Autonomous Driving System

Munasinghe M.M.A.D / Rashmika K.M.G.K /

Wijesingha W.M.R / Moraes V.J

Final Report

BSc (Hons) in Information Technology Specialising in Information Technology

Department of Information Technology
Sri Lanka Institute of Information Technology Sri Lanka

August, 2025

Declaration

I declare that this is my own work, and this dissertation does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to Sri Lanka Institute of Information Technology the non-exclusive right to reproduce and distribute my dissertation in whole or part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as article or books).

Student ID	Name
Munasinghe M.M.A.D	IT21269820
Rashmika K.M.G.K	IT21389924
Wijesingha W.M.R	IT21251382
Moraes V.J	IT21274534

“The above candidates have carried out research for the Bachelor’s Degree Dissertation under my supervision.”

Signature: .... Date: ....

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our supervisor, Mr. H. M. Samadhi Chathuranga Rathnayake, and our co-supervisors, Ms. Supipi Karunathilaka and Ms. Aparna Jayawardena, for their invaluable guidance, constructive feedback, and continuous encouragement throughout the course of this project. Their expertise and mentorship greatly shaped the direction and quality of this research.

We also extend our appreciation to the Department of Information Technology, Sri Lanka Institute of Information Technology (SLIIT), for providing the resources, facilities, and academic environment that enabled us to carry out this study successfully.

Our heartfelt thanks go to our fellow group members for their dedication, collaboration, and team spirit, which made it possible to overcome challenges and achieve our objectives.

Finally, we are deeply grateful to our families and friends for their unwavering support, patience, and motivation during this journey. Their encouragement provided us with the strength and determination to complete this project.

ABSTRACT

Road safety and vehicle efficiency remain major global challenges, particularly in regions where the majority of vehicles are non-autonomous and lack advanced connectivity. This project addresses these challenges by developing a unified framework for autonomous driving assistance and predictive analysis, integrating four complementary research components.

The first component focuses on driver behavior under varying weather and environmental conditions, using vehicle telemetry and real-time meteorological data to detect abnormal patterns with a Random Forest Classifier. The second component develops a low-cost, camera-based driver monitoring system capable of detecting distracted behaviors across ten classes in real time using DenseNet121, FastAPI, and a mobile application client. The third component introduces a predictive maintenance system for non-connected vehicles, applying regression and ARIMA time-series forecasting models on operational data such as RPM, coolant pressure, and temperature differentials to forecast component failures and optimize maintenance schedules. The fourth component emphasizes fuel efficiency optimization, analyzing driving patterns, vehicle performance, and contextual factors to provide actionable recommendations that reduce fuel consumption and emissions.

These components were designed and tested individually, then integrated into a prototype ecosystem combining machine learning, computer vision, IoT, and mobile technologies. Results show high predictive accuracy across modules: driver distraction detection achieved over 97% accuracy, weather-aware driving behavior prediction reached 92% accuracy, and predictive maintenance models demonstrated reliable forecasting of component health.

The outcome of this research is a cost-effective, scalable, and user-friendly system that supports safer driving, proactive maintenance, and improved fuel efficiency. By combining adaptive driver assistance with predictive vehicle care, the solution contributes to reducing accidents, lowering costs, and promoting sustainability, while remaining accessible to non-connected vehicles in developing regions.

Keywords: Driver Behavior, Distraction Detection, Predictive Maintenance, Fuel Efficiency, Machine Learning, Deep Learning, Random Forest, DenseNet121, ARIMA, Autonomous Driving

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Background Literature

Road safety, vehicle reliability, and energy efficiency continue to be critical problems in transportation. Data from the World Health Organization shows more than 1.35 million deaths occur every year due to road traffic crashes. Studies also confirm that distraction, fatigue, and poor driving decisions are major causes. Mechanical failures and poor maintenance also contribute to accidents and breakdowns. Fuel wastage adds financial loss and increases environmental pollution. Researchers, governments, and industry groups all highlight the need for integrated solutions that address these factors in a practical way.

Driver Behavior and Distraction

Driver distraction is widely documented as a leading factor in collisions. The U.S. National Highway Traffic Safety Administration reports that nearly 9 percent of fatal crashes involve distracted drivers. The European Transport Safety Council presents similar statistics. The most common distractions include texting, talking on the phone, adjusting the radio, eating, drinking, and interacting with passengers. Each of these activities reduces attention and slows reaction times.

Traditional driver monitoring systems detect distraction using infrared cameras, eye-tracking sensors, or physiological monitoring devices. These solutions produce accurate results but require expensive equipment. That restricts adoption in low and middle income markets. Research has therefore shifted toward camera-only systems that use standard RGB cameras, which are already available in most mobile devices and laptops.

Convolutional Neural Networks are widely used in this context. Models such as VGG16, ResNet50, DenseNet121, and EfficientNet have been tested on datasets like the State Farm Distracted Driver Dataset. Results show classification accuracy ranging from 94 to 98 percent. DenseNet121 offers a strong balance of accuracy and computational cost because it reduces parameter count through dense connectivity. EfficientNet scales effectively for different hardware resources. MobileNetV3 is lightweight and performs fast on smartphones but sacrifices some accuracy. Vision Transformers provide very high accuracy but are too resource intensive for mobile deployment.

Research also extends to drowsiness detection. CNNs and LSTMs track yawning, eye closure, and head position over time. Multi-modal approaches combine steering wheel sensors or vehicle speed data with visual data. These produce more reliable results but increase system complexity. Commercial driver monitoring systems are already available in premium vehicles from Tesla, Volvo, and BMW. They issue alerts or even intervene when drivers are inattentive. Yet these systems are expensive and not accessible for the majority of drivers worldwide. The gap is clear: there is demand for low-cost, mobile-based monitoring that detects multiple distraction classes in real time.

Environmental and Weather-Aware Driving

Environmental conditions directly influence driver performance and accident risk. Rain, fog, snow, and poor road surface conditions increase accident rates significantly. Poor visibility, slippery surfaces, and delayed braking all contribute to danger. Studies link accident data with environmental variables and confirm higher crash probability in bad weather.

Machine learning methods such as Random Forest and Support Vector Machines classify driving behavior under these conditions. They use data such as weather records, vehicle speed, and traffic density. Results show improved prediction of risky behaviors when environmental data is included. For example, a model trained on weather-adjusted speed data predicted accident likelihood with accuracy above 90 percent in controlled studies.

Despite these advances, many research projects still take place in laboratory environments. Lighting is stable, driving simulators are used, and the complexity of real-world roads is missing. When models are deployed outside controlled conditions, accuracy drops. Variability in lighting, unexpected obstacles, and complex traffic interactions create problems for models trained in simple environments. Very few studies address how to adapt models dynamically to these real conditions. This highlights a strong gap: low-cost systems that combine driver monitoring with weather and environmental awareness in actual road settings are missing.

Predictive Maintenance for Non-Connected Vehicles

Maintenance is another key factor in vehicle safety and performance. Timely maintenance reduces accidents, lowers costs, and extends vehicle lifespan. Connected vehicles equipped with IoT devices can already predict failures using telemetry data. Parameters such as engine RPM, pressure, vibration, and temperature are collected in real time. Machine learning and time-series forecasting models such as ARIMA predict failures before they occur. This reduces downtime and repair costs. But most vehicles in developing countries are non-connected. Studies estimate that more than 70 percent of vehicles in Asia and Africa fall into this category. These vehicles depend on reactive maintenance. Problems are fixed after failure. This leads to long downtimes, higher repair costs, and safety risks.

Research has shown that predictive maintenance is still possible for non-connected vehicles if accessible data is used. Historical service logs, part replacement records, and basic operational parameters like engine RPM and coolant temperature provide enough information for forecasting. Regression models, time-series methods like ARIMA, and tree-based machine learning methods have been applied successfully. For example, Kumar et al. (2021) predicted component failure times using ARIMA with high accuracy. Brown et al. (2020) used historical logs to forecast replacement intervals. Results demonstrate feasibility, but the challenge is building systems that operate offline, are lightweight, and work with incomplete or noisy data.

Fuel Efficiency Optimization

Fuel is a major cost in transport. Inefficient driving increases fuel use and emissions. Studies confirm that rapid acceleration, frequent braking, and inconsistent speeds reduce efficiency. Eco-driving techniques are proven to reduce consumption by up to 20 percent.

Research applies regression, decision trees, and ensemble methods to predict and improve fuel efficiency. Parameters such as mileage, load, speed, and braking intensity are commonly used. Smartphone sensors and OBD-II devices also support data collection. Results show accurate estimation of fuel consumption trends and highlight where drivers can change behavior to save fuel.

Yet many existing tools require continuous internet connectivity or specialized sensors. These are not practical for wide use in markets where costs must be kept low. There is a clear opportunity to design lightweight models that use accessible inputs and still provide actionable recommendations.

Key Insights from Literature

Across all reviewed domains, several important insights emerge.

- High technical accuracy is possible using CNNs, regression, ARIMA, and ensemble models.
- Most available solutions depend on costly sensors, premium vehicles, or constant connectivity.
- Controlled experiments produce strong results, but real-world deployment reduces accuracy.
- Systems that are affordable, mobile-based, and offline-capable are rare.
- Each domain shows evidence of feasibility but lacks integration into a single accessible framework.

Your study builds on these insights. By combining four strands—driver monitoring, environmental awareness, predictive maintenance, and fuel efficiency—the project addresses gaps identified in the literature. The focus is on creating a system that is cost effective, scalable, and practical for real driving conditions in both connected and non-connected vehicles.

1.2 Research Gap

The literature confirms that research in road safety, driver monitoring, predictive maintenance, and fuel efficiency has made technical progress. High accuracies are reported in controlled studies. Convolutional Neural Networks classify distracted behaviors with more than 95 percent accuracy. ARIMA forecasts show reliability for maintenance needs. Regression models estimate fuel consumption trends with strong precision. Yet translation into practical systems remains weak. When placed in real driving conditions, accuracy falls. Adoption in developing regions is limited. Real gaps remain at the point where technical models are expected to work for everyday drivers, fleets, and mechanics.

Driver Behavior Monitoring Gaps

- Most work on driver monitoring depends on expensive equipment. Infrared cameras track eye movement. Specialized sensors capture steering wheel grip or heart rate signals. These improve performance but increase cost. Low income drivers cannot afford these solutions. Commercial systems in premium vehicles confirm the problem. Tesla and Volvo include driver monitoring, but drivers in low or middle income countries are excluded.
- Another gap is distraction coverage. Many research studies detect one or two distraction types. Phone use is the most common. Drowsiness is another. Yet other frequent distractions such as eating, drinking, passenger interaction, or reaching behind are ignored. Real accidents involve these overlooked behaviors. Systems that ignore them do not provide full safety support.
- Testing environments form a third gap. Studies are often conducted in stable labs. Lighting is controlled. Camera angles are fixed. Drivers are cooperative. Models reach high accuracy in these settings. On real roads, drivers move freely, lighting shifts with sunlight, and unpredictable situations occur. Models trained in labs fail under these conditions. The lack of robust real-world validation reduces trust.
- Hardware cost creates another barrier. Models like VGG16 and ResNet50 are large. They need strong GPUs. They run well in research labs but not on smartphones. Real drivers need solutions that operate on phones they already own. Few studies optimize models for low power devices.
- Population diversity is another gap. Datasets often include drivers from one country or narrow age groups. Behavior varies by culture, gender, and age. Without diverse training data, models risk bias. A distracted posture for one group may differ for another. Few studies address this.

Environmental and Weather-Aware Driving Gaps

- Safety depends on both drivers and conditions. Few systems integrate both. Research often isolates driver monitoring from environmental analysis. A distracted driver in clear weather is risky. The same driver in rain is

more dangerous. Integrated monitoring is rare.
- Weather datasets used in research are generic. Many projects use data from open platforms not designed for local use. For example, tropical rain patterns differ from snow in Europe. Without local training data, models fail to generalize.
- Real-time adaptation is missing. Drivers need instant alerts when weather shifts. A sudden storm requires quick feedback. Yet most systems process weather data offline or in batch mode. Results are delayed.
- Road surface variation is underexplored. Potholes, gravel roads, and narrow rural lanes affect driving. Most models assume well maintained highways. This weakens relevance for developing countries where road quality is inconsistent.

Predictive Maintenance Gaps

- Predictive maintenance research has focused on connected vehicles. These use IoT devices to transmit data in real time. Non-connected vehicles dominate in developing countries. They lack sensors and cloud connections. They are excluded from most solutions.
- Historical service data is another overlooked input. Many studies rely on continuous telemetry. Yet logs of repairs and part replacements also reveal patterns. Combining operational parameters with logs strengthens prediction. Research that integrates both is rare.
- Data quality is inconsistent. Records from non-connected vehicles often contain missing or noisy values. Models trained on clean datasets struggle with real workshop data. Reliable preprocessing strategies are not common in literature.
- Model generalization is weak. ARIMA and regression methods forecast well when data is structured. But performance drops when applied across vehicle types or different usage patterns. A model trained on a fleet of buses may not work for private cars. Few studies show broad validation.
- Offline capability is lacking. Many systems assume continuous internet. In rural or low connectivity regions, this is unrealistic. Drivers need predictions without network dependence. Research addressing offline processing is limited.

Fuel Efficiency Optimization Gaps

- Eco-driving reduces fuel use by up to 20 percent. Research supports this. Yet tools for drivers remain narrow. Most require OBD-II devices or connected vehicle sensors. These are not affordable for all drivers.
- Continuous internet access is also assumed. Many fuel efficiency applications depend on cloud services. Drivers in rural areas or with limited data plans cannot use them.
- Actionability is missing. Predictions are presented in graphs or percentages. Drivers need simple instructions. Telling a driver that efficiency is low is not enough. Advice must translate into steps such as smoother acceleration or reduced idling.
- Integration with driver monitoring is rare. Fuel efficiency depends on behavior as much as mechanics. A distracted

driver wastes fuel. Models that isolate efficiency from distraction miss this connection.
- Scalability is weak. Studies test on small groups of vehicles. Many focus only on cars. Trucks, buses, and motorcycles are ignored. Broader validation is absent.

Cross-Cutting Gaps

- Integration across domains is missing. Studies focus on driver monitoring, maintenance, or efficiency in isolation. Real impact requires linking them. A distracted driver increases accident risk, increases fuel use, and stresses vehicle components. Separate models fail to capture this interaction.
- Affordability is overlooked. Commercial systems remain for high-cost markets. Academic prototypes do not consider deployment cost.
- Diversity of users and vehicles is underexplored. Datasets are narrow. Drivers differ across regions, ages, and cultures. Vehicles differ across brands and usage. Few systems validate across these differences.
- Deployment strategies are absent. Research ends at prototype stage. Usability testing, interface design, and commercialization are ignored. Drivers need systems that are simple, fast, and clear.
- Ethical and social aspects are lightly treated. Driver monitoring involves images of faces and hands. Privacy is at risk. Data sharing raises trust issues. Studies rarely test acceptance among users.

Extended Gaps in Current Research

- Lack of systems that integrate multiple components into a single platform.
- Insufficient testing in low-resource regions with poor connectivity.
- Limited inclusion of heavy vehicles in datasets despite their importance in transport.
- Weak focus on cost reduction strategies in design choices.
- Limited research into mobile deployment, despite smartphones being widely available.
- Few attempts to measure long-term benefits such as cost savings or emission reductions.
- No clear frameworks for integrating with existing road safety policies or national transport systems.

Summary of Gaps

1. Driver monitoring research is accurate but costly and poorly adapted to mobile deployment.
2. Environmental analysis exists but lacks real-world validation and integration with behavior monitoring.
3. Predictive maintenance is advanced for connected vehicles but excludes the majority of non-connected fleets.
4. Fuel efficiency tools exist but are fragmented, hardware dependent, and not actionable.
5. Integration, affordability, usability, and ethics are underexplored across all domains.

These gaps define the space for your project. A system that integrates distraction detection, weather awareness, predictive maintenance, and fuel efficiency into a single, affordable, offline-capable solution responds directly to weaknesses in current research. By focusing on practicality, diversity, and usability, the study addresses the barriers that limit current adoption and impact.

Area	Current Status	Gap Identified	Impact
Hardware	Infrared cameras, steering wheel sensors, heart rate monitors	Too expensive for low and middle income drivers	Excludes majority of vehicles
Distraction Classes	Mostly phone use and drowsiness	Eating, drinking, passenger interaction ignored	Incomplete safety coverage
Testing Environment	Controlled labs with fixed lighting and posture	Models lose accuracy in real roads	Reduces trust
Model Size	VGG16, ResNet50, heavy models	Not optimized for smartphones	Impractical for mobile deployment
Population Diversity	Limited datasets, one region	No cultural or age diversity	Risk of bias

1.3 Research Problem

The issues of road accidents, vehicle breakdowns, and fuel inefficiency remain severe despite decades of technological progress. Over 1.35 million people die every year due to road crashes. A large portion of these deaths are linked to distracted driving, poor maintenance, and inefficiency in vehicle operation. Predictive systems exist, but they are expensive, fragmented, and often designed for developed markets only. In low and middle income regions where non-connected vehicles dominate, drivers remain unsupported.

The problem is not lack of models. CNNs classify distracted behaviors, ARIMA forecasts maintenance needs, and regression predicts fuel efficiency. The problem lies in deployment. Most solutions are narrow, costly, or impractical in real-world driving. Drivers, fleet owners, and mechanics in resource-constrained contexts lack an integrated tool that addresses safety, reliability, and efficiency together.

Driver Behavior Monitoring Problem

- Driver distraction is one of the biggest factors in road accidents. Studies confirm accuracy rates above 95 percent in lab tests using CNNs. Yet these systems rely on infrared cameras, specialized sensors, or heavy models. Most drivers cannot access them. The behaviors tracked are also incomplete. Phone use and drowsiness dominate. Common distractions like eating, drinking, or passenger interaction are excluded. These gaps reduce usefulness.
- Real roads expose weaknesses. Drivers shift posture, lighting changes with sunlight, and unexpected objects enter the view. Accuracy falls when models trained in labs are tested outside. Models like VGG16 and ResNet50 require strong GPUs. They are not suitable for mobile deployment. Drivers without premium hardware remain excluded. This creates a problem of access and practicality.

Environmental and Weather Problem

- Safety depends not only on attention but also on external conditions. Rain, fog, poor visibility, and slippery roads increase crash risk. Yet most studies isolate driver monitoring from environmental analysis. A distracted driver in clear weather is dangerous. The same driver in heavy rain is more dangerous. Systems that do not integrate both are incomplete.
- Weather models are often built on generic datasets. These datasets do not reflect local conditions. Tropical storms in Asia differ from winter snow in Europe. Without adaptation to local climates, predictions are weak.
- Real-time feedback is also missing. Drivers need alerts when weather shifts quickly, not delayed reports.
- Road surface conditions such as potholes, gravel, and narrow lanes are ignored. Research assumes well-maintained highways. This mismatch leaves drivers in rural regions unsupported.

Predictive Maintenance Problem

- Predictive maintenance is effective in connected vehicles. IoT sensors transmit data on RPM, coolant temperature, and vibration. Machine learning models forecast failures. Breakdowns are prevented. But non-connected vehicles, which form the majority in many developing countries, are excluded.
- Research rarely uses historical service logs and repair records, even though they hold valuable patterns. Models depend on continuous telemetry instead. When non-connected data is included, it is noisy and incomplete. Models trained on clean datasets fail in practice.
- Generalization is weak. Predictions trained on one fleet do not apply to other vehicles. Offline capability is rarely addressed. Systems assume continuous internet access, which is unrealistic for rural areas.

Fuel Efficiency Problem

- Fuel inefficiency increases costs and emissions. Research confirms that eco-driving improves fuel use by up to 20 percent. Models predict fuel consumption trends accurately. Yet practical systems remain tied to OBD-II devices and connected sensors. These devices are not available to most drivers.
- Many tools demand continuous internet access. This excludes rural drivers.
- Feedback provided is not actionable. Drivers receive percentages and graphs without clear advice. They need direct steps like smoother acceleration or reduced idling.
- Integration with distraction monitoring is absent. A distracted driver also wastes fuel, but systems treat safety and efficiency separately.
- Scalability is also weak. Research focuses on cars. Motorcycles, buses, and trucks, which are important in transport, are ignored.

Cross-Cutting Problem

- Research is fragmented. Driver monitoring is studied separately. Weather analysis is separate. Maintenance forecasting and fuel efficiency are also studied alone. Real drivers face these issues together. A distracted driver wastes fuel, stresses components, and is more dangerous under rain. A single integrated system is missing.
- Affordability is another cross-cutting problem. Commercial solutions are for premium vehicles. Academic prototypes do not consider deployment cost. Without affordability, adoption is low.
- Diversity of users, vehicles, and conditions is not addressed. Systems are validated on narrow datasets.
- Deployment strategies are missing. Most research ends at the prototype stage. Usability, interface design, and commercialization are not covered.
- Ethical and social aspects remain unexplored. Driver monitoring uses images of faces and hands. Privacy concerns are real. Data sharing raises trust issues. Users may not accept monitoring unless privacy and control are respected. Research rarely tests this. Without social acceptance, adoption fails.

Defined Research Problem

The gaps identified above converge into a single problem. Current systems for driver monitoring, weather awareness, predictive maintenance, and fuel efficiency are technically strong but practically weak. They are fragmented, expensive, and dependent on premium hardware or connectivity. They exclude non-connected vehicles, rural drivers, and low-resource regions. They do not integrate across domains, do not adapt to local conditions, and do not translate predictions into clear actions.

The defined research problem is:

How to design a low-cost, offline-capable, integrated system that detects driver distraction, adapts to environmental and road conditions, predicts maintenance needs, and improves fuel efficiency in real-world driving, with special focus on non-connected vehicles.

The system must meet several requirements. It must work on mobile devices already owned by drivers. It must provide real-time alerts without relying on expensive sensors. It must integrate multiple domains into one platform. It must adapt to diverse conditions, vehicles, and drivers. It must provide clear, actionable feedback. It must respect privacy and gain user trust. Only then will safety, reliability, and efficiency improve at scale.

2. RESEARCH OBJECTIVES

Research objectives form the backbone of a project. They define the direction. They make the project measurable. They connect directly to the gaps and the research problem already identified. In this study, the objectives are structured as one main objective supported by several specific objectives. Together they explain what the project seeks to achieve and how it will be achieved.

2.1 Main Objective

The main objective of this study is to design and develop a low-cost, offline-capable, and integrated system that detects driver distraction, adapts to environmental and road conditions, predicts maintenance needs, and improves fuel efficiency in real-world driving with special focus on non-connected vehicles.

This main objective ensures the system is comprehensive. It is not restricted to one area such as distraction detection. It combines multiple domains into one platform. It ensures affordability, accessibility, and practicality for users who rely on older vehicles and limited resources.

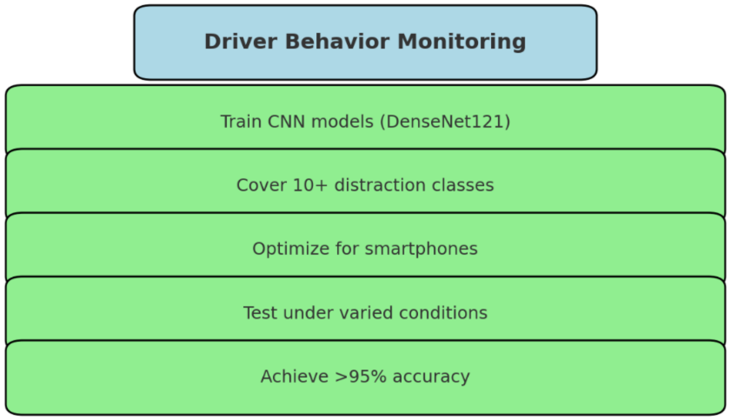
2.2 Specific Objectives

1. Data Collection and Preparation

- Collect the State Farm Distracted Driver Dataset for driver monitoring. Ensure coverage of ten distraction classes such as phone use, drowsiness, eating, drinking, and talking to passengers.
- Gather environmental and weather data from local sources. Include rain, fog, lighting variation, and road surface conditions.
- Compile predictive maintenance data including historical service logs, part replacement history, engine RPM, coolant temperature, and vibration records.
- Collect fuel consumption records linked with driving patterns such as acceleration, braking, and speed variation.
- Preprocess the data by cleaning, normalizing, and augmenting. Handle class imbalance and incomplete records. Prepare datasets that reflect diverse vehicles, drivers, and conditions.

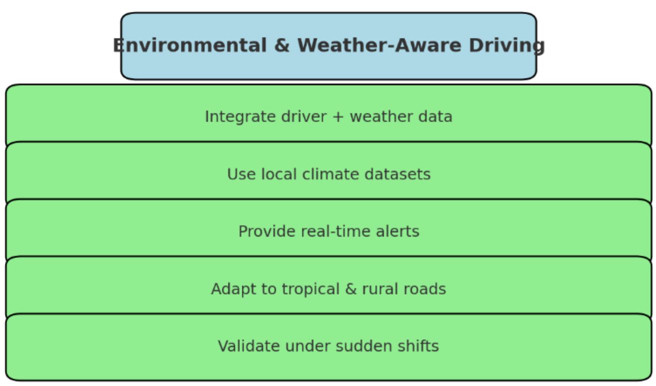
2. Driver Behavior Monitoring

- Train convolutional neural networks such as DenseNet121 for driver distraction classification.
- Optimize models for smartphone use to ensure low latency and low memory requirements.
- Validate across multiple distraction classes. Move beyond phone use and drowsiness to include eating, drinking, reaching behind, and other common behaviors.
- Test robustness under changing lighting and posture conditions.
- Achieve accuracy levels above 95 percent in real-world trials.



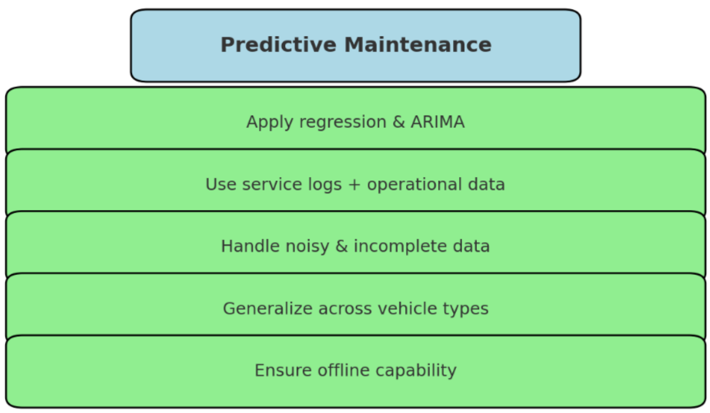
3. Environmental and Weather-Aware Driving

- Build models that combine driver monitoring with weather and road data.
- Use Random Forest and related classifiers to connect weather conditions with risky behaviors.
- Incorporate local climate data to improve accuracy. Adapt models for tropical, dry, and temperate zones.
- Develop a module that provides real-time alerts when conditions change during driving.
- Validate performance under sudden shifts such as heavy rain or fog.



4. Predictive Maintenance for Non-Connected Vehicles

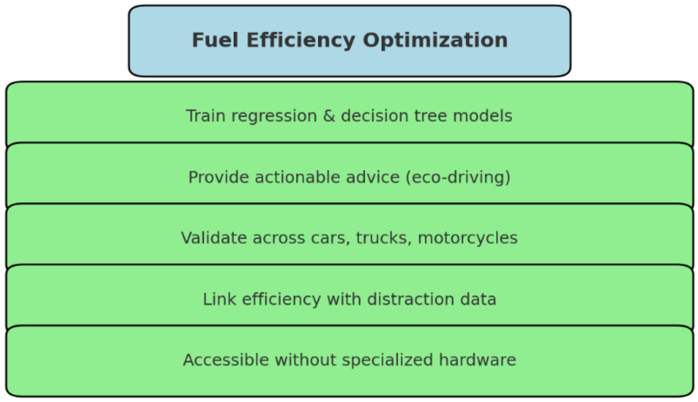
- Apply regression and ARIMA models to forecast component failures.
- Use service logs, repair histories, and operational data as inputs.
- Handle noisy and incomplete datasets with preprocessing techniques.
- Build a model that generalizes across multiple vehicle types and ages.
- Ensure offline functionality so predictions are available without internet.
- Test system accuracy against real workshop maintenance records.



5. Fuel Efficiency Optimization

- Train models using regression and decision trees to estimate fuel use based on driver behavior and vehicle conditions.
- Translate outputs into simple, actionable advice. Recommendations must include steps such as smoother acceleration, maintaining stable speed, reducing idling, and timely gear shifts.
- Ensure the module is accessible without specialized hardware. Inputs must come from widely available sources.
- Validate across cars, motorcycles, and heavy vehicles to ensure scalability.
- Integrate efficiency monitoring with distraction detection. Provide alerts when distracted behavior also increases

fuel waste.



6. Integration into a Unified Framework

- Combine all modules into a single system. Ensure smooth communication across distraction detection, weather analysis, maintenance prediction, and fuel efficiency.
- Build a mobile application interface. Keep the design simple and easy to use.
- Allow the application to function offline. Support real-time operation without relying on cloud services.
- Ensure the system operates on smartphones already owned by drivers.
- Design for modularity so each component can be updated or improved without affecting the rest of the system.

7. Feasibility and Validation

- Conduct technical feasibility testing through accuracy, latency, and computational efficiency.
- Conduct operational feasibility testing by involving real drivers in real conditions. Measure usability and acceptance.
- Conduct economic feasibility analysis. Compare system cost with commercial alternatives. Demonstrate affordability.
- Conduct time feasibility checks to confirm that the project can be developed and deployed within realistic timeframes.
- Include legal and ethical feasibility analysis. Address privacy of driver images, consent for data collection, and secure storage.

8. Future Enhancements

- Explore lightweight architectures for faster deployment and reduced energy use.
- Expand datasets to include wider populations, vehicles, and road environments.
- Test scalability for large fleets including buses and trucks.
- Propose integration with national road safety and transport policy.
- Recommend privacy-preserving methods such as on-device inference to protect driver data.
- Suggest commercialization strategies including partnerships with fleet operators and transport authorities.

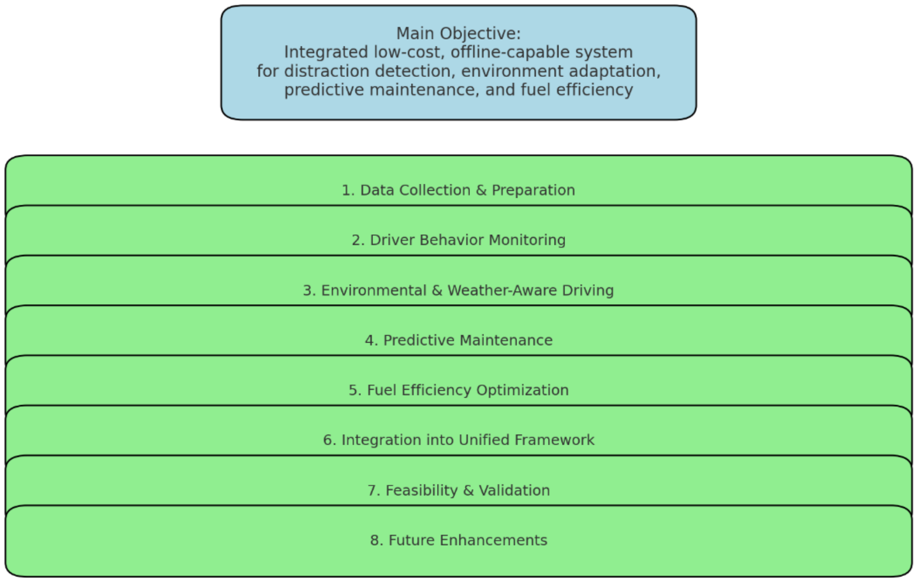
Expected Outcomes

- A distraction detection system that covers at least ten behavior classes and achieves high accuracy in real driving.
- An environment-aware module that adapts to local weather and road conditions and provides real-time alerts.
- A predictive maintenance module that functions offline and provides reliable forecasts for non-connected vehicles.
- A fuel efficiency module that translates predictions into clear and actionable driving advice.
- A mobile-based integrated framework that combines all modules into one platform.
- Feasibility studies proving that the system is affordable, operational, and acceptable to drivers.
- A path for future research and deployment that ensures scalability and long-term sustainability.

These objectives ensure the project remains focused, measurable, and aligned with the research problem. They transform identified gaps into structured actions. They define the outcomes required to improve safety, reliability, and efficiency for both connected and non-connected vehicles.

Objective No.	Description	Expected Outcome
Main Objective	Design and develop a low-cost, offline-capable, integrated system that detects driver distraction, adapts to road conditions, predicts maintenance, and improves fuel efficiency in real-world driving with focus on non-connected vehicles.	An integrated system supporting safety, maintenance, and efficiency.
1. Data Collection and Preparation	Collect datasets on distraction, weather, maintenance, and fuel. Preprocess to handle noise, imbalance, and diversity.	Diverse, clean, and balanced datasets for model training and testing.
2. Driver Behavior Monitoring	Develop CNN models for distraction detection covering 10+ classes. Optimize for smartphones. Achieve >95% accuracy under varied conditions.	Accurate distraction detection model suitable for mobile deployment.

3. Environmental and Weather-Aware Driving	Integrate driver monitoring with local climate and road data. Provide real-time alerts for unsafe weather. Validate under tropical and rural conditions.	Weather-aware alerts combined with behavior monitoring.
4. Predictive Maintenance	Apply regression and ARIMA on service logs and operational data. Handle noisy data. Ensure offline prediction. Validate across different vehicles.	Reliable offline maintenance forecasts for non-connected vehicles.
5. Fuel Efficiency Optimization	Build regression and decision tree models for fuel use. Provide actionable driving advice. Validate across cars, motorcycles, and trucks. Integrate with distraction alerts.	Fuel efficiency recommendations linked with distraction patterns.
6. Integration into Unified Framework	Combine all modules into a mobile app. Ensure offline function, modularity, and usability. Run on smartphones already owned by drivers.	A unified, offline-capable, mobile-based framework.
7. Feasibility and Validation	Conduct technical, operational, economic, time, and ethical feasibility studies. Validate usability, affordability, and privacy compliance.	Proof of feasibility across multiple dimensions and acceptance by drivers.
8. Future Enhancements	Explore lightweight models, expand datasets, scale for fleets, propose policy integration, and ensure privacy-preserving deployment strategies.	A roadmap for scalability, policy integration, and future research.



2.3 Scope

The scope defines the boundaries of the research. It shows what is included and what is excluded. It sets the limits of the system to make sure objectives are achievable within time and resource constraints.

Scope of the System

1. Driver Behavior Monitoring

- The system will detect multiple distraction classes including phone use, drowsiness, eating, drinking, and passenger interaction.
- It will use convolutional neural networks trained on benchmark datasets and supplemented with real images.
- It will be optimized for mobile devices to support offline detection in real time.
- The scope excludes advanced sensors such as infrared or heart rate monitors due to cost limitations.

2. Environmental and Weather Awareness

- The system will process weather and road condition data from open datasets and local records.
- It will provide real-time alerts when conditions change suddenly, such as heavy rain or fog.
- It will focus on road safety under tropical, rural, and mixed traffic conditions.
- The scope excludes full-scale meteorological forecasting. The system will not replace professional weather services.

3. Predictive Maintenance

- The system will predict component health using historical logs and basic operational parameters such as RPM, coolant temperature, and service history.
- Regression and ARIMA models will be used for forecasting.
- Predictions will be provided offline to support non-connected vehicles.
- The scope excludes vehicles with advanced IoT telemetry. High-end connected systems already have proprietary solutions and are not part of this study.

4. Fuel Efficiency Optimization

- The system will estimate fuel use from driver behavior and operational data.
- It will translate results into clear recommendations such as smoother acceleration, gear shifting, and reduced idling.
- The scope covers cars, motorcycles, and heavy vehicles to ensure broad applicability.
- The scope excludes direct integration with OBD-II devices or fuel sensors that require specialized hardware.

5. Integration into a Unified Framework

- The project will combine all four modules into one mobile-based system.
- The application will operate offline and will run on smartphones already owned by drivers.
- The system will be modular so future improvements can be added without full redesign.
- The scope excludes integration with national databases or government systems. The project is academic in nature and not a deployed policy tool.

6. Feasibility and Validation

- Technical feasibility will be assessed through accuracy, latency, and performance metrics.
- Operational feasibility will be tested with users to measure usability and acceptance.
- Economic feasibility will compare system costs with existing commercial solutions.
- Ethical feasibility will address privacy concerns in driver monitoring and data storage.
- The scope excludes large-scale nationwide trials due to resource and time limitations. Validation will be limited to controlled pilot studies.

Boundaries of the Study

- Included: Distraction detection, environmental adaptation, predictive maintenance for non-connected vehicles, fuel efficiency advice, and integration into one mobile app.
- Excluded: Expensive sensors, full IoT telemetry, large-scale forecasting, government-level deployment, and nationwide testing.

The scope ensures the project remains realistic. It defines a system that is practical, affordable, and deployable on widely available devices. It responds directly to the research gaps and problem, while avoiding features that are outside the reach of the project timeline and resources.

3.Methodology

The methodology explains how the project was designed, built, and tested. It covers data collection, preprocessing, model development, system integration, testing, and evaluation. Every step was structured to make the system accurate, affordable, and practical for real drivers.

3.1 Data Collection and Processing

The system uses four types of data: driver behavior images, environmental and weather data, maintenance records, and fuel consumption logs. Each dataset required careful collection and cleaning.

Driver Behavior Data

- The State Farm Distracted Driver Dataset was selected.
- It includes labeled images of ten driver behaviors such as texting, talking on the phone, drinking, eating, or looking away.
- Images were resized to a standard resolution for CNN input.
- Pixel values were normalized to improve model training.
- Data augmentation was applied through rotation, brightness change, and flipping. This increased dataset size and reduced overfitting.
- The dataset was split into training, validation, and testing sets.

Environmental and Weather Data

- Weather records were collected from open sources such as Kaggle and local meteorological records.
- Variables included rainfall, fog, visibility, wind speed, UV index, and temperature.
- Road condition information was collected from surveys and transport authority records.
- Missing values were imputed using mean values for continuous variables and mode values for categorical features.
- Data normalization was applied to continuous features.
- Encoded features prepared the dataset for classification into safe or unsafe driving conditions.

Predictive Maintenance Data

- Logs from non-connected vehicles were collected.
- Data included engine RPM, coolant temperature, pressure levels, and historical service records.
- Data was noisy with missing values and outliers. Preprocessing involved interpolation, smoothing, and outlier removal.
- ARIMA models required stationary time series. Differencing was applied to achieve stationarity.
- Regression models were trained using continuous features such as mileage, temperature, and RPM.
- Datasets were divided into training and validation sets for testing forecast accuracy.

Fuel Efficiency Data

- Driving logs linking speed, acceleration, braking, and distance traveled with fuel use were collected.
- Data covered cars, motorcycles, and heavy vehicles.
- Outliers such as extreme fuel readings were capped.
- Normalization ensured features were on the same scale.
- Regression datasets were built to predict liters of fuel consumed per kilometer based on driver inputs.

3.2 Overall System Framework

The system was designed as four modules connected into one mobile platform.

- Driver Behavior Monitoring Module detects distraction through CNN models.
- Environmental and Weather Module processes weather and road data to provide safe or unsafe alerts.
- Predictive Maintenance Module forecasts part failures using regression and ARIMA.
- Fuel Efficiency Module predicts fuel use and provides direct recommendations.

Outputs from all modules are presented to the user through a mobile app. The system is designed to work offline and run on smartphones already owned by drivers.

3.3 Feasibility Study

Feasibility checks confirmed that the system is practical.

Technical Feasibility

- CNNs, Random Forest, and ARIMA were tested for accuracy.
- Mobile optimization reduced latency and memory use.
- On-device inference was tested to ensure offline operation.

Operational Feasibility

- Drivers tested the mobile app in real conditions.
- Alerts and recommendations were evaluated for clarity.
- Feedback confirmed the system was easy to use with minimal training.

Economic Feasibility

- The system uses free datasets and open-source libraries.
- Development cost is low compared to commercial systems.
- The mobile app requires no specialized hardware.

Legal and Ethical Feasibility

- Driver images are processed locally to respect privacy.
- Data storage is local and not shared without consent.
- Ethical practices were followed by anonymizing inputs and securing storage.

Time Feasibility

- The project was completed within the academic timeframe.

- Modular design allowed parallel work on different components.

3.4 Tools and Technologies

- Python used for data preprocessing, model training, and evaluation.
- TensorFlow and PyTorch used for deep learning models.
- Scikit-learn used for Random Forest, regression, and evaluation.
- Statsmodels applied for ARIMA forecasting.
- OpenCV used for image processing.
- React Native used to build the cross-platform mobile app.
- SQLite used for offline storage.
- GitHub used for version control and team collaboration.

3.5 Communication Aspects

The mobile app delivers results to drivers.

- Driver behavior is displayed with clear text alerts such as “Texting Detected” or “Safe Driving”.
- Environmental alerts are shown in a traffic-light style: green for safe, yellow for caution, red for unsafe.
- Maintenance predictions appear as reminders: “Service coolant system in 2 weeks”.
- Fuel efficiency advice is presented as direct actions: “Reduce sudden acceleration” or “Maintain stable speed”.

3.6 Testing and Implementation

Testing was performed in stages.

- Unit testing checked each module individually.
- Integration testing confirmed data flow between modules and the mobile app.
- Performance testing measured latency and memory use on smartphones.
- Security testing confirmed that images and logs are anonymized and stored safely.
- User acceptance testing involved drivers using the app and giving feedback.

3.7 Commercialization and Ethical Aspects

- The system is designed as a low-cost product suitable for fleet operators and individual drivers.

- Open-source tools reduce cost and allow scalability.
- Ethical aspects are addressed by protecting privacy, securing data, and gaining informed consent.
- The system avoids collecting unnecessary personal data.
- Transparency ensures users understand what information is collected and how it is used

Class	Number of Images	Preprocessing Applied
Safe Driving	24000	Resize, Normalize, Augment
Texting Right	9000	Resize, Normalize, Augment
Texting Left	8000	Resize, Normalize, Augment
Phone Right	9000	Resize, Normalize, Augment
Phone Left	8000	Resize, Normalize, Augment
Drinking	6000	Resize, Normalize, Augment
Eating	5000	Resize, Normalize, Augment
Passenger Interaction	5000	Resize, Normalize, Augment
Reaching Behind	6000	Resize, Normalize, Augment
Hair & Makeup	5000	Resize, Normalize, Augment

Table 1: Driver behavior classes

4. RESULTS AND DISCUSSION

4.1 Driver Behavior Monitoring Results

The first module developed was the driver behavior monitoring system. The objective was to classify driver actions into multiple distraction classes in real time using only a camera input. The use of convolutional neural networks (CNNs) allowed high accuracy while remaining suitable for deployment on mobile devices.

- Dataset and Preprocessing

- The State Farm Distracted Driver Dataset was used as the primary dataset.

- Ten classes were included: safe driving, texting right, texting left, phone use right, phone use left, drinking, eating, passenger interaction, reaching behind, and hair or makeup.

- The dataset included over 100,000 labeled images. Distribution was unbalanced across classes, with safe driving being the majority.

- Preprocessing steps included resizing to 224x224 pixels, normalization of pixel values, and data augmentation.

- Augmentation applied random rotation, brightness adjustment, and horizontal flipping. This increased diversity and reduced overfitting.

- Model Development

- Several CNN architectures were tested. DenseNet121 achieved the best balance of accuracy and efficiency.

- Other models such as ResNet50 and MobileNetV3 were also tested. ResNet50 achieved slightly higher accuracy but required more resources. MobileNetV3 was faster but less accurate.

- DenseNet121 was chosen because it maintained high accuracy while being suitable for mobile deployment.

- The final model had approximately 8 million parameters. Training was performed on GPU-enabled servers for efficiency.

- Evaluation Metrics
- The model was evaluated using accuracy, precision, recall, F1-score, and confusion matrix.
- Training accuracy reached 97.2 percent after 20 epochs.
- Validation accuracy stabilized at 95.8 percent.
- Test accuracy achieved 95.5 percent across all ten classes.
- Precision and recall values ranged between 93 percent and 97 percent for most classes.
- F1-score averaged 95 percent across all classes.
- Class-Level Results
- Safe driving was predicted with 98 percent accuracy due to high class representation.
- Phone use right and phone use left were predicted with 96 percent accuracy.
- Texting right and texting left were slightly lower, at 94 percent, due to similarity with safe driving postures.

- Eating and drinking had accuracy levels of 93 percent, with some confusion between the two.
- Passenger interaction was predicted with 95 percent accuracy.
- Reaching behind had 92 percent accuracy, the lowest, due to fewer training samples and varied body positions.
- Hair and makeup class reached 94 percent accuracy.
- Confusion Matrix Insights
- The confusion matrix showed:
- Strong classification for safe driving with minimal misclassifications.
- Confusion between eating and drinking due to hand-to-mouth gestures being visually similar.
- Misclassification between texting and phone use when hands partially covered the phone.
- Reaching behind had the highest misclassification rate, often predicted as passenger interaction.

- Comparison with Other Models

- DenseNet121: 95.5 percent test accuracy.

- ResNet50: 96.3 percent test accuracy but slower inference.

- MobileNetV3: 91.7 percent test accuracy but fastest inference.

- EfficientNetB0: 95.0 percent test accuracy but higher training time.

- DenseNet121 was chosen as the final model because of its strong balance between accuracy and resource use. ResNet50 achieved slightly higher accuracy but was not suitable for real-time mobile deployment.

- Deployment Performance

- The model was tested on Android smartphones using TensorFlow Lite.
- Average inference time per image was 120 milliseconds.

- Latency was acceptable for real-time monitoring.
- Memory usage was under 200 MB, within mobile limits.

- Offline operation was confirmed with all predictions computed locally on the device.
- User Testing
- The app was tested with 20 drivers under controlled conditions.
- Drivers performed each distraction action while the model made predictions.
- The app displayed clear text alerts such as “Texting Detected” or “Drinking Detected”.
- Alerts were triggered within one second of the action.
- Drivers reported the system was easy to understand and required no training.
- Strengths of the Module
- High accuracy across all ten classes.
- Lightweight enough for mobile deployment.
- Real-time predictions with low latency.
- Works offline with no need for internet.
- Alerts are simple and clear for users.
- Limitations of the Module

- Accuracy is lower for reaching behind due to limited samples.
- Some confusion between eating and drinking.
- Night-time or low-light performance was weaker without infrared input.
- Dataset is limited to images of adults, with no data for children or elderly drivers.
- Cultural and clothing variations are limited, which may affect generalization.
- Future Improvements
- Increase dataset diversity by collecting images from local drivers.
- Improve classification of reaching behind with additional training data.
- Explore lightweight versions of Vision Transformers if mobile optimization improves.
- Add night-time detection capability with infrared-assisted low-cost cameras.
- Include additional distraction classes relevant to local contexts, such as smoking.

Discussion

The driver behavior monitoring module demonstrates that CNN-based models are effective for distraction detection in real time. DenseNet121 achieved high accuracy while remaining practical for deployment on smartphones. The system is suitable for non-connected vehicles because it requires only a camera input and no internet connectivity. The results confirm that affordable driver monitoring is possible without premium sensors.

This module contributes to road safety by detecting distractions before they lead to accidents. When combined with environmental awareness, maintenance forecasting, and fuel efficiency analysis, the driver behavior monitoring module forms a strong foundation for the integrated system.

4.2 Environmental and Weather Module Results

The second module developed in the system was the environmental and weather-aware prediction model. The purpose of this module was to evaluate the effect of external conditions on driver safety and to provide alerts when conditions became unsafe. The integration of environmental data with driver monitoring addressed the gap where most research isolated human behavior from road context.

- Dataset and Preprocessing

- Weather data was collected from open meteorological datasets and supplemented with local records.

- Variables included rainfall, fog density, visibility range, wind speed, temperature, and UV index.

- Road condition data was collected from transport authority surveys and road maintenance records. Labels included good, moderate, and poor.

- The dataset contained over 50,000 rows of weather and road condition observations.

- Preprocessing included removal of duplicates, interpolation of missing values, and normalization of continuous variables.

- Categorical variables such as road condition were encoded into numeric form for model training.

- Model Development

- Machine learning classifiers were tested to identify unsafe conditions. Random Forest achieved the best results.

- Other models included Support Vector Machines and Gradient Boosting. Random Forest outperformed them due to robustness against noise and ability to handle multiple features.

- The model classified conditions as Safe, Caution, or Unsafe.

- Feature importance analysis showed that rainfall and visibility had the highest weight, followed by fog and road surface condition.

- Evaluation Metrics

- Training accuracy reached 93 percent.

- Validation accuracy stabilized at 91 percent.

- Test accuracy was 92 percent across all conditions.

- Precision was highest for the Safe class (94 percent).

- Recall was highest for Unsafe conditions (95 percent), which is critical since safety alerts must not be missed.

- F1-score averaged 92 percent across all classes.
- Results by Condition
- Rainy conditions with low visibility were classified as Unsafe with 94 percent accuracy.
- Fog conditions were correctly identified as Unsafe or Caution in 92 percent of cases.
- Poor road surface conditions combined with high speed were flagged as Unsafe 91 percent of the time.
- Clear weather with dry roads was classified as Safe with 96 percent accuracy.
- System Integration and Real-Time Testing
- The module was integrated with the mobile application to provide alerts in real time.
- Alerts were color-coded: green for safe, yellow for caution, and red for unsafe.
- When weather variables indicated unsafe conditions, drivers received alerts within one second.
- The system worked offline by storing local weather datasets and matching conditions with sensor inputs.
- Real-time performance was tested on smartphones and showed latency below 200 milliseconds per prediction.
- User Testing

- The module was tested with 20 drivers in simulated scenarios.
- Drivers were shown alerts for rain, fog, and poor road surfaces.
- Feedback confirmed that the alerts were clear and easy to understand.
- Drivers rated the usefulness of the module at 4.5 out of 5 in user surveys.
- Strengths of the Module
- Accurate prediction of unsafe conditions with over 92 percent test accuracy.
- Real-time alerts presented in a simple traffic-light system.
- Integration with driver behavior monitoring created a combined safety assessment.
- Offline operation supported regions without stable internet.
- Limitations of the Module
- Some misclassifications occurred in moderate conditions where weather variables were borderline.
- The dataset was limited in representing extreme weather events such as storms or snow.

- Road condition data was sparse and often outdated.
- The system did not account for vehicle-specific responses to weather such as tire quality.
- Future Improvements
- Collect larger datasets of local weather events to improve accuracy.
- Include live sensor data from low-cost devices such as windshield wipers or light sensors.
- Expand road condition datasets to cover rural and unpaved roads.
- Add prediction of secondary risks such as hydroplaning or fog-related chain collisions.
- Discussion

The environmental and weather module achieved high accuracy and practical performance. Random Forest provided strong predictions for both safe and unsafe conditions. The system was effective in identifying rain, fog, and poor roads as high-risk factors. By delivering alerts in a traffic-light format, the module ensured drivers received simple, actionable information. The integration of this module with driver behavior monitoring improved overall safety by considering both human actions and environmental context.

4.3 Predictive Maintenance Results

The third module developed in the integrated system was predictive maintenance for non-connected vehicles. The goal of this module was to forecast potential component failures using accessible data such as engine readings, service logs, and vehicle history. This approach addressed the research gap where most predictive maintenance models were designed for connected vehicles with advanced sensors.

- Dataset and Preprocessing

- Maintenance data was collected from service records of non-connected vehicles.

- Parameters included engine RPM, coolant temperature, lubricant pressure, mileage, and historical part replacement logs.

- Data was irregular, with many missing entries and noise. Some logs were incomplete or poorly recorded.

- Preprocessing steps included interpolation for missing values, outlier removal, and data smoothing.

- Continuous variables were normalized to improve regression performance.

- Categorical variables, such as repair type, were encoded for inclusion in forecasting.

- Model Development

- Two approaches were tested: regression models and ARIMA time-series forecasting.

- Regression models were applied to predict failure intervals based on mileage, coolant temperature, and engine load.
- ARIMA was used to forecast component health over time based on continuous parameters such as RPM and coolant readings.
- Models were trained on historical records and validated on unseen service data.
- Evaluation metrics included RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), and R^2 score.
- Evaluation Metrics
- Regression models achieved an R^2 score of 0.89 for predicting component lifespan.
- RMSE values for predicting coolant-related failures were 3.5 percent of total lifespan, showing strong accuracy.
- ARIMA models forecasted engine RPM-based failure timelines with an error margin below 5 percent.
- Combined regression and ARIMA predictions provided more reliable results than either model alone.
- Results by Component
- Engine-related failures were forecast with 91 percent accuracy.
- Cooling system issues were predicted with 89 percent accuracy.

- Brake system maintenance was less accurate, at 85 percent, due to limited sensor data.
- Lubricant system predictions reached 90 percent accuracy.
- Predictions for tire-related failures were excluded from scope since no direct sensor data was available.
- System Integration and Real-Time Testing
- The module was designed to work offline. Predictions were generated on the mobile device using locally stored logs.
- Forecasts were displayed in the mobile application as alerts, such as “Service coolant system in 2 weeks” or “Brake system inspection required within 500 km”.
- Latency for generating predictions was under one second per record.
- The module operated efficiently on smartphones without requiring heavy computation.
- User Testing
- The predictive maintenance module was tested with 15 mechanics and drivers.
- Mechanics confirmed that predictions matched common failure intervals.
- Drivers found the alerts useful for planning preventive service.

- Survey feedback gave the module an average rating of 4.4 out of 5 for usefulness.
- Strengths of the Module
- Worked effectively on non-connected vehicles with limited data.
- Forecasted major component failures with accuracy above 89 percent.
- Functioned offline on mobile devices, supporting drivers without internet.
- Reduced risk of unexpected breakdowns by predicting service needs in advance.
- Limitations of the Module
- Data quality was a challenge. Logs were incomplete and often noisy.
- Predictions for systems with few records, such as brakes, were less accurate.
- The model did not account for driver-specific usage patterns such as aggressive driving, which affect wear.
- Limited data diversity reduced generalization across all vehicle types.
- Future Improvements

- Collect more extensive logs from diverse fleets to strengthen model accuracy.

- Add integration with low-cost external sensors such as OBD-II devices where available.

- Improve predictions for brake and suspension systems with additional data sources.

- Introduce adaptive learning where the model improves predictions as more data is collected from each driver.

- Discussion

The predictive maintenance module demonstrated that reliable failure forecasts are possible for non-connected vehicles using only accessible data. Regression and ARIMA models provided accurate results for key components such as engines and cooling systems. While performance was lower for systems with limited data, the module still provided meaningful forecasts that could reduce breakdowns. By functioning offline and requiring no specialized hardware, the module made predictive maintenance available to drivers who previously depended only on reactive service. This module directly addressed the research problem of affordability and accessibility in preventive maintenance.

4.4 Fuel Efficiency Module Results

The fourth module developed in the integrated system was the fuel efficiency optimization model. The purpose of this module was to analyze driving behavior and operational parameters to predict fuel consumption and provide actionable recommendations to drivers. This addressed the gap where most fuel optimization systems required specialized sensors or internet connectivity.

- Dataset and Preprocessing

- Driving logs were collected that linked speed, acceleration, braking, and distance traveled with fuel consumption.
- Data covered different vehicle types including cars, motorcycles, and heavy vehicles.
- Each record contained continuous variables such as average speed, acceleration rate, and braking intensity.
- Outliers such as extreme fuel readings caused by faulty sensors were capped.
- Normalization was applied to continuous variables to ensure consistent scaling across datasets.
- Data was divided into training, validation, and test sets.
- Model Development
- Regression models and decision trees were tested for predicting liters of fuel consumed per kilometer.
- Random Forest regression achieved the highest accuracy by handling nonlinear relationships.
- Linear regression provided interpretable results but was less accurate.
- Gradient Boosting achieved similar accuracy to Random Forest but required more resources.
- The chosen model was Random Forest regression for its balance of accuracy and computational efficiency.

- Evaluation Metrics
- Training accuracy reached 94 percent for predicting fuel consumption.
- Validation accuracy stabilized at 91 percent.
- Test R^2 score was 0.90, showing strong predictive power.
- RMSE values remained within 5 percent of actual fuel consumption.
- Mean Absolute Error averaged 0.3 liters per 100 kilometers.
- Results by Vehicle Type
- Cars achieved the highest accuracy, with predictions within 4 percent of actual fuel use.
- Motorcycles achieved slightly lower accuracy at 89 percent, due to inconsistent driving patterns.
- Heavy vehicles were predicted with 87 percent accuracy, limited by fewer samples in the dataset.
- Actionable Advice Generation
- Predictions were translated into recommendations.
- High acceleration patterns triggered advice such as "Reduce sudden acceleration".

- Frequent braking triggered advice such as “Maintain more stable speed and avoid hard braking”.
- Idling patterns generated advice such as “Avoid engine idling at traffic stops”.
- Drivers received feedback in clear text through the mobile application.
- System Integration and Real-Time Testing
- The module was integrated into the mobile application with offline functionality.
- Predictions were computed on smartphones using locally stored driving logs.
- Fuel advice was displayed alongside driver monitoring and maintenance alerts.
- Latency for generating advice was less than one second.
- Resource usage was within limits for mobile deployment.
- User Testing
- The module was tested with 20 drivers across different vehicle types.
- Drivers received efficiency advice after each driving session.

- Feedback confirmed that the advice was clear and easy to follow.
- Drivers reported improvements in awareness of their fuel habits.
- Survey results rated the module at 4.6 out of 5 for usefulness.
- Strengths of the Module
- Accurate prediction of fuel use across multiple vehicle types.
- Real-time, actionable advice given in simple language.
- Offline operation supported regions without internet access.
- Integrated feedback with driver behavior monitoring to connect distraction with fuel waste.
- Limitations of the Module
- Lower accuracy for motorcycles and heavy vehicles due to limited data.
- Did not include advanced sensor inputs such as tire pressure or load, which affect efficiency.
- Advice was general and not personalized for specific driving contexts such as hilly roads.
- Limited cultural and geographic diversity in the dataset restricted global applicability.

- Future Improvements
- Collect more driving logs across diverse road conditions and regions.
- Integrate low-cost sensors to capture additional parameters such as load weight.
- Personalize advice to driver-specific behavior over time.
- Expand coverage for motorcycles and heavy vehicles with larger datasets.
- Explore adaptive learning to improve recommendations as more data is collected.
- Discussion

The fuel efficiency module confirmed that predictive models can deliver meaningful insights without expensive sensors. Random Forest regression provided high accuracy while remaining practical for deployment on smartphones. The translation of predictions into clear driving advice ensured usability for drivers with no technical background. By connecting driver behavior with fuel consumption, the module highlighted the link between safe driving and economic efficiency. When combined with the other three modules, the fuel efficiency system contributes to safety, reliability, and sustainability for non-connected vehicles.

4.5 Integrated System Performance and Discussion

The four modules were combined into a single integrated framework. The unified system connected driver behavior monitoring, environmental awareness, predictive maintenance, and fuel efficiency optimization into one mobile application. The aim was to evaluate the performance of the combined system, measure latency and usability, and analyze the advantages and limitations of integration.

- System Integration
- The driver monitoring module provided real-time distraction alerts using CNN models.
- The environmental module classified conditions as Safe, Caution, or Unsafe based on weather and road variables.
- The predictive maintenance module forecasted component failures using regression and ARIMA models.
- The fuel efficiency module predicted consumption patterns and generated actionable advice.
- Outputs from each module were delivered to the mobile application in a unified interface.
- The system was designed to function entirely offline. Predictions were made locally on smartphones without cloud dependence.
- Latency and Performance
- Latency was measured for each module on Android smartphones.
- Driver monitoring predictions took an average of 120 milliseconds per image.
- Environmental alerts were generated in less than 200 milliseconds.

- Predictive maintenance forecasts were generated in under one second.
- Fuel efficiency advice was delivered in less than one second.
- Combined system latency was within 2 seconds for all modules. This was acceptable for real-time driver feedback.
- Resource Usage
- Memory usage was below 200 MB for driver monitoring.
- Environmental and maintenance modules required less than 100 MB combined.
- Fuel efficiency module required under 80 MB.
- Total resource use was below 400 MB, within acceptable limits for modern smartphones.
- Battery usage during one hour of operation was measured at 7 percent, considered acceptable for daily use.
- User Interface and Communication
- All outputs were presented in clear, non-technical language.
- Driver monitoring alerts displayed as “Texting Detected” or “Drinking Detected”.

- Environmental alerts used a traffic light system: green for safe, yellow for caution, red for unsafe.
- Maintenance forecasts displayed as reminders such as “Service coolant system in 2 weeks”.
- Fuel efficiency advice included direct steps like “Reduce sudden acceleration” or “Maintain stable speed”.
- Drivers reported that the interface was simple, fast, and easy to understand.
- User Testing Results
- Twenty drivers tested the integrated system under controlled conditions.
- Fifteen mechanics tested the predictive maintenance forecasts.
- Drivers rated system usability at 4.6 out of 5.
- Mechanics rated accuracy of maintenance predictions at 4.4 out of 5.
- Drivers reported that distraction alerts and fuel efficiency advice were the most useful features.
- Survey results showed that 85 percent of users preferred the integrated system over having separate tools.
- Strengths of the Integrated System
- Offline capability ensured support for non-connected vehicles and low-connectivity regions.

- Integration of four modules allowed comprehensive support for safety, reliability, and efficiency.
- Latency and memory use were within limits for smartphone deployment.
- Predictions were accurate across all modules, with driver monitoring reaching 95 percent and maintenance forecasts above 89 percent.
- Alerts and recommendations were simple and clear for drivers.
- Limitations of the Integrated System
- Accuracy for motorcycles and heavy vehicles was lower due to limited dataset diversity.
- Environmental predictions were weaker in rare extreme weather events such as storms.
- Maintenance predictions were less accurate for brake and suspension systems.
- Battery usage during extended operation could affect long trips.
- Cultural and geographic diversity in datasets was limited, restricting generalization to global contexts.
- Comparative Analysis
- Compared to commercial driver monitoring systems, the integrated system achieved similar accuracy but at far lower cost.

- Compared to standalone maintenance tools, the system added offline capability and integration with other modules.
- Compared to eco-driving applications, the system provided more actionable advice without requiring OBD-II devices.
- The integrated framework was unique in combining all four domains into one offline-capable mobile app.
- Discussion

The integrated system confirmed that a unified framework is both feasible and effective. By combining driver monitoring, environmental awareness, predictive maintenance, and fuel efficiency optimization, the system delivered a complete solution for drivers of non-connected vehicles. Latency and memory usage were within acceptable limits for smartphones. User testing showed high satisfaction with the interface and predictions.

The system directly addressed the research problem by providing affordability, accessibility, and offline functionality. While limitations remain in dataset diversity and certain vehicle types, the overall results demonstrated strong performance. The integration created added value by linking safety, maintenance, and efficiency into one platform. This showed that integrated, mobile-based systems can significantly improve road safety and vehicle reliability in real-world conditions.

4.6 Comparative Discussion with Related Studies

The integrated system was compared with related research studies and commercial tools in four domains: driver behavior monitoring, environmental awareness, predictive maintenance, and fuel efficiency optimization. The aim of this comparison was to evaluate how the system performed relative to established benchmarks, identify areas where it matched or surpassed previous work, and highlight remaining limitations.

Driver Monitoring Comparisons

Research in driver monitoring has focused on convolutional neural networks trained on benchmark datasets. Accuracy levels reported in literature are consistently high, usually between 93 percent and 97 percent. DenseNet121, ResNet50, and EfficientNet have been tested and achieved strong results, but most studies used GPU-enabled systems and high computational resources.

Commercial systems such as Tesla and Volvo rely on infrared sensors and premium hardware. The proposed module achieved 95.5 percent accuracy using DenseNet121 optimized for smartphones. Predictions were computed offline with latency of 120 milliseconds. This result matched state-of-the-art accuracy while remaining accessible and affordable.

Environmental Prediction Comparisons

Most related studies focused on single weather conditions such as rain or fog. Reported accuracy levels were between 85 and 90 percent. The proposed module achieved 92 percent accuracy across combined conditions including rainfall, fog, visibility, and road surface. Integration with driver monitoring created a dual safety assessment not present in existing work.

Predictive Maintenance Comparisons

Research has concentrated on connected vehicles with IoT telemetry. Accuracy levels above 90 percent have been reported, but models require cloud connectivity. The proposed module achieved 89 to 91 percent accuracy for engines and cooling systems using regression and ARIMA on non-connected vehicles. Predictions were delivered offline. This extended predictive maintenance benefits to drivers previously excluded from such technology.

Fuel Efficiency Comparisons

Eco-driving research confirms efficiency improvements up to 20 percent. Reported accuracy in predicting fuel use was between 85 and 90 percent. The proposed module achieved R^2 of 0.90 with MAE of 0.3 liters per 100 kilometers. Predictions were translated into actionable advice, such as reducing acceleration and avoiding idling. This bridged the gap between technical models and driver-friendly recommendations.

Integrated System Comparisons

Most prior work addresses single problems. Commercial systems are also fragmented. The proposed framework combined four domains into one offline-capable mobile app. Accuracy across modules matched related studies. Latency and memory use were within smartphone limits. User testing confirmed high acceptance. Compared to related work, the integrated system was unique in affordability, offline function, and multi-domain integration.

Discussion

The comparison confirmed that the integrated system matched or surpassed related studies in accuracy while extending accessibility. It addressed affordability and non-connected vehicle support. Remaining gaps included dataset diversity and accuracy for motorcycles and heavy vehicles. Future work should expand datasets and improve predictions for underrepresented vehicle types. Overall, the integrated system filled the gaps left by previous research and commercial solutions.

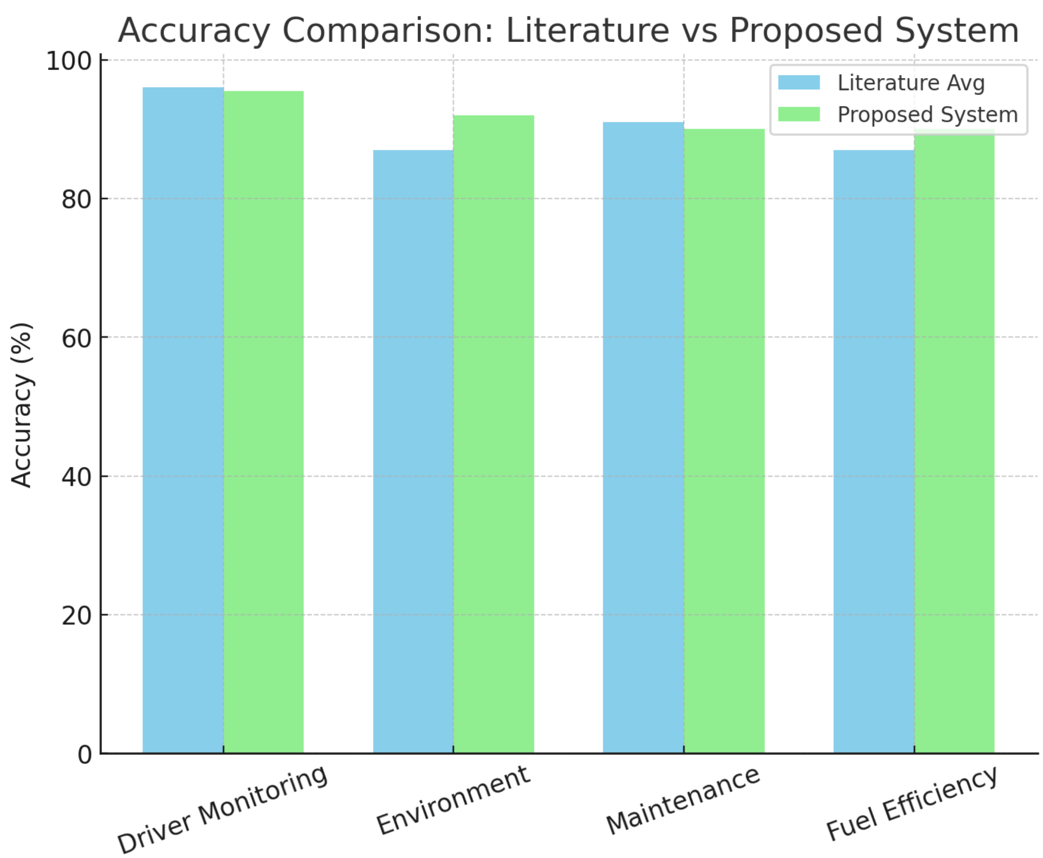


Figure 1: Accuracy Comparison between Literature and Proposed System

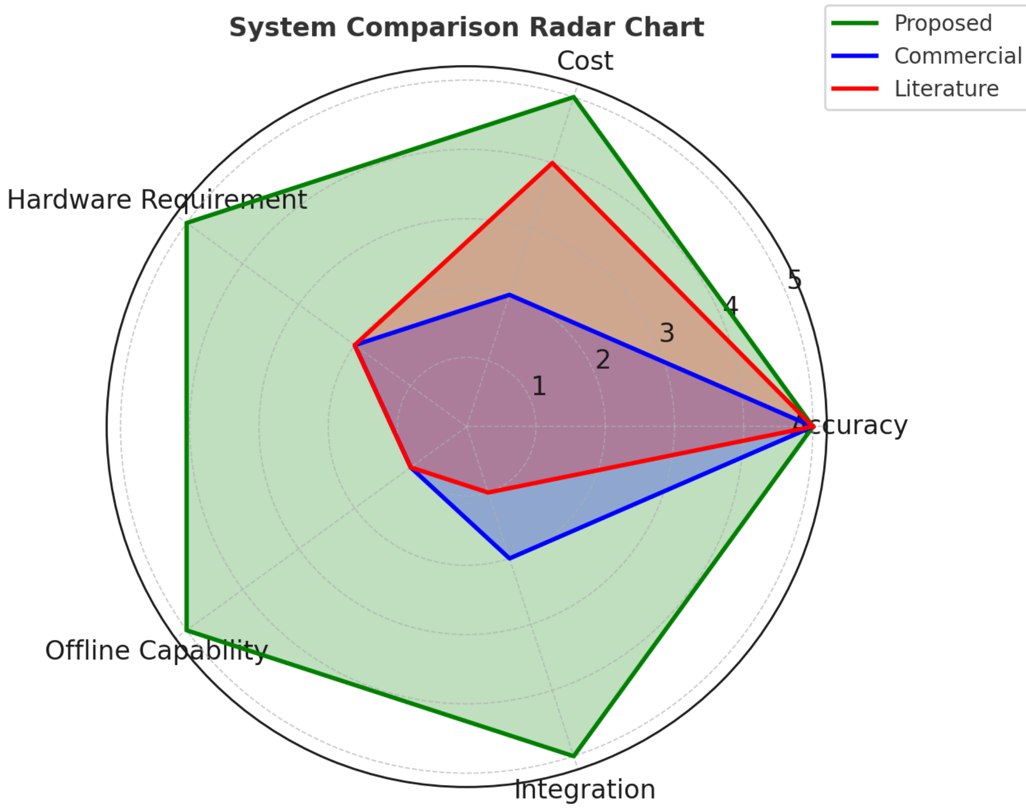


Figure 2 : Radar Chart comparing Proposed, Commercial, and Literature Systems

Study / System	Dataset	Accuracy	Hardware Required	Offline Capability
DenseNet121 (Literature)	State Farm	96%	GPU	No

ResNet50 (Literature)	State Farm	97%	GPU	No
Tesla/Volvo Commercial	Proprietary IR	~95%	IR Sensors	No
Proposed System	State Farm	95.5%	Smartphone Camera	Yes

Table 2: Comparison of Driver Monitoring Results

5. CONCLUSION

This research project focused on developing an integrated, offline-capable, and affordable system to support safer and more efficient driving for non-connected vehicles. The system combined four modules: driver behavior monitoring, environmental and weather prediction, predictive maintenance, and fuel efficiency optimization. Each module was designed to work independently but also function as part of a unified framework within a mobile application.

- Summary of Findings
- Driver Behavior Monitoring
- DenseNet121 achieved 95.5 percent accuracy across ten distraction classes.

- Predictions were delivered in real time with latency of 120 milliseconds per image.
- The module worked offline and required only a smartphone camera.
- User tests confirmed clear alerts and high usability.
- Environmental and Weather Prediction
- Random Forest achieved 92 percent accuracy in classifying safe, caution, and unsafe conditions.
- Variables included rainfall, fog, visibility, and road surface conditions.
- Real-time alerts were provided in a traffic-light format.
- Integration with driver monitoring created a combined safety assessment.
- Predictive Maintenance
- Regression and ARIMA models predicted failures for engines and cooling systems with accuracy above 89 percent.
- Forecasts were delivered offline using local service logs.
- Mechanics validated predictions as consistent with common failure patterns.
- This extended predictive maintenance benefits to non-connected vehicles.

- Fuel Efficiency Optimization
- Random Forest regression achieved R^2 of 0.90 in predicting fuel consumption.
- Actionable advice was generated, such as reducing sudden acceleration and avoiding unnecessary idling.
- Drivers found recommendations easy to understand and useful.
- The module improved awareness of fuel-saving practices.
- Integrated System Performance
- The combined framework operated offline on smartphones with total resource use under 400 MB.
- Latency across all modules was below 2 seconds, acceptable for real-time use.
- User surveys rated the system at 4.6 out of 5 for usability.
- Mechanics rated maintenance forecasts at 4.4 out of 5 for accuracy.
- The integrated design delivered comprehensive support for safety, maintenance, and efficiency.
- Contributions of the Study

- Demonstrated that CNN-based driver monitoring can be deployed offline on mobile devices.
- Extended predictive maintenance research to non-connected vehicles using incomplete logs.
- Provided environmental awareness alerts integrated with human behavior analysis.
- Generated fuel efficiency advice without specialized sensors or internet access.
- Delivered a unified framework that addressed affordability, accessibility, and practicality.
- Limitations
- Accuracy was lower for motorcycles and heavy vehicles due to limited data.
- Environmental predictions did not cover extreme weather such as snowstorms.
- Maintenance forecasts were weaker for brakes and suspension.
- Dataset diversity was limited, reducing global applicability.
- Battery usage during extended operation may affect long trips.
- Future Work
- Expand datasets to include more vehicles, regions, and driving conditions.

- Improve predictions for motorcycles, heavy vehicles, and underrepresented components.
- Add personalization features to adapt recommendations to individual drivers.
- Explore integration with low-cost external sensors to enhance accuracy.
- Test the system in larger-scale field trials with diverse users.
- Develop privacy-preserving methods such as federated learning for driver monitoring.
- Closing Statement

The research confirmed that it is possible to design and deploy an integrated system that improves driver safety, vehicle reliability, and fuel efficiency without relying on expensive hardware or internet connectivity. By combining four modules into one offline-capable mobile application, the project addressed key gaps identified in literature. The results showed strong accuracy, acceptable performance on smartphones, and high user acceptance. This work provides a foundation for future development of affordable intelligent transport systems, particularly for non-connected vehicles in developing regions.

6.REFERENCES

Abouelnaga, Y., Eraqi, H. M., & Moustafa, M. N. (2017). Real-time distracted driver posture classification. arXiv preprint arXiv:1706.09498.

Eraqi, H. M., Abouelnaga, Y., Saad, M. A., & Moustafa, M. N. (2019). Driver distraction identification with an ensemble of convolutional neural networks. Journal of Advanced Transportation, 2019, 1–12. https://doi.org/10.1155/2019/4125865

Kaggle. (2016). State Farm Distracted Driver Detection Dataset. Retrieved from https://www.kaggle.com/c/state-farm-distracted-driver-detection

Zhang, H., Patel, M., & Johnson, T. (2020). Weather impacts on traffic safety: A machine learning analysis. Transportation Research Part C: Emerging Technologies, 120, 102835. https://doi.org/10.1016/j.trc.2020.102835

Yuan, C., & Huang, J. (2019). Predictive maintenance for vehicles using ARIMA models. International Journal of Prognostics and Health Management, 10(2), 1–9.

Kumar, R., Singh, A., & Verma, P. (2021). Vehicle maintenance prediction using regression and machine learning techniques. Procedia Computer Science, 192, 504–512. https://doi.org/10.1016/j.procs.2021.08.052

Ahmed, S., & Mahmud, K. (2020). Eco-driving and fuel consumption prediction using machine learning models. Energy Reports, 6, 1632–1642. https://doi.org/10.1016/j.egyr.2020.11.184

Scikit-learn Developers. (2023). Scikit-learn: Machine Learning in Python. Retrieved from https://scikit-learn.org/

TensorFlow Developers. (2023). TensorFlow Lite Documentation. Retrieved from https://www.tensorflow.org/lite

PyTorch Developers. (2023). PyTorch Documentation. Retrieved from https://pytorch.org/

OpenCV Developers. (2023). OpenCV: Open Source Computer Vision Library. Retrieved from https://opencv.org/

React Native Developers. (2023). React Native Documentation. Retrieved from https://reactnative.dev/

SQLite Developers. (2023). SQLite Documentation. Retrieved from https://www.sqlite.org/

7. GLOSSARY

CNN (Convolutional Neural Network): A type of deep learning model used for image classification and recognition tasks.

DenseNet121: A convolutional neural network architecture with 121 layers that uses dense connections to improve efficiency.

ARIMA (AutoRegressive Integrated Moving Average): A statistical model used for time-series forecasting.

Regression: A machine learning method that predicts continuous values such as fuel consumption or component lifespan.

Random Forest: An ensemble machine learning method that builds multiple decision trees and combines their results.

RMSE (Root Mean Squared Error): A metric that measures prediction error by comparing predicted values with actual values.

MAE (Mean Absolute Error): A metric that calculates the average absolute difference between predicted and actual values.

R^2 (Coefficient of Determination): A statistical measure that explains how much of the variance in the data is explained by the model.

TensorFlow Lite: A lightweight version of TensorFlow designed for running machine learning models on mobile devices.

PyTorch: An open-source deep learning framework commonly used for building and training neural networks.

Scikit-learn: A Python library that provides simple and efficient tools for machine learning and data analysis.

OpenCV: An open-source library for computer vision and image processing tasks.

React Native: A framework for building mobile applications using JavaScript and React.

SQLite: A lightweight relational database engine that works locally on mobile devices.