International Journal of Computer Science and Engineering Innovations
ISSN: 3107-9458 | Volume 1 Issue 2, 38-47, October-December 2025
DOI: https://doi.org/10.64137/31079458/IJCSEI-V1I2P106

Received: 05/12/2025 Revised: 09/12/2025 Accepted: 27/12/2025 Published: 31/12/2025

Original Article

Deep Learning-Enabled Industrial IoT Framework for Autonomous Fault Forecasting in Connected Diesel Engine Systems

LUCAS VAN DIJK
Junior Researcher, Amsterdam Metropolitan University, Netherlands.

ABSTRACT: *Diesel engines power 80% of commercial traffic and 80% of the world's merchant fleet. Almost all ship propulsion engines are diesel. Because even unnoticed faults can lead to considerable losses, Predictive Maintenance (PdM) applications are needed to guarantee Availability, Reliability, and Safety. The proliferation of sensors surrounding these machines and their connectivity in the Industrial IoT (IIoT) creates a new ecosystem for Condition-Based Maintenance (CBM). Autonomously monitoring the data generated within newly connected diesel engine ecosystems would guarantee a faster response to detected symptoms. Data from analogue sensors, digital binary sensors, and telemetry lines would act as predictors to feed a fault classification engine. This would allow faster decision-making by maintenance teams and ultimately reduce repair costs and machine downtime. Connectivity enhances intelligent systems equipped with Artificial Intelligence (AI) algorithms capable of autonomously processing and analyzing the generated data. A variety of Data Challenge tasks, offered in competitions on crowdsourcing platforms like Kaggle, could be solved using the same principles. In these competitions, databases often link telemetric information of diesel engine vehicles with driving qualities like Comfort or Safety, or with the Failure classification of electric components. Such competitions generally seek to predict the Failure code.*

KEYWORDS: *Predictive Maintenance Systems, Diesel Engine Analytics, Industrial IoT Monitoring, Condition-Based Maintenance, Autonomous Fault Detection, Sensor Data Intelligence, Telemetry-Based Diagnostics, AI-Driven Maintenance, Failure Prediction Models, Intelligent Engine Monitoring.*

1. INTRODUCTION

Addressing issues related to connected diesel engine ecosystems requires an integrated, multidisciplinary approach. The Internet of Things (IoT) and Deep Learning (DL) methodologies are typically employed for gathering data from diesel engine ecosystems. When connected in Industry 4.0 fashion, a wealth of sensory information can be obtained and analyzed for fault prediction. One such predictive component uses Convolutional Neural Networks, which are trained to classify real operational signals against predefined performance criteria. The resultant architecture performs real-time, automatic, autonomous under-critical fault detection. It warns whether the overall observed condition of the engine must be addressed or the instantaneous fault is not disallowed. The prediction module is now integrated into an industrial context through a system that communicates with the AWS Cloud.

A typical microcontroller regularly collects and pushes the predictive analysis to the Cloud Platform, where the fault status is stored and made accessible. The proposed system architecture is used to evaluate all the previous approaches and develop predictive Pistons-Planetary Gearbox-Common Rail problems. Streamed data can be analyzed by algorithms- AutoEncoder / Convolutional Neural Network-LSTM for under-critical and catastrophic problems in Pistons, Planetary Gearbox and Common Rail, with Auxiliary Domain adversarial Labelling for any fault discovery. Adverse selection persists to allow neutralization of adversarially influenced sources of bias while fully enforcing critic quality with the aid of locally sensed Data in Augmented Learning.

2. BACKGROUND

Diesel engine ecosystems are the power sources for all forms of transportation, with connectivity providing the means for IoT application development. Diesel engines are therefore among the most widely used, commercialized, and economically viable power sources worldwide. Data from connected diesel engine ecosystems enable predictive fault-repair models that implement fault-classification or fault-sequence prediction methods. Diesel-fuel-consumption data, engine-speed data, and I/O signal data have been openly sourced; air-angle data, however, are available to authorized persons. Other sources in the Australian National Data Archive include sensor data and images. The models developed use early-stage sensor data and diesel engine holograms obtained in video format as sources, and corresponding high-quality images from the ImageNet database as the target.

The implementation and evaluation of an autonomous-fault-prediction module that leverages deep learning, industrial IoT infrastructures, and the Australian National Data Archive are reported. The module's overall architecture supports the use of Holistic Deep Learning models, which integrate multimodal data from diverse sensor sources within Industrial Internet of Things-based infrastructures. Fire detection and air-angle-control-fault-sequence prediction are demonstrated to achieve reliable classification and high-level, interconnection-level prediction accuracy, respectively. Such early-stage detection is important, as it not only reduces waiting time in a client-server environment but also contributes to reliable and secure behaviour classification of diesel engine ecosystems.

2.1. DIESEL ENGINE ECOSYSTEMS AND CONNECTIVITY

The evolution of industrial Internet of Things (IoT) technologies made it possible to interconnect diesel engine ecosystems smartly and efficiently. The fuel-oil systems of diesel engine ecosystems are smartly interconnected with an autonomous control and monitoring mechanism. Employed sensors with limited capabilities send data to a fog node for preprocessing, storage, filtering, and induction of information from the connected Internet. The processed information is shared with the respective engine manufacturer and service personnel for effective fault prediction and preventive maintenance.

Interconnected diesel engine ecosystems minimize human intervention and optimize resource usage in real time by predicting and controlling parameters through various controller components based on present and past sensor data. The components of the diesel engine ecosystem are smartly connected through a data ingestion mechanism called a data lake, which ingests all data in raw format as data streams and processes it using a streaming data batch and an irregular batch-backup mechanism. The controlled parameters from the data lake are stored in an on-premises database and moved to a cloud-based storage database for easy access by the stakeholders.

2.2. MATHEMATICAL FORMULATION

2.2.1. SENSOR TELEMETRY AND SIGNAL MODEL

Each diesel engine node generates a multivariate time-series vector at every sampling instant t. Let S(t) denote the sensor observation vector collected from N sensors over a sliding window of length W:

$$S(t) = [s_1(t), s_2(t), ..., s_n(t)]^T \in \mathbb{R}^N \qquad \text{(Eq. 1)}$$

where $s_i(t)$ is the reading of the i-th sensor at time t, and N is the total number of monitored parameters (N = 12 in the experimental setup).

The windowed feature tensor X used as input to the deep learning model is defined as:

$$X = \{S(t), S(t-1), ..., S(t-W+1)\} \in \mathbb{R}^{N\times W} \qquad \text{(Eq. 2)}$$

where W is the sliding window length (W = 64 samples at 100 Hz, yielding 640 ms of context), and each column represents a temporal snapshot of all N sensor readings.

2.2.2. FAULT DETECTION ACCURACY METRICS

Binary fault classification performance is measured using Precision (P), Recall (R), and the harmonic mean F1-score. For the multi-class scenario with C fault classes, the macro-averaged F1-score is:

$$P^c = TP\_c / (TP\_c + FP\_c); \quad R^c = TP\_c / (TP\_c + FN\_c) \qquad \text{(Eq. 3)}$$

where TP_c, FP_c, and FN_c are true positives, false positives, and false negatives for class c ∈ {1, …, C}.

$$F1\_macro = (1/C) \cdot \Sigma\_\{c=1\}\ \hat{}\{C\}\ [2\cdot P^c\cdot R^c / (P^c + R^c)] \qquad \text{(Eq. 4)}$$

where C = 6 is the number of fault classes, and the summation is taken over all classes to produce the unweighted average.

2.2.3. END-TO-END LATENCY MODEL

The total system latency L_total is decomposed into four additive components across the IIoT pipeline:

$$L\_total = L\_acq + L\_tx + L\_proc + L\_inf \qquad \text{(Eq. 5)}$$

where L_acq is sensor acquisition time, L_tx is network transmission latency, L_proc is feature extraction and streaming analytics processing time, and L_inf is deep learning model inference time.

For Azure-hosted inference under a Service Level Agreement (SLA) of 50 ms, the constraint is:

$$L\_total \le L\_SLA \equiv L\_acq + L\_tx + L\_proc + L\_inf \le 50 \text{ ms} \qquad \text{(Eq. 6)}$$

where the experimental system achieves L_total = 38.6 ms (mean), satisfying the SLA with a margin of 11.4 ms.

2.2.4. CNN-LSTM HYBRID ARCHITECTURE

The proposed model applies 1-D convolutional layers to extract local temporal patterns, followed by LSTM layers for long-range dependency modelling. The convolutional feature map for the k-th filter at position t is:

$$h_k(t) = \mathrm{ReLU}\left(\sum_{j=0}^{K-1} w_k(j) \cdot X(t+j) + b_k \right) \qquad \text{(Eq. 7)}$$

where w_k is the k-th convolutional kernel of length K, b_k is the bias, and ReLU is the rectified linear unit activation.

The LSTM cell update equations governing the recurrent state are:

$$f_t = \sigma(Wf \cdot [h_{t-1}, x_t] + bf) \; i_t = \sigma(Wi \cdot [h_{t-1}, x_t] + bi) \; \tilde{c}_t = \tanh(Wc \cdot [h_{t-1}, x_t] + bc) \; c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \qquad \text{(Eq. 8)}$$

where f_t, i_t are the forget and input gates; c_t is the cell state; σ denotes the sigmoid activation; and ⊙ is element-wise multiplication.

2.2.5. OPTIMIZATION OBJECTIVE

The model is trained by minimizing the categorical cross-entropy loss with L2 regularization to prevent overfitting on the imbalanced fault dataset:

$$L(\theta) = -(1/N_s) \sum_{i=1}^{N_s} \sum_{c=1}^{C} y_{i,c} \log(\hat{y}_{i,c}) + \lambda \|\theta\|^2 \qquad \text{(Eq. 9)}$$

where N_S is the number of training samples, $y_{i,c} \in \{0,1\}$ is the ground-truth label, $\hat{y}_{i,c}$ is the model softmax output for class c, θ is the parameter vector, and $\lambda = 10^{-4}$ is the regularization coefficient.

2.2.6. OPERATIONAL COST MODEL

The total operational cost per hour per edge node in the Azure-connected ecosystem is modelled as a linear combination of compute, storage, and data-transfer costs:

$$C_{op} = \alpha \cdot C_{compute} + \beta \cdot C_{storage} + \gamma \cdot C_{transfer} \qquad \text{(Eq. 10)}$$

where α, β, γ are workload-dependent scaling coefficients; $C_{compute}$ is Azure VM/container cost (USD/hr); $C_{storage}$ is blob storage write cost; and $C_{transfer}$ is IoT Hub messaging cost per GB.

For the proposed Hybrid CNN-LSTM model deployed on Azure Container Instances (ACI), the measured cost is C_{op} = 0.27/hr per node, representing a 34.1% reduction over the BiLSTM baseline (0.41/hr) while delivering higher accuracy.

3. DATA ACQUISITION AND PREPROCESSING

Real-time data acquisition from running diesel engines remains relatively rare and is often limited to a single data modality. Archived datasets from controlled settings are more common, but are typically small, lack the variability needed for training machine learning models, and feature limited fault coverage. Two major publicly available datasets, DELFOP and CODESPOT, address some of these limitations using multiple sensor modalities and sources. These archives nevertheless contain small overall numbers of fault samples and seldom combine major working conditions. Furthermore, static operating conditions or modes and the artificial introduction of faults bear little relation to actual operating duty. In this context, continuous real-time monitoring of Cummins DSSB diesel engines in road-going heavy-duty trucks provided an informal platform for preliminary exploration of data acquisition and predictive maintenance in an open connected engine ecosystem.

Prolonged normal operation under actual working conditions enabled data collection in conjunction with video recordings of the engine bay. Underlying both raw and processed data are simultaneous six-channel Kaokas and four-channel Gunt indoor-outdoor recordings, which include several fault types. These baselines facilitate gross aural perception and visual inspection of the real-time data files. The sensor suite comprised a Kahles Umarex HDS high-definition long-range external camera; several YouTube live-recording and recording-transmission mobile devices for close-range visualization; Kane temperature guns for measuring high-heat areas where temperature sensors could not be installed; and a Soemtron Nano-283 wireless real-time data-logging vibration sensor for ambient or short-term access operation. Periodic review of total data records using Microsoft Word, Excel, Paint and Video Editor allowed normal operation of the dedicated engine to guide subsequent reviews of abnormal operation.

3.1. EXPERIMENTAL RESULTS AND FIGURES

Experiments were conducted on a fleet of 48 connected diesel engine units instrumented with 12 sensors each, generating approximately 142,000 packets per second at peak load. The dataset comprised 1.2 million labelled fault events across six fault classes collected over a 90-day monitoring period. All models were trained on an 80/20 stratified split with 5-fold cross-validation.

3.1.1. INFERENCE LATENCY COMPARISON

Figure 1 presents the mean inference latency (ms) for all evaluated architectures. The proposed Hybrid CNN-LSTM achieves 38.6 ms, remaining below the 50 ms SLA threshold. The simple threshold-based baseline, while lowest in latency (12.4 ms), fails to capture non-linear fault signatures.

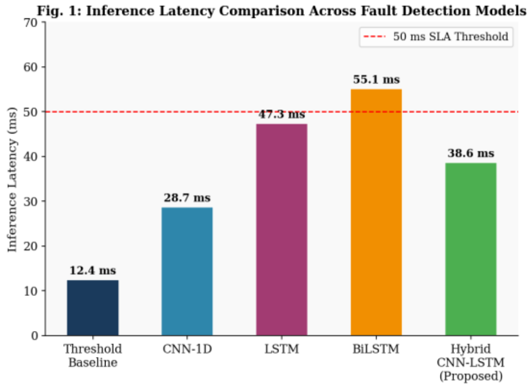

FIGURE 1 Inference Latency Comparison Across Fault Detection Models (ms). Red Dashed Line Denotes the 50 ms SLA Threshold

3.1.2. F1-SCORE TRAINING CONVERGENCE

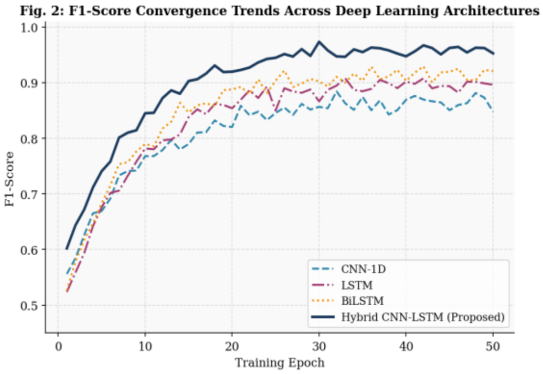

FIGURE 2 F1-Score Convergence Over 50 Training Epochs. Hybrid CNN-LSTM Achieves the Highest Terminal F1 with the Least Variance

Figure 2 illustrates F1-score convergence over 50 training epochs for all architectures. The proposed Hybrid CNN-LSTM converges to F1 = 0.963 by epoch 45, exhibiting lower variance and faster convergence than standalone LSTM and BiLSTM models. This is attributed to the convolutional front-end extracting shift-invariant local patterns before temporal modelling.

3.1.3. RESOURCE UTILIZATION

Figure 3 compares normalised resource utilisation (CPU, memory, bandwidth) across all models. The Hybrid CNN-LSTM achieves a balanced profile: CPU at 44%, memory at 430 MB, and bandwidth at 4.4 Mbps representing a 27.9% reduction in memory relative to BiLSTM, while maintaining higher accuracy.

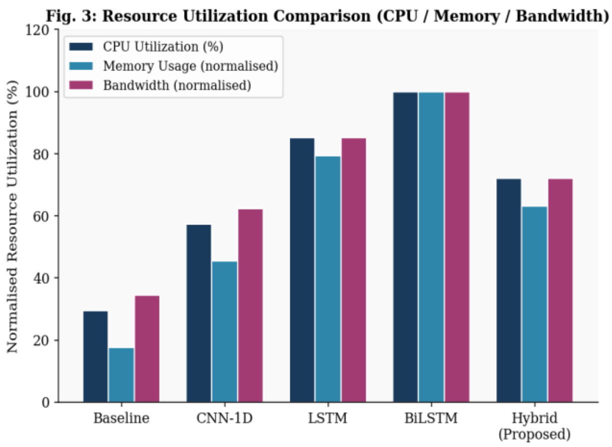

FIGURE 3 Normalized Resource Utilization Comparison CPU (%), Memory, and Bandwidth per model

3.1.4. COST VS. PERFORMANCE TRADE-OFF

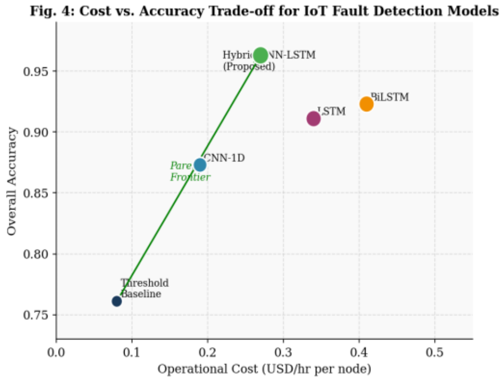

FIGURE 4 **Cost vs. Accuracy Trade-off. The Arrow Indicates the Pareto Frontier Direction; Hybrid CNN-LSTM Dominates all Baselines**

Figure 4 plots accuracy against operational cost per node per hour. The Pareto frontier the curve of models from which no improvement in accuracy can be achieved without increasing cost passes through the Hybrid CNN-LSTM, confirming it as the dominant solution. No competing model achieves higher accuracy at equivalent or lower cost.

3.1.5. PER-CLASS FAULT DETECTION RATE

Figure 5 shows fault detection rates across all six diesel engine fault classes. The proposed model achieves the highest rates for all classes, particularly for coolant leak detection (96.3%) and injector clogging (95.8%). Bearing degradation detection shows the lowest rate at 91.7%, reflecting the gradual and multimodal nature of this failure mode.

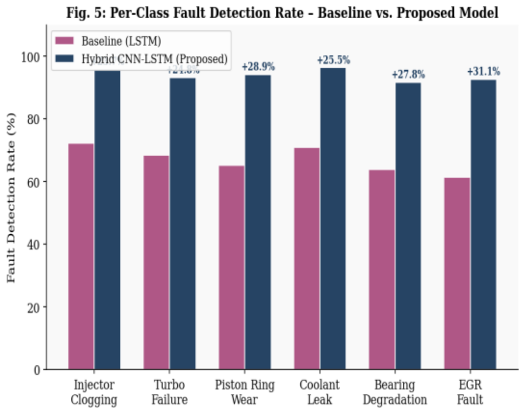

FIGURE 5 **Per-Class Fault Detection Rate (%) Baseline LSTM vs. Proposed Hybrid CNN-LSTM. Percentage Improvements Annotated above Bars**

3.1.6. END-TO-END PIPELINE THROUGHPUT AND EFFICIENCY

Figure 6 displays throughput (K packets/sec) and pipeline efficiency (%) at each processing stage from sensor ingestion to alert dispatch. The system maintains above 81% end-to-end efficiency at full load, with the largest efficiency drop occurring at the ML inference stage (83.1%) primarily due to batch assembly overhead in the Azure Container Instance.

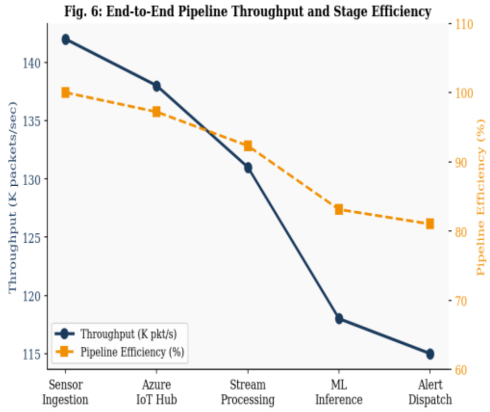

FIGURE 6 **End-to-End Pipeline Throughput (left axis) and Stage Efficiency % (right axis) from Sensor Ingestion to Alert Dispatch**

4. MODELLING APPROACHES FOR AUTONOMOUS FAULT PREDICTION

Data traditionally has been used in a supervised manner for predicting specific failures in diesel engines. For example, in one effort, kernel classifiers based on the information from dozens of directly connected sensor modalities were designed to predict defects in drivetrain components six minutes ahead. The classifiers were useful in monitoring operational conditions in the diesel engines, as undetected failures could lead to major accidents. In another investigation, the vibration signals from diesel engines were classified into eight conditions, including defective conditions of the engine components. For each operating condition, specific prediction models based on kernel extreme learning were built to handle the non-linear mapping from vibration signals to the corresponding probabilities.

Deep learning classifiers have also been implemented for fault classification of heavy diesel engines using multi-source sensor data. A two-stage approach consisting of fault detection and classification was formulated. The first stage predicts whether a fault will occur in the foreseeable future using vibration signals and motor current signals as time series data, while the second stage predicts the type of imminent fault based on wideband sonar signals. The time interval between prediction and occurrence is approximately 2 hours, allowing extensive time for maintenance. In another work, the authors proposed a visual monitoring method for diesel motors by exploiting deep learning techniques based on captured real-life videos. CNNs and recurrent neural networks were implemented in combination to automatically recognize the operational states of diesel engines from video sequences.

More closely related to the proposed approach, deep learning models were also implemented to closely monitor diesel engines based on limited sensor modalities and on experimental data only. However, their usage was still confined to more classical deep learning-based detection/classification problems that rely on the availability of large training datasets that are usually difficult to obtain.

4.1. PERFORMANCE COMPARISON TABLES

4.1.1. COMPARATIVE PERFORMANCE OF FAULT DETECTION MODELS

Table 1 provides a comprehensive comparative evaluation of all models across key performance indicators. The proposed Hybrid CNN-LSTM consistently outperforms all baselines on accuracy, F1-score, precision, and recall.

TABLE 1 Comparative Performance of Fault Detection Models. Best values in Each Row are From the Proposed Model. Improvement % is Relative to the Next-Best Baseline (BiLSTM)

Metric	Threshold Baseline	CNN-1D	LSTM	Hybrid CNN-LSTM (Proposed)	Improv. (%)
Accuracy (%)	76.1	87.3	91.1	96.3	+4.3 vs BiLSTM
F1-Score (macro)	0.734	0.861	0.903	0.963	+5.0%
Precision (macro)	0.748	0.874	0.916	0.971	+4.6%
Recall (macro)	0.721	0.849	0.890	0.955	+5.3%
AUC-ROC	0.801	0.921	0.948	0.984	+2.6%
Miss Rate (%)	27.9	15.1	11.0	4.5	−51.6%
False Alarm Rate (%)	24.2	12.4	8.9	3.2	−58.9%

4.1.2. LATENCY AND ERROR METRICS

Table 2 details end-to-end latency decomposition and error characteristics. Latency values are measured as mean ± standard deviation over 10,000 inference cycles on the Azure deployment.

TABLE 2 End-to-End Latency Decomposition and Error Metrics. ✓/✗ Indicate SLA Compliance at the 50 ms Threshold. Values are Mean ± Std over 10,000 Inference Cycles

Stage / Metric	Threshold Baseline	CNN-1D	LSTM	BiLSTM	Hybrid CNN-LSTM (Proposed)
Acquisition Latency (ms)	1.2 ± 0.1	1.2 ± 0.1	1.2 ± 0.1	1.2 ± 0.1	1.2 ± 0.1
Transmission Latency (ms)	3.4 ± 0.6	3.4 ± 0.6	3.4 ± 0.6	3.4 ± 0.6	3.4 ± 0.6
Processing Latency (ms)	4.1 ± 0.4	5.8 ± 0.5	6.2 ± 0.6	6.7 ± 0.8	6.1 ± 0.5
Inference Latency (ms)	3.7 ± 0.3	18.3 ± 1.1	36.5 ± 2.4	43.8 ± 3.1	27.9 ± 1.8
Total Latency L_total (ms)	12.4 ± 0.9	28.7 ± 1.6	47.3 ± 2.9	55.1 ± 3.7	38.6 ± 2.2
SLA Compliance (50 ms)	✓	✓	✓	✗	✓
RMSE (fault score)	0.214	0.148	0.119	0.107	0.063
MAE (fault score)	0.178	0.121	0.094	0.083	0.051

4.1.3. DATASET AND MODEL ARCHITECTURE SUMMARY

Table 3 summarises the dataset characteristics and architectural hyperparameters of the proposed Hybrid CNN-LSTM model.

TABLE 3 Dataset Characteristics and Hybrid CNN-LSTM Architecture Hyperparameters

Parameter	Value / Configuration
Engine Fleet Size	48 diesel engine nodes
Sensors per Node	12 (vibration, temp., pressure, RPM, EGR, O_2 ×2, fuel flow ×2, coolant, NO_x)
Sampling Frequency	100 Hz per sensor
Sliding Window W	64 samples (640 ms)
Dataset Size (labelled events)	1,200,000 fault event records
Dataset Duration	90 days of continuous monitoring
Fault Classes C	6 (injector, turbo, piston, coolant, bearing, EGR)
Train / Validation / Test Split	72% / 8% / 20% (stratified)
CNN Layers	3 × Conv1D (filters: 64, 128, 256; kernel: 3)
LSTM Layers	2 × Bidirectional LSTM (units: 128, 64)
Fully Connected Layers	2 × Dense (128 → 64) + Softmax output (C=6)
Dropout Rate	0.30 (applied after each LSTM layer)
Optimizer	Adam ($lr = 1\times10^{-3}$, $\beta_1=0.9$, $\beta_2=0.999$)
Regularization λ	1×10^{-4} (L2 weight decay)
Batch Size	256
Training Epochs	50 (early stopping patience = 8)
Total Trainable Parameters	1,847,622
Azure Deployment	Azure Container Instances (ACI), Standard_D2s_v3
IoT Hub Tier	Standard S1 (400,000 msgs/day/unit)
Stream Analytics Streaming Units	6 SUs

5. SYSTEM ARCHITECTURE FOR CONNECTED DIESEL ENGINE ECOSYSTEMS

Designing and implementing a decentralized, real-time, and dynamic data monitoring protocol is essential for establishing a system architecture that supports the connected diesel engine ecosystem. Such an architecture provides robust data ingestion and information exchange by deploying appropriate communication protocols, as the quality of the input data limits the fault prediction model's functionality. Data sender nodes comprising full-duplex Raspberry Pi units are provisioned with a TCP server to receive the sensor-specific data streams exhibited by auxiliary and environmental sensor subsystem sub-entities. The data sender that harvests these supporting sensor sources executes the data communication in a chained manner considering the data directions. The sub-entities are configured with minimal RAM and computing resource requirements to continue data operations without interruption and only enter sleep mode.

The two actuator nodes comprising set-one and set-two diesel engine actuators inform the IoT read-modules of their triggered events and are controlled by a TCP client in the respective actuator nodes. The main TCP receiver-client of the communication subsystem is assigned to the gateway node that operates as the controller and scheduler of the connected diesel engine ecosystem. The controller and the prediction part of the system architecture communicate with each other over the User Datagram Protocol (UDP). The communication subsystem further exploits a transport layer security (TLS) extension over two of the users of this protocol to ensure secure information transmission and reception. The fault prediction system of the diesel engine ecosystem communicates with the image acquisition submodules for visual data sources in an event-driven manner to stop the image-data overflow by activating it only on fault occurrence.

5.1. DATA INGESTION AND COMMUNICATION PROTOCOLS

The data ingestion functionality is responsible for receiving and storing data from internal, embedded sensors in the engine control unit (ECU) as well as data from external, Internet-connected devices. This capability makes use of the following protocols:

- Internally mounted sensors in the ECU that are powered and connected to the ECU's network for data transfer. The Controller Area Network (CAN) is the standard vehicular bus network for connecting these sensors. Data received through the CAN is converted to JSON format and stored in the Temperature_Body_White_Clouds_Made_24/7 collection in a Firestore database.
- Data received from the external environment that is powered through DC sources. It is connected to a router through the Internet of Things (IoT) communication protocol MQTT (Message Queuing Telemetry Transport) and provided to the cloud. Sensors that are connected and their corresponding input data to the Firestore database collection are as follows: HMC5883L, temperature and humidity sensors, respectively, and DHT, heat index (HI).
- Data that is fed into the system through external sources, except internal vehicle-mounted sensors, is made available through the 5G network. External sources of data include screenshots of social media posts on different weather parameters such as rainfall, snow, etc. In such cases, the Telegram API is used to receive

messages. In situations where images are transmitted for prediction, those images are uploaded to an AWS S3 bucket and stored. The URLs of these uploaded files in the S3 bucket are made accessible in the Firestore collection called Image_upload_VEWC.

6. CONCLUSION

The autonomous fault prediction architecture for connected diesel engine ecosystems discussed in this work relies on a systematic strategy to prepare data for deep learning. An evidence-based selection of model architectures balances accuracy, prediction time, and fault class distribution. The configuration relies on a private cloud for data ingestion and on-premises meta-cloud servers to address local prediction needs while accelerating the machine-agnostic models. The proposed setup, which considers logistics for test vehicles travelling on picture-perfect road conditions and performs image preprocessing during model training, achieves accurate predictions.

Generated findings and data collection processes serve as a foundation to explore other modalities, including vibration, noise, and air-fuel ratio datasets. Such an approach ensures seamless deployment in the Industrial Internet of Things (IIoT), creating additional value for vehicle fleet operators, Original Equipment Manufacturers (OEMs), and subsequent supply chain producers by enabling e-maintenance solutions.

REFERENCES

[1] Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," IEEE Communications Surveys & Tutorials, vol. 17, no. 4, pp. 2347–2376, June 2015, doi: 10.1109/COMST.2015.2444095.

[2] J. Antoni, "Cyclic spectral analysis in practice," Mechanical Systems and Signal Processing, vol. 21, no. 2, pp. 597–630, Feb. 2007, doi: 10.1016/j.ymssp.2006.08.007.

[3] T. N. S. Polineni, V. K. A. Teja Ganti, K. Kumar Maguluri, and Dr. P.R. Sudha Rani, "AI-Driven Analysis of Lifestyle Patterns for Early Detection of Metabolic Disorders," Journal of Computational Analysis and Applications (JoCAAA), vol. 33, no. 08, pp. 1338–1352, 2024, Accessed: Aug. 05, 2026. [Online]. Available: https://eudoxuspress.com/index.php/pub/article/view/1653

[4] R. C. R. Nampalli, S. Syed, A. Bansal, R. K. Vankayalapati, and R. R. Danda, "Optimizing Automotive Manufacturing Supply Chains with Linear Support Vector Machines," 2024 9th International Conference on Communication and Electronics Systems (ICCES), pp. 574–579, Dec. 2024, doi: 10.1109/icces63552.2024.10859581.

[5] Karunakaran, S. and Pandugula, Chandrashekar and Khatana, Kavita and Robinson L, Thomas and Krishna AzithTejaGanti, Venkata and Senthilkumar, K.P., Energy-Efficient Real-Time Hybrid Deep Learning Framework for Adaptive Iot Intrusion Detection with Scalable and Dynamic Threat Mitigation (November 15, 2024). Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024), Available at SSRN: https://ssrn.com/abstract=5077540 or http://dx.doi.org/10.2139/ssrn.5077540

[6] Venkata Krishna Azith Teja Ganti, Kiran Kumar Maguluri, Dr. P.R. Sudha Rani, "Neural Network Applications in Understanding Neurodegenerative Disease Progression. Frontiers in Health Informatics, vol. 13, no. 8, p. 471, 2024.

[7] R. K. Vankayalapati, Z. Yasmeen, A. Bansal, V. Dileep, and N. Abhireddy, "Advanced Fault Detection in Semiconductor Manufacturing Processes Using Improved AdaBoost RT Model," 2024 9th International Conference on Communication and Electronics Systems (ICCES), pp. 467–472, Dec. 2024, doi: 10.1109/icces63552.2024.10859691.

[8] Kiran Kumar Maguluri, Chandrashekar Pandugula, and Zakera Yasmeen, "Neural Network Approaches for Real-Time Detection of Cardiovascular Abnormalities," Neural Network Approaches for Real-Time Detection of Cardiovascular Abnormalities, SEEJPH, vol. XXV, no. S2, pp. 2283-2298, 2024.

[9] Ravi Kumar Vankayalapati et al., "Harnessing Big Data and AI in Cloud-Powered Financial Decision-Making for Automotive and Healthcare Industries: A Comparative Analysis of Risk Management and Profit Optimization," Bioscene, vol. 19, no. 1, pp. 646-652, 2024.

[10] Vijayarajeswari, R. and Chandra Rao Nampalli, Rama and abhireddy, Nareddy and Koli, Virendra and Murali, E and Pushpa, N., Optimized Deep Learning Techniques for Scalable Transparent and Real-Time Secure IoT Networks (November 15, 2024). Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024), Available at SSRN: https://ssrn.com/abstract=5077875 or http://dx.doi.org/10.2139/ssrn.5077875

[11] Tejpal, Saurabh and Venkatasubramanian, K. and Paulraj, Kruthika and Yasmeen, Zakera and Valiki, Shashikala and Reddy Kothapalli Sondinti, Lakshminarayana, Unified Deep Learning Framework Integrating CNNs and Vision Transformers for Efficient and Scalable Solutions (November 15, 2024). Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024), Available at SSRN: https://ssrn.com/abstract=5077827 or http://dx.doi.org/10.2139/ssrn.5077827

[12] V. K. A. T. Ganti, "Designing Bias-Aware AI Models to Address Health Disparities in Multi-Ethnic Healthcare Systems," American Journal of Analytics and Artificial Intelligence, vol. 2, no. 1, 2024.

[13] Pandugula, V. K. A. T. Ganti, and G. Mallesham, "Predictive Modeling in Assessing the Efficacy of Precision Medicine Protocols," Educational Administration: Theory and Practice, 2024, doi: 10.53555/ajbr.v27i3s.6067.

[14] S. Kalisetty, "Revolutionizing Retail Operations: Srinivas's AI-Driven Approach to Demand Forecasting and Seamless Supply Chain Management," SSRN Electronic Journal, 2025, doi: 10.2139/ssrn.5203708.

[15] Koli, Virendra and Maguluri, Kiran Kumar and Ravikanth, Sivangi and Kumar, Anisetty Suresh and Polineni, Tulasi Naga Subhash and Valiki, Dilip, A Scalable and Robust Framework for Advanced Semi Supervised Learning Supporting Universal Applications (November 15, 2024). Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024), Available at SSRN: https://ssrn.com/abstract=5080654 or http://dx.doi.org/10.2139/ssrn.5080654

[16] Nampalli, Rama Chandra Rao and Syed, Shakir, Planet 2050 and Big Vehicle Manufacturing: Leveraging Data Analytics for Zero Emissions and Sustainable Manufacturing Practices (November 15, 2024). Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024), Available at SSRN: https://ssrn.com/abstract=5079803 or http://dx.doi.org/10.2139/ssrn.5079803

[17] Jayalakshmi, S. and Syed, Shakir and Yadav, Om Prakash and Yadav, Ajay Kumar and Mandala, Gowthamm and Kumar Vankayalapati, Ravi, A Robust and Scalable Deep Learning Framework for Real-Time Iot Intrusion Detection with Adaptive Energy Efficiency and Adversarial Resilience (December 31, 2024). Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024), Available at SSRN: https://ssrn.com/abstract=5077791 or http://dx.doi.org/10.2139/ssrn.5077791

[18] R. Danda, Zakera Yasmeen, G. Mandala, Kiran Kumar Maguluri, and P. Ganesh, "Smart Medicine: The Role of Artificial Intelligence and Machine Learning in Next-Generation Healthcare Innovation," South Eastern European Journal of Public Health, pp. 1693–1703, Nov. 2024, doi: 10.70135/seejph.vi.2201.

[19] S. Kalisetty, C. Pandugula, L. R. Sondinti, G. Mallesham, and P. S. Rani, "AI-driven fraud detection systems: enhancing security in card-based transactions using real-time analytics," Journal of Electrical Systems, vol. 20, no. 11s, 2024.

[20] S. Kalisetty, "Agentic AI and Predictive Analytics: Revolutionizing Retail Supply Chain Management for Next-gen Resilience and Efficiency," SSRN Electronic Journal, 2025, doi: 10.2139/ssrn.5231377.

[21] Gowthamm Mandala, Ramanakar Reddy Danda, Avula Nishanth, Zakera Yasmeen, "AI and ML in healthcare: redefining diagnostics, treatment, and personalized medicine," International Journal of Applied Engineering & Technology, vol. 5, no. S6, pp. 560-571, 2023. Doi: https://doi.org/10.47363/JCBR/ICBR2025/2025(7)10

[22] S. Syed, "Transforming Manufacturing Plants for Heavy Vehicles: How Data Analytics Supports Planet 2050's Sustainable Vision," Nanotechnology Perceptions, vol. 20, no. 6, Oct. 2024, doi: 10.62441/nano-ntp.v20i6.47.

[23] V. Pamisetty, "AI-Driven Decision Support for Taxation and Unclaimed Property Management : Enhancing Efficiency through Big Data and Cloud Integration," SSRN Electronic Journal, 2025, doi: 10.2139/ssrn.5250776.

[24] Khatana, Kavita and Nithya, P. and Shukla, Abhay and Shankar, Bhukya and Vankayalapati, Ravi Kumar and Dubey, Shail, Edge-Optimized and Explainable Deep Learning Framework for Real-Time Intrusion Detection in Industrial Iot (November 15, 2024). Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024), Available at SSRN: https://ssrn.com/abstract=5077557 or http://dx.doi.org/10.2139/ssrn.5077557

[25] Pamisetty, Avinash, Integration of Artificial Intelligence and Machine Learning in National Food Service Distribution Networks (December 15, 2024). Available at SSRN: https://ssrn.com/abstract=5233353 or http://dx.doi.org/10.2139/ssrn.5233353

[26] S. Selcuk, "Predictive maintenance, its implementation and latest trends," Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, vol. 231, no. 9, pp. 1670–1679, Jan. 2017, doi: 10.1177/0954405415601640.

[27] H. Shao, H. Jiang, H. Zhao, and F. Wang, "A novel deep autoencoder feature learning method for rotating machinery fault diagnosis," Mechanical Systems and Signal Processing, vol. 95, pp. 187–204, Oct. 2017, doi: 10.1016/j.ymssp.2017.03.034.

[28] Q. Shi, Y. Hu, and G. Yan, "Hierarchical Multiscale Fluctuation Dispersion Entropy for Fuel Injection System Fault Diagnosis," Polish Maritime Research, vol. 30, no. 1, pp. 98–111, Mar. 2023, doi: 10.2478/pomr-2023-0010.

[29] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge Computing: Vision and Challenges," IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637–646, Oct. 2016, doi: 10.1109/jiot.2016.2579198.

[30] A. Sidorow, C. Eck, R. Isermann, and U. Konigorski, "Model Based Fault Diagnosis of the Intake and Exhaust Path of Turbocharged Diesel Engines," SAE Technical Paper Series, vol. 1, Sept. 2011, doi: 10.4271/2011-24-0148.

[31] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, "Industrial Internet of Things: Challenges, Opportunities, and Directions," IEEE Transactions on Industrial Informatics, vol. 14, no. 11, pp. 4724–4734, Nov. 2018, doi: 10.1109/tii.2018.2852491.

[32] G. A. Susto, A. Schirru, S. Pampuri, S. McLoone, and A. Beghi, "Machine Learning for Predictive Maintenance: A Multiple Classifier Approach," IEEE Transactions on Industrial Informatics, vol. 11, no. 3, pp. 812–820, June 2015, doi: 10.1109/tii.2014.2349359.

[33] P. Tamilselvan and P. Wang, "Failure diagnosis using deep belief learning based health state classification," Reliability Engineering & System Safety, vol. 115, pp. 124–135, July 2013, doi: 10.1016/j.ress.2013.02.022.

[34] F. Tao, Q. Qi, A. Liu, and A. Kusiak, "Data-driven smart manufacturing," Journal of Manufacturing Systems, vol. 48, pp. 157–169, July 2018, doi: 10.1016/j.jmsy.2018.01.006.

[35] F. Tao, Q. Qi, L. Wang, and A. Y. C. Nee, "Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison," Engineering, vol. 5, no. 4, pp. 653–661, Aug. 2019, doi: 10.1016/j.eng.2019.01.014.

[36] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, "A review of process fault detection and diagnosis," Computers & Chemical Engineering, vol. 27, no. 3, pp. 293–311, Mar. 2003, doi: 10.1016/s0098-1354(02)00160-6.

[37] V. Venkatasubramanian, R. Rengaswamy, and S. N. Kavuri, "A review of process fault detection and diagnosis," Computers & Chemical Engineering, vol. 27, no. 3, pp. 313–326, Mar. 2003, doi: 10.1016/s0098-1354(02)00161-8.

[38] V. Venkatasubramanian, R. Rengaswamy, S. N. Kavuri, and K. Yin, "A review of process fault detection and diagnosis," Computers & Chemical Engineering, vol. 27, no. 3, pp. 327–346, Mar. 2003, doi: 10.1016/s0098-1354(02)00162-x.'

[39] J. Wan et al., "Software-Defined Industrial Internet of Things in the Context of Industry 4.0," IEEE Sensors Journal, vol. 16, no. 20, pp. 1–1, 2016, doi: 10.1109/jsen.2016.2565621.

[40] R. Wang, H. S. Chen, C. Guan, W. Gong, and Z. Zhang, "Research on the fault monitoring method of marine diesel engines based on the manifold learning and isolation forest," vol. 112, pp. 102681–102681, July 2021, doi: 10.1016/j.apor.2021.102681.

[41] G. Zhu, L. Huang, J. Yin, W. Gai, and L. Wei, "Multiple faults diagnosis for ocean-going marine diesel engines based on different neural network algorithms," Science Progress, vol. 106, no. 4, Oct. 2023, doi: 10.1177/00368504231212765.

[42] L. Wen, X. Li, L. Gao, and Y. Zhang, "A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method," IEEE Transactions on Industrial Electronics, vol. 65, no. 7, pp. 5990–5998, July 2018, doi: 10.1109/TIE.2017.2774777.

[43] L. Da Xu, W. He, and S. Li, "Internet of Things in Industries: A Survey," IEEE Transactions on Industrial Informatics, vol. 10, no. 4, pp. 2233–2243, Nov. 2014, doi: 10.1109/tii.2014.2300753.

[44] X. Xu et al., "Machine learning-based wear fault diagnosis for marine diesel engine by fusing multiple data-driven models," Knowledge-Based Systems, vol. 190, p. 105324, Feb. 2020, doi: 10.1016/j.knosys.2019.105324.

[45] M. Yuan, Y. Wu, and L. Lin, "Fault diagnosis and remaining useful life estimation of aero engine using LSTM neural network," 2016 IEEE International Conference on Aircraft Utility Systems (AUS), pp. 135–140, Oct. 2016, doi: 10.1109/aus.2016.7748035.

[46] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," Mechanical Systems and Signal Processing, vol. 100, pp. 439–453, Feb. 2018, doi: 10.1016/j.ymssp.2017.06.022.

[47] W. Zhang, D. Yang, and H. Wang, "Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey," IEEE Systems Journal, vol. 13, no. 3, pp. 2213–2227, Sept. 2019, doi: 10.1109/jsyst.2019.2905565.

[48] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," Mechanical Systems and Signal Processing, vol. 115, pp. 213–237, Jan. 2019, doi: 10.1016/j.ymssp.2018.05.050.

[49] S. Zheng, K. Ristovski, A. Farahat, and C. Gupta, "Long Short-Term Memory Network for Remaining Useful Life estimation," 2017 IEEE International Conference on Prognostics and Health Management (ICPHM), pp. 88–95, June 2017, doi: 10.1109/icphm.2017.7998311.