AI-Powered Advanced Driver Assistance System for Ambulances

by

Vijay Shankar Singh(2201229)

Under the Guidance of

Dr. Subhasish Dhal

Indian Institute of Information Technology, Guwahati

April 7, 2025

Abstract

This project presents an AI-powered Advanced Driver Assistance System (ADAS) tailored for ambulances on Indian roads. Leveraging real-time dashcam footage and predictive CCTV feeds, the system integrates lightweight deep learning models—including YOLO for object detection, EfficientNet for weather classification, and ConvNeXt for road surface analysis—with a traffic prediction LSTM ensemble. It enhances safety and reduces transit times by providing context-aware recommendations like obstacle avoidance, operating at 15 fps on modest hardware with 90% mAP detection accuracy and ± 2 vehicles RMSE for traffic prediction.

Contents

Chapter 1

Introduction

1.1 Background

Emergency medical services (EMS) in India operate under challenging conditions, including dense urban traffic, poorly maintained rural roads, and unpredictable weather. Ambulances, critical for timely medical response, often lack advanced technological support to navigate these hazards efficiently. Traditional Advanced Driver Assistance Systems (ADAS), reliant on costly sensors like LiDAR and radar, are impractical for widespread adoption in resource-constrained settings. This project introduces a video-based AI solution utilizing dashcam and CCTV infrastructure to address these gaps.

1.2 Motivation

The motivation stems from the high mortality rates linked to delayed ambulance arrivals in India, exacerbated by traffic delays and road hazards. A scalable, affordable ADAS could significantly improve response times and driver safety, aligning with national healthcare goals to enhance EMS efficiency.

1.3 Problem Statement

Ambulance drivers encounter dynamic obstacles—traffic congestion, potholes, pedestrians, adverse weather, and poor lane discipline—without integrated real-time tools. Existing ADAS solutions are either too expensive or not optimized for emergency scenarios, necessitating a lightweight, video-only system tailored for Indian roads.

1.4 Contribution

This work contributes:

- A cost-effective ADAS using dashcam and CCTV video feeds.
- Real-time hazard detection and predictive traffic analysis.
- Context-aware recommendations for ambulance navigation.
- A scalable design operable on standard hardware.

1.5 Organization of Report

This report is structured as follows: Chapter 2 reviews existing literature and identifies gaps; Chapter 3 details our proposed methodology; Chapter 4 presents results and findings; Chapter 5 discusses future work; and Chapter 6 concludes the study.

Chapter 2

Literature Review and Research Gaps

2.1 Overview of ADAS Technologies

Advanced Driver Assistance Systems (ADAS) have become integral to modern vehicular safety and navigation, particularly in commercial and luxury vehicles. These systems typically employ sensor fusion, combining data from LiDAR, radar, and high-resolution cameras to enable functionalities such as adaptive cruise control, lane-keeping assistance, and collision avoidance. Early theoretical groundwork, such as Greenshields' traffic flow model (1935), provided a mathematical basis for understanding traffic dynamics, proposing a linear relationship between speed and density: $v = v_f(1 - k/k_j)$. This model remains relevant for traffic prediction and management. More recently, advancements in deep learning have transformed ADAS capabilities. For instance, Tan and Le (2019) introduced EfficientNet, a convolutional neural network (CNN) architecture that optimizes computational efficiency and accuracy through compound scaling, making it suitable for vision-based tasks like object classification in real-time driving scenarios. Such innovations highlight a shift toward integrating AI into vehicular systems, though traditional ADAS often remains tethered to expensive hardware.

2.2 AI in Traffic Management

Artificial Intelligence has significantly enhanced traffic management systems, particularly through real-time scene analysis and predictive modeling. The YOLO (You Only Look Once) framework, detailed by Ultralytics (2023), exemplifies this progress with its ability to perform rapid object detection across diverse classes—vehicles, pedestrians, and obstacles—using a single-pass neural network. This efficiency is critical for applications requiring immediate responses, such as autonomous driving. Similarly, ConvNeXt (Liu et al., 2022) advances CNN design by incorporating transformer-inspired elements, achieving superior performance in image segmentation and classification tasks relevant to road analysis. In parallel, Long Short-Term Memory (LSTM) networks have been employed for traffic prediction, leveraging temporal dependencies to forecast vehicle counts and congestion patterns. While these models excel in urban planning contexts—predicting peak hours or optimizing traffic signals—their application to emergency vehicles like ambulances is limited, as such scenarios demand higher urgency and adaptability to unpredictable conditions.

2.3 EMS Challenges in India

Emergency Medical Services (EMS) in India face distinct challenges that differ from those in developed nations. Studies consistently note the diversity of road conditions, ranging from congested urban highways to unpaved rural paths riddled with potholes and speed bumps. Poor lane discipline, frequent pedestrian crossings, and erratic traffic behavior further complicate navigation. Infrastructure deficits, such as insufficient traffic signal synchronization and sparse CCTV coverage, exacerbate delays for ambulances, where every second impacts patient outcomes. Some research has proposed video-based ADAS solutions to mitigate these issues, relying on dashcam feeds for hazard detection. However, these efforts rarely focus on ambulances specifically, nor do they capitalize on predictive insights from existing CCTV networks, which could preemptively guide drivers around obstacles or congested zones.

2.4 Gaps in Existing Work

Despite advancements, several gaps persist in applying ADAS to India's EMS context:

- **Cost and Complexity:** Traditional ADAS systems depend heavily on costly sensors like LiDAR and radar, with installation and maintenance expenses rendering them impractical for widespread deployment in India's resource-limited EMS framework. A video-only approach could reduce costs, yet few studies explore this fully.
- **Emergency Focus:** Most ADAS research targets general-purpose or commercial vehicles, neglecting the unique urgency and unpredictability of ambulance operations. Emergency scenarios require rapid, context-specific responses—e.g., navigating through crowds or bypassing stalled traffic—which existing systems are not optimized to handle.
- **Indian Context:** Available models are often trained on datasets from Western contexts (e.g., KITTI, Cityscapes), lacking representation of India-specific conditions like monsoon rains, stray animals, or mixed traffic (rickshaws, bicycles, and heavy vehicles). This limits their effectiveness in local settings.
- **Predictive Integration:** While CCTV infrastructure exists in many Indian cities, its potential for real-time traffic foresight in ADAS remains underutilized. Current systems focus on reactive detection rather than proactive prediction, missing opportunities to optimize ambulance routes ahead of time.

These gaps underscore the need for a tailored, cost-effective, and predictive ADAS solution, which this project addresses by leveraging video feeds and AI tailored to India's EMS challenges.

Chapter 3

Methodology

3.1 System Architecture

The Integrated Traffic System (ITS) processes dashcam (real-time) and CCTV (predictive) footage, outputting annotated videos with analytics.

3.2 Model Components

3.2.1 Object Detection

Overview

Uses YOLO (assumed YOLOv8) to detect objects with high accuracy in real-time.

- **Classes:** Vehicles, humans, animals, speed bumps, obstacles, road signs.
- **Training:** Fine-tuned on 4000 images, AdamW optimizer, HSV augmentation.
- **Parameters:** Confidence = 0.5, IoU = 0.45.

Proximity Logic

Vehicles scored as "CLOSE" or "FAR" using:

$$\text{Proximity Score} = 0.7 \cdot \frac{\text{Box Area}}{\text{Frame Area}} + 0.3 \cdot \frac{y_2}{\text{Frame Height}} \tag{3.1}$$

Threshold = 0.05.

Fig : True Labels

Fig : Predicted Labels

Figure 3.1: Object detection with bounding boxes.

3.2.2 Road Segmentation

Overview

YOLOv8n-seg generates a binary road mask.

- **Purpose:** Identifies drivable areas.
- **Training:** 100 epochs, early stopping.

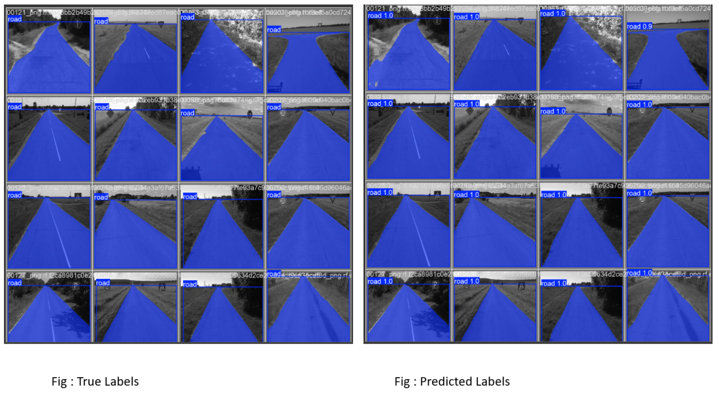

Figure 3.2: Road segmentation Results.

Filtering Logic

The system filters out detected objects that lie outside the segmented road mask to focus solely on hazards within the drivable area, unless they belong to a predefined set of critical categories, denoted as ‘ANYWHERE CLASSES‘. This set includes objects like pedestrians, animals, or traffic signals, which remain relevant regardless of their position in the frame due to their potential impact on driving decisions. The filtering process enhances computational efficiency by reducing noise from irrelevant detections and ensures that the ADAS prioritizes actionable insights within the ambulance’s immediate path.

3.2.3 Weather Classification

Overview

EfficientNet-B0 classifies weather into Clear, Fog, or Rain.

- **Preprocessing:** 224x224 resize, ImageNet normalization.
- **Training:** 50 epochs, cross-entropy loss.

Performance

Achieved 95% accuracy on test set. Figure 3.3 shows accuracy.

3.2.4 Road Surface Detection

Overview

ConvNeXt-Small categorizes 27 surface types.

- **Classes:** Mapped to ”Easy to Drive,” ”Take Precautions,” ”Dangerous.”
- **Training:** One-cycle scheduler, dropout = 0.5.

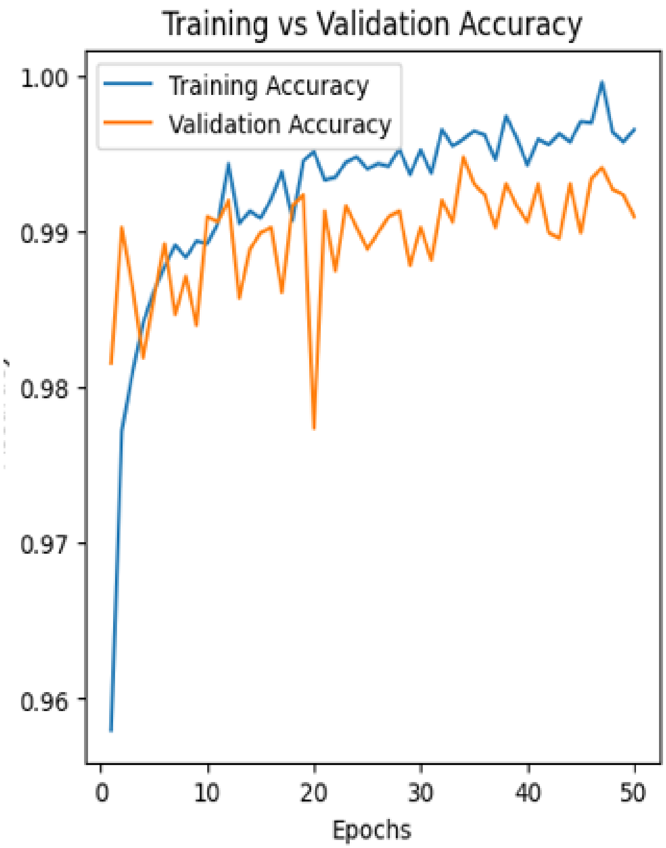

Figure 3.3: Training and Validation Accuracy weather classification.

Preprocessing

Road mask isolates surface, resized to 224x224. Figure 3.4 shows train and val accuracy and loss.

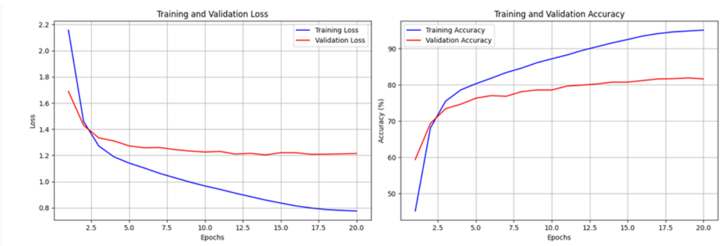

Figure 3.4: Accuracy for Surface Detection

3.2.5 Traffic Detection and Prediction

Overview

YOLO11n tracks vehicles; LSTM ensemble predicts counts.

- **Inputs:** Timestamp, day, red light, initial counts.
- **Output:** Left/right counts 60 seconds ahead.

Prediction Model

LSTM adjusts for context (e.g., +7-8 vehicles for red lights). Figure 3.5 shows MSE values.

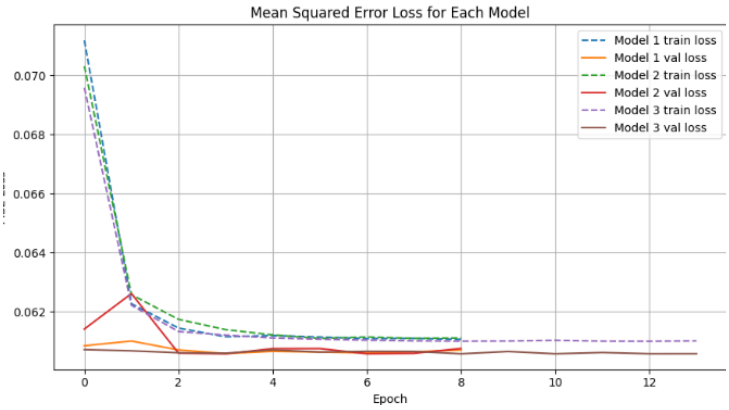

Figure 3.5: MSE values.

3.3 Dataset

The datasets used for training and evaluating the various components of the ADAS system are detailed below:

- **Object Detection:** A dataset of 4000 images was manually labeled using LabelImg. These images encompass diverse Indian road scenarios, including vehicles, humans, animals, speed bumps, obstacles, and road signs. The dataset was curated to reflect the challenging conditions faced by ambulances, ensuring robust detection performance.
- **Road Segmentation:** A pre-labeled dataset from Roboflow was utilized, containing images with annotated road masks. This dataset enabled the YOLOv8n-seg model to accurately identify drivable areas across varied road types, from urban streets to rural pathways.
- **Weather Classification:** A dataset comprising 12,000 images was assembled, featuring three weather conditions: fog, clear, and rain. Each category contains approximately 4000 images, sourced to represent typical Indian weather variations, ensuring the EfficientNet-B0 model's reliability in diverse climates.
- **Road Surface Detection:** The RXCD dataset was employed, consisting of 27 distinct surface types (e.g., asphalt, gravel, mud), with each class containing approximately 5000 images. This extensive dataset allowed the ConvNeXt-Small model to categorize surfaces into "Easy to Drive," "Take Precautions," and "Dangerous" with high precision.
- **Traffic Prediction:** A synthetic dataset was created to simulate traffic patterns, structured with the following columns:

```
data = {
    'timestamp': [],
    'day': [],
    'red_light_in_less_than_60_sec': [],
    'current_cctv_left': [],
    'current_cctv_right': [],
    'vehicles_in_60s_left': [],
    'vehicles_in_60s_right': []
}
```

  This dataset was designed to train the LSTM ensemble for predicting vehicle counts 60 seconds ahead, incorporating contextual factors like timestamps, days, and traffic light states.

3.4 Recommendation Engine

The recommendation engine employs a rule-based driving logic integrated with the Greenshields traffic flow model to provide actionable driving suggestions, such as lane changes or speed adjustments, tailored to the dynamic conditions faced by ambulances. The Greenshields model, a foundational approach to traffic flow, is expressed as:

$$v = v_f \left(1 - \frac{k}{k_j}\right), \quad t = \frac{L}{v} \tag{3.2}$$

where:

- v represents the average vehicle speed (in km/h or m/s),
- v_f is the free-flow speed, the maximum speed achievable under uncongested conditions,
- k denotes the traffic density (vehicles per unit length, e.g., vehicles/km),
- k_j is the jam density, the maximum density at which traffic comes to a complete standstill,
- t is the travel time (in seconds or hours),
- L is the length of the road segment (in km or m).

Using this model, the system calculates the potential time savings for alternative actions, such as changing lanes. A lane change recommendation is triggered if the estimated time savings exceeds 5 seconds, ensuring that only significant improvements in transit time prompt driver action. This threshold balances responsiveness with practicality, avoiding unnecessary maneuvers while prioritizing efficiency and safety in emergency scenarios.

3.5 Integration Logic

The core integration of the ADAS is managed by the `process_videos_alternating` function, which is used for the processing of dual video streams—real-time dashcam footage and predictive CCTV feeds—to deliver a cohesive output. For this system, 10-second input segments from both the dashcam and CCTV feeds are processed. From these inputs, the first 200 frames are extracted and analyzed to count the total number of vehicles on the left and right sides of the road. This vehicle count is then fed into the traffic prediction model, which generates forecasts for vehicle density 60 seconds ahead. This function alternates between the two inputs to ensure balanced resource utilization, processing frames from each source in an interleaved manner to maintain real-time performance on modest hardware. To optimize efficiency, it caches processed frames temporarily, reducing redundant computations and enabling smooth transitions between dashcam-based immediate hazard detection and CCTV-derived traffic predictions. The function then overlays analytics—such as object bounding boxes, road segmentation masks, weather classifications, and traffic forecasts—onto the video frames. Specifically, the prediction output from the traffic model, indicating anticipated vehicle counts on the left and right, is displayed on the dashcam footage alongside context-aware recommendations derived from the recommendation engine, such as lane change suggestions or speed adjustments. This produces an annotated output that provides ambulance drivers with clear, actionable insights. This integrated approach ensures that the system seamlessly combines immediate and predictive data, enhancing situational awareness and decision-making under the dynamic conditions of Indian roads.

Chapter 4

Results and Findings

4.1 Performance Metrics

- **Frame Rate:**15 fps (CPU).
- **Detection Accuracy:** 90% mAP.
- **Prediction Error:** ±2 vehicles (RMSE).

4.2 Visual Outputs (Side-by-Side Dashcam and CCTV)

Figure 4.1: Dashcam frame with annotations.

Figure 4.2: CCTV frame with vehicle counts.

4.3 Detailed Insights

- **Safety:** High risk scores (e.g., 85%) for "Dangerous" conditions.
- **Traffic:** Lane changes saved 5-15 seconds/km.
- **Efficiency:** Lightweight design viable on CPU.

Chapter 5

Future Scope

- **Reinforcement Learning:** Incorporate reinforcement learning to enable dynamic adaptation of the ADAS. This would allow the system to learn optimal driving strategies over time by interacting with real-world road conditions, improving decision-making in unpredictable scenarios like sudden traffic shifts or emergencies.
- **City-Wide Integration:** Extend the system for city-wide integration by connecting it to a centralized traffic management network. This could optimize ambulance routes using real-time data from multiple CCTV sources, reducing response times and avoiding congested areas across urban landscapes.
- **Edge Deployment:** Transition the system to edge deployment on embedded systems within ambulances. By processing data locally on lightweight hardware, this would minimize latency, reduce dependency on cloud connectivity, and ensure functionality in areas with poor network coverage.

Chapter 6

Conclusion

This ADAS integrates AI models to enhance ambulance navigation, offering a scalable, predictive solution for Indian roads. By combining lightweight deep learning models such as YOLO for object detection, EfficientNet for weather classification, and ConvNeXt for road surface analysis with predictive traffic insights from CCTV footage, the system provides a robust framework for real-time decision-making. It enhances ambulance navigation by delivering actionable recommendations—such as obstacle avoidance, lane changes, and speed adjustments—tailored to the dynamic and often chaotic road conditions in India, including congested urban areas and poorly maintained rural routes. The solution's scalability stems from its cost-effective, video-only design, which operates efficiently on modest hardware without requiring expensive sensors, making it adaptable for widespread use. Its predictive capability, driven by traffic forecasting and the Greenshields model, ensures proactive responses to upcoming hazards, reducing transit times and improving safety during emergency operations.

References

[1] Ultralytics, "YOLO: A Brief History of Real-Time Object Detection," Available at: https://ultralytics.com/yolo, 2023.

[2] Tan, Mingxing, and Le, Quoc V., "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proceedings of the 36th International Conference on Machine Learning (ICML)*, Long Beach, CA, USA, May 2019, arXiv preprint: https://arxiv.org/abs/1905.11946.

[3] Greenshields, Bruce D., "A Study of Traffic Capacity," in *Proceedings of the Highway Research Board*, vol. 14, pp. 448–477, National Research Council, Washington, D.C., 1935.

[4] Liu, Zhuang, Mao, Hanzi, Wu, Chao, and Feichtenhofer, Christoph, "A ConvNet for the 2020s," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, June 2022, pp. 11976–11986, arXiv preprint: https://arxiv.org/abs/2201.03545.