKSHMI ENGINEERING COLLEGE

Reviva - Real Time Yawning Detection for Driver Distraction Monitoring using Deep Learning and IoT

Team Members:

- Sriram K (221501144),
- Sudharsan S (221501149),
- Tamilarasan D (221501157)

Supervisor Name:

- Ms. J. Jasmine Gabriel M.E.,(Ph.D).,

Department of Artificial Intelligence & Machine Learning

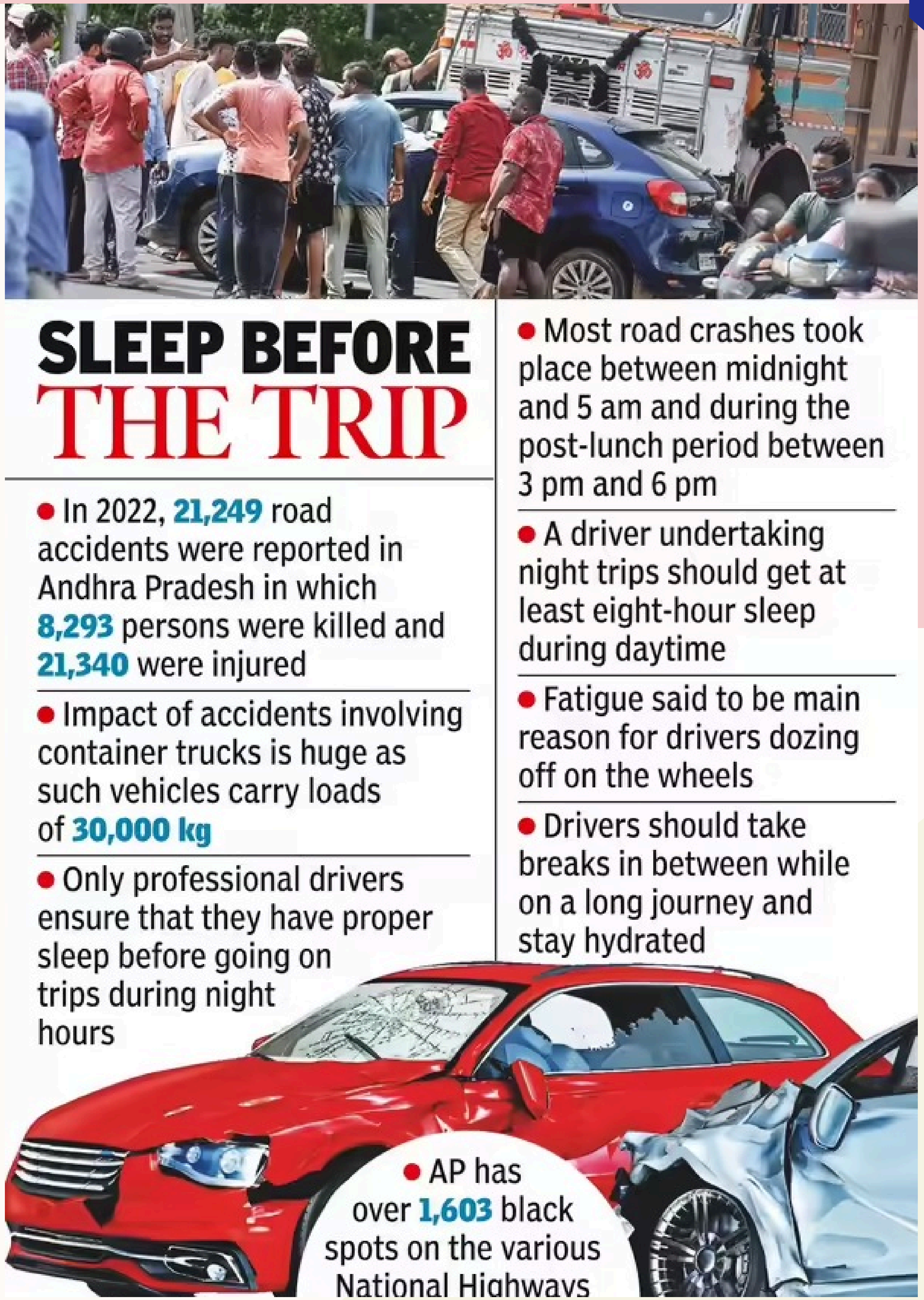
AI19711 – Phase I Project

PROBLEM STATEMENT

In today's transportation environment, driver fatigue has become a major cause of road accidents and fatalities worldwide. Long driving hours and lack of rest often lead to drowsiness, with yawning serving as one of the earliest and most noticeable indicators. However, such signs of fatigue often go unnoticed until they result in dangerous lapses in attention. Traditional monitoring systems either rely on manual observation or complex setups that are costly and inefficient for real-time use. There is a growing need for an intelligent, automated system that can continuously monitor driver alertness, detect early signs of drowsiness such as yawning, and help prevent fatigue-related accidents through timely intervention.

WHY THIS PROJECT?

- Drowsy and distracted driving is a major cause of increasing road accidents.
- Yawning is an early and visible sign of fatigue that can help detect driver distraction.
- Combining Deep Learning with IoT edge devices offers a smart, real-time solution for continuous driver monitoring and early alerts.
- This project aims to improve driver safety and support the development of intelligent transportation systems.



OBJECTIVES

- Design an AI-driven vision system capable of identifying yawning behavior through real-time image processing.
- Integrate an embedded platform with the detection model to enable continuous facial monitoring.
- Implement an automated alert mechanism that responds instantly when yawning is detected.
- Evaluate system performance based on detection accuracy, responsiveness, and consistency under different conditions.
- Develop a compact and efficient framework suitable for real-time deployment in practical environments.

ABSTRACT

Driver fatigue is one of the main causes of road accidents, so detecting it early is crucial for improving road safety. This project presents an intelligent system that identifies yawning, which is a clear sign of drowsiness, using continuous video monitoring and deep learning techniques. A live camera captures the driver's facial expressions, and a trained model processes the video in real time to detect yawning patterns. The system runs entirely on local hardware, ensuring quick responses without depending on cloud services. Designed to be lightweight, cost-effective, and suitable for real-world driving conditions, this solution helps improve road safety through smart and reliable driver monitoring.

LITERATURE REVIEW - 1

S.No	Paper Title	Author & Year	Dataset	Methodology	Inference	Limitations
1	Bias Remediation in Driver Drowsiness Detection Systems Using Generative Adversarial	Mkhuseli Ngxande, Jules-Raymond Tapamo, Michael Burke (2020)	NTHU-drowsy, DROZY, CEW + additional African faces	ResNet CNN with GAN-based targeted augmentation to reduce ethnic bias	Accuracy : 85%	Requires some training data for all ethnic groups, residual bias
2	Multimodal System to Detect Driver Fatigue Using EEG, Gyroscope, and Image Processing	Naveen Senniappan Karuppusamy, Bo-Yeong Kang (2020)	Own data from 5 subjects in driving simulator	EEG + gyroscope + vision modules; deep neural networks (LSTM + NN), multimodal fusion	Combined system accuracy : 93.91%; unimodal modules : 98-99%	Small dataset, need more real-world testing, variability in individuals
3	Early Identification and Detection of Driver Drowsiness by Hybrid Machine Learning	Ayman Altameem, Ankit Kumar, Ramesh Chandra Poonia, Sandeep Kumar, Abdul Khader Jilani Saudagar (2021)	Own and public datasets, 50 subjects	SVM with image segmentation, facial landmarks, multiple facial expressions	Accuracy : 83-95%	Reduced accuracy in low light, camera distance impacts detection
4	Privacy-Preserving Federated Transfer Learning for Driver	Linlin Zhang, Hideo Saito, Liang Yang, Jiajie Wu (2022)	NTHU-DDD and YAWDD driver drowsiness video	Federated transfer learning with encrypted communication, CNN	Accuracy : 83.5%,	Still communication overhead, non-IID data challenges
5	IoT-Based Non-Intrusive Automated Driver Drowsiness Monitoring Framework for Logistics and Public Transport Applications	M. Adil Khan, Tahir Nawaz, Umar S. Khan, Amir Hamza, Nasir Rashid (2023)	Custom + public datasets, 50 subjects	Edge computing + facial landmark detection (EAR, MAR, ENED) with cloud backend and Android app	Accuracy : 90%, real-time alerts with 100% success	Landmark detection affected by glasses/reflections

LITERATURE REVIEW - 2

S.No	Paper Title	Author & Year	Dataset	Methodology	Inference	Limitations
6	DrowsyDetectNet: Driver Drowsiness Detection Using Lightweight CNN With Limited Training Data	Madduri Venkateswarlu, Venkata Rami Reddy Ch (2024)	Dataset-1 (324 images categorized into open/closed eyes) and Dataset-2 (Kaggle YawnEyeDatasetNew with 1452 images labeled open/closed eyes)	Shallow CNN on eye regions extracted via 68-point landmarks; transfer learning comparisons	Accuracy : 99.2%	Limited datasets, less tested in night conditions
7	Biosignals Monitoring for Driver Drowsiness Detection Using Deep Neural Networks	Jose Alguindigue, Amandeep Singh, Apurva Narayan, Siby Samuel (2024)	Simulated driving, 30 participants (HRV, EDA, eye data)	SNN for HRV, 1D-CNN for EDA, CRNN for eye tracking, multisensor fusion	HRV model achieved precision 98.28%, recall 98%, F1-score 98%	Eye-based data imbalance, limited real-world variability
8	A Real-Time Vision Transformers-Based System for Enhanced Driver Drowsiness Detection and Vehicle	Anwar Jarndal, Hissam Tawfik, Ali I. Siam, Imad Alsyouf, Ali Cheaitou (2024)	MRL Eye dataset, NTHU-DDD, CEW	Vision Transformers (ViT, Swin Transformer) + fine-tuned transfer learning CNNs	Accuracy : 99%	Needs large datasets, complex model
9	Deep Learning-Based Drowsiness Detection System for Driver's Safety	Sindhu Vidyanathan Dixith, Shrikant Jadhav, Youngsoo Kim, Naveenkumar Jayakumar (2025)	NTHU-DDD, CEW	CNN-based facial expression extraction, deep classifiers, multi-feature fusion	Accuracy : 94%	High computing cost, privacy concerns
10	Deep Learning-Based Drowsiness Detection System for Driver's Safety	Dixith et al. (2025)	Kaggle four-class dataset, MRL eye dataset	Hybrid CNN + SVM classifier with EAR and MAR features	Accuracy : 99%	High computational cost; limited dataset diversity

INSIGHTS FROM PREVIOUS STUDIES

- Various studies explored AI and deep learning for detecting driver fatigue and yawning.
- Common methods include CNN, SVM, LSTM, and Vision Transformers for facial and eye-based analysis.
- Datasets such as NTHU-DDD, YAWDD, MRL Eye, and CEW are widely used for training and evaluation.
- Most systems achieve high accuracy (85 - 99%), proving the reliability of AI-based visual detection.
- However, limitations include low-light sensitivity, dataset imbalance, and high computational cost.
- Recent research trends focus on real-time performance, lightweight models, and IoT integration for alerts.

SCOPE

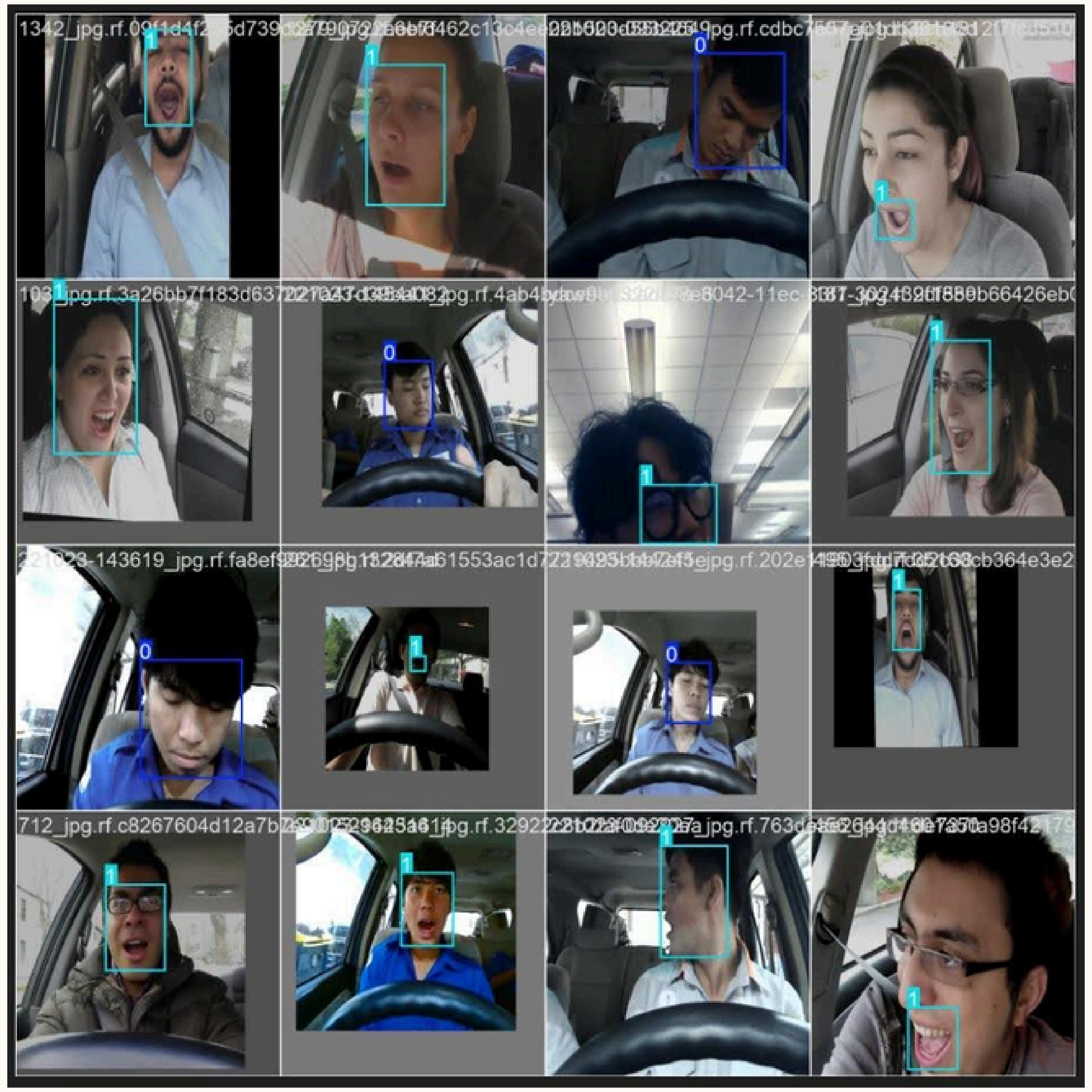
1. Detects yawning in real-time using AI and image processing.
2. Monitors facial activity through an embedded system.
3. Gives instant alert when yawning is detected.
4. Designed to be simple, efficient, and easy to deploy.

LIMITATIONS

1. Works best in good lighting conditions.
2. Detection may drop if the face is not clear or partially covered.
3. Focuses only on yawning, not other facial expressions.
4. Requires fixed camera position for accuracy.

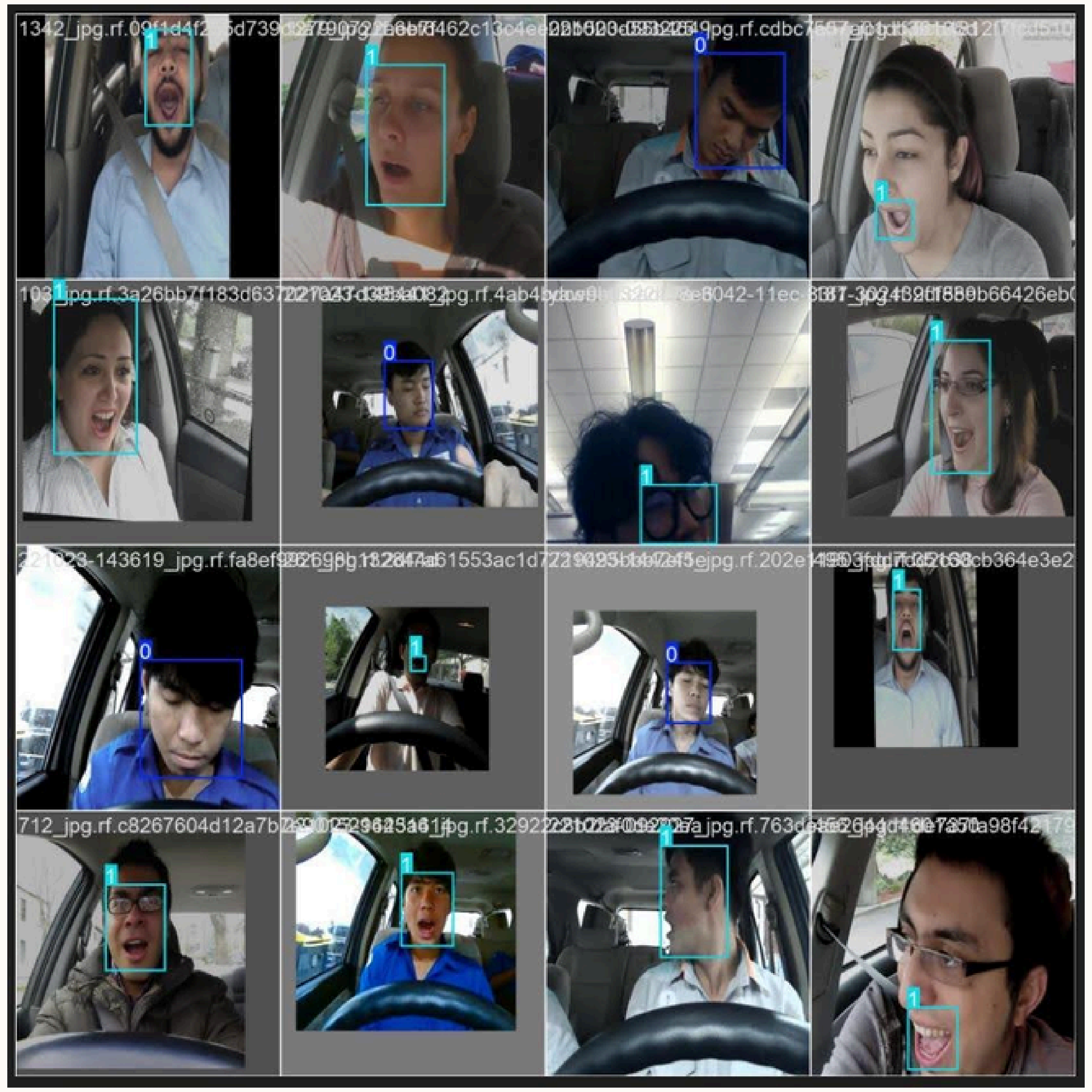
DATASET DESCRIPTION

- The dataset contains 15,533 labeled images specifically designed for yawning detection.
- Every image is annotated with bounding boxes and class labels, making it ready for training detection and classification models.
- The dataset focuses on yawn-related facial expressions, mainly for applications like driver drowsiness and distraction monitoring.
- It includes multiple individuals with different backgrounds, lighting conditions, face angles, and expressions, improving model robustness in real-world environments.
- The diversity of the dataset helps the model generalize well beyond controlled settings, ensuring accurate performance in live video scenarios.

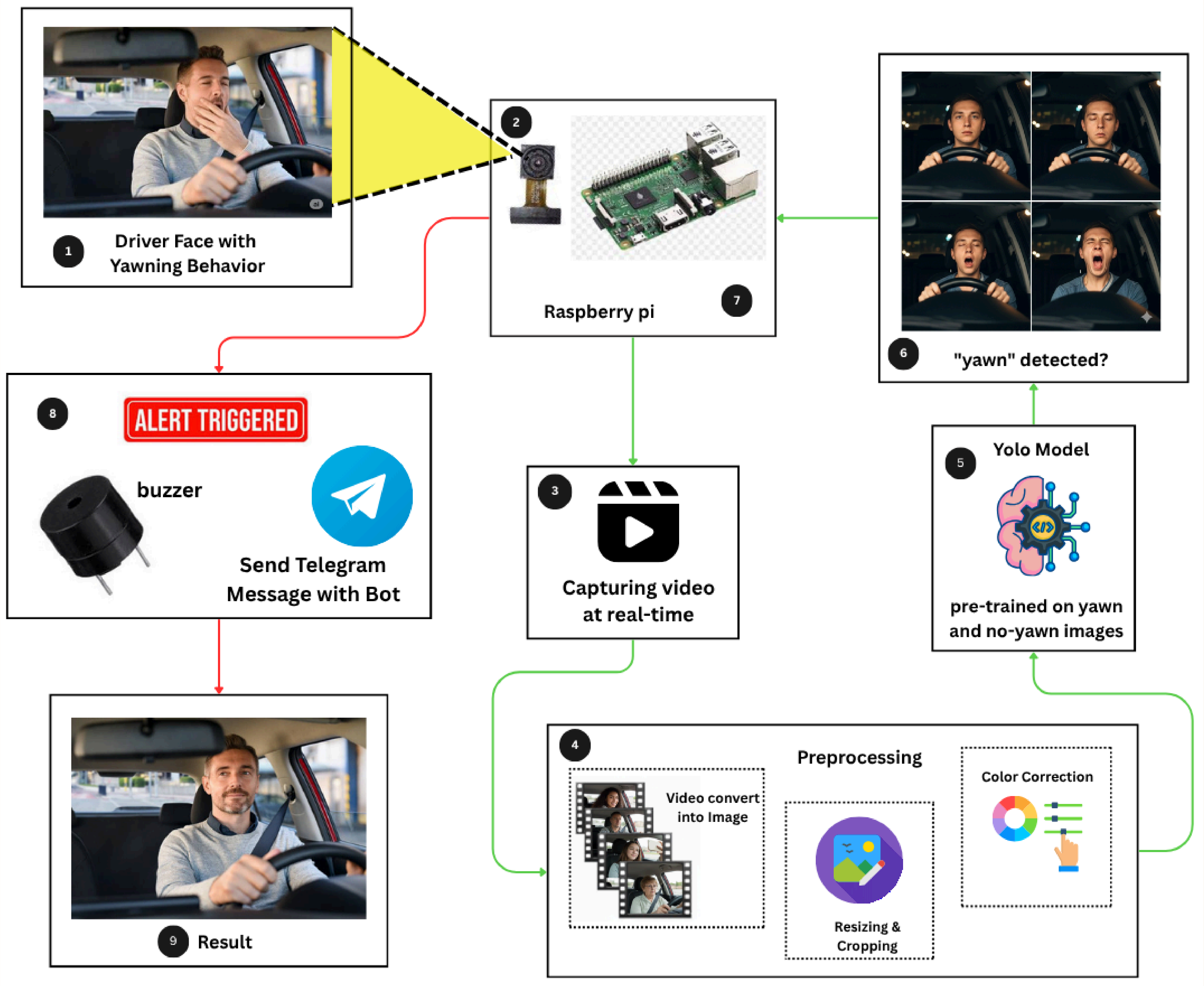


TRAIN-TEST SPLIT

- Our dataset from Roboflow was divided into three parts to train and evaluate the yawning detection model:
- 70% – Training set 10,873 images
- Used to train the YOLO model to learn yawning vs non-yawning patterns.
- 20% – Validation set 3,106 images
- Used during training to tune hyperparameters and prevent overfitting.
- 10% – Test set (1,554 images)
- Used after training to evaluate the final model performance on unseen data.
- This split ensures balanced learning, model stability, and accurate real-world prediction.



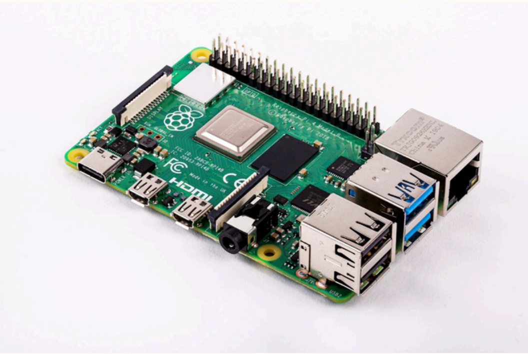
REVIVA ARCHITECTURE DIAGRAM



TOOLS AND TECHNOLOGIES USED

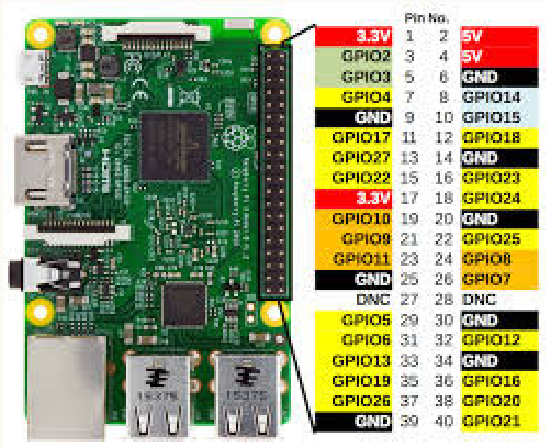
Hardware Requirements:

- Raspberry Pi 4 (main controller)
- USB Webcam (to capture video)
- Buzzer (to give alert sound)
- SD Card (16 GB or more)
- Power Adapter (5V / 3A)
- Monitor (optional for testing)



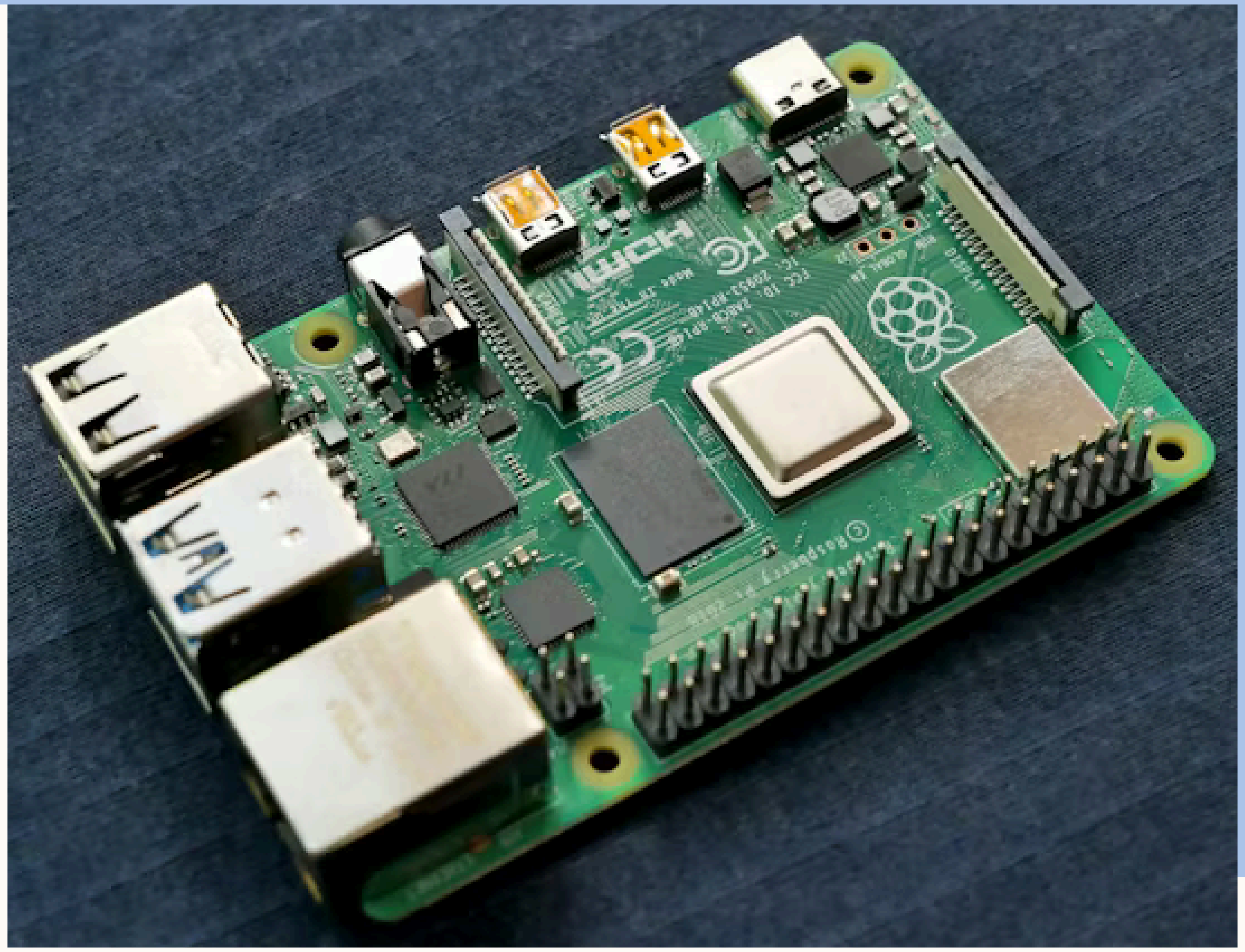
Software Requirements:

- Python 3
- OpenCV – for image and video processing
- NumPy – for calculations
- GPIO Library – to control buzzer
- Termius – to transfer files remotely to raspberry pi
- VS Code – for coding and debugging

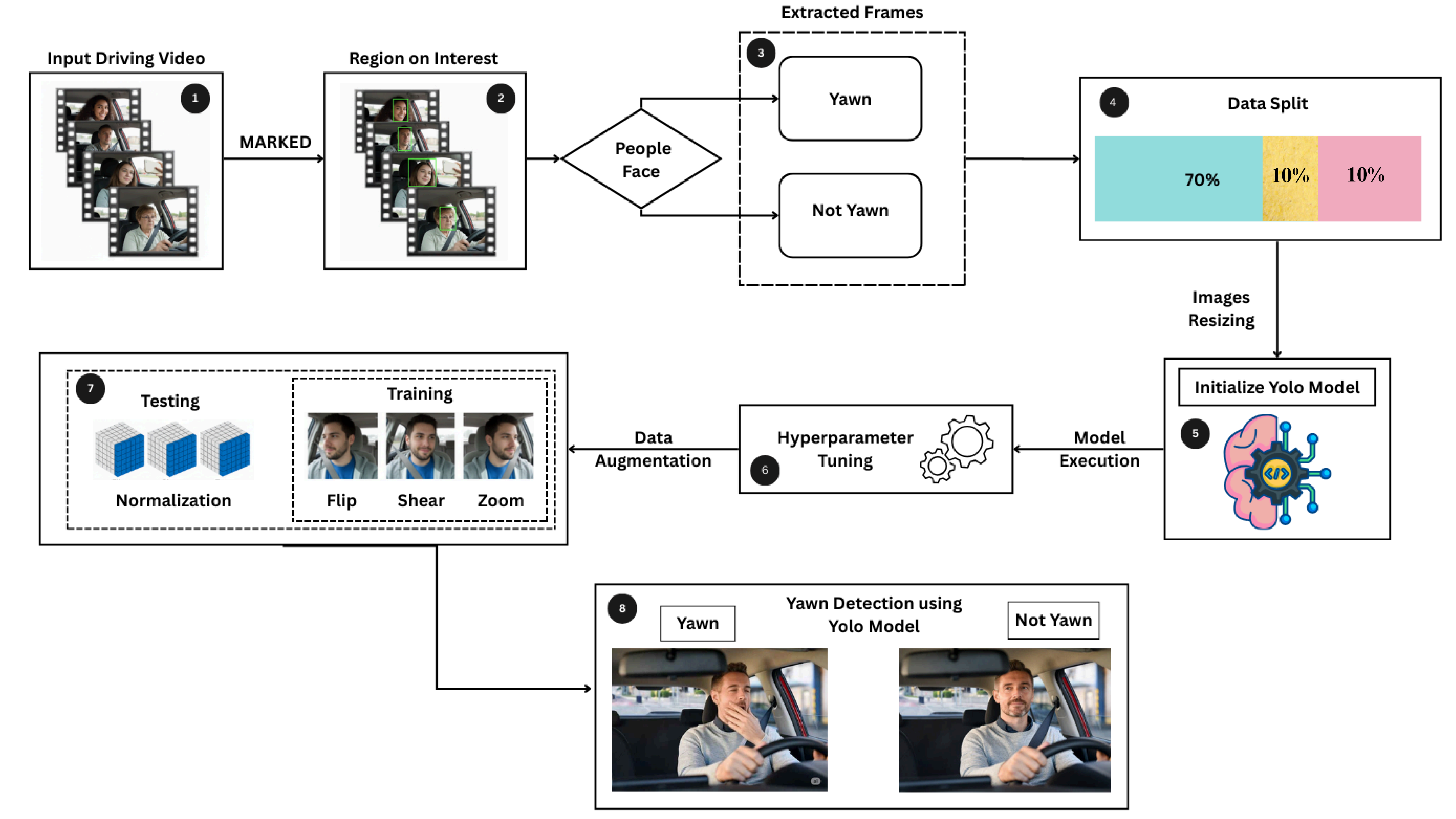


RASPBERRY PI 5

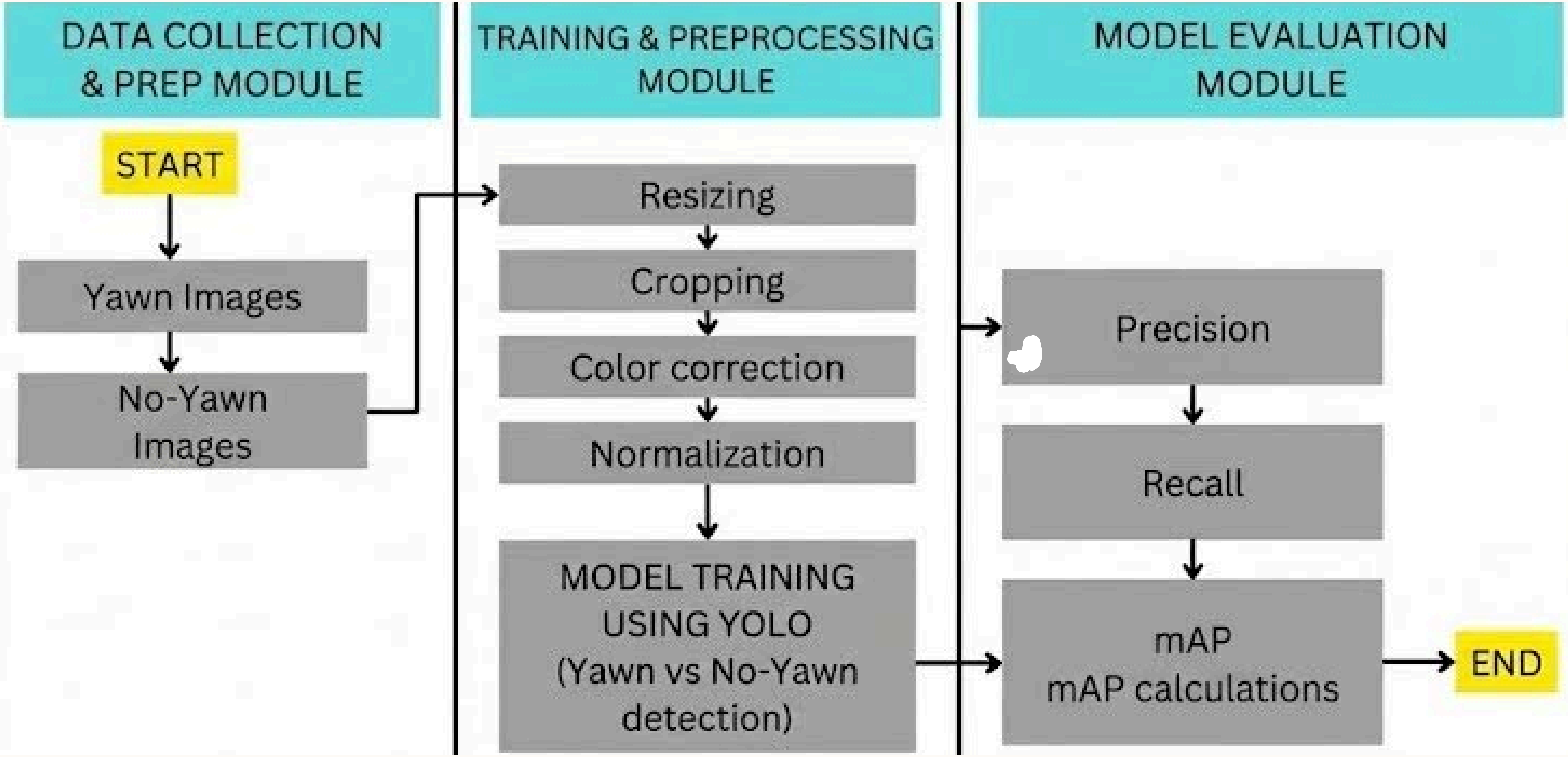
- Raspberry Pi is a small computer that works just like a normal PC but is much cheaper, compact, and power-efficient.
- It can run programs, connect with sensors, camera, and other devices, which makes it perfect for projects in IoT, automation, and AI.
- In our project, Raspberry Pi is used as the brain of the system.
- It captures the driver's live video, detects yawning in real time, and immediately triggers alerts.



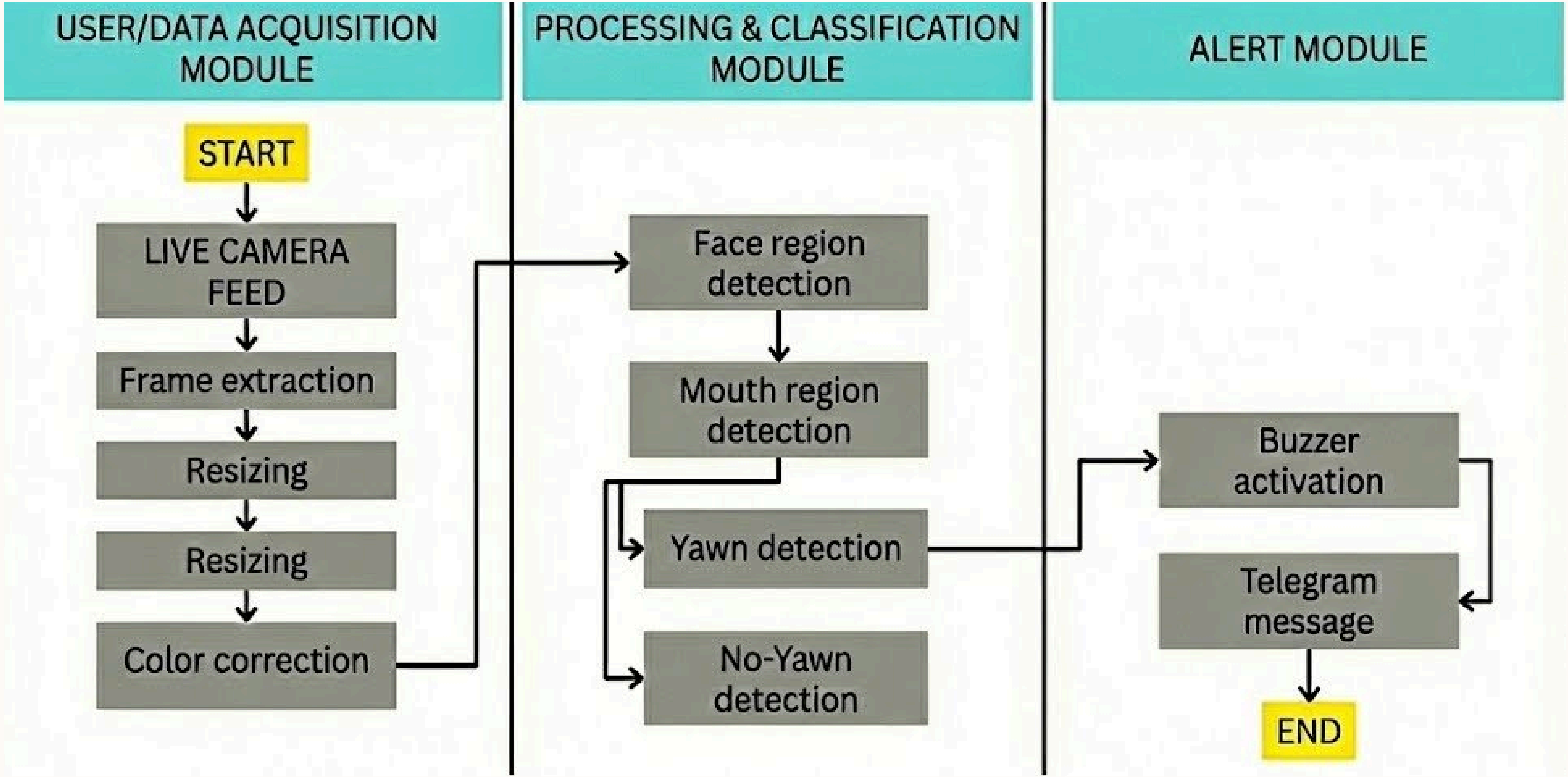
REVIVA ARCHITECTURE FLOW



REVIVA MODEL TRAINING



USER INPUT CLASSIFICATION



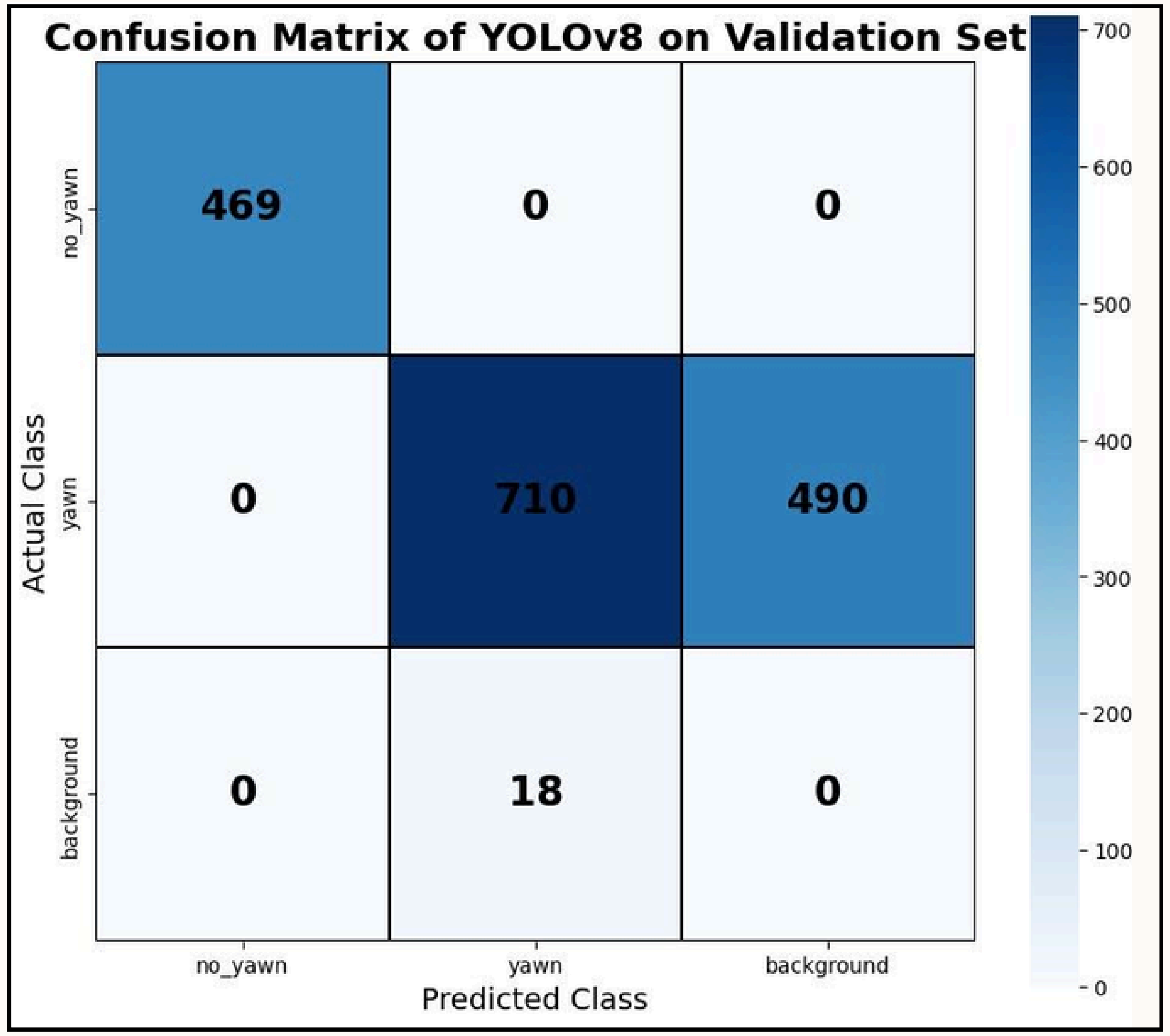
MODULE DESCRIPTION

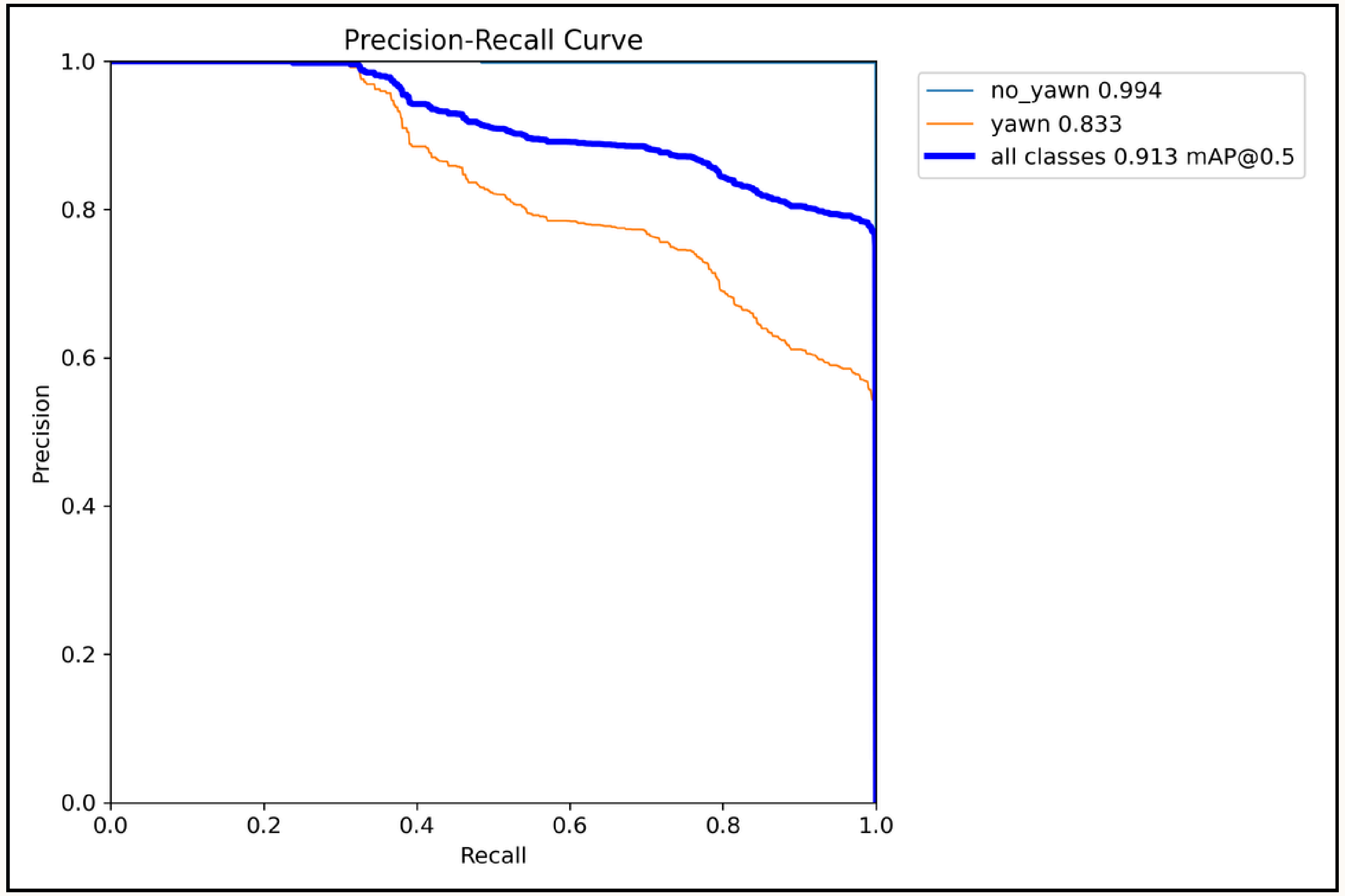
1. **Frame extraction and preprocessing:** The system captures live driving video through a Raspberry Pi camera, converts it into frames, and applies resizing, cropping, and color correction to ensure image consistency.
2. **Feature detection and classification:** The YOLO model detects the driver's face and classifies each frame as Yawn or Not Yawn based on pre-trained weights and extracted facial features.
3. **Alert generation:** When a yawn or drowsy behavior is detected, an alert is triggered via a buzzer and a Telegram bot notification to warn the driver in real time.
4. **Real-time execution:** The integrated system performs continuous monitoring and instant decision-making, enabling on-device, low-latency detection of driver fatigue.

YOLOV8N MODEL AND ITS DETECTION MECHANISM

1. **Real Time One Stage Detection Framework:** YOLOv8n uses a one stage detection method that performs feature extraction, localization, and classification in one fast step. This makes yawn detection extremely quick and suitable for real time use on devices like the Raspberry Pi.
2. **Bounding Box Localization + Class Prediction Together:** For every prediction, YOLO outputs the bounding box, confidence score, and class probability (yawn / no yawn), ensuring fast and accurate localization with fewer false detections.
3. **Filtering Predictions Using NonMaximum Suppression (NMS):** YOLO generates several overlapping bounding boxes for the same region. NMS filters these by keeping only the highest-confidence box, removing duplicates, and producing a clean final detection such as a single accurate box around the mouth during a yawn.

EVALUATION METRICS

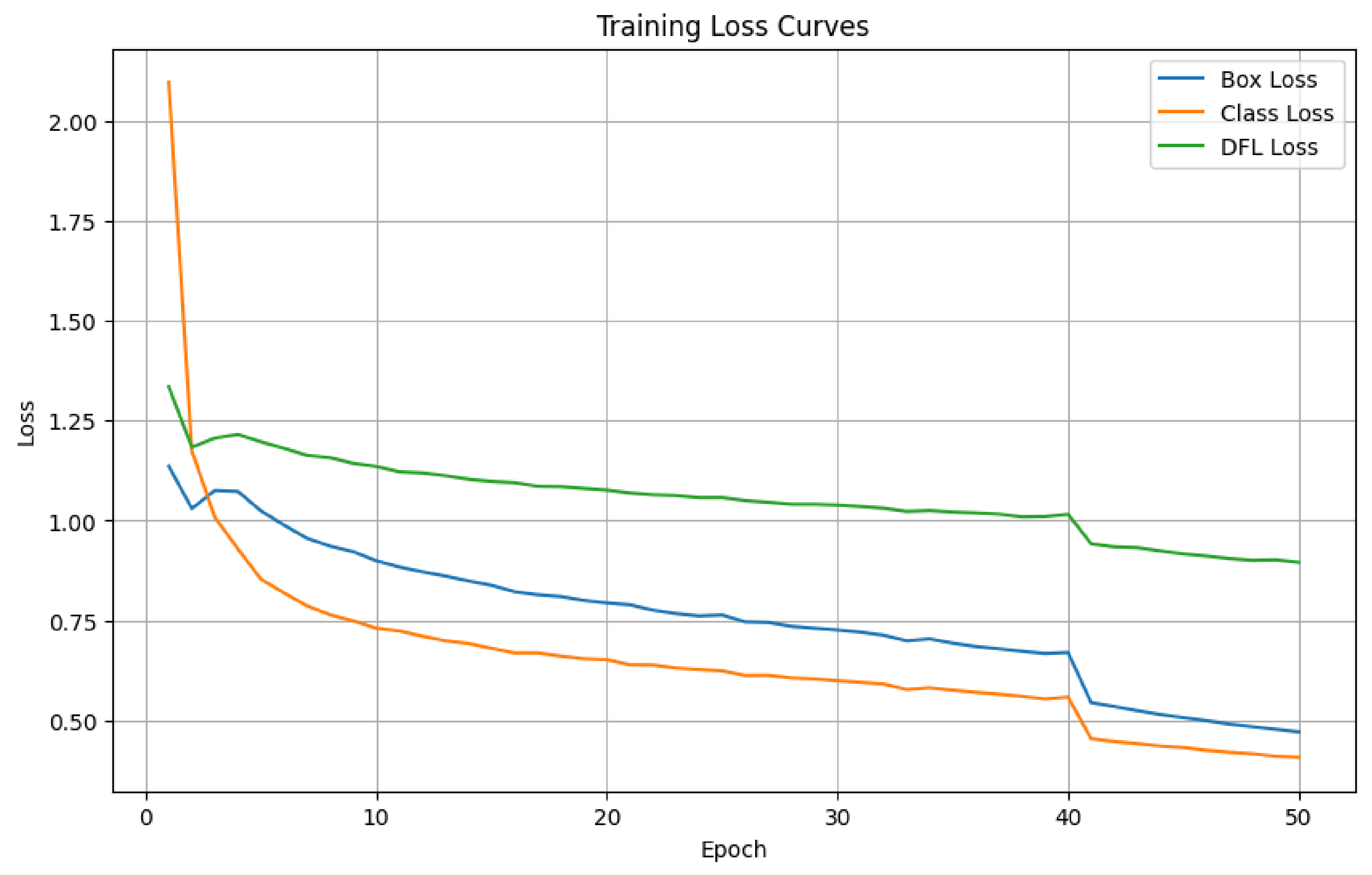


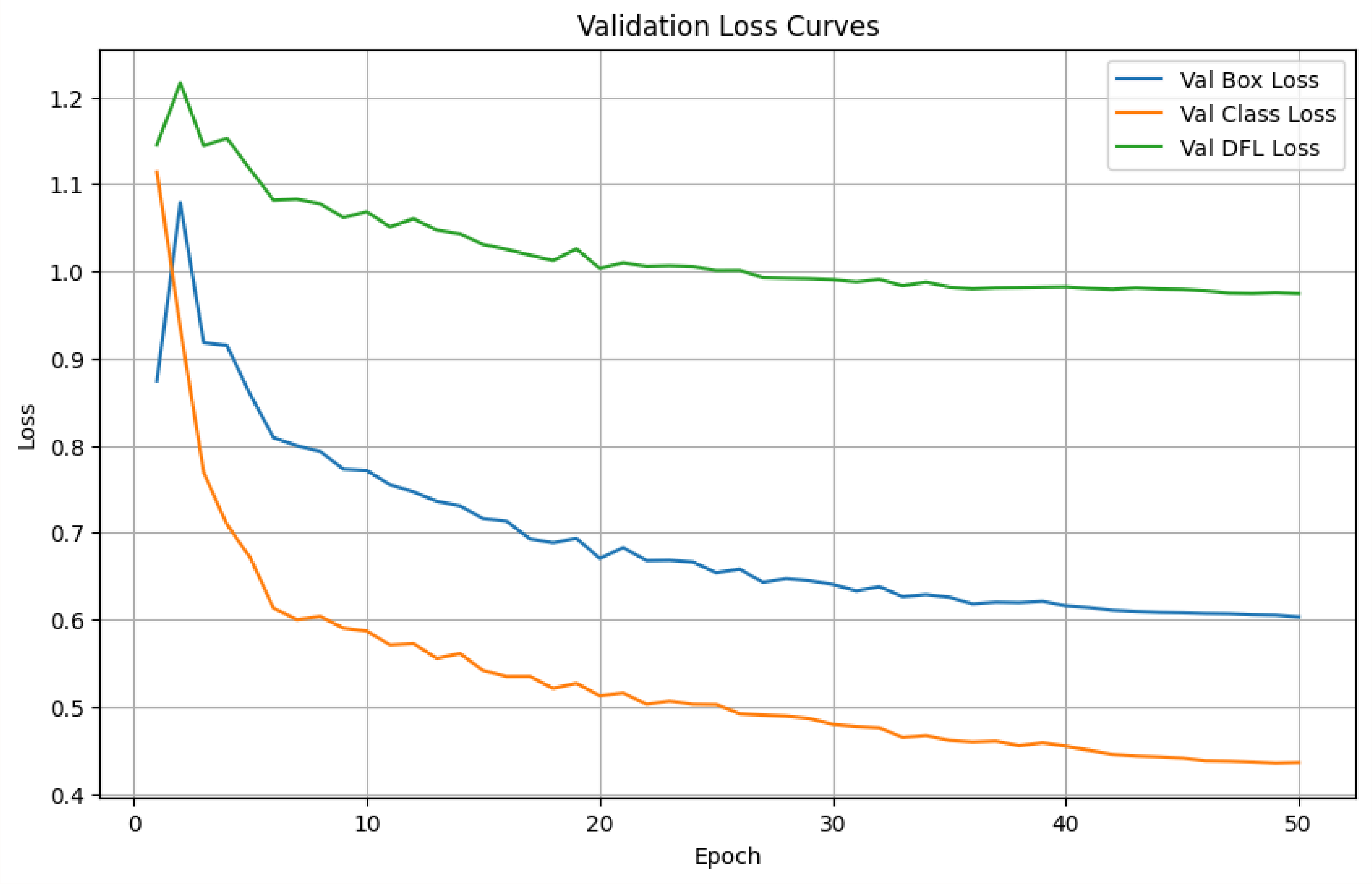


- This matrix compares predicted and actual results for each class.
- It shows that the model correctly detects most yawns but sometimes confuses them with the background.

- This graph shows how well the model distinguishes between yawning and non-yawning.
- The high precision and recall indicate strong detection performance with 91.3% overall accuracy.

EVALUATION METRICS

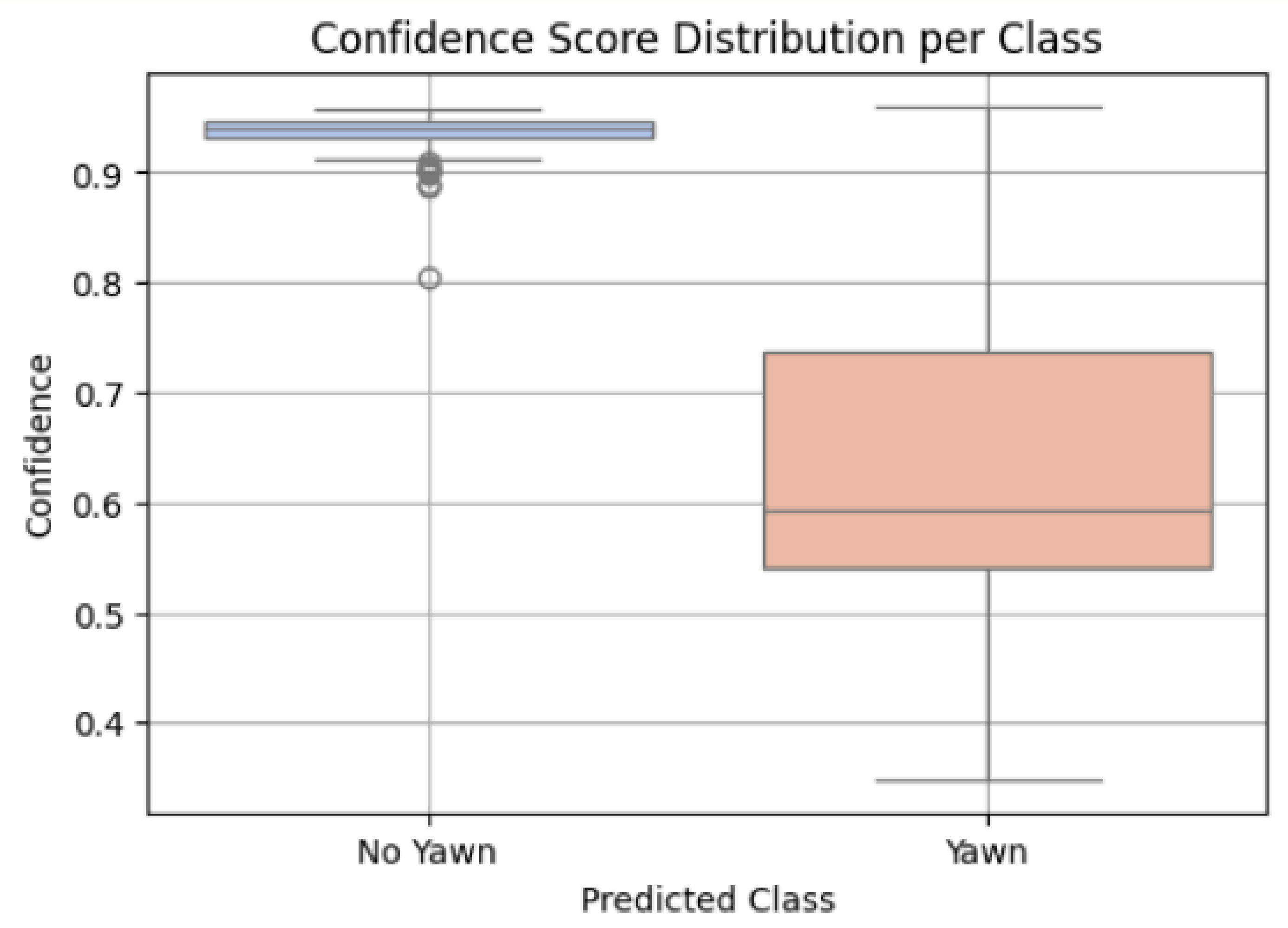




- The Box, Class, and DFL losses consistently decrease across epochs, indicating stable YOLOv8 training.
- The steep drop in early epochs reflects rapid feature learning and convergence.
- By the final epochs, all losses flatten, showing the model has reached a well-optimized state.

- Validation losses decrease smoothly across epochs, demonstrating good convergence on external data.
- Initial spikes settle rapidly as the model learns stable feature patterns.
- Final low loss values show that the model performs well on validation frames and generalizes effectively.

EVALUATION METRICS

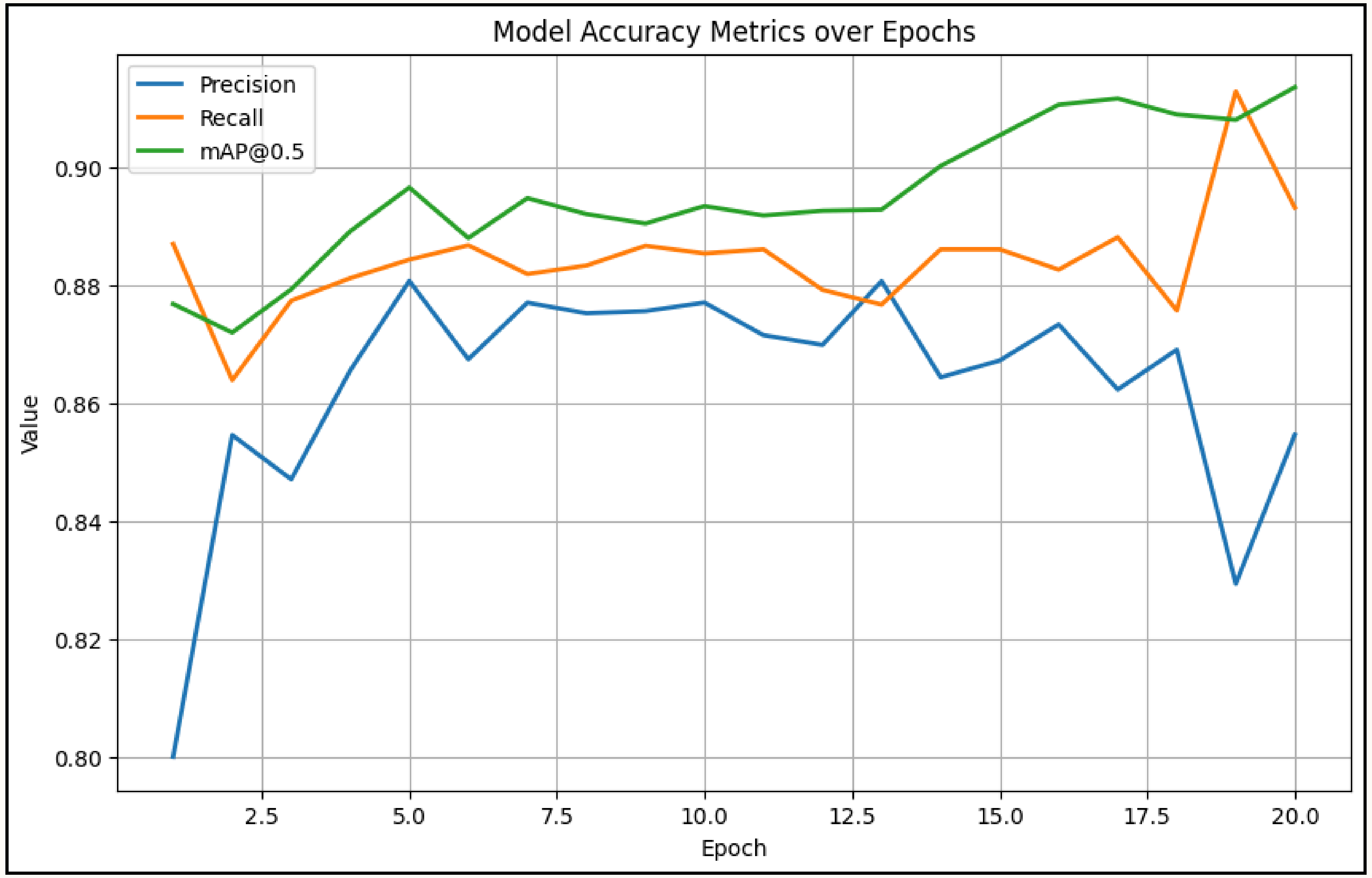


- No-Yawn predictions maintain consistently high confidence, showing strong model reliability.
- Yawn predictions have broader confidence variation but remain centered at a stable mid-high range.
- This distribution confirms that the model is confident in normal states and reasonably discriminates yawn events.

Platform	**FPS**	**Latency (ms)**	**Detection Accuracy**	**Notes**
Laptop / PC	30	140	Same as trained model	Fast, smooth inference
Raspberry Pi 4	32	148	Same as trained model	Stable real time performance

- Both the Laptop/PC and Raspberry Pi 5 deliver real-time inference with comparable FPS and latency, showing the model is lightweight and deployment-ready.
- Detection accuracy remains identical across platforms, confirming that hardware differences do not affect prediction quality.
- The Raspberry Pi achieves stable performance despite limited compute, validating its suitability for embedded yawn-monitoring applications.

FINAL MODEL RESULTS

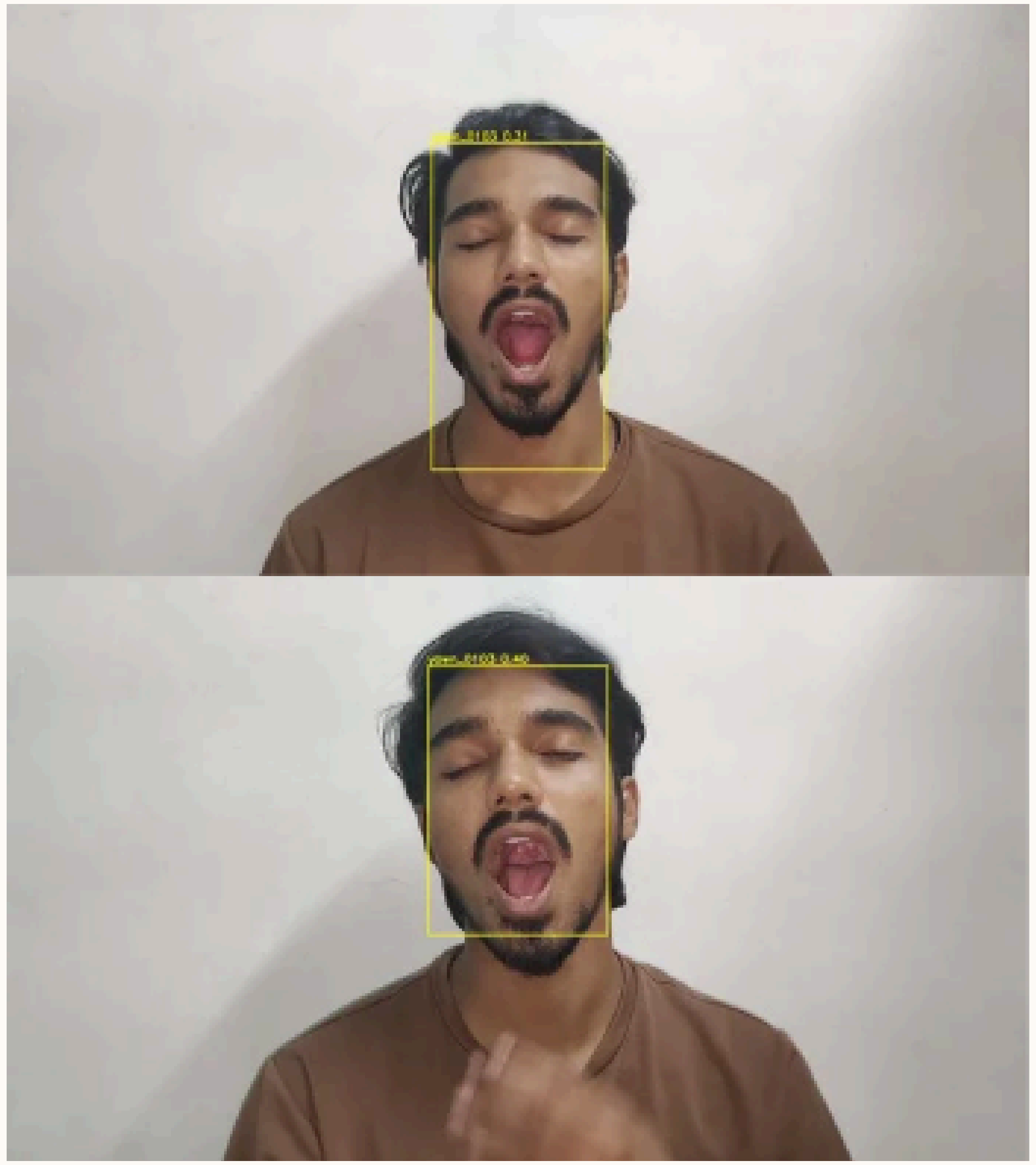


- The graph shows how the YOLO model’s performance improved **over 20 epochs**, with precision, recall, and mAP@0.5 steadily increasing.
- It highlights that the model became more accurate and consistent in detecting yawns as training progressed.

```
Final Model Accuracy Summary:
Precision          : 85.46%
Recall             : 89.31%
mAP@0.5 (Accuracy): 91.35%
mAP@0.5-0.95       : 81.52%
```

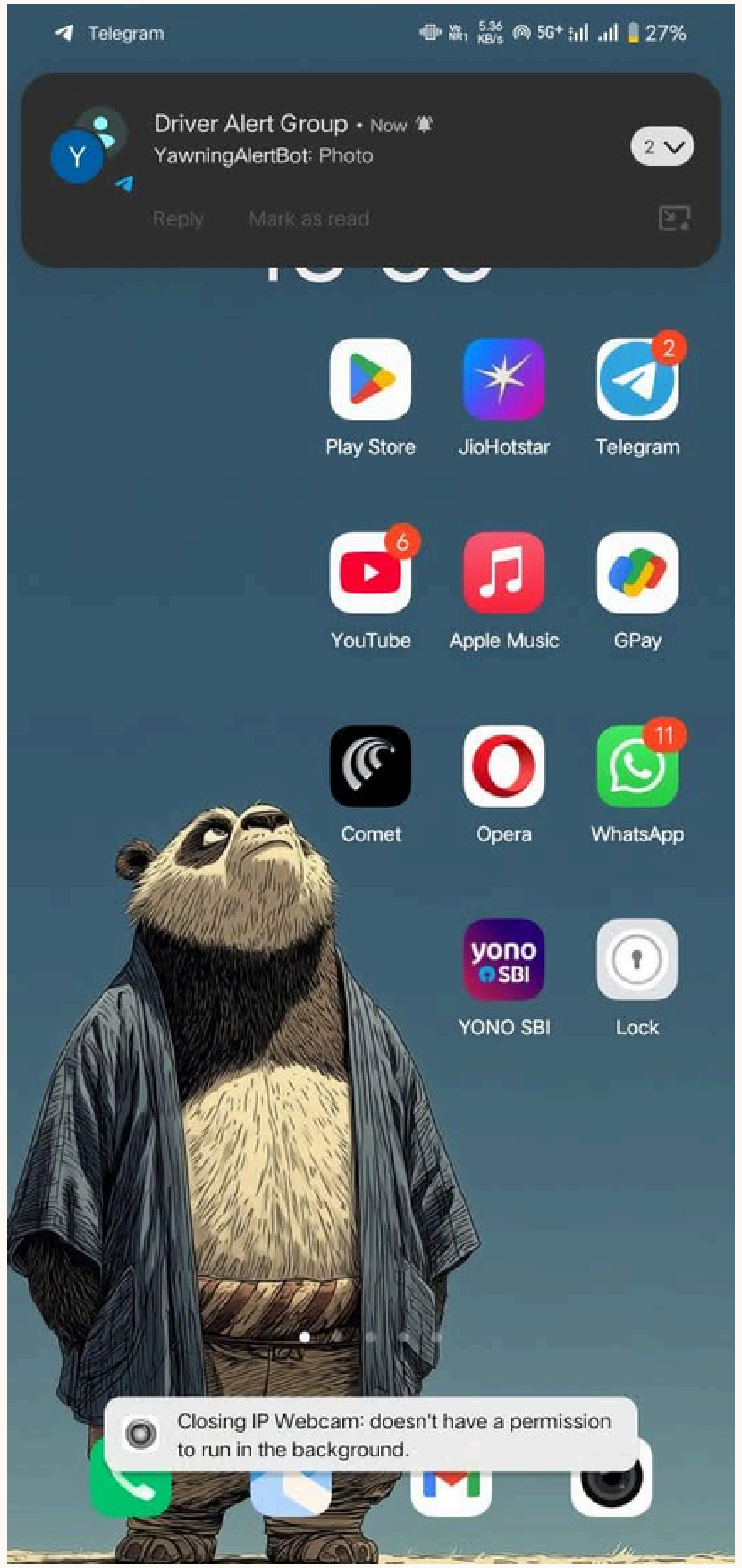
- This graph shows how well the model distinguishes between yawning and non-yawning.
- The high precision and recall indicate strong detection performance with **91.3% overall accuracy**.The final YOLO model achieved strong performance with **85.46% precision and 89.31% recall**, showing a good balance between accurate and complete yawn detection.
- The overall detection accuracy (mAP@0.5) reached 91.35%, while the more strict metric (mAP@0.5–0.95) scored 81.52%, indicating robust and reliable model performance across different thresholds.

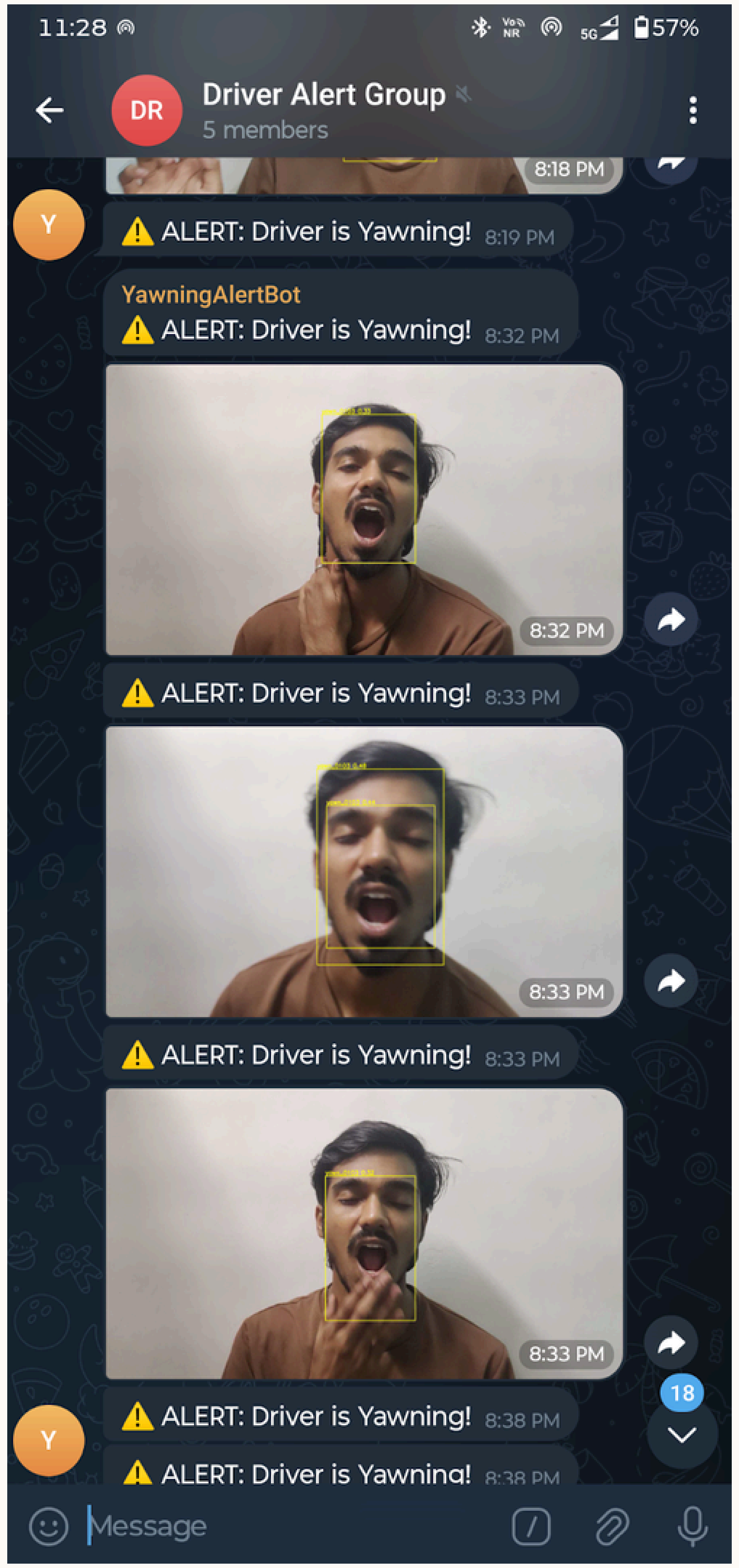
OUTPUT SCREENSHOTS



```
Vaults    SFTP    homelab
(.venv) admin@homelab:~/yawn $ python3 yawn_eye_alert_2.py
[🚗] Yawn monitor active (with Telegram alerts)
[⚠] Press CTRL + C to stop
[SAVED] yawn_captures/yawn_2025-11-22_13-21-42.jpg
[INFO] Telegram alert sent
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[INFO] Cooldown active - no Telegram message
[SAVED] yawn_captures/yawn_2025-11-22_13-21-57.jpg
[INFO] Telegram alert sent
```

OUTPUT SCREENSHOTS





THANK YOU

Team Members:

- Sriram K
  221501144@rajalakshmi.edu.in
- Sudharsan S
  221501149@rajalakshmi.edu.in
- Tamilarasan D
  221501157@rajalakshmi.edu.in