

Maharaja Surajmal Institute Technology



Bachelor of Technology (2022-2026)
Department of Information Technology

Project - Presentation On

Real-Time Driver Inattention Detection and Alert Generation System

Mentor: Dr. Sitender Malik

Team Members:

Lakshya Prabhakar (02196303122)
Gautam Dahiya (35296303122)
Lakshay Joshi (02296303122)

INTRODUCTION

This project falls under the domain of Artificial Intelligence and Computer Vision, specifically focusing on driver monitoring systems. Using deep learning and image processing, the system detects various forms of driver inattention such as distraction, drowsiness, drinking, yawning, and dangerous driving behaviors.

Relevance to Real-World Problem or Application -

- This project directly contributes to road safety and accident prevention.
- It can be integrated into Advanced Driver Assistance Systems (ADAS) and smart vehicles.
- Helps in designing low-cost, lightweight, and real-time driver monitoring solutions suitable for both commercial and personal vehicles.

PROBLEM STATEMENT

Driver distraction and inattention are major causes of road accidents. Existing monitoring systems typically detect only a single behaviour (e.g., drowsiness) and are not suitable for lightweight, real-time deployment. This project aims to develop a fast and accurate vision based system that can automatically detect and classify multiple types of driver distraction from live camera input prevent potential accidents.

SCOPE

- The system will detect and classify multiple types of driver inattention (safe, distracted, dangerous, drinking, sleepy, yawning) from live camera input.
- Implementation will be based on computer vision and deep learning techniques to ensure accuracy and speed.

OBJECTIVES

1. Classify driver distraction into six behaviour classes: Safe Driving, Distracted Driving, Dangerous Driving, Drinking, Sleepy Driving, and Yawn.
2. Build a vision-based system that identifies these behaviours from camera input (prototype capable of running on recorded or live video).
3. Ensure high classification accuracy on the held-out test set through appropriate preprocessing, training, and evaluation.

LITERATURE SURVEY

- N. Mutlu, A. Sharma, and T. Daim, “Driver Distraction Detection Methods: A Literature Review and Framework,” IEEE Access, vol. 10, pp. 118523–118545, 2022.

The paper focuses on three main types:

Visual – eyes off the road

Manual – hands off the steering wheel

Cognitive – mind not focused on driving

LITERATURE SURVEY

- G. S. Krishna, K. Supriya, J. Vardhan, and M. Rao K, “Vision Transformers and YOLOv5 Based Driver Drowsiness Detection Framework,” Proceedings of the IEEE Conference on Data Science and Artificial Intelligence, IIIT Naya Raipur, India, 2022.

This paper introduced a two-stage model:

- **Stage 1:** YOLOv5 for face detection (extracting Region of Interest from driver video frames).
- **Stage 2:** Vision Transformer (ViT) for binary classification — “drowsy” vs. “alert.”

PROPOSED SOLUTION

The proposed solution is to build a machine learning–based system for recognising driver distraction and inattention from visual input. The approach includes:

- 1. Data-Driven Learning:** Using a labelled dataset of driver behaviours to train a model capable of distinguishing six classes (safe, distracted, dangerous, drinking, sleepy, and yawning).
- 2. Pre-Processing:** Applying standardisation and augmentation to improve model generalisation across diverse real-world conditions.
- 3. Deep Learning Framework:** Employing a Convolutional Neural Network (EfficientNet-B0) that balances accuracy with efficiency, making it suitable for potential real-time use.
- 4. Behaviour Classification:** The trained model classifies driver states, forming the basis of a monitoring system that can later be extended to raise alerts for unsafe behaviours.

PROPOSED SOLUTION

1. Use the trained model on frames extracted from videos to pre-process and pass through the network, which outputs a six-dimensional softmax probability vector corresponding to the behaviour classes: Dangerous Driving, Distracted Driving, Drinking, Safe Driving, Sleepy Driving, and Yawn.
2. An attention score is calculated :
 - a. Attention Score $\geq 70\%$ — Safe
 - b. $40\% \leq$ Attention Score $< 70\%$ — Mild Risk
 - c. Attention Score $< 40\%$ — Unsafe

MODEL COMPARISONS

EfficientNet outperforms other CNNs by scaling depth, width, and resolution more uniformly, giving better accuracy with far fewer parameters.

Below Is the accuracy-latency comparisons for the CNN models

Accuracy @ Latency		Accuracy @ Latency	
Resnet	86% @ 0.554s	Yolo – V5	88% @ 0.78s
EfficientNet – B0	90% @ 0.098s	EfficientNet – B0	90% @ 0.098s

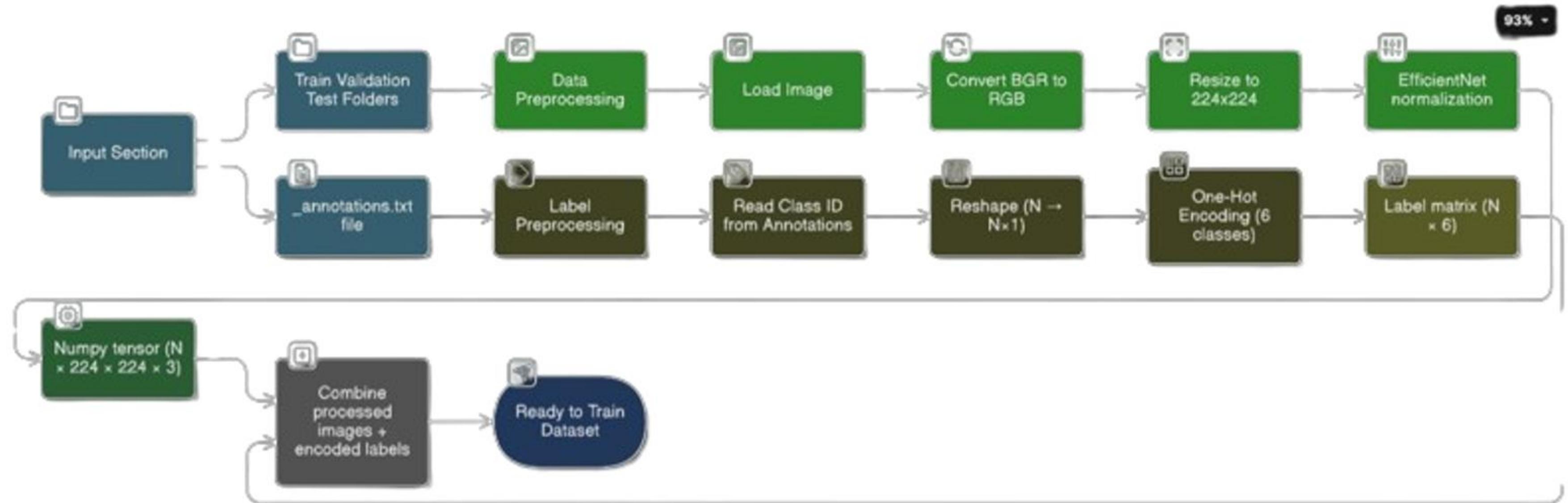
METHODOLOGY

1. The proposed solution is to build a machine learning-based system for recognising driver distraction and inattention from visual input. The approach includes:
 2. Data-Driven Learning: Using a labelled dataset of driver behaviours to train a model capable of distinguishing six classes (safe, distracted, dangerous, drinking, sleepy, and yawning).
 3. Pre-Processing: Applying standardisation and augmentation to improve model generalisation across diverse real-world conditions.
 4. Deep Learning Framework: Employing a Convolutional Neural Network (EfficientNet-B0) that balances accuracy with efficiency, making it suitable for potential real-time use.
 5. Behaviour Classification: The trained model classifies driver states, forming the basis of a monitoring system that can later be extended to raise alerts for unsafe behaviours.

TECHNOLOGY USED

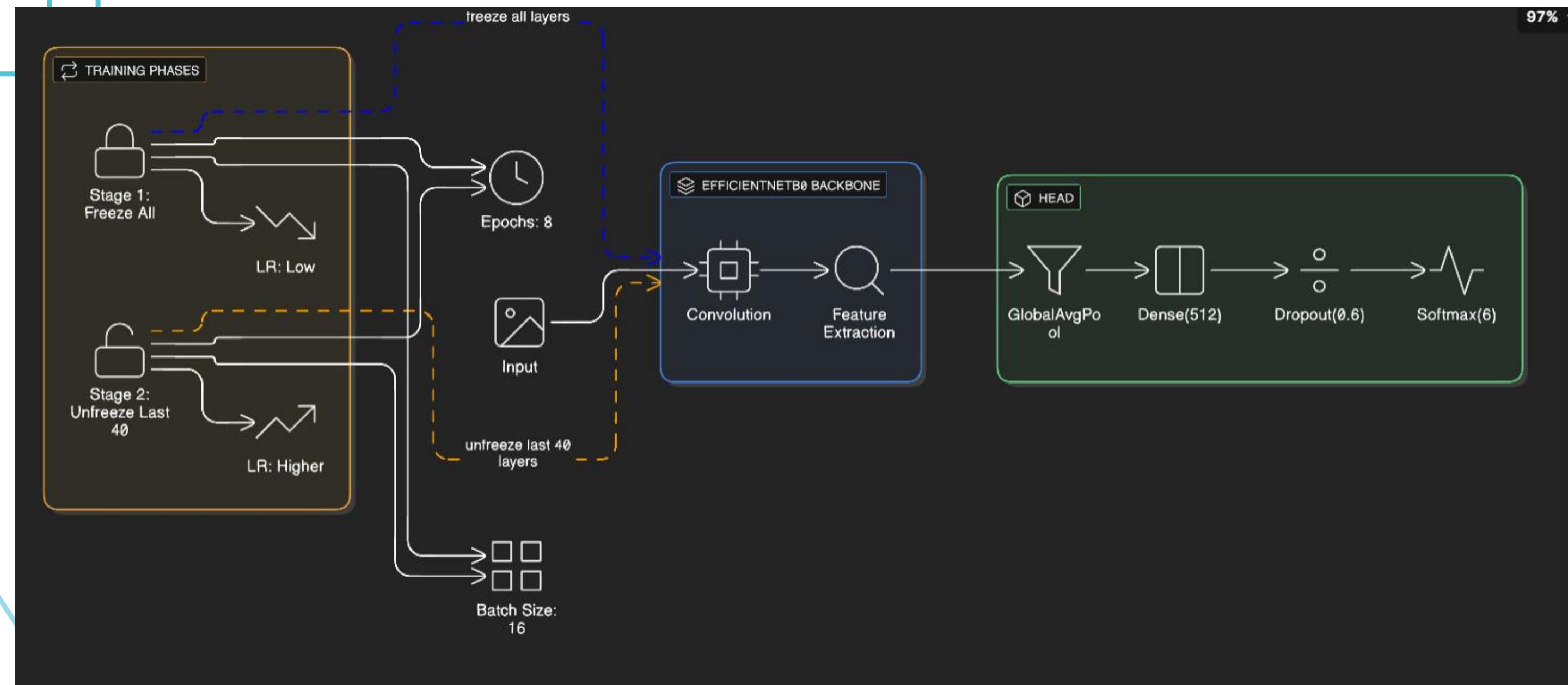
Component	Tool/Technology Used
Programming Language	Python
Inattention Detection	EfficientNet-B0
ML/DL Framework	PyTorch or TensorFlow
Development Environment	Google Collab

FLOW DIAGRAM



Pre Processing

MODEL ARCHITECTURE



MODEL ARCHITECTURE

Layer (type)	Output Shape	Param #
efficientnetb0 (Functional)	(None, 7, 7, 1280)	4,049,571
global_average_pooling2d (GlobalAveragePooling2D)	(None, 1280)	0
dense (Dense)	(None, 256)	327,936
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 6)	1,542

Total params: 5,038,007 (19.22 MB)

Trainable params: 329,478 (1.26 MB)

Non-trainable params: 4,049,571 (15.45 MB)

Optimizer params: 658,958 (2.51 MB)

MODEL ARCHITECTURE

Layer (type)	Output Shape	Param #
efficientnetb0 (Functional)	(None, 7, 7, 1280)	4,049,571
global_average_pooling2d (GlobalAveragePooling2D)	(None, 1280)	0
dense (Dense)	(None, 256)	327,936
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 6)	1,542

Total params: 5,038,007 (19.22 MB)

Trainable params: 329,478 (1.26 MB)

Non-trainable params: 4,049,571 (15.45 MB)

Optimizer params: 658,958 (2.51 MB)

DATASET

Driver Inattention Detection Dataset

The dataset is organized into three main directories:

Training Set (train): This directory contains 11,942 gray-scale images, carefully curated and labeled across the six classes.

Validation Set (validation): With 1,922 gray-scale images, this subset provides a means for fine-tuning models and evaluating their performance during development.

Test Set (test): Comprising 985 gray-scale images, this directory is reserved for final model evaluation and benchmarking.

The dataset encompasses six classes of driving behaviors:

1- Dangerous Driving: Gray-scale images capturing instances of reckless or hazardous driving behavior, such as speeding or erratic lane changes.

2- Distracted Driving: Instances where the driver's attention is diverted away from the road, possibly due to smartphone usage, eating, or interacting with passengers.

3- Drinking: Gray-scale images depicting drivers consuming alcoholic beverages while behind the wheel, highlighting the dangers of driving under the influence.

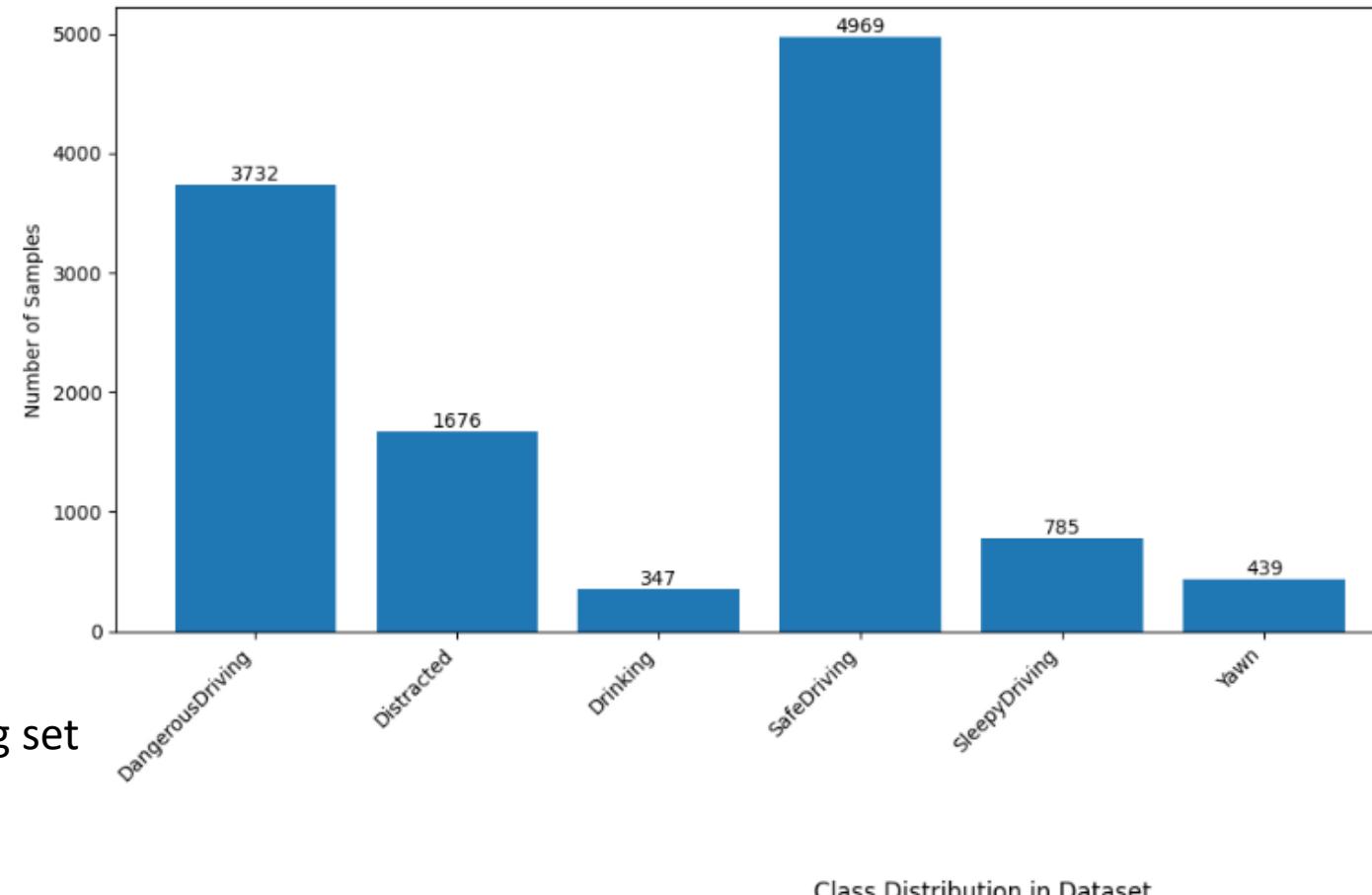
4- Safe Driving: Examples of responsible and cautious driving behavior captured in gray-scale, including obeying traffic laws, maintaining safe distances, and using turn signals.

5- Sleepy Driving: Instances where drivers exhibit signs of drowsiness or fatigue, posing a significant risk of accidents due to reduced alertness, depicted in gray-scale.

6- Yawn: Gray-scale images capturing drivers in the act of yawning, often indicative of fatigue or tiredness, which can impair driving performance.

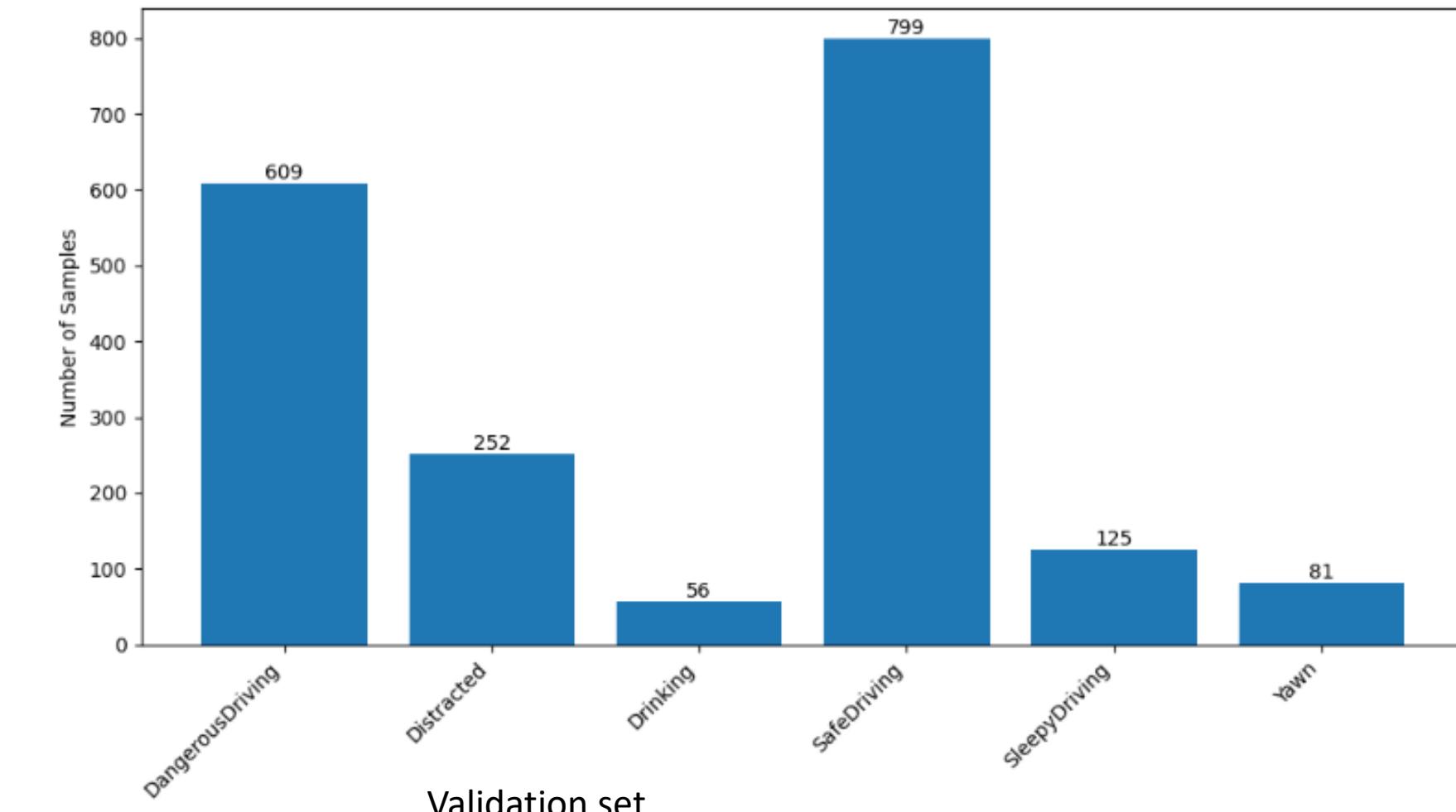
DATASET

Class Distribution in Dataset



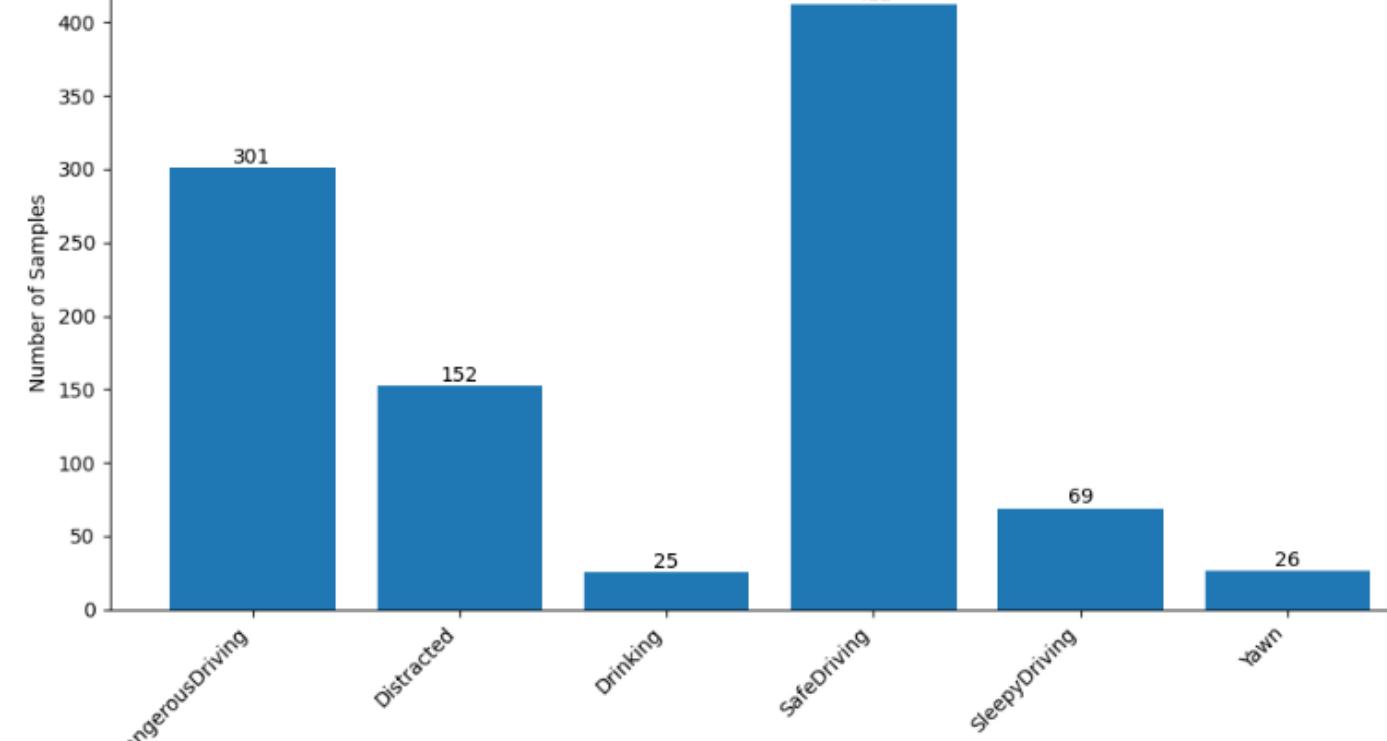
Training set

Class Distribution in Dataset



Validation set

Class Distribution in Dataset



Test set

WORK DONE

Label Preprocessing

1) Extract Class IDs -

- Each annotation contains a `class_id` representing driver action:
- e.g., safe driving, phone, eating, etc.

2) Reshape Labels -

- Labels reshaped from $(N,)$ $\rightarrow (N, 1)$
- Required for scikit-learn encoders.

3) One-Hot Encoding

- Convert integer class labels into one-hot vectors using `OneHotEncoder`.
- Example (for 6 classes):
- $2 \rightarrow [0, 0, 1, 0, 0, 0]$.

4) Consistency Across Splits-

- Same encoding logic applied to train, validation, and test labels.

5) Final Label Shape -

- Stored in matrix form: $(N, 6)$
- Compatible with softmax cross-entropy loss.

WORK DONE

Model Architecture (EfficientNet-B0)

1) Base Model: EfficientNet-B0 -

- Each Pretrained on ImageNet (1.2M images).
- Used as a feature extractor.
- Lightweight and fast; ideal for mobile/real-time environments.

2) Base Model Configuration-

- `include_top = False` to remove ImageNet classifier.
- Input size: $224 \times 224 \times 3$.
- Outputs rich feature maps via convolutional layers.

3) Added Custom Classification Head-

- `GlobalAveragePooling2D` and converts feature maps into a single 1D vector.
- `Dense(512, ReLU)` and adds learnable high-level abstraction.
- `Dropout(0.6)` and Reduces overfitting by randomly dropping neurons.
- `Dense(6, Softmax)` and Final classification layer for 6 driver-action classes.

4) Regularization -

- L2-regularization on dense layer to avoid overfitting.

WORK DONE

Training Stage 1 (Feature Extraction)

1) Freeze Entire EfficientNet-B0 -

- All convolutional layers frozen.
- Only custom head layers get trained.

2) Optimizer & Hyperparameters-

- Adam optimizer, LR = 0.001.
- Loss: categorical cross-entropy.

3) Data Augmentation Applied-

- Rotation, shifting, zoom, flipping.
- Helps generalization with limited data.

4) Callbacks -

- ReduceLROnPlateau for dynamic learning rate.
- EarlyStopping to prevent overfitting..

WORK DONE

Fine-Tuning Stage 2 (Unfreezing Top Layers)

1) Unfreeze Last Few Layers -

- EfficientNet-B0 has ~237 layers.
- Unfreezing the last ~40 layers allows learning deeper features.

2) Very Low Learning Rate -

- LR = 2e-5 (to avoid damaging pretrained weights).
- Only slight adjustments to convolutional filters.

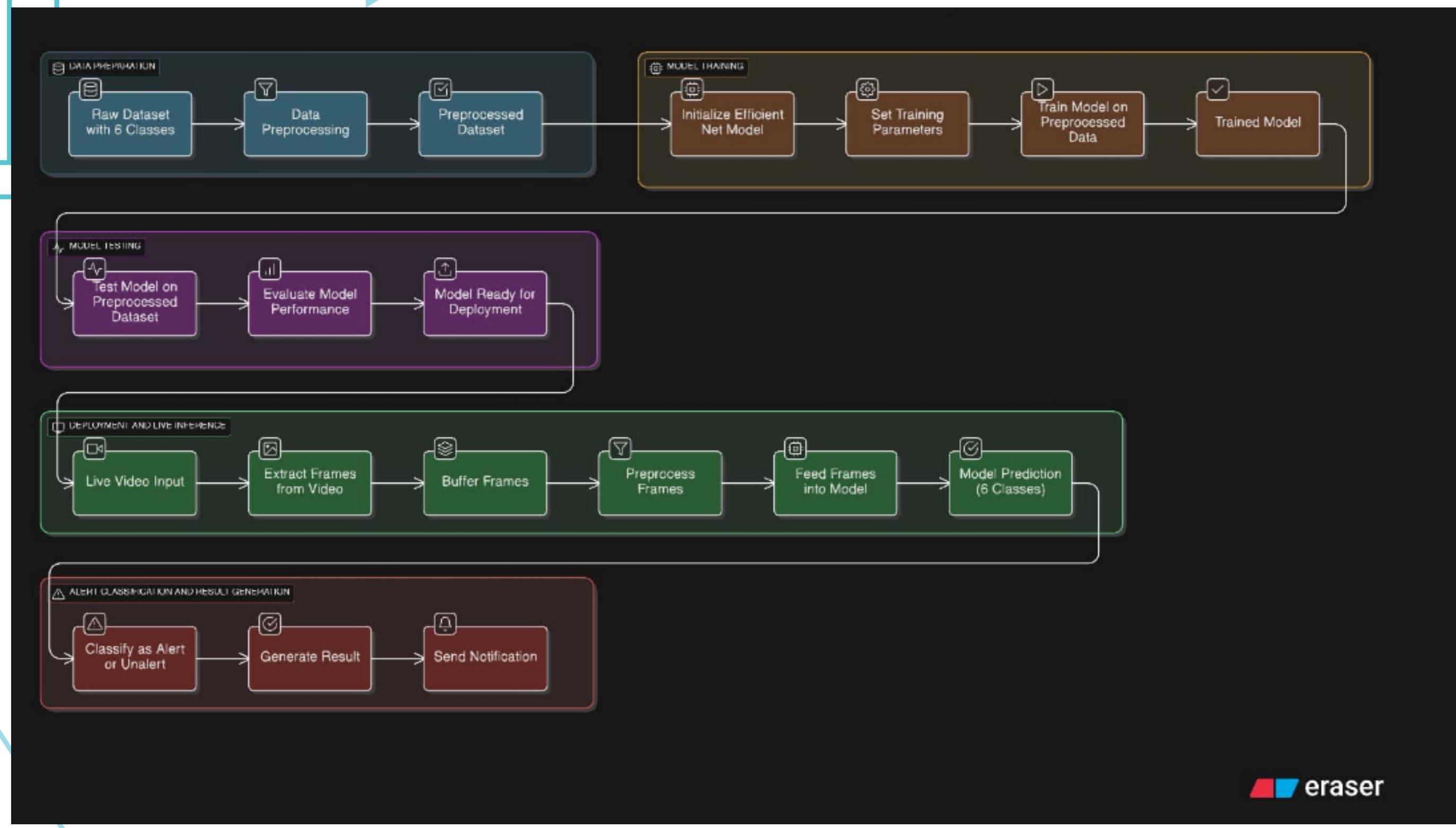
3) Continue Training with Augmentation-

- Same datagen flow used to maintain variation.
- Same class weights applied for imbalance correction

4) Benefits of Fine-Tuning -

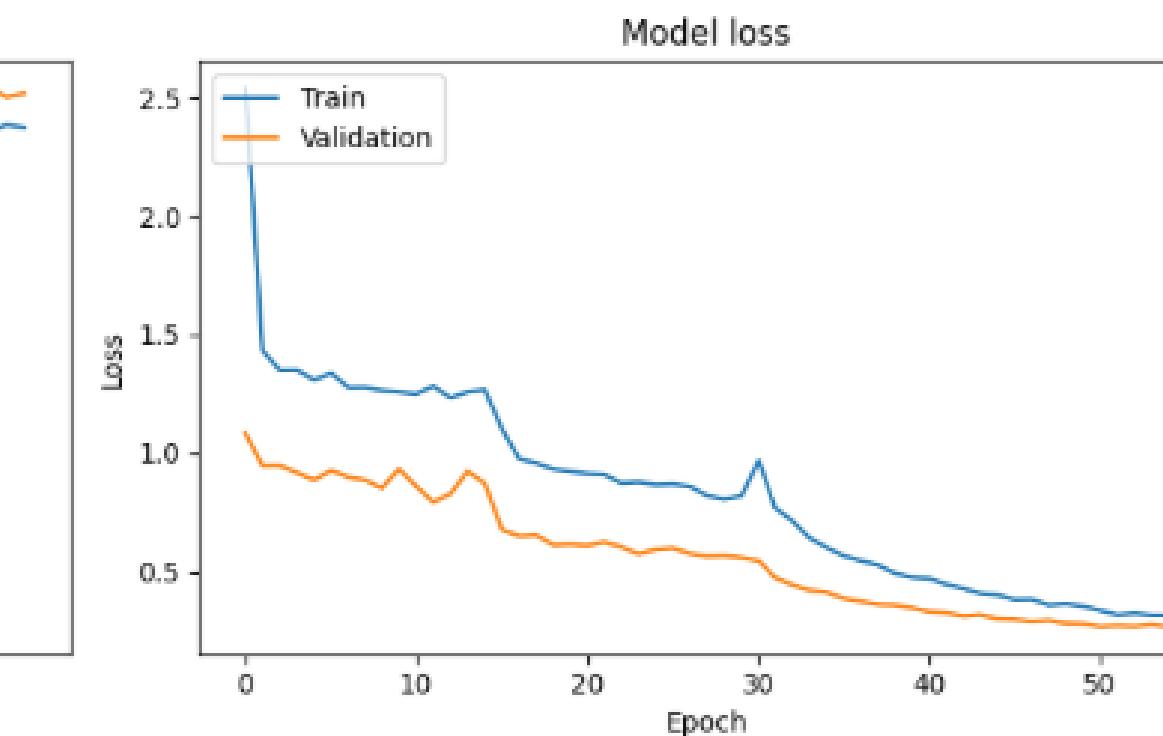
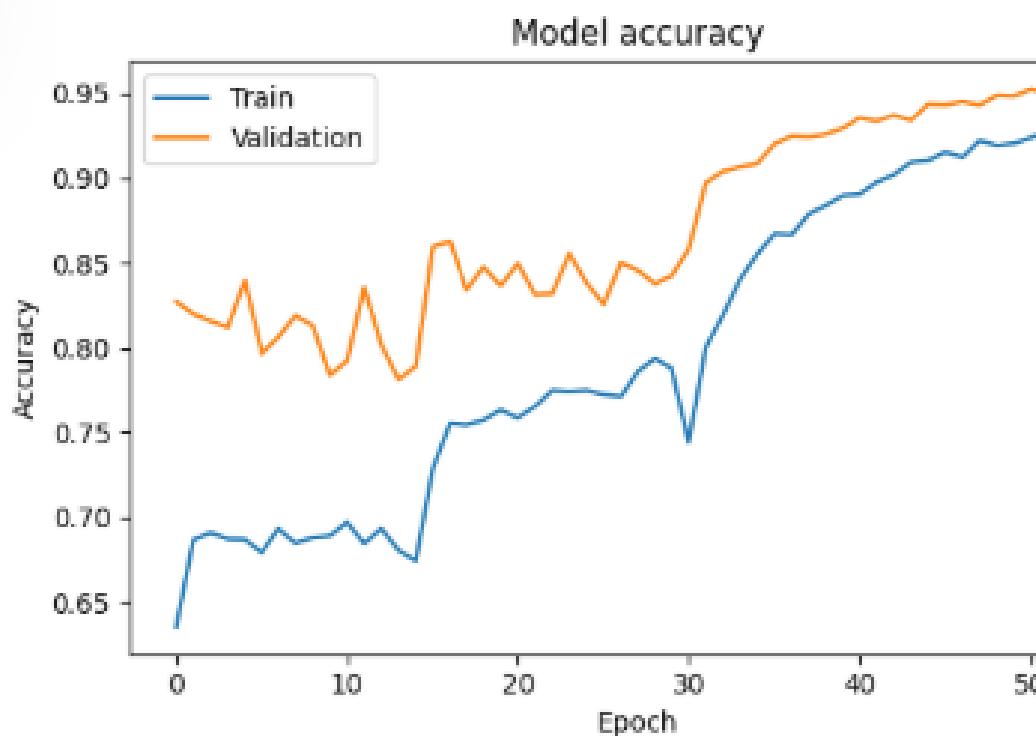
- Improves feature representation for your custom dataset.
- Boosts accuracy and reduces validation loss.

SYSTEM ARCHITECTURE



EVALUATION SCORES

Accuracy: 0.9036
Precision (macro): 0.8615
Recall (macro): 0.8958
F1-score (macro): 0.8772



RESULT



RESULT



RESULTS



CHALLENGES FACED

- **Model Selection:** Finding a deep learning model that is both accurate and lightweight enough for real-time use on onboard hardware, not just a high-performing theoretical model.
- **Dataset Issues:** Class imbalance and noisy annotations can affect training quality and model generalisation.
- **Preprocessing Variability:** Differences in lighting, driver posture, and camera placement may reduce robustness.
- **Finetuning :** Finetuning the model to balance both accuracy and latency for real time system

CONCLUSION

The project successfully identifies various driver activities and classifies it into inattention and attention based on scores generated. It generated warnings based on the danger level

REFERENCES

- C. Chitraranjan, V. Vipulanathan, and T. Sritharan, “Vision-Based Collision Warning Systems with Deep Learning: A Systematic Review,” *Journal of Imaging*, vol. 11, no. 64, pp. 1–25, 2025.
- A. Abutahoun, S. Glaser, M. Elhenawy, A. Rakotonirainy, N. Watson-Brown, and T. I. Alhadidi, “Technological Interventions to Improve Novice Driver Safety: A Review of Current Technologies and Future Directions,” *IEEE Transactions on Intelligent Transportation Systems*, 2025.
- G. S. Krishna, K. Supriya, J. Vardhan, and M. Rao K, “Vision Transformers and YOLOv5 Based Driver Drowsiness Detection Framework,” *Proceedings of the IEEE Conference on Data Science and Artificial Intelligence*, IIIT Naya Raipur, India, 2022.
- N. Mutlu, A. Sharma, and T. Daim, “Driver Distraction Detection Methods: A Literature Review and Framework,” *IEEE Access*, vol. 10, pp. 118523–118545, 2022.
- I. Kotseruba and J. K. Tsotsos, “Attention for Vision-Based Assistive and Automated Driving: A Review of Algorithms and Datasets,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 23456–23480, 2022.

THANK YOU