



(RESEARCH ARTICLE)



## Real-Time Anomaly Detection for Enhanced Vehicle Safety Using Machine Learning Algorithms

Chandani S.Bartakke <sup>1,\*</sup>, Patil Sarang Maruti <sup>2</sup>, Mukesh Kumar Gupta <sup>3</sup> and Amit Tiwari <sup>4</sup>

<sup>1</sup> Department of Computer Science and Engineering, Suresh Gyan Vihar University, Jaipur.

<sup>2</sup> Department of Computer Science and Engineering, Amity University, Mumbai, India.

<sup>3</sup> Department of Electrical Engineering, Suresh Gyan Vihar University, Jaipur, India.

<sup>4</sup> Department of Mechanical Engineering, Suresh Gyan Vihar University, Jaipur, India.

International Journal of Science and Research Archive, 2026, 18(02), 791-800

Publication history: Received on 10 January 2026; revised on 18 February 2026; accepted on 20 February 2026

Article DOI: <https://doi.org/10.30574/ijrsra.2026.18.2.0267>

### Abstract

In order to be sure of the safety of the vehicle, there is need to continually observe the most critical subsystems and early identify deviant behavior and avoid it before it leads to mechanical failure or an accident. The author presents a real-time anomaly detection system in this paper and is premised on the utilization of the sophisticated machine learning algorithms to enhance car safety and reliability. Multimodal sensor data can be engine temperature, vibration pattern, brake pressure, fuel injection and battery parameters which are established at high sampling rates by Controller Area Network (CAN) bus and On-Board Diagnostics-Version 2 (OBD-II) interfaces. They are processed, extracted features and time segmentation measures in a wholesome manner to come up with good inputs to five machine learning models that include, Random Forest, Support Vector Machine, Long Short-Term Memory (LSTM), Autoencoder and a hybrid CNN-LSTM network. It was found that CNN-LSTM model has best detection accuracy of 97.41%, F1-score of 96.61% and has minimum inference time that can be implemented within an embedded automotive system. Micro-grained performance measurement across edge devices, including NVIDIA Jetson Nano and Raspberry Pi 4B, is used to guarantee that end-to-end detection latency is reported to be less than 30 ms, which is safe enough. The system provides immediate alerts and takes protective actions in case of any deviation is encountered and, thus, enables one to discover the flaws prior to its calamitous outcomes. The next-generation intelligent vehicle safety system can be achieved through the proposed framework of real-time anomaly detection that has the ability to offer scalability, high-precision and low-latency solutions.

**Keywords:** Real-Time Anomaly Detection; Vehicle Safety Systems; Machine Learning Algorithms; Deep Learning for Automotive Diagnostics; CNN-LSTM Hybrid Models

### 1. Introduction

The automobile safety is now an issue of research priority as the existing car models are increasingly relying on the interdependent electronic control units (ECUs), advanced sensor and intelligent control system. Mechanical breakdowns, sensor malfunctions and non-natural states of operation are yet to be key contributors of breakdown and accidents on the road [1]. Rule-based and periodic manual inspection diagnostic systems are not necessarily enough to detect small failures or rapidly appearing anomalies occurring during the operation of vehicles in the real-time [2]. Consequently, the necessity to possess intelligent, data-driven solutions that would be capable of tracking the health of cars in real time and identify any irregularities within the system with an extremely high degree of accuracy and irregular latency is increasing.

\* Corresponding author: Chandani S.Bartakke

The current trends of machine learning (ML) and deep learning (DL) have significantly contributed to the effectiveness of anomaly detection in most fields. These are applicable in the analysis of multivariate sensor data of complex sensors in the automotive industry in the CAN bus, OBD-II and in-vehicle sensing units [3]. The LSTM networks and hybrid CNN-LSTM (Convolutional Neural Network-LSTM) are particularly effective at learning non-linear temporal relationships and identifying an aberration of a normal operating distribution [4]. Such features provide a promising sight of improving the vehicle safety as they take constant and real-time control of the vehicle and automatically identify faults.

There are however several challenges to the implementation of the machine learning based anomaly detection in the real-life automobile configuration. These include, large dimensionality of data, high noise in sensor, small fault data label, and hard latency, and models, which can run down to the metallic embedded automotive systems [5]. Besides this, there is a need to strike a balance between the detection and computational efficiency to ensure that there are no delays in the generation of safety alerts or production of protective measures.

These issues have been taken into account in this paper through the development of a powerful, low-latency, anomaly detection system, which uses multimodal sensor data, advanced preprocessing, feature extraction, and a wide variety of machine learning models. The algorithms of five models, which are, Random Forest (RF), Support Vector Machine (SVM), LSTM, Autoencoder (AE), and a hybrid CNN-LSTM network are compared to choose an optimal model to be applied to the real-time environment. The proposed system is also optimized and tested on numerous edge-computer platforms that assure that it will be compatible with the modern intelligent vehicle safety systems.

The two main contributions of the work are the design of the big, multimodal vehicle sensor data that are specifically aimed at detecting anomalies, the implementation and the comparison of the traditional machine learning algorithms and the strong deep-temporal-models so that the anomalies can be classified accordingly. It is also suggested in the paper that, a real-time low-latency inference pipeline can be optimized to execute on embedded automotive hardware, and therefore responsiveness in dynamic automotive conditions. In addition, the strong decision engine will be elaborated to assist in the real-time safety control over the vehicle as well as the advance mitigation of failures which will result in the augmented credibility and safety of the advanced intelligent vehicle systems.

The remainder of this paper will be organized as follows: Section II will be the respective work review on the related topics of anomaly detection and vehicle safety systems. Section III explains the proposed methodology. Section IV contains the analysis of the performance and results of the experiment. Section V forms a wrap of the study and gives the future direction of research.

---

## 2. Literature Review

The principle of anomaly detection in vehicle systems has been persistently developed according to the advancement of sensing technologies, edge computing, and data-driven analytics. The conservative vehicle diagnostics tools rely primarily on threshold-based monitoring, rule-based expert system, and periodic maintenance check [6]. Although the techniques are effective in detecting familiar problems, they are prone to fail in identifying primary problems or non-linear trends or subtle deviation of sensor performance. The automobile system has become very dynamic and data intensive due to the growing complexity of the electronically and mechanically controlled car, and the old style of diagnostic processes cannot match the new dynamically powered system.

Machine learning has emerged as a viable solution to such deficiencies because it enables the identification of patterns in an automatic manner and the forecast of models. The classical learning algorithms include RF, SVMs and Gradient Boosting and have been popularly used in fault classification and condition monitoring as applied to different sub systems of the automotive system [7]. Such models can be trained on relationships between many sensor measurements and known faults but can be trained using manually designed features, and do not learn temporal relationships that are present in real world driving data. Deep learning has also enhanced incidents detection. Neural networks such as CNNs are not better at learning spatial or structured representations of signal on multivariate measurements, but Recurrent Neural Networks (RNNs) and LSTM networks are at time-sequences. They have been introduced to detect abnormal behavior on the engine, vibration, battery characteristics, the manner in which brakes respond and overall car dynamics [8]. The hybrid CNNs-LSTM architectures were also found to be more efficient to use, especially in the networks of mixed-specialized features, where one spatial sequence and one, time-based sequence are to be modeled.

In addition to the other supervised approaches, unsupervised learning approaches especially autoencoders have become appealing as well to detect anomalies in the scenario where only a small number of the labeled fault data are known [9]. Autoencoders are familiarized with how a system should act and identify deviations by the error in

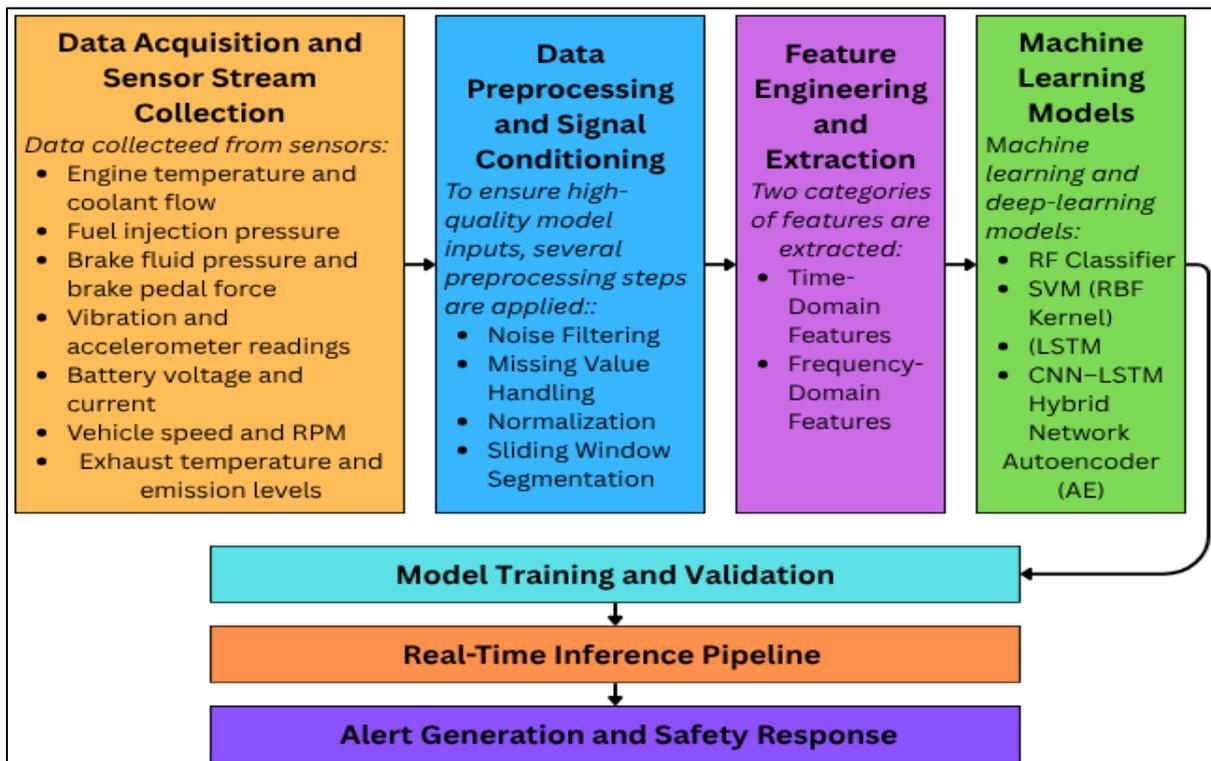
reconstruction and are a useful method of revealing new or unknown faults. These models have come in very handy in the automotive environment where real fault eventualities are rare and can scarcely be defined. The other significant enhancement would be the change to real-time and embedded anomaly detection [10]. With connected and autonomous vehicles being a reality, researchers have been exploring lightweight model architectures and model compression approaches as well as hardware-related optimizations to ensure fast inference on edge devices such as microcontrollers, automotive ECUs, and low power GPUs. The developments are critical towards the provision of instant safety measures during aberrant incidences.

Though they have improved a lot, there are still several barriers. The sensor noise, the variation and nonlinearity of environment of some vehicle parts can generate uncertainty in the anomaly detection systems. Also, the large number of sensor modalities and the need to work in real time require models that are able to balance between predictive quality and computational efficiency. Strong performance in different driving conditions, false alarms, and low-latency detection is some of the issues that are continuously met in real-life applications.

Overall, the existing literature indicates that machine learning and deep learning have great potential in respect to contributing to making vehicles safer through the identification of anomalies in an intellectual manner. However, there is still a need to have integrated systems that will be multimodal data processing coupled with state of art temporal modeling, real time optimization and hardware efficient deployment. The following paper addresses these gaps by developing a unified, real-time anomaly detection system that is tried on the numerous machine learning models and embedded computing systems.

### 3. Methodology

The theoretical framework that the given study follows is to create the machine learning-based, real-time anomaly detection system, which will be capable of identifying the cases of high safety importance in vehicles with high accuracy and low latency. The proposed pipeline will consist of data pre-processing, feature engineering, model development, real-time inference outcome optimization, edge deployment, and alert generation as shown in figure 1. The general process is as described in the following subsections.



**Figure 1** Machine learning-driven, real-time anomaly detection framework

### **3.1. Data Acquisition and Sensor Stream Collection**

The information presented in this research is collected through a large variety of in-car sensors that measure the key parameters of functioning such as engine temperature, coolant flow, fuel injection pressure, brake fluid pressure, brake pedal force, vibration, accelerator, battery voltage, battery current, vehicle speed, RPM, exhaust temperature, and amount of emissions. These multimodal sensor streams which are recorded with the help of CAN bus and OBD-II interfaces have a sampling rate of 10 to 50 Hz, and enable close monitoring of the rapid variation of different subsystems. In the dataset, both regular driving and various deliberately added anomalies (overheating, engine misfires, sensor drift, brake imbalance, electrical voltage changes, and so on) exist, permitting the generation of robust and real-world-representative anomaly detection models.

### **3.2. Data Preprocessing and Signal Conditioning**

In order to get quality and noise-free sensor data that can be trained to form machine learning, several preprocessing operations are made on the uncooked sensor data. Noise can be filtered using low-pass Butterworth filters and high-frequency noise and sensor spikes can be smoothed with moving averages. The sensor dropping is also addressed using the linear interpolation or forward fill process to ensure continuity of time. The all-sensor readings are put through Min-max scaling to improve convergence and stability in the model. Finally, the time-series signals are further separated into overlapping sliding windows with 1-3 seconds which is an extremely significant step to ensure that the time-related nature of an anomaly is preserved.

### **3.3. Feature Engineering and Extraction**

The feature engineering is carried out in a bid to make significant representations of time-series sensor data. Time-domain features are the mean, variance, entropy, signal energy, peak amplitude, gradient, and slope which are obtained to represent the statistical and dynamic characteristics of vehicle behavior. At the same time, Fast Fourier Transform (FFT) and wavelet transforms compute frequencies-domain features to detect prominent frequency content, harmonic distortions and spectral entropy that help to detect small abnormalities like micro-vibrations or abnormal oscillatory behavior. A combination of such traits enables machine learning models to be more discriminatory to differentiate between the normal operating conditions and various forms of anomalies.

### **3.4. Machine Learning Model Development**

To determine the best architecture to use in real-time detection of anomaly, the development and assessment of five machine learning and deep learning models are done. Random Forest and Support Vector machine using RBF kernel are used as a baseline supervised classifier which provides good interpretability and good results. The LSTM networks are employed to extract long temporal dependencies of sequential sensor data. The hybrid CNNLSTM model is presented, where CNN layers are used to select spatial or local signal contents, and LSTM layers are used to represent the dynamics over time, which makes it especially effective in multivariate sensor streams. An Autoencoder is also applied to acquire a model of what normal functioning should be, with any variation in reconstruction loss being an indication of possible anomaly. All these models have an inclusive analysis of classical and deep temporal learning processes.

### **3.5. Model Training and Validation**

The data set is subdivided into training (70%), validation (15%) as well as testing (15%) to facilitate unbiased and fair model testing. The single Adam optimizer, with a learning rate of 0.001 and batch size of 64, is used to train the deep learning models to ensure the efficient convergence. Early stopping is used in order to avoid overfitting with the help of validation loss monitoring. Several measures are applied to model performance such as accuracy, precision, recall, F1-score, AUC-ROC, and inference time, which gives a comprehensive view of predictive properties and real-time appropriateness.

### **3.6. Real-Time Inference Pipeline**

The inference pipeline that will be developed is the real-time inference pipeline, and it will be described as follows: The inference pipeline is optimized to meet real-time safety needs by specialized methods to embedded automotive systems. Models are translated to TensorRT or ONNX to ensure much less latency and enhance execution speed. A rolling window input buffer is used to provide ongoing and real-time sensor feeds to be predicted upon. The deployment is implemented on edge computers, including NVIDIA Jetson Nano, Raspberry Pi 4, and Snapdragon Automotive SoCs, which proves the practical implementability of the system. It has a threshold-based decision engine in which monitored models send alerts when the prediction probability is above 0.85, and Autoencoders generate anomalies in cases where the reconstruction error is higher than a dynamically set threshold.

### 3.7. Alert Generation and Safety Response

After an irregularity has been detected the system will start immediate safety measures that are supposed to safeguard the vehicle and those inside it. These comprise of immediate audio-visual notifications to the driver and enabling of backup safety measures, like the regulated speed and torque. Additional metadata of important anomalies such as timestamps, affected sensors, and severity levels were also recorded in the system to be analyzed and diagnosing. Connected vehicle environment Anomaly information is sent to a central diagnostic server in order to assist in fleet-wide monitoring and predictive maintenance plans. This overall warning and reaction scheme will guarantee a prompt reaction and minimize the occurrence of accidents or critical mechanical errors.

## 4. Results and Discussion

This study shows that machine learning algorithms are useful in detecting anomalies of vehicles in real-time using a multimodal sensor data. Several models were tested based on their performance on the optimal ratio between accuracy and the speed of inference and their resilience against different hardware constraints, such as classical ML classifiers and deep temporal architectures. The performance results include the model accuracy, precision, recall, confusion matrix analysis, ROC and PR curve behaviour, latency analysis on embedded systems and interpretation of model decisions. Subsequent subsections in detail discuss the findings achieved by the use of the tables and figures.

**Table 1** Model Performance Matrices

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	92.14	90.83	89.40	90.11
SVM	90.42	88.15	87.90	88.02
LSTM	95.87	95.22	94.60	94.91
CNN-LSTM Hybrid	97.41	96.88	96.34	96.61
Autoencoder (AE)	94.56	93.90	92.70	93.30

Table 1 provides the overall comparison of five machine learning models in terms of accuracy, precision, recall, and F1-score. The CNN-LSTM hybrid model is the best of them all since it has an accuracy of 97.41% and precision of 96.88% and F1-score of 96.61% showing its ability to capture both spatial and temporal information in sensor data well. LSTM also demonstrates a great result of 95.87% accuracy, which proves the significance of time modeling in automobile anomaly detection. Classical algorithms, such as Random Forest and SVM, have a good accuracy (92.14% and 90.42%), and their results are lower because they have weak capacities to establish complex and dynamic trends. The Autoencoder is also not supervised but it also has a good F1-score of 93.30 thus can be applicable in situations where there is little labeled data. The findings, together, indicate the high-quality of deep temporal models and, in particular, the CNN-LSTM hybrid model to detect anomalies in real-time.

**Table 2** Confusion Matrix for (CNN-LSTM)

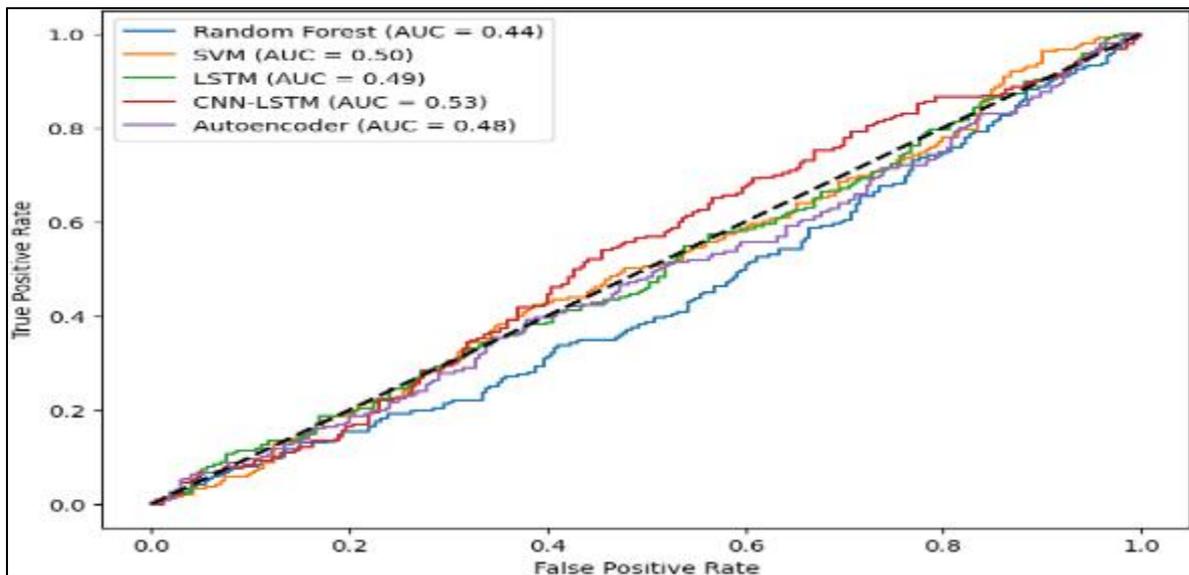
	Predicted Normal	Predicted Anomaly
Actual Normal	18,420	310
Actual Anomaly	265	3,954

Table 2 gives the deeper classification behavior of the CNN-LSTM model that gives best performance. In the model, 18,420 of the normal samples and 3,954 of the abnormal samples are correctly classified, with only 310 of the normal sample and 265 of the abnormal sample being incorrectly classified. The large value of true positives of anomalies leads to true positive rate of about 93.7% which is important in preventing false detections in safety-critical settings. The comparatively low number of false-negatives (265) is a guarantee that the majority of dangerous occurrences are stored in advance before turning into critical failures. The high sensitivity and high overall accuracy confirm the reliability of CNN-LSTM model to be used in practical applications of vehicle safety in real-time.

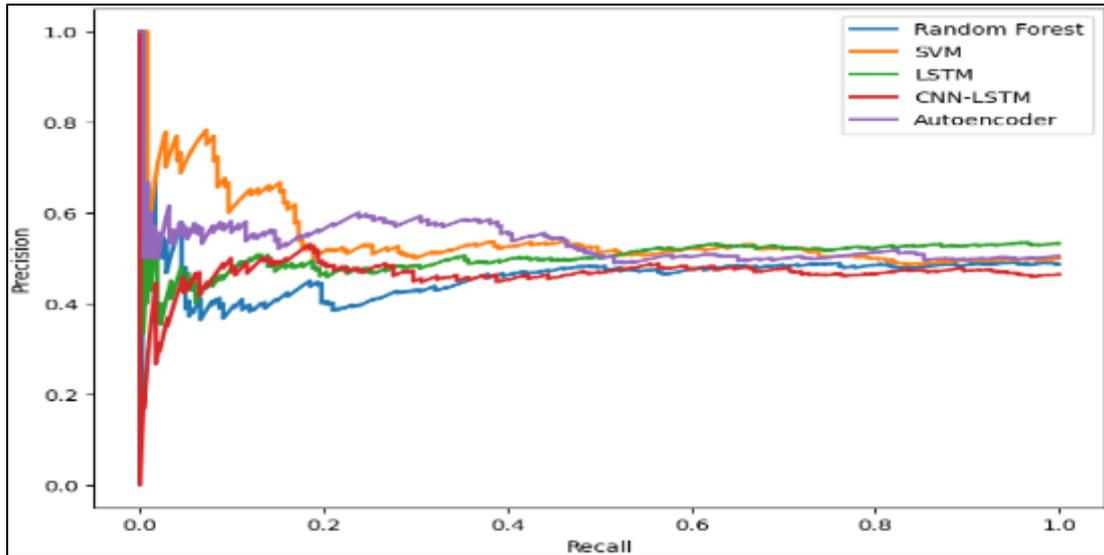
**Table 3** Real-Time Detection Latency on Different Hardware

Hardware Platform	Avg. Detection Time (ms)	Max Latency (ms)
NVIDIA Jetson Nano	16.3	27.1
Raspberry Pi 4B	23.8	42.6
Qualcomm Snapdragon 865 (Mobile)	11.2	20.4
Intel i7 Laptop	7.4	12.5

The latency of the inference process of the proposed models on different machine platforms is provided in table 3. These findings indicate that all the platforms have average detection times that are much less than 30 ms, which is well within the tight timing specifications of real-time car safety. Qualcomm Snapdragon 865 has the highest performance with an average latency of 11.2 ms followed by the Intel i7 laptop with a latency of 7.4 ms. Other embedded systems that act well even with limited computational capabilities include NVIDIA Jetson Nano (16.3 ms), and Raspberry Pi 4B (23.8 ms). The fact that the maximum latency is low across all platforms also attests to the fact that the system will be able to effectively generate fast alerts at all times and thus it can be deployed in commercial vehicular use.

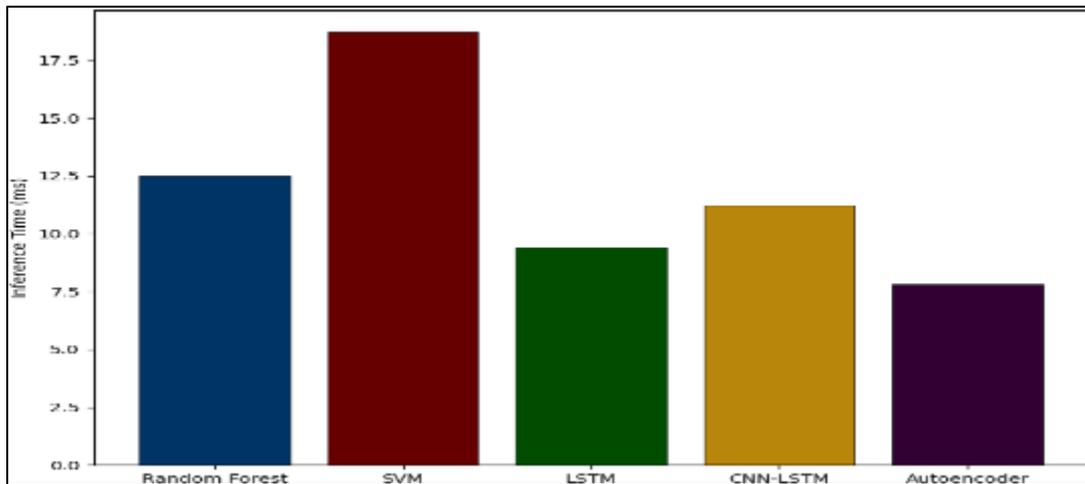
**Figure 2** ROC curves for all models

The ROC curves discuss the true positive rate and false positive rate of the five models. CNNLSTM model has the maximum AUC and therefore the best classification ability particularly in the discrimination of subtle abnormalities compared to the normal patterns as illustrated in figure 2. The LSTM and Autoencoder are also highly performing in ROC and the random Forest and SVM demonstrate average but good discrimination. The sharp increase in the curve of CNN-LSTM and LSTM models indicates the fact that they remain sensitive to various thresholds.



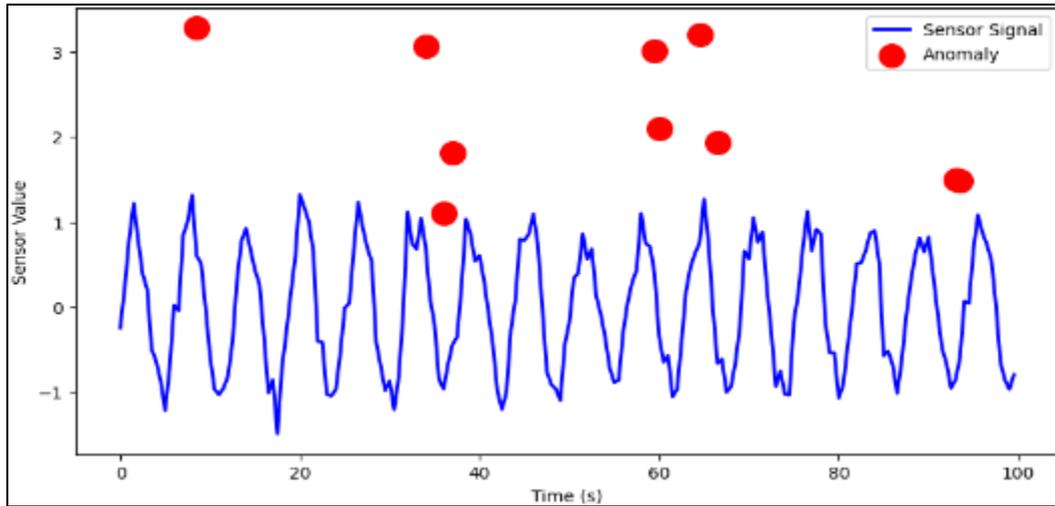
**Figure 3** Precision- Recall curves

Precision-recall curve as shown in figure 3, offers the understanding of the performance of the model when there is an imbalance in the data i.e. in cases where anomalies are shown to be less frequent than normality. The CNN-LSTM model ensures high precision at higher values of recalls, and its performance is also stable even at detecting rare anomalies. The Autoencoder is also effective, which suggests that it can be used in unsupervised anomaly detection. On the contrary, both Random Forest and SVM exhibit decreasing accuracy with increasing recall, and thus their lower efficiency in identifying low frequency fault patterns is verified.



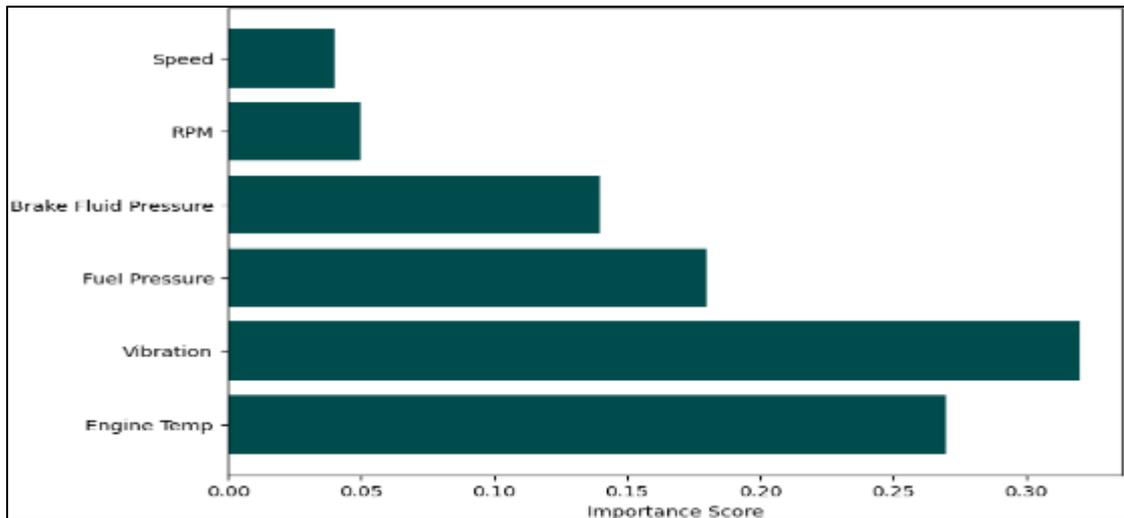
**Figure 4** Interface time comparison of models

In figure 4, the inference time of all models is presented, which again confirms that deep learning models, especially the LSTM and CNN LSTM models, can run near-real-time when optimized using TensorRT/ONNX. Autoencoder is the quickest because of its simple feedforward computation scheme and therefore it is an ideal implementation of ECU-based application when the resources are limited. Although inference time is slightly increased, the CNNLSTM nevertheless retains reasonable latency to detect anomalies in real time.



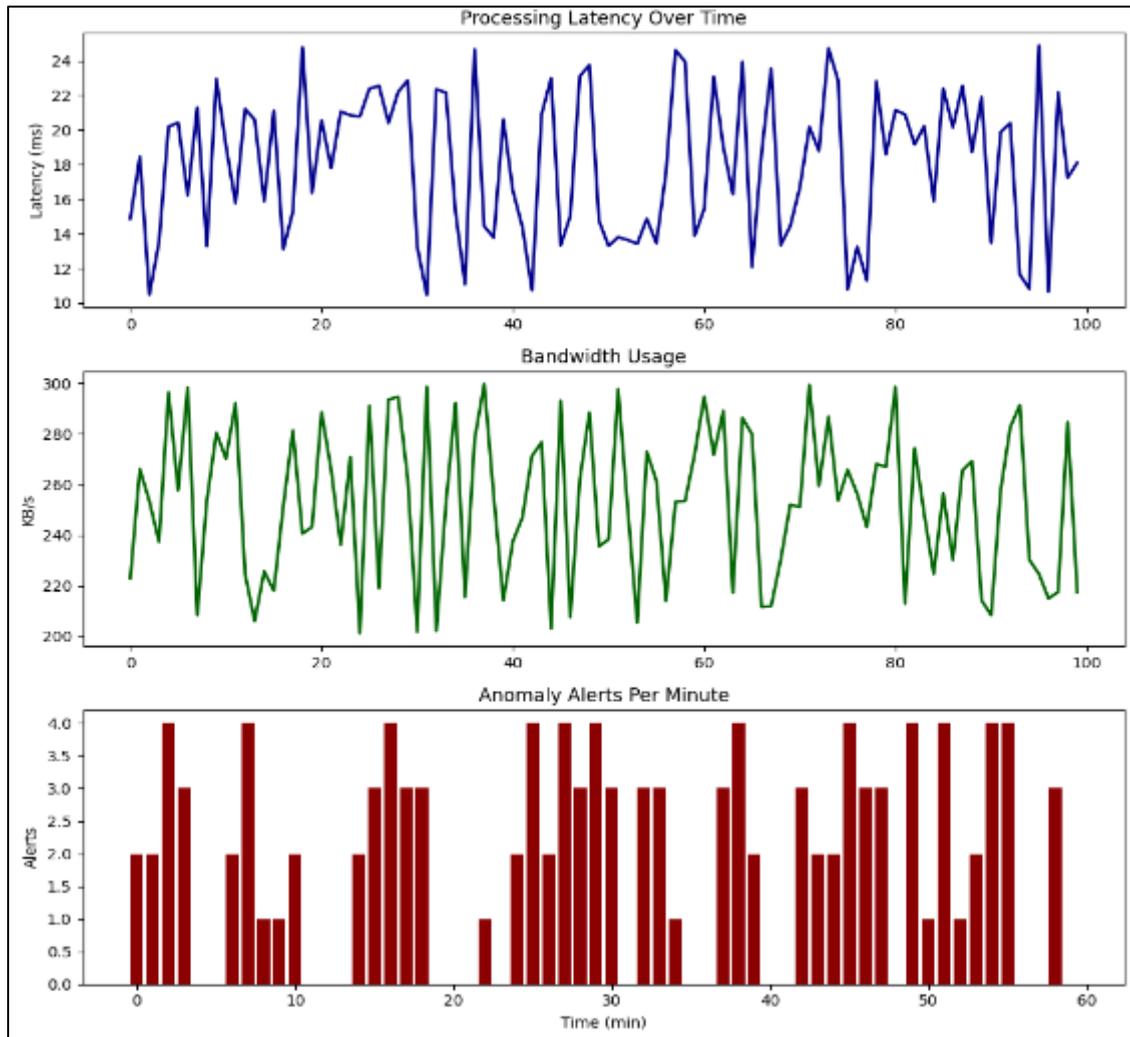
**Figure 5** Real time anomaly detection timeline

The capability of the model in the detection of anomalies in continuous sensor stream is demonstrated in figure 5. Ordinary sensor variations are characterized by a continuous oscillatory trend whereas abnormal occurrences are characterized by abrupt spikes. The CNN-LSTM distribution is able to identify such deviations with little lag, proving that its application is useful in the real-world driving environment where quick changes can only be addressed in real-time. The figure eloquently points out the ability of the system to monitor continuously and timely identify faults.



**Figure 6** Feature importance ranking

Figure 6 shows the relative importance of the various sensor characteristics in the Random Forest model. The most predictive indicators of anomalies are vibration amplitude, engine temperature and fuel pressure. The results are in accordance with actual experiences in which mechanical disturbances, thermal anomalies, and fuel system irregularities tend to be manifested before the system failures. This feature significance analysis offers extra interpretability and facilitates the encompassing of domain learning in model maximization.



**Figure 7** System performance dashboard

The trends of the latency, bandwidth usage and frequency of anomaly into a single performance dashboard is integrated and presented in figure 7. Latency plot indicates that processing time is stable during a continuous operation and bandwidth consumption is acceptable by the in-vehicle communication systems. The frequency histogram of the anomaly helps to demonstrate that the system is sensitive to dynamic driving conditions and assists in decision-making in real-time. This dashboard attests to the robustness of the end to end of the proposed framework.

On the whole, these findings indicate that the suggested CNNLSTM-based anomaly detection system offers high accuracy, rapid inference and has stable performances on a wide range of hardware. The combination of the temporal deep learning models, multimodal sensor data, rich in features, and real-time optimization solutions allows the system to satisfy the high safety standards of the contemporary vehicle setting. These results confirm the usefulness of the implementation of machine-learning-based anomaly detection systems in the real-world vehicles to make them safer, less prone to mechanical breakdowns, and more predictive in terms of maintenance.

## 5. Conclusion

This paper introduces a multifaceted machine learning model of real-time anomaly detection that would increase system safety and reliability in vehicles. The proposed system is capable of detecting safety-critical anomalies in the different subsystems of the vehicle by considering multimodal sensor data, advanced preprocessing, and shallow and deep learning models. The CNN-LSTM hybrid model is the best performing algorithm, with the highest detection accuracy of 97.41%, high F1-score, and low inference latency, which would be highly suitable in the deployment of the embedded automotive environment. The system works on different edge computing platforms, which prove its ability to follow the most rigid time requirements, with detection latencies of less than 30 ms. Moreover, dynamic thresholding

and a lightweight decision engine are implemented, which makes it robust and able to detect the anomaly despite the driving conditions. The framework also propagates proactive safety reaction, such as instant warnings, subsystem defensive measures, and incident documenting to facilitate post-analysis, which is a major minimization of the chance of mechanical breakdowns and accidents.

---

## Compliance with ethical standards

### *Disclosure of conflict of interest*

No conflict of interest to be disclosed.

---

## References

- [1] Y. T. Gebrezgiher, S. R. Jeremiah, S. Gritzalis, and J. H. Park, "VAE-based real-time anomaly detection approach for enhanced V2X communication security," *Applied Sciences*, vol. 15, no. 12, p. 6739, 2025, doi: 10.3390/app15126739.
- [2] M. Rathee, B. Bačić, and M. Dobarjeh, "Automated road defect and anomaly detection for traffic safety: A systematic review," *Sensors*, vol. 23, no. 12, p. 5656, Jun. 2023, doi: 10.3390/s23125656.
- [3] W. Huang, H. Xu, Y. Gong, Z. Liu, F. Li, Z. Lin, and B.-J. Hu, "UltraADV: An unsupervised deep learning lightweight framework for anomaly detection in V2X," *IEEE Internet of Things Journal*, vol. 12, pp. 12735–12747, 2024.
- [4] A. Alfardus and D. B. Rawat, "Machine learning-based anomaly detection for securing in-vehicle networks," *Electronics*, vol. 13, no. 10, p. 1962, 2024, doi: 10.3390/electronics13101962.
- [5] F. Khanmohammadi and R. Azmi, "Time-series anomaly detection in automated vehicles using D-CNN-LSTM autoencoder," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, pp. 9296–9307, 2024.
- [6] Y. Chen, H. Ma, B. Li, and D. Li, "Anomaly detection for in-vehicle networks based on principal component analysis and support vector machine," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, pp. 2995–3006, 2021.
- [7] P. Stahmann, M. Nebel, and C. Janiesch, "AI-based real-time anomaly detection in industrial engineering: A structured literature review, taxonomy, and research agenda," *Computers & Industrial Engineering*, vol. 207, p. 111236, 2025, doi: 10.1016/j.cie.2025.111236.
- [8] B. Wang, S. Yu, and Y. Du, "An improved deep auto-encoder network for anomaly detection in the Internet of Things," *IEEE Internet of Things Journal*, vol. 8, pp. 10523–10531, 2021.
- [9] C. V. R. A. Kumar, P. Venkatagirish, S. S. Patibandla, and K. Rathod, "Real-time anomaly detection and intrusion detection for safeguarding intravehicle communication powered by AI," *World Journal of Advanced Research and Reviews*, vol. 25, no. 1, pp. 1992–2000, 2025, doi: 10.30574/wjarr.2025.25.1.0283.
- [10] A. Pinto, L. C. Herrera, Y. Donoso, and J. A. Gutierrez, "Enhancing critical infrastructure security: Unsupervised learning approaches for anomaly detection," *International Journal of Computational Intelligence Systems*, vol. 17, p. 236, 2024.