DEGGENDORF INSTITUTE OF TECHNOLOGY

AUTOMOTIVE DRIVER DROWSINESS DETECTION USING ON-SENSOR AI

This project presents an embedded driver drowsiness detection system using on-sensor Edge AI.
A YOLO-based model deployed on the Sony IMX500 performs real-time driver monitoring on a Raspberry Pi, providing visual and audio alerts to enhance automotive safety.

Presented By:

Deep Bharatbhai Savaliya (12501180)

Yutsav Hari Bhagat (12500192)

Supervisor:

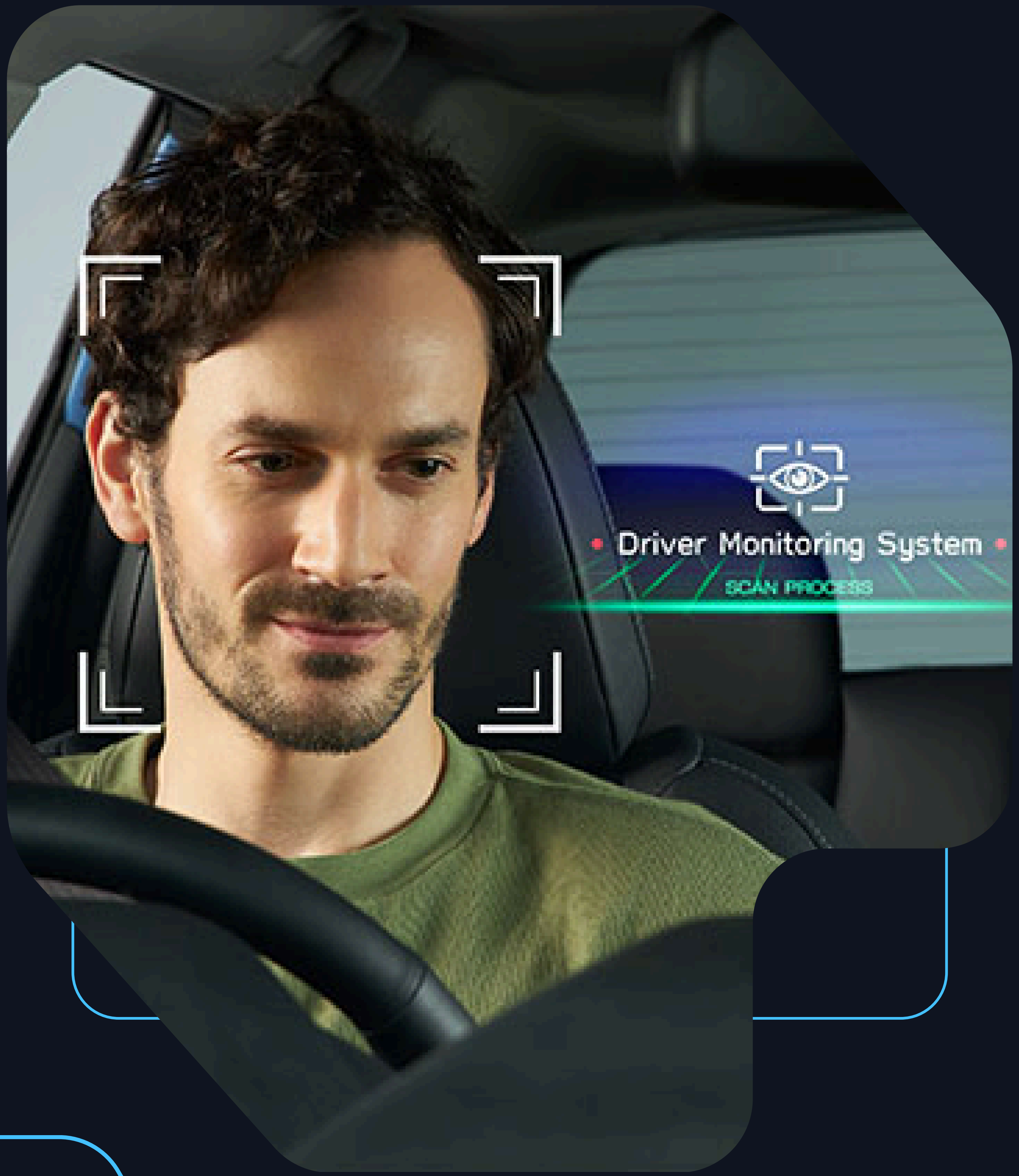
Prof. Thomas Ewender

SOURCE IMAGES

AGENDA

- Project Overview
- Motivation and Concept
- Dataset and Training Setup
- Evaluation and Results
- Limitations and Future Improvements

SOURCE IMAGES



source image

Project Overview

- Automotive driver drowsiness detection system
- Real-time monitoring using embedded Edge AI
- Deployed on Sony IMX500 AI camera and Raspberry Pi
- Designed for Driver Monitoring Systems (DMS)

Objectives

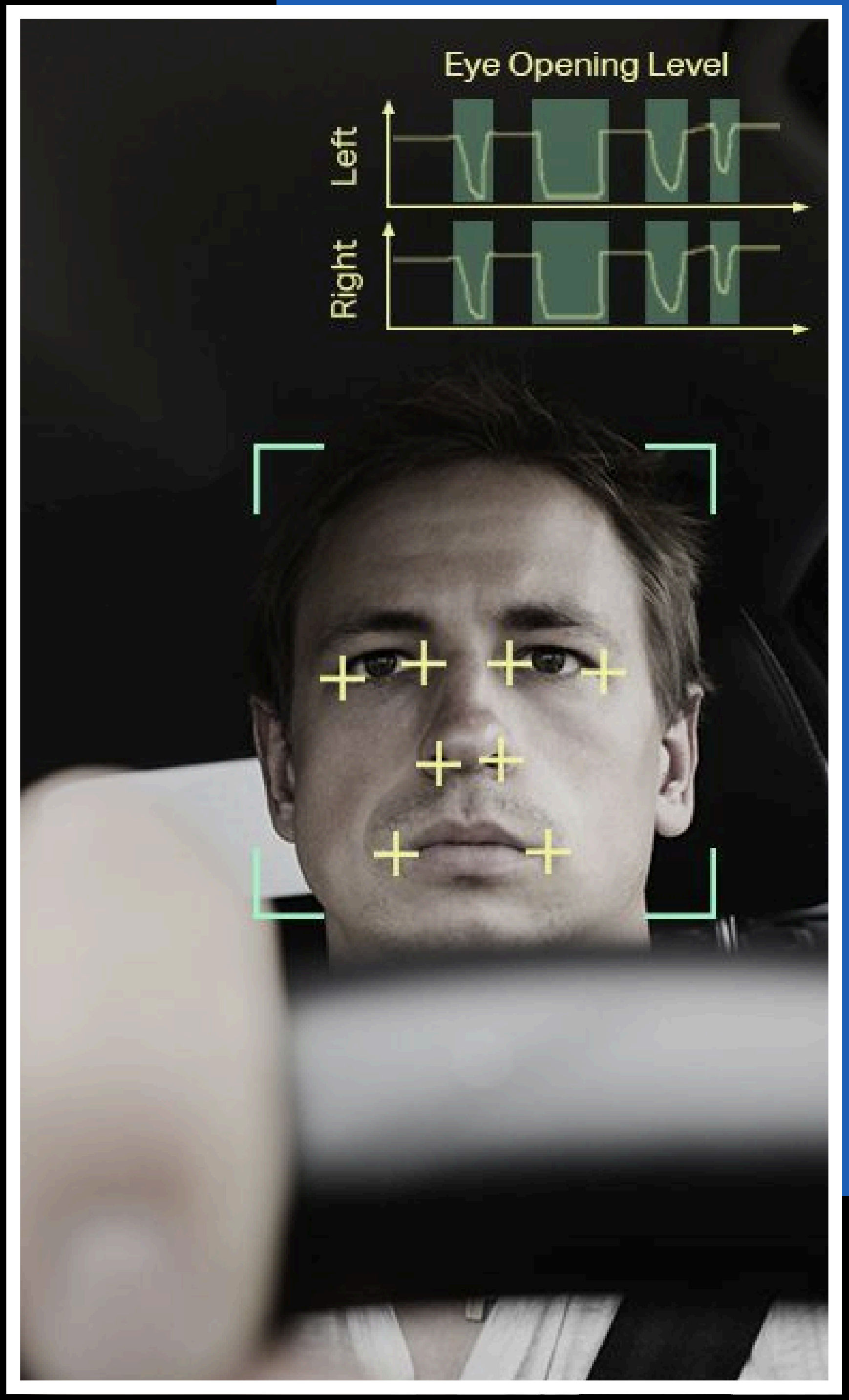
- The primary objective of this project is to detect driver alertness in real time using computer vision techniques.
- The system classifies the driver’s facial state directly on embedded hardware without cloud dependency.
- Visual and audio alerts are generated to warn the driver and enhance driving safety under drowsy conditions.

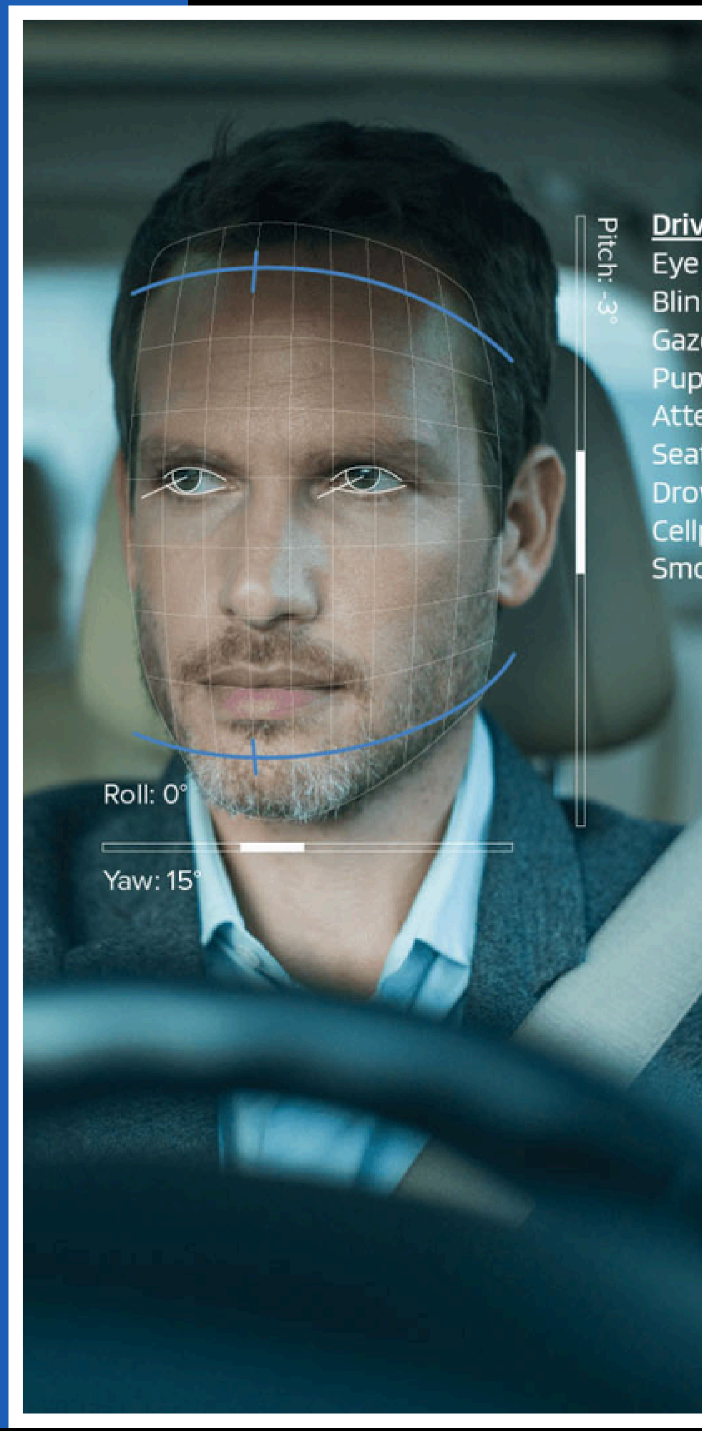
Workflow

- The project follows a complete end-to-end embedded AI workflow starting from data collection and annotation.
- A YOLO-based model is trained and optimized for embedded deployment on the Sony IMX500.
- The trained model is deployed on a Raspberry Pi to perform real-time inference and driver safety alerting.

MOTIVATION: ROAD SAFETY & ADAS CONTEX

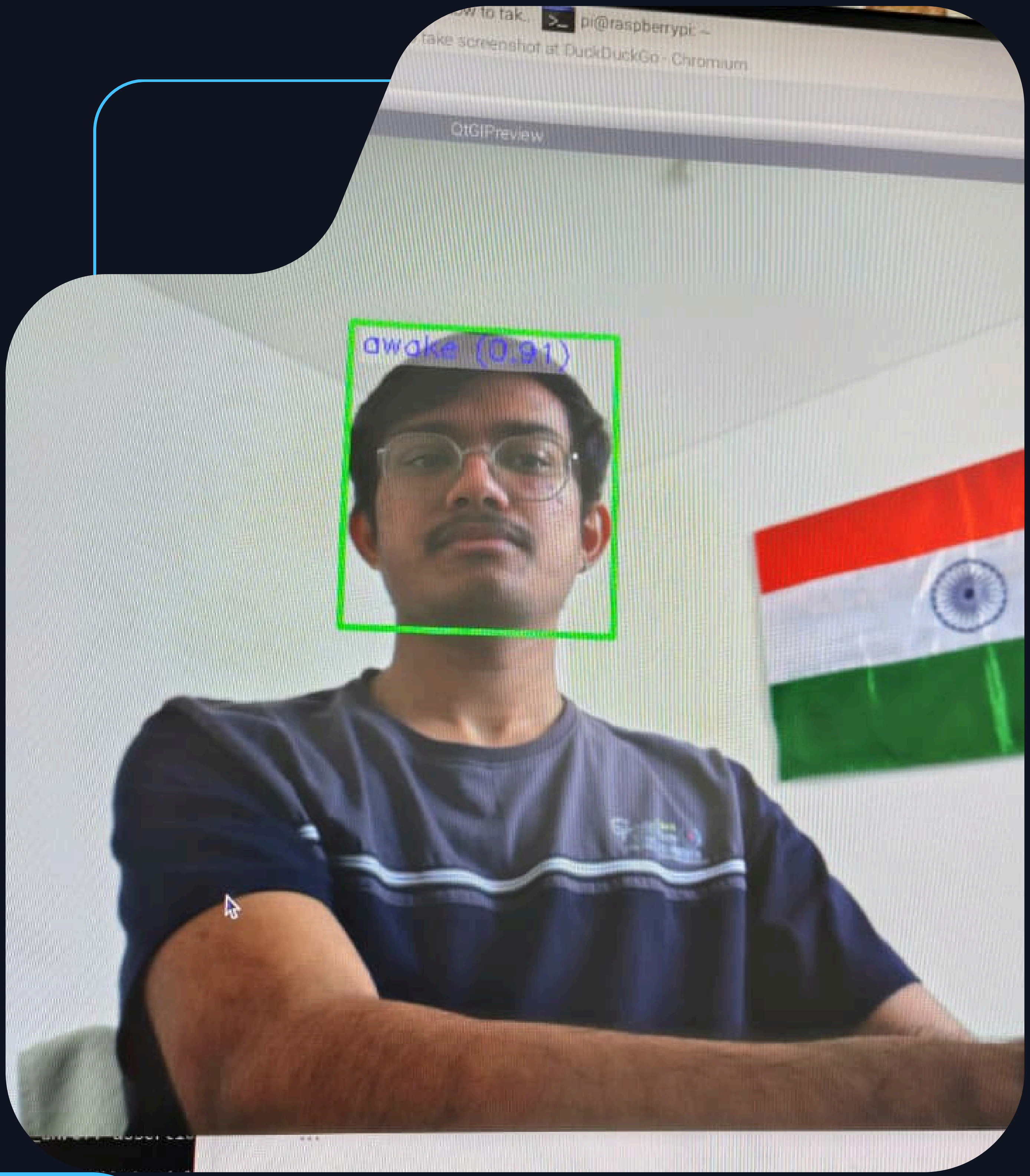
- Driver drowsiness is a major cause of traffic accidents, as fatigue significantly reduces reaction time and situational awareness.
- Micro-sleep events often occur without warning, making manual detection unreliable.
- To address this risk, modern vehicles integrate Driver Monitoring Systems as part of Advanced Driver Assistance Systems.
- Real-time, embedded, and privacy-preserving driver monitoring is therefore essential for improving road safety.



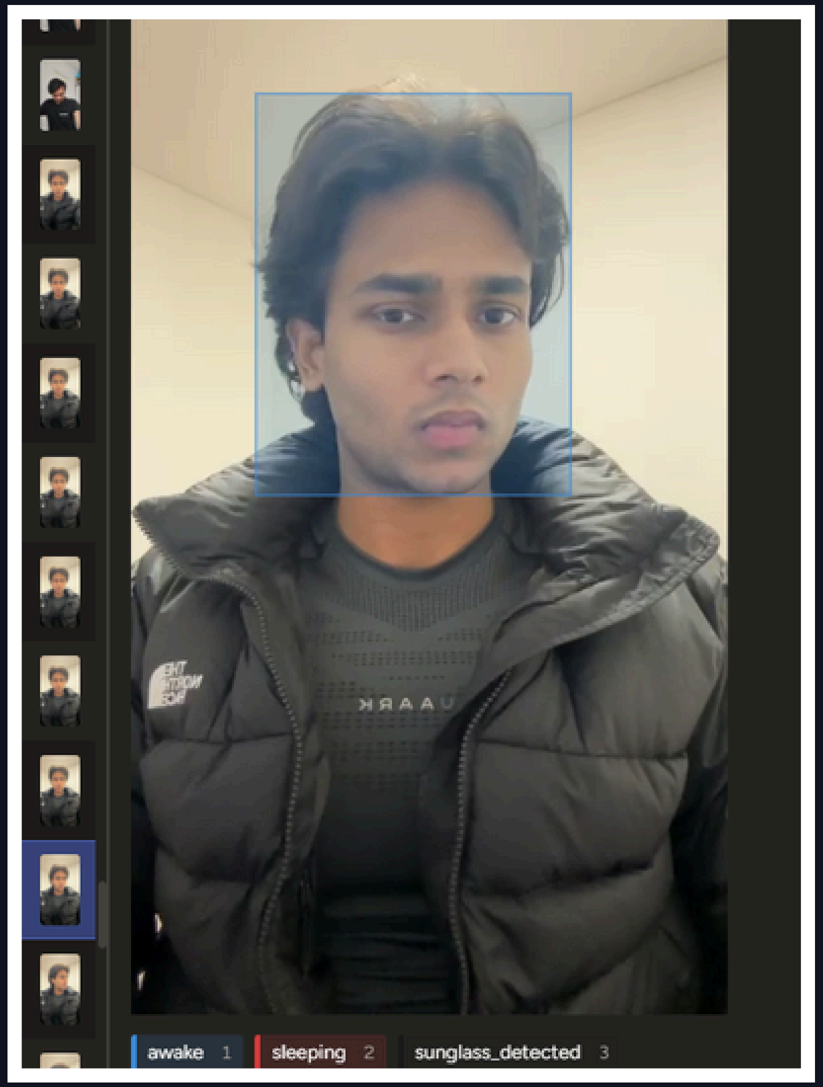


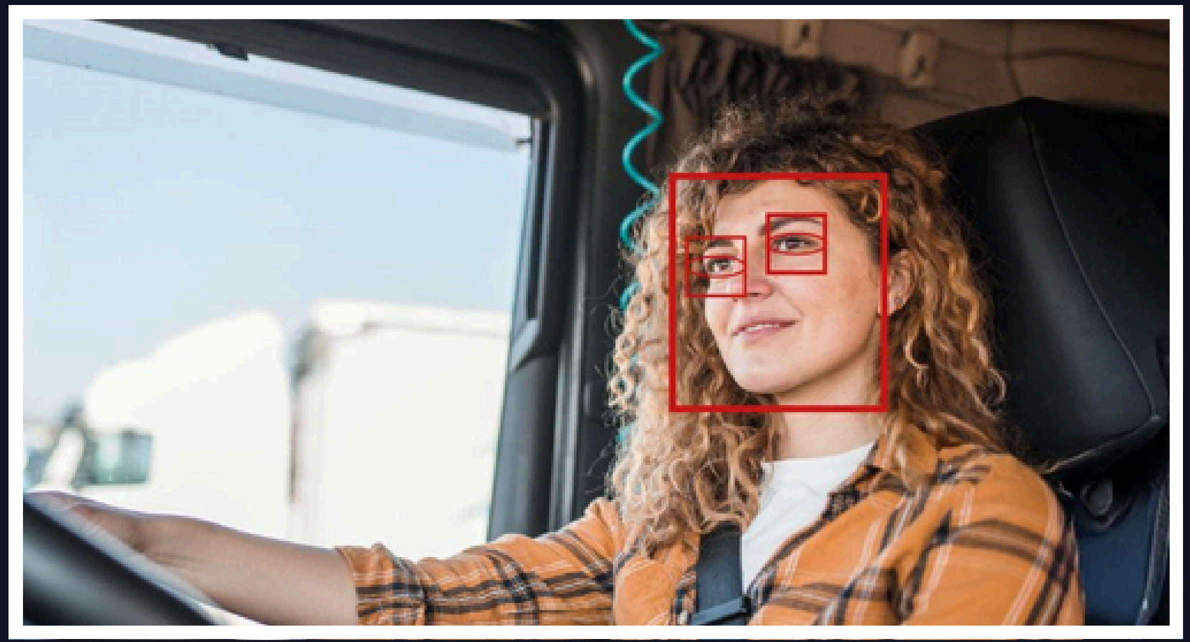
source image

Problem Statement, Concept & System Overview



- The objective of this project is to detect driver drowsiness accurately and in real time while operating on embedded automotive hardware.
- The system uses a vision-based approach with a YOLO object detection model to classify the driver’s facial state directly on the Sony IMX500 using on-sensor inference.
- The camera continuously captures the driver’s face, and the AI model displays bounding boxes and status information in the live preview.
- Visual warnings and audio alerts are generated, with danger escalation triggered during prolonged drowsy conditions, ensuring low latency and high reliability without cloud dependency.





source image

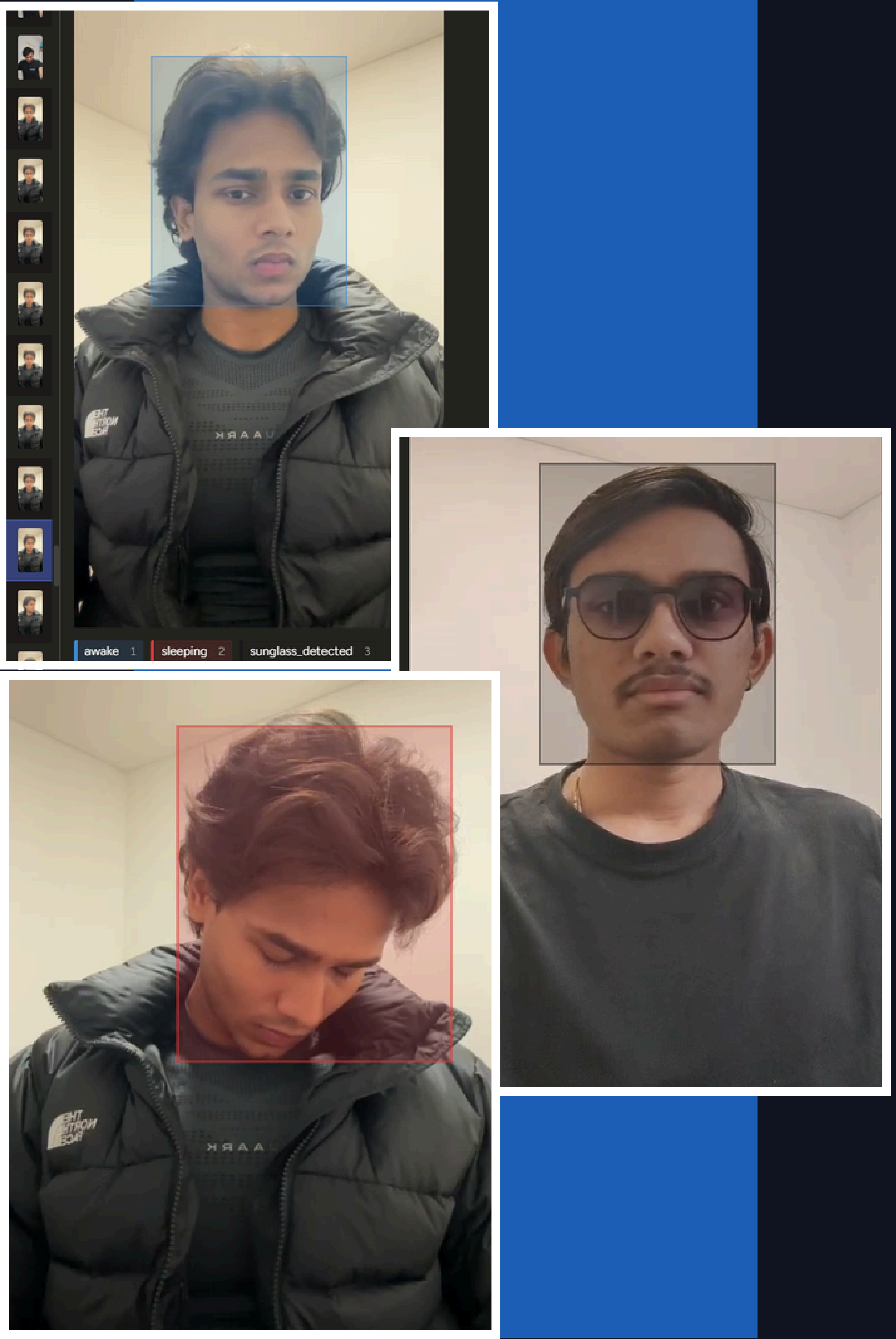
Dataset Collection & Annotation

Dataset Collection

- The dataset was manually collected using a fixed in-cabin camera setup, simulating a real-world driver monitoring scenario.
- Images were captured under different lighting conditions and head poses to ensure the model's robustness across diverse driving situations.
- Real human facial data was used to represent realistic drowsiness detection conditions in the automotive environment.

Annotation Process

- The dataset was annotated using Label Studio, an open-source tool for image annotation.
- Bounding boxes were drawn around the driver's face in each image, and the annotations followed the YOLO-compatible format.
- Each annotation was stored as normalized bounding box coordinates, making the data ready for model training.



Dataset Classes, Validation & Split

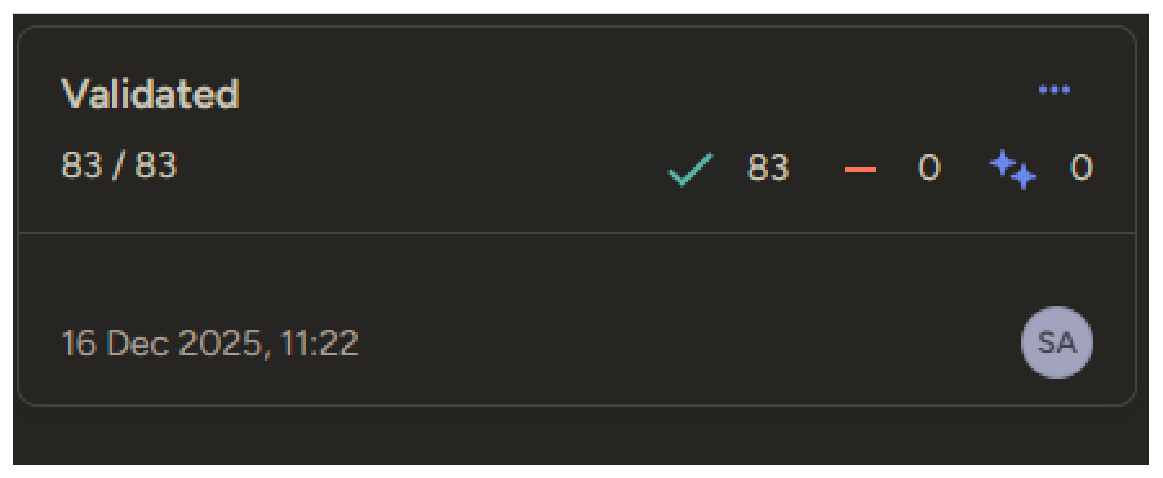
Dataset Classes

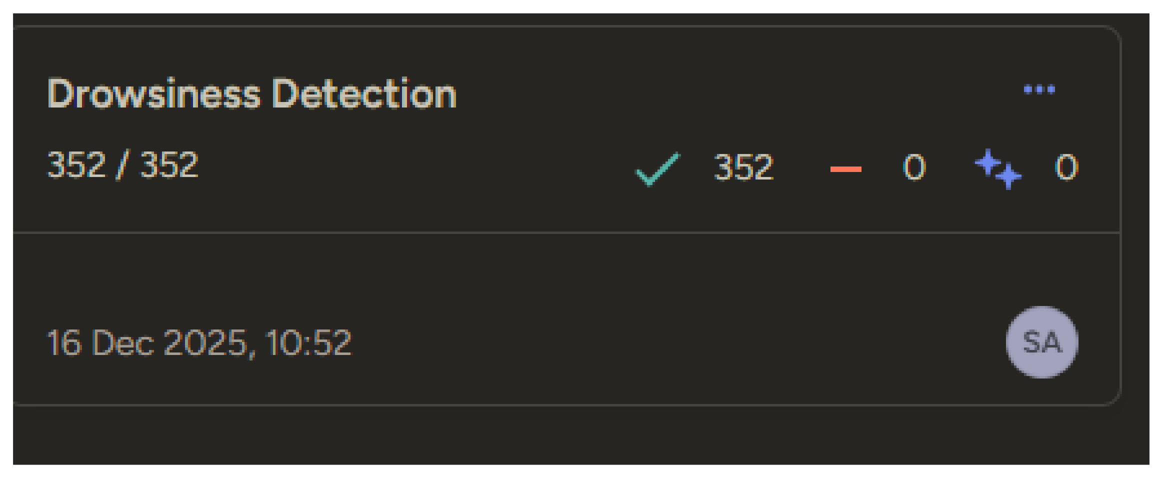
- The dataset contains three distinct driver state classes used for model training:
  - Awake: Represents an alert and attentive driver.
  - Sleeping: Indicates drowsiness or closed eyes.
  - Sunglass Detected: Identifies reduced eye visibility due to sunglasses.
- These classes allow the model to learn to classify the driver’s state based on facial features.

Dataset Validation

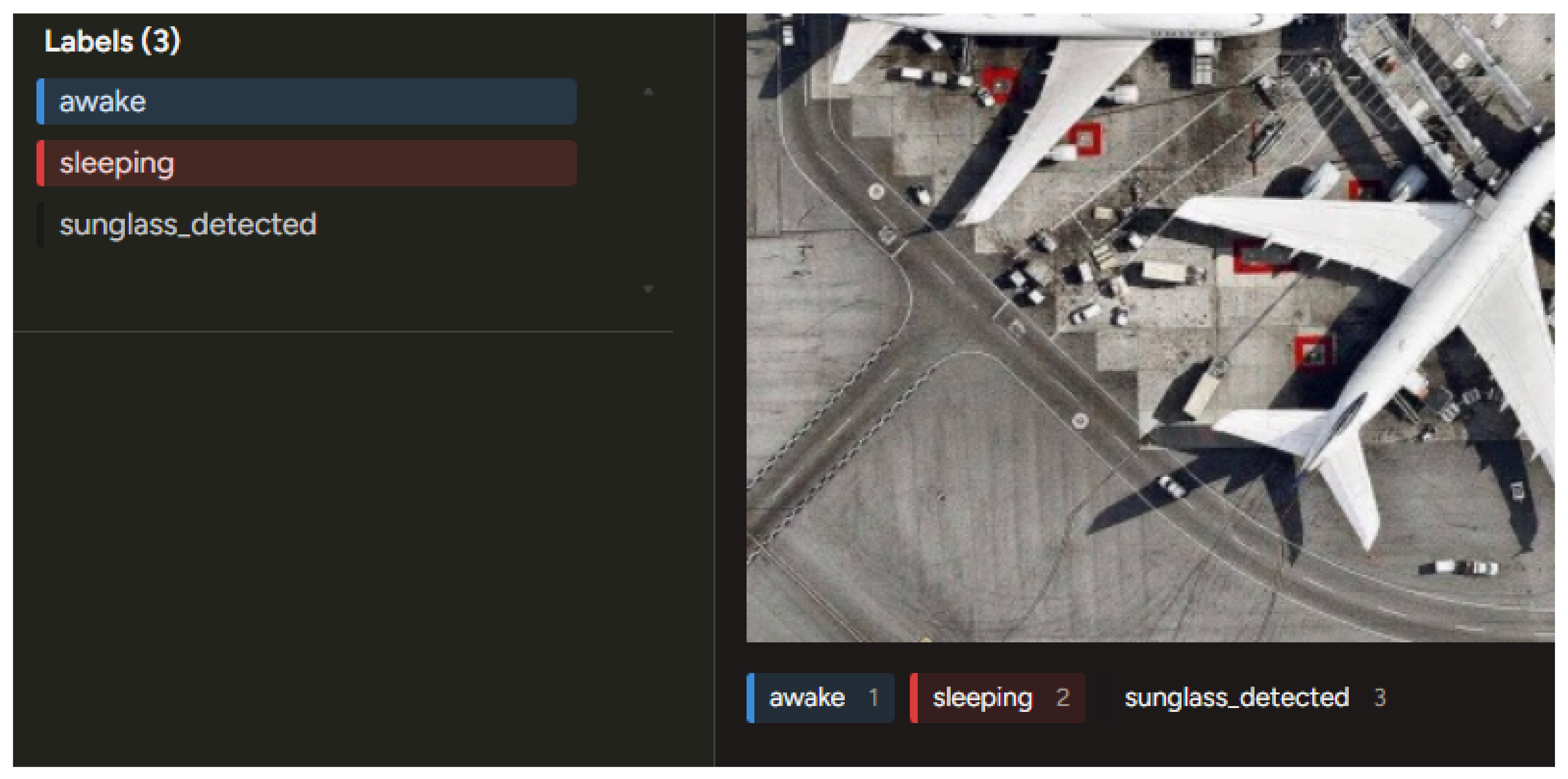
- After annotation, the dataset was fully validated within Label Studio.
- All annotation tasks were completed and reviewed successfully, ensuring data quality.
- This validation process reduces label noise and guarantees high-quality training data.

Dataset Split

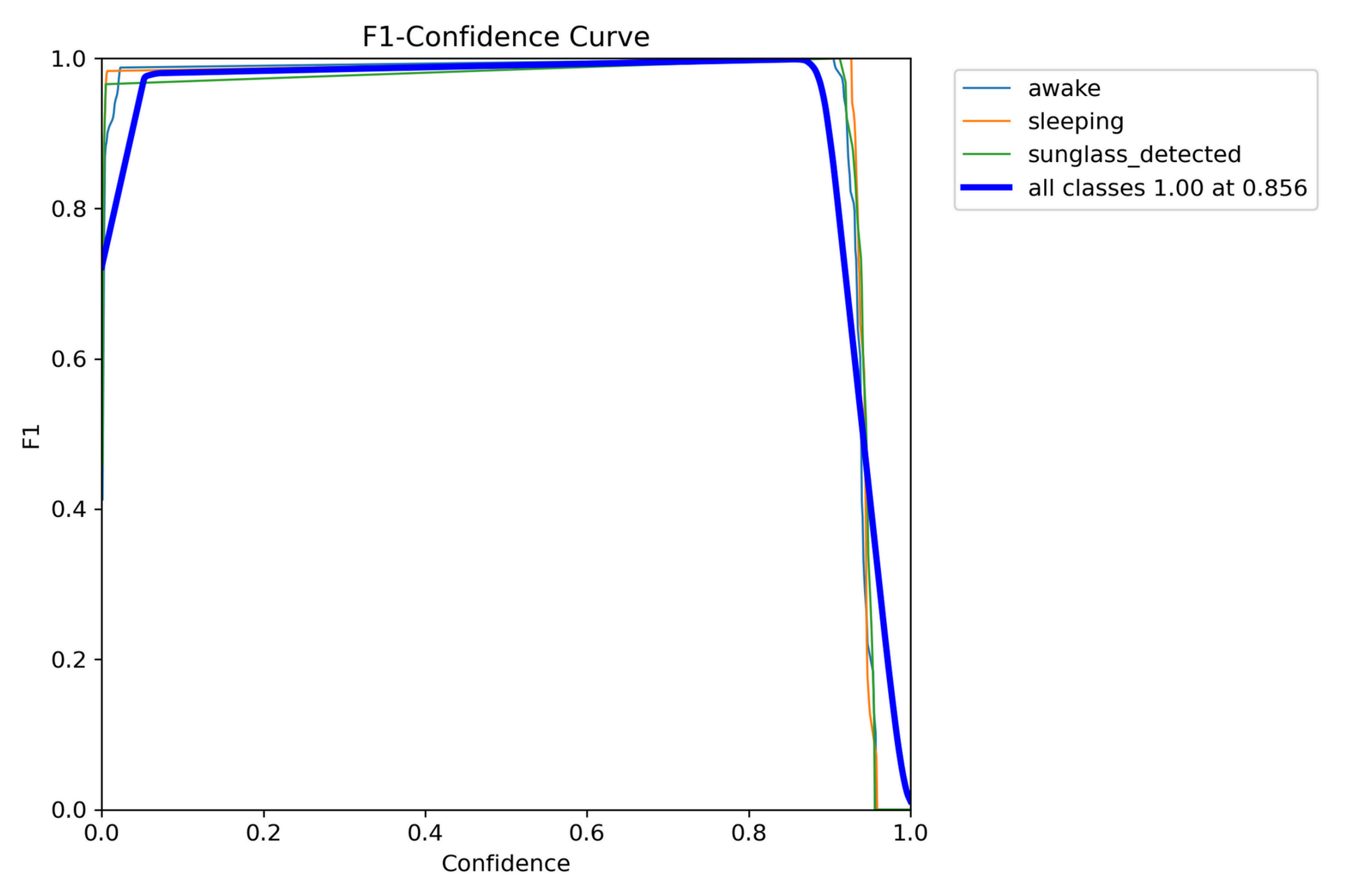




- The validated dataset was split into training and validation sets:
- The training set was used for learning model parameters.
- The validation set was used to evaluate the model's performance and generalization ability.
- The split was performed class-balanced to ensure fairness in training and evaluation.

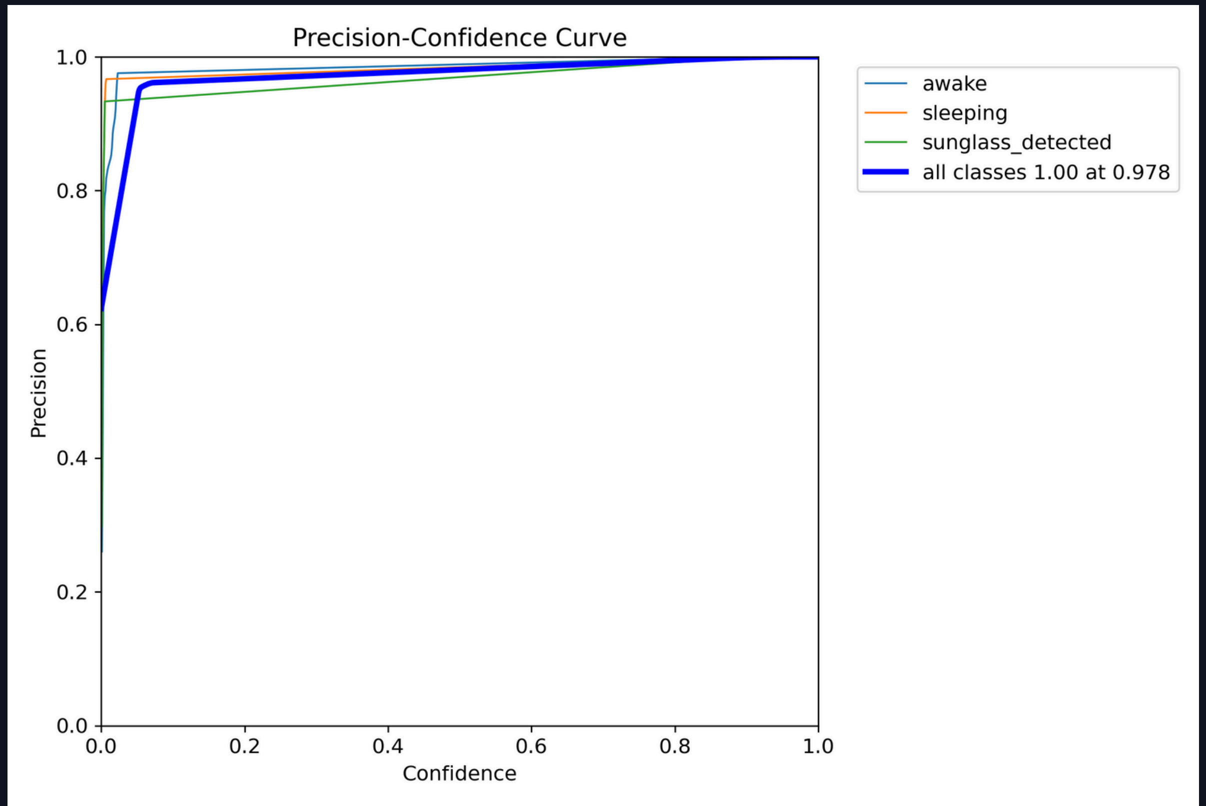


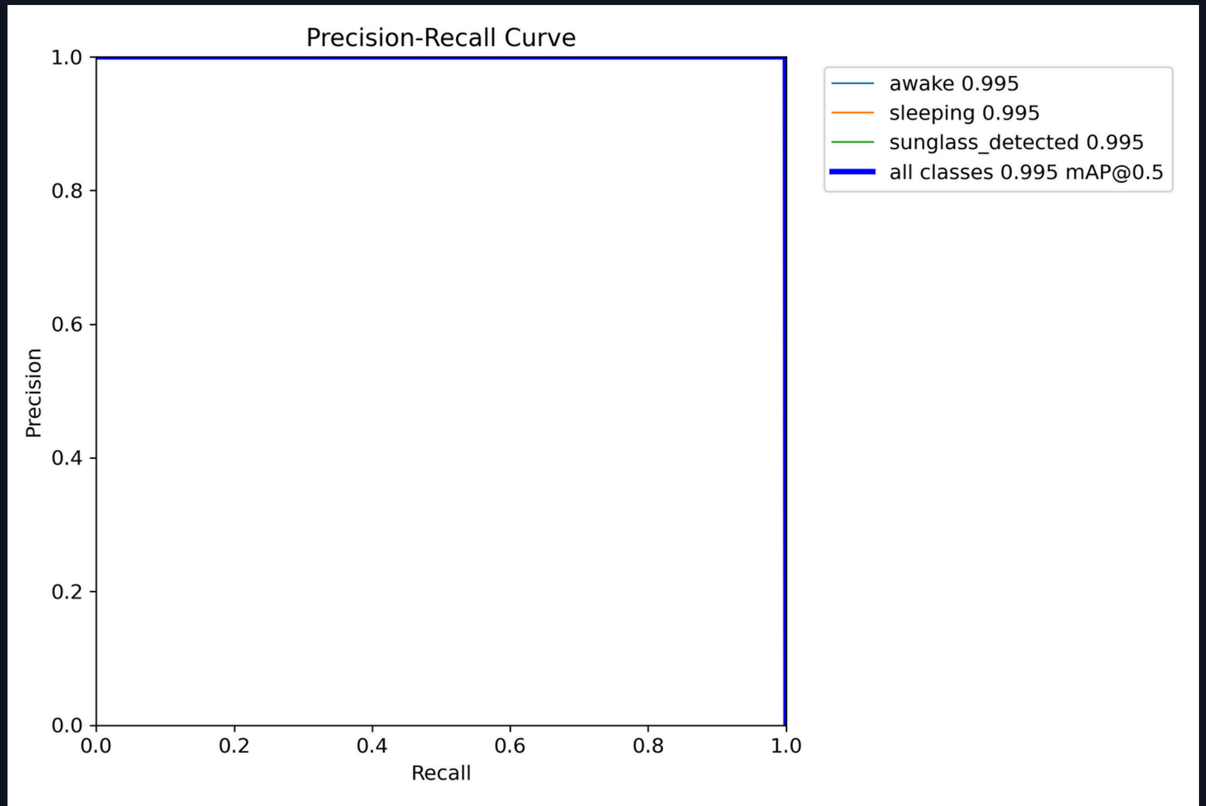
Model Training Setup



- Model: YOLOv11 Nano
  - Chosen for its lightweight architecture, perfect for embedded deployment.
  - Balances accuracy and efficiency.
- Input Image Size: 640 × 640
  - Standardized input size for the YOLO model to capture detailed features of the driver's face.
- Training Epochs: 100
  - Trained for 100 epochs to ensure sufficient learning.
- Batch Size: 16
  - Batch size of 16 used to balance memory and processing time during training.

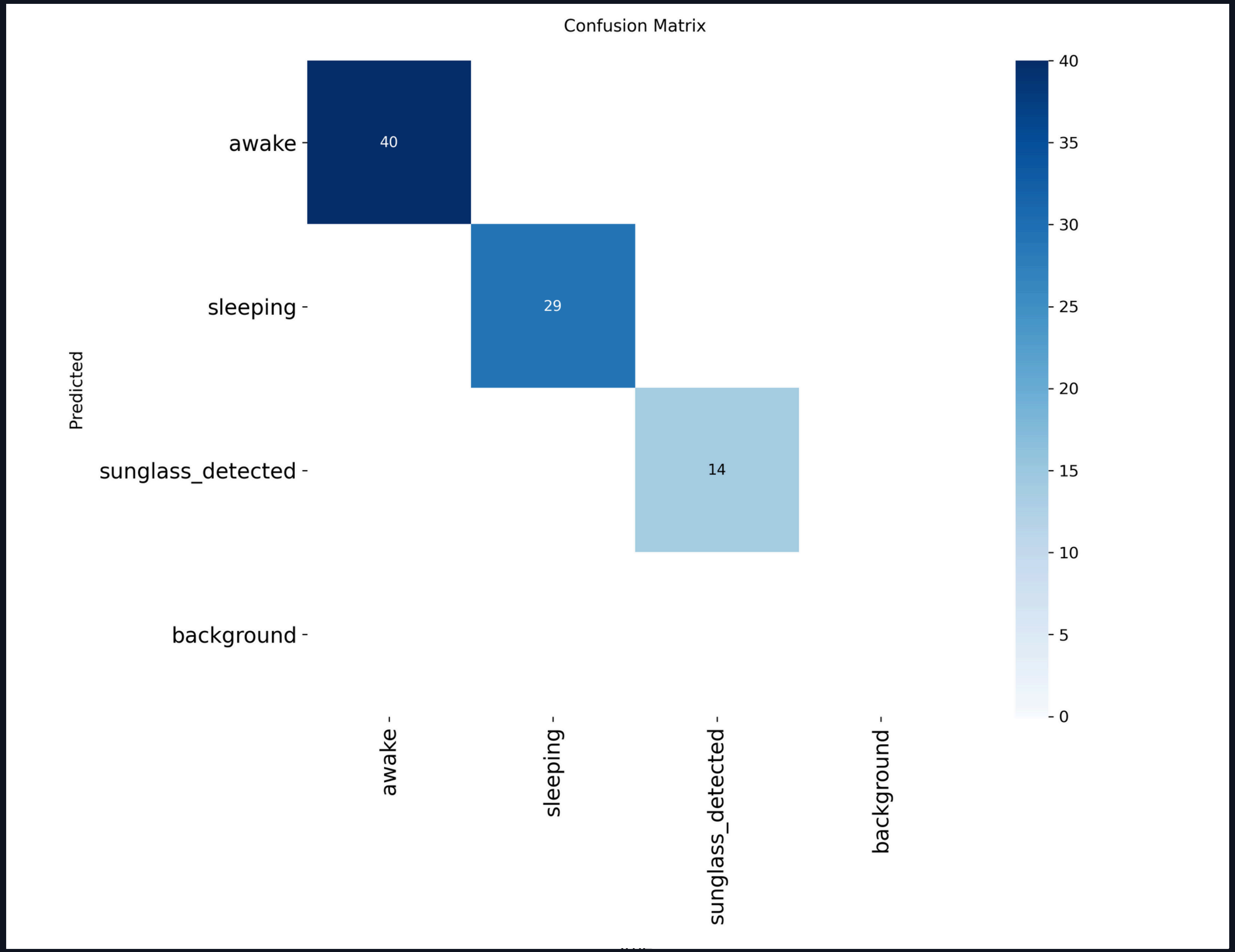
EVALUATION & RESULTS





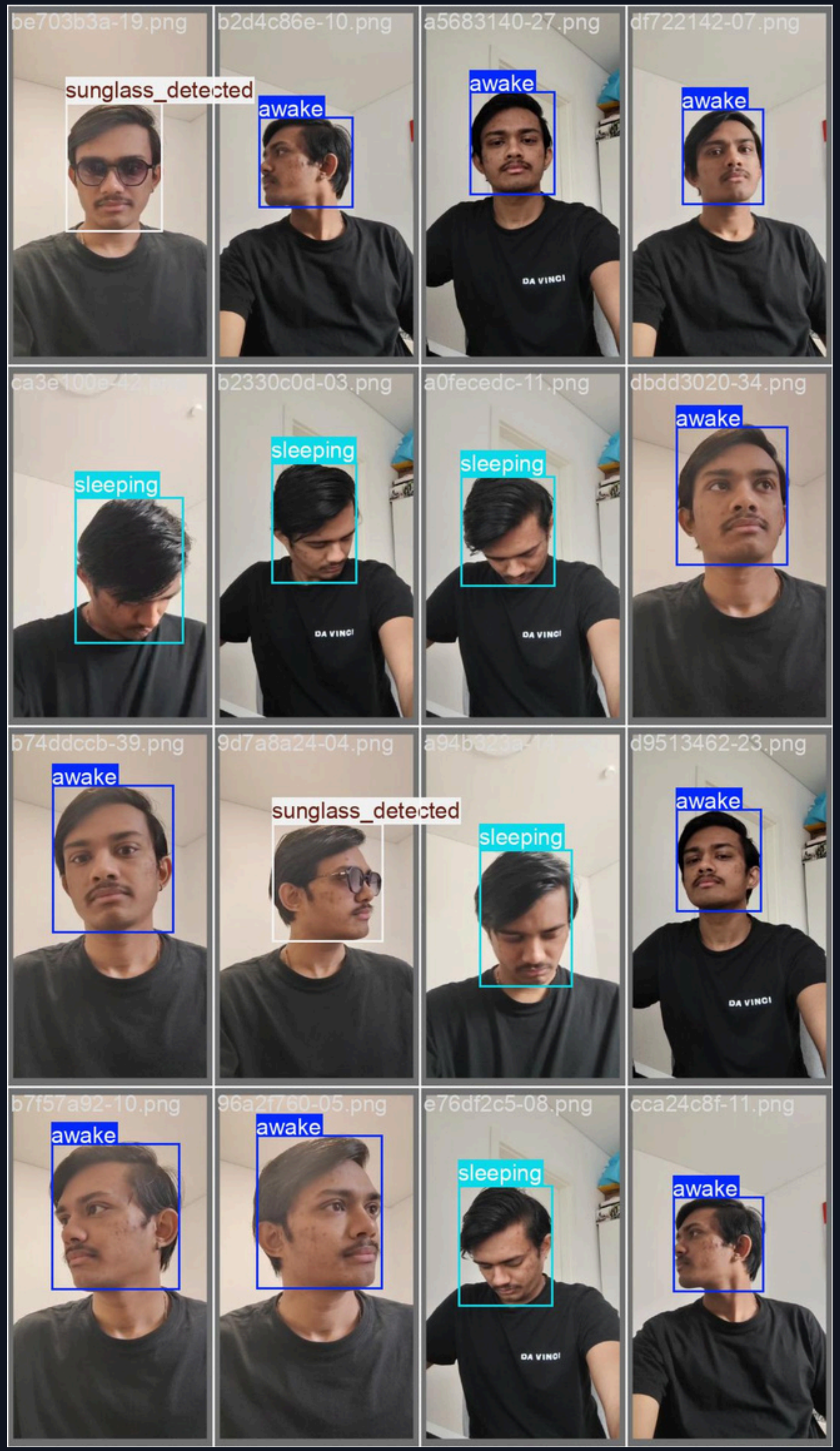
- **Evaluation Metrics**
  - Precision: Measures the correctness of positive predictions.
  - Recall: Measures the completeness of positive class detection.
  - mAP@0.5: Mean Average Precision at an IoU threshold of 0.5, used to evaluate overall model accuracy.
  - Loss Curves: Training and validation loss curves show how the model learns across epochs.
- **Training Behaviour**
  - The training process showed stable convergence, with a steady decrease in both training and validation loss.
  - No overfitting was observed, as the losses remained closely aligned, ensuring robust learning.
  - The model demonstrated strong generalization over multiple epochs, confirming its effectiveness.

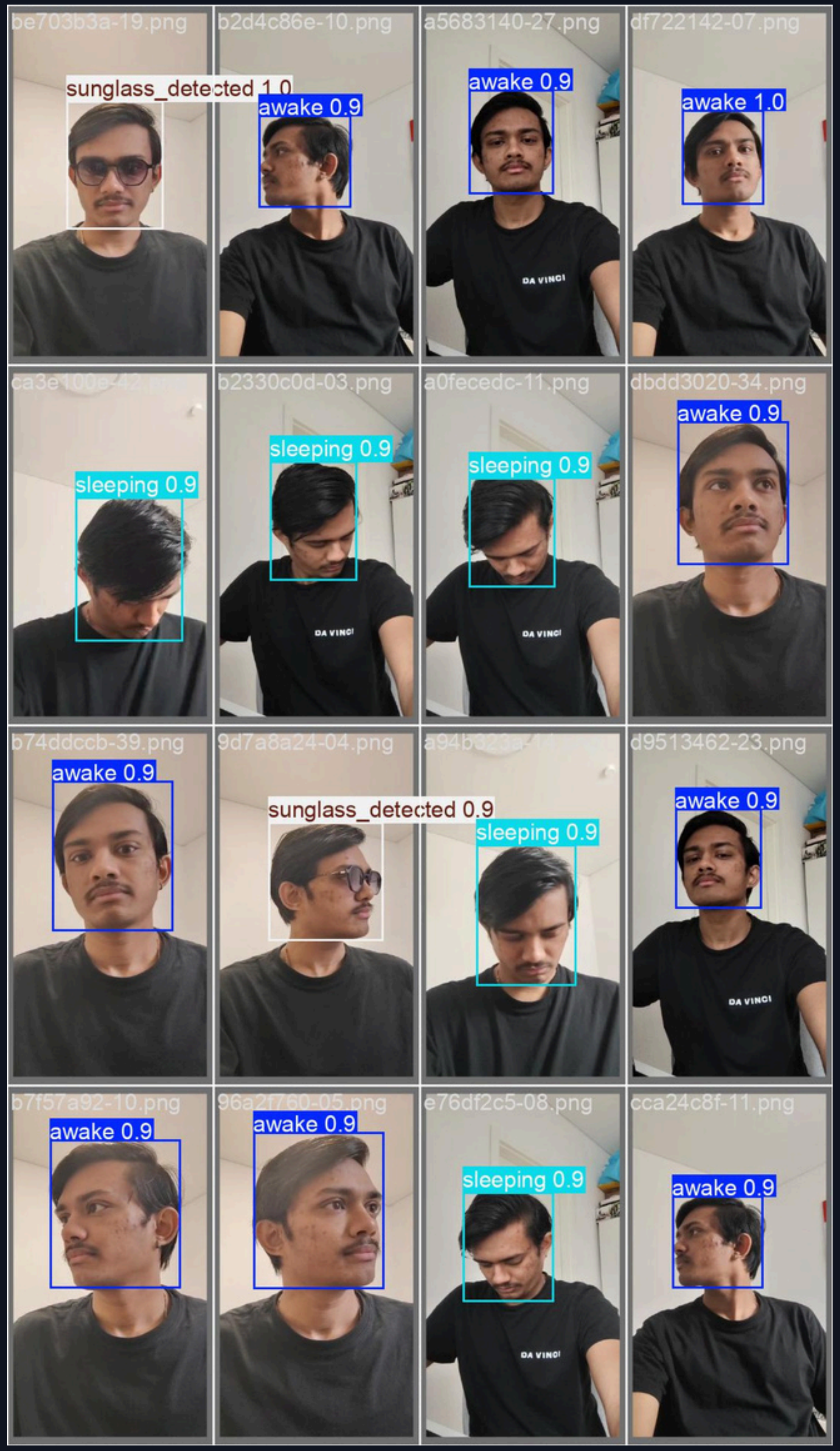
Confusion Matrix & Qualitative Results

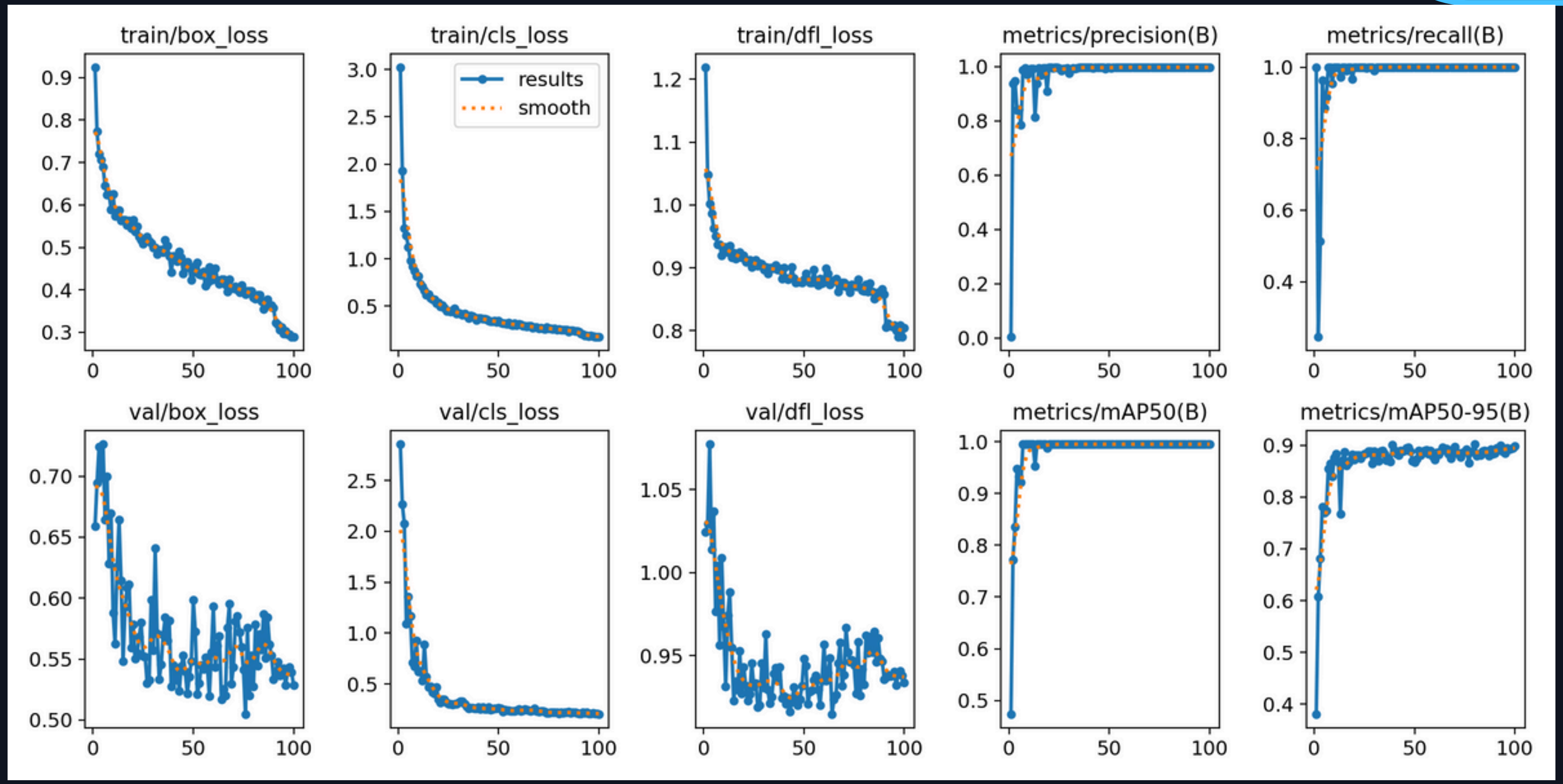


- **Confusion Matrix**
  - The confusion matrix shows strong diagonal dominance, indicating accurate classification for Awake and Sleeping states.
  - Minor confusion was observed between the Sleeping and Sunglass Detected classes due to partial eye visibility under certain conditions.
- **Qualitative Results**
  - Bounding box accuracy was consistently high, with precise localization of the driver's face in various scenarios.
  - The model demonstrated robust performance across different head poses and lighting conditions, ensuring that it could generalize well to unseen data.

Qualitative Results & .csv







	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	epoch	time	train/box_los	train/cls_los	train/dfl_loss	metrics/precisio	metrics/recall(I	metrics/mAP5(	metrics/mAP5(	val/box_loss	val/cls_loss	val/dfl_loss	lr/pg0	lr/pg1	lr/pg2
2	1	2.59884	0.92452	3.024	1.21938	0.00332	1	0.47389	0.38097	0.65905	2.85676	1.02443	0.00030009	0.00030009	0.00030009
3	2	4.52144	0.77264	1.93101	1.04838	0.93596	0.24538	0.77214	0.60773	0.69428	2.26333	1.02956	0.000608387	0.000608387	0.000608387
4	3	6.40659	0.71954	1.32144	1.00211	0.94763	0.51225	0.83562	0.6808	0.7239	2.07688	1.07732	0.000910459	0.000910459	0.000910459
5	4	8.33815	0.70657	1.24557	0.98656	0.83865	0.96326	0.94826	0.78149	0.69847	1.09331	1.01373	0.00120631	0.00120631	0.00120631
6	5	10.1831	0.68884	1.12253	0.96234	0.84116	0.88538	0.93321	0.76992	0.72625	1.35673	1.03667	0.00137241	0.00137241	0.00137241
7	6	12.0245	0.64489	0.9822	0.94995	0.78391	0.91639	0.92177	0.774	0.66423	1.16851	0.9764	0.00135826	0.00135826	0.00135826
8	7	13.836	0.62432	0.91906	0.93754	0.98675	1	0.995	0.85447	0.69984	0.70843	1.00411	0.00134412	0.00134412	0.00134412
9	8	15.6682	0.62536	0.87434	0.9369	0.99382	1	0.995	0.8641	0.6281	0.67335	0.95644	0.00132997	0.00132997	0.00132997
10	9	17.5492	0.58968	0.8134	0.91984	0.97027	0.95411	0.995	0.83989	0.66947	0.92502	1.00868	0.00131582	0.00131582	0.00131582
11	10	19.3734	0.62607	0.81641	0.93342	0.98575	1	0.995	0.87551	0.58801	0.62217	0.96774	0.00130168	0.00130168	0.00130168
12	11	21.1919	0.57324	0.73736	0.93305	0.99083	1	0.995	0.88309	0.56244	0.62525	0.93126	0.00128753	0.00128753	0.00128753
13	12	23.0264	0.58794	0.70445	0.92509	0.99158	1	0.995	0.86876	0.62589	0.53479	0.97405	0.00127338	0.00127338	0.00127338
14	13	24.8376	0.58753	0.66552	0.93457	0.813	0.97184	0.95279	0.76716	0.66403	0.88564	0.98803	0.00125923	0.00125923	0.00125923
15	14	26.6687	0.56316	0.62099	0.91699	0.93683	1	0.995	0.87118	0.61439	0.5715	0.95466	0.00124509	0.00124509	0.00124509
16	15	28.5418	0.56488	0.62217	0.9241	0.99496	1	0.995	0.88679	0.54819	0.47171	0.92296	0.00123094	0.00123094	0.00123094
17	16	30.3898	0.56431	0.58619	0.91439	0.96777	0.99071	0.995	0.8611	0.59256	0.47483	0.93453	0.00121679	0.00121679	0.00121679
18	17	32.1984	0.55275	0.57241	0.91695	0.99339	0.99841	0.995	0.86524	0.60837	0.42091	0.92937	0.00120265	0.00120265	0.00120265
19	18	34.0216	0.56261	0.574	0.92432	0.9949	1	0.995	0.87053	0.61145	0.40596	0.95287	0.0011885	0.0011885	0.0011885
20	19	35.8267	0.54541	0.5325	0.91614	0.90794	0.96619	0.98782	0.88205	0.55926	0.4676	0.92723	0.00117435	0.00117435	0.00117435
21	20	37.7168	0.56528	0.53687	0.91966	0.99562	1	0.995	0.87168	0.57848	0.34588	0.943	0.00116021	0.00116021	0.00116021
22	21	39.529	0.53809	0.49449	0.90912	0.99656	1	0.995	0.88122	0.55052	0.31503	0.93187	0.00114606	0.00114606	0.00114606
23	22	41.3615	0.54948	0.51248	0.91091	0.99338	1	0.995	0.88022	0.55483	0.3503	0.9227	0.00113191	0.00113191	0.00113191
24	23	43.1774	0.52507	0.48499	0.91206	0.99554	1	0.995	0.8742	0.56816	0.34998	0.92659	0.00111776	0.00111776	0.00111776
25	24	45.0025	0.51787	0.44796	0.90096	0.99567	1	0.995	0.88126	0.57999	0.31581	0.94549	0.00110362	0.00110362	0.00110362
26	25	46.8226	0.50836	0.44943	0.90609	0.99517	1	0.995	0.88272	0.55265	0.30348	0.93143	0.00108947	0.00108947	0.00108947
27	26	48.7028	0.5142	0.43868	0.912	0.98525	0.99741	0.995	0.88404	0.55187	0.30587	0.93316	0.00107532	0.00107532	0.00107532
28	27	50.5197	0.52509	0.45441	0.90271	0.99285	1	0.995	0.88844	0.53018	0.29737	0.91883	0.00106118	0.00106118	0.00106118
29	28	52.3534	0.51473	0.47602	0.90723	0.99118	1	0.995	0.88788	0.53277	0.30317	0.92014	0.00104703	0.00104703	0.00104703
30	29	54.168	0.50977	0.42276	0.90662	0.99348	1	0.995	0.864	0.59868	0.30546	0.9453	0.00103288	0.00103288	0.00103288
31	30	55.9959	0.49912	0.42483	0.89655	0.97569	0.98977	0.995	0.88875	0.55711	0.31787	0.93048	0.00101873	0.00101873	0.00101873
32	31	57.8832	0.49902	0.42313	0.90102	0.99053	1	0.995	0.88057	0.64091	0.33184	0.96272	0.00100459	0.00100459	0.00100459
33	32	59.7047	0.48372	0.40872	0.89123	0.9951	1	0.995	0.86989	0.57078	0.3294	0.92938	0.00099044	0.00099044	0.00099044

Real-Time Deployment

IMX500 Model Export

- The trained YOLO model was exported to the Sony IMX500 format using the yolo_export.py script.
- INT8 quantization was applied to make the model efficient for embedded processing.

RPK Deployment

- The model was converted into RPK format using the imx500-package tool for Raspberry Pi deployment.
- Command used:
  - imx500-package -i packerOut.zip -o network.rpk

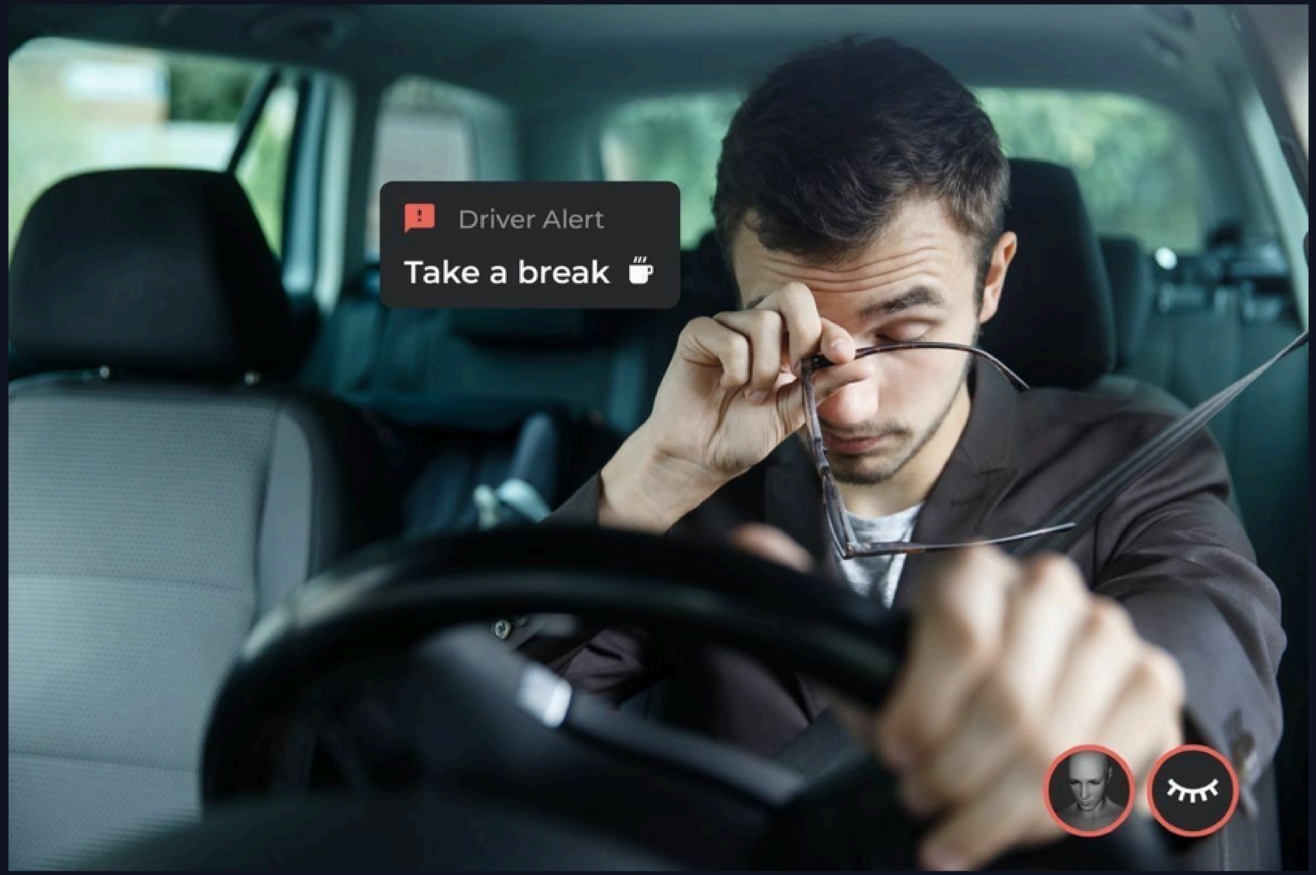
Raspberry Pi Integration

- The RPK model was deployed on the Raspberry Pi for real-time driver monitoring.
- Picamera2 was used for camera interfacing and live feed processing.

source image

Limitations

- Limited dataset size: The dataset is relatively small, which may affect the model's ability to generalize to unseen data in real-world situations.
- No infrared (IR) support: The system does not currently support infrared vision, which limits its effectiveness in low-light or night-time driving conditions.
- Sunglasses issue: Sunglasses can obscure eye features, reducing detection accuracy and causing false negatives.
- Single camera setup: The current system relies on a single camera, which can limit the field of view and reduce robustness under certain driving scenarios.



source image

Future Improvements

- Eye Aspect Ratio (EAR)-based detection: Implementing EAR can improve eye-closure detection, especially in cases of partial eye closure during drowsiness.
- Infrared camera support: Adding infrared cameras would enable the system to operate in night-time driving conditions, increasing robustness.
- Temporal models (LSTM / GRU): Introducing LSTM or GRU models would allow the system to track driver alertness over time, improving detection during continuous driving sessions.
- Vehicle CAN / Car2X integration: Integrating with CAN (Controller Area Network) or Car2X systems would allow the driver monitoring system to communicate with the vehicle's overall safety mechanisms, enhancing functionality.

Conclusion

- This project successfully demonstrates a complete end-to-end Edge AI pipeline for driver drowsiness detection.
- The system is highly automotive relevant, offering a real-time embedded solution that operates entirely on edge hardware, ensuring low latency and high efficiency.
- The YOLO-based model deployed on the Sony IMX500 provides an effective solution for driver monitoring systems, which can be integrated into Advanced Driver Assistance Systems (ADAS).

DEGGENDORF INSTITUTE OF TECHNOLOGY

THANK YOU

source image