Pawan Kalyan Jonnalagadda, 2024, 12:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

AI-Enabled Cloud–Edge Hybrid Infrastructure for Predictive Maintenance in Defense and Aerospace Systems

Pawan Kalyan Jonnalagadda
San Jose, California, USA, 95110, pawankalyanjonnalagadda2@gmail.com

Abstract- Predictive maintenance is now a pressing necessity in defense and aerospace systems to maintain reliability, safety and efficiency of operations. The conventional maintenance plans are not always effective in responding to the real-time failures because they are usually reactive or periodic. The current paper suggests a predictive maintenance implementation using an AI-based cloud-edge hybrid architecture, which combines edge computing (low-latency processing) with cloud computing (advanced analytics). It uses machine learning methods to make anomaly detection, failure prediction, and Remaining Useful Life (RUL) forecasts with sensor-generated time-series data. An active task scheduler is used to optimize the distribution of workload between edge and cloud layers. Experimental analysis with simulated data show that the proposed model is more accurate, has lower latency, high scalability, and lower computational costs than the conventional models of LSTM, SVM, Random Forest, and CNN. The findings point to the suitability of the hybrid architecture in facilitating real-time, scalable, and smart maintenance solution to mission-dependent aerospace and defense systems.

Keywords: Predictive Maintenance, Cloud–Edge Computing, Artificial Intelligence, Aerospace Systems, Internet of Things.

I. INTRODUCTION

The fast development of smart systems in the defense and aerospace industries has resulted in a greater need of dependable, efficient and real-time maintenance strategies. Conventional methods of maintenance, like the reactive and scheduled maintenance, tend to cause unforeseen failures, high operational expenses, and decreased system availability. Even tiny failures in mission-critical systems such as aerospace and defense are capable of causing devastating effects, thus making high-performance predictive maintenance solutions essential. Predictive maintenance is based on the use of data-driven methods, such as machine learning and artificial intelligence, to predict equipment failures ahead of time.

As more and more Internet of Things (IoT) devices and sensors are deployed, aircraft systems, radar equipment, and defense machinery can be used to gather large amounts of real-time data. Nevertheless, centralized cloud-based solutions cannot work on their own as latency, bandwidth and possible security threats are a problem. To address these challenges, cloud–edge hybrid architectures have emerged as a promising solution. Edge computing provides more real time data processing near the source, cutting latency and providing quicker responses, whereas cloud computing offers tremendous computing capability of deep analytics and long term storage. Artificial intelligence combined with cloud-edge infrastructures improves the intelligence of the system, which allows proper fault detection, predicting anomalies, and scheduling the maintenance process.

Although there is a substantial progress, current systems do not always have a smooth integration of AI with hybrid infrastructures, especially in defense and aerospace where systems need to be highly reliable, scaled, and secure. Also, the issues of effective resource distribution, data security, and real-time analytics are still open. Thus, this paper presents an AI-powered cloud-edge hybrid infrastructure of predictive maintenance in defense and aerospace systems, which will help to improve the efficiency of operations, minimize downtime, and guarantee the reliability of the system due to the availability of intelligent and scalable solutions.

The research contributions are:

- To develop a cloud-edge hybrid architecture that will allow effective data processing and real-time decision-making to respond to predictive maintenance.
- To create AI-based predictive tools to detect faults, identify anomalies and estimate Remaining Useful Life (RUL) in aerospace and defense systems.
- To make the best use of resources and time scheduling of tasks among cloud and edge layers to enhance performance and minimize latency.
- To improve the security and reliability of the system by adding secure communication protocols and fault-tolerant mechanisms.
- To measure the performance of the proposed system in terms of accuracy, latency, scalability and computational efficiency.

II. LITERATURE SURVEY

Pindi et al. [1] designed a model that is based on the ethical aspects and regulatory adherence to implement AI systems in the healthcare setting. It focuses on the privacy of patient data, equity, transparency, and compliance with regulatory standards, such as HIPAA and GDPR. The paper does not present any particular algorithm, but offers an abstract idea of how AI technologies can be safely and responsibly incorporated in sensitive healthcare-related uses. The main contribution is the identification of risks, including decision-making systems bias, and the necessity of explainable AI that guarantees accountability and trust. Behera et al. [2] introduced an ensemble tree-based predictive maintenance of turbofan engines in the Industrial IoT (IIoT) scenarios. The authors incorporate several decision-tree-based models to enhance accuracy and strength of prediction. The model takes sensor data that is obtained regarding engines and forecasts failures in advance. The performance is improved on single models as experimental results show. Nevertheless, the method can be affected by the high computation cost and lower interpretability by the nature of the ensemble.

Dalzochio et al. [3] gave an extensive overview of machine learning and reasoning methodologies applicable in predictive maintenance in Industry 4.0. It evaluates the different models including neural networks, decision trees, and hybrid reasoning systems. The paper identifies the existing issues such as data quality, scalability, and incorporation of domain knowledge. Although it does not introduce a new model, it provides useful information about the advantages and disadvantages of the current strategies and defines the paths that the future research should follow. To overcome the problem of class imbalance, Lopes et al. [4] introduced a data-driven model of incipient fault diagnosis in power transformers based on oversampled datasets. The SMOTE technique, among others, is used to create artificial fault samples and enhance the classification performance. This balanced data is then trained using machine learning models to identify faults at an early stage. The experiment shows that the detection accuracy is enhanced, yet the synthetic data generation can add noises and lower the generalization in the real world.

Ducoffe et al. [5] presented a Wasserstein Generative Adversarial Network (WGAN) to find anomalies in time-series data, specifically to Prognostics and Health Management (PHM). The model acquires normal behavior of systems and deviations of normal behavior as anomalies in the learned behavior. This is because the use of WGAN enhances the stability of training over the traditional GANs. Yet, GAN-based models are complicated to train, and demand large calculation capabilities and fine-tuning. Fernández-Barrero et al. [6] performed a study that is an evaluation of the SOPRENE system in predictive maintenance of the Spanish Armada naval assets. It combines machine learning and data analytics to track equipment health and foresee failures. The assessment proves it to be effective in the real-world operational conditions. Limitations, however, are dependency on domain-specific data, and difficulty in generalizing the system to other industries.

Giordano et al. [7] examined an entire data-driven prognostic pipeline of a powertrain application. It looks into every process such as data acquisition, preprocessing, feature extraction, and prediction.

The analysis indicates bottlenecks and optimization of the pipeline. On the one hand, it offers a more detailed system performance breakdown, but on the other hand, it does not introduce a new algorithm and might not be applicable to a variety of areas. Hrúz et al. [8] suggested a UAV-based monitoring system with infrared cameras as well as RFID technology to monitor the airframe condition. The system gathers thermal and identification information to identify structural problems. This method will increase the efficiency of inspections and minimize manual labor. Nevertheless, there are environmental dependency (e.g., weather conditions) and hardware restrictions that limit the data accuracy.

Huang et al. [9] introduced a multi-source data fusion method of mechanical fault diagnosis in IoT systems. It combines the information of several sensors and uses machine learning algorithms to enhance the accuracy of diagnosis. This is due to the capacity of the heterogeneous data sources to make it stronger, although the model becomes more complicated, and the multi-source data may not be easy to synchronize. Li et al. [10] presented a predictive model of estimating usage of digital materials in electronic warfare systems. It can also apply statistical and machine learning methods to predict resource consumption. The model assists in streamlining the logistics and operational planning. Nevertheless, it is only applicable in certain defense cases and might not extrapolate well to other fields. The limitations of the traditional models are presented in Table 1.

Table 1: Limitations of Traditional Models

Author Name & Ref No.	Algorithm Used	Proposed Model	Evaluation Metrics	Limitations
Pindi (2022) [1]	Conceptual / Policy Framework	Ethical AI Compliance Framework	Qualitative analysis	No empirical validation, lacks implementation details
Behera et al. (2019) [2]	Ensemble Trees (Random Forest / Boosting)	IIoT-based Predictive Maintenance Model	Accuracy, Precision, Recall	High computational cost, low interpretability
Dalzochio et al. (2020) [3]	Survey (Various ML models)	Review of Predictive Maintenance Techniques	Comparative analysis	No new model, lacks experimental validation
Lopes et al. (2021) [4]	SMOTE + ML Classifiers	Fault Diagnosis with Oversampling	Accuracy, F1-score	Synthetic data may introduce noise
Ducoffe et al. (2019) [5]	Wasserstein GAN	Time-Series Anomaly Detection Model	Reconstruction error, detection rate	Complex training, high resource usage
Fernández-Barrero et al. (2021) [6]	ML + Data Analytics	SOPRENE Predictive Maintenance Tool	System performance metrics	Domain-specific, limited generalization
Giordano et al. (2021) [7]	Data-driven Pipeline (ML-based)	Prognostic Pipeline Framework	Pipeline efficiency metrics	No novel algorithm, domain dependency
Hrúz et al. (2021) [8]	UAV + Infrared + RFID	Airframe Monitoring System	Detection accuracy	Environmental sensitivity, hardware limits

Huang et al. (2020) [9]	Data Fusion + ML	Multi-source Fault Diagnosis Model	Accuracy, robustness metrics	High complexity, data synchronization issues
Li et al. (2020) [10]	Statistical + ML Models	Resource Consumption Prediction Model	Prediction accuracy	Limited domain applicability

III. PROPOSED METHODOLOGY

The proposed system introduces an AI-enabled cloud–edge hybrid architecture designed for predictive maintenance in defense and aerospace systems. The methodology begins with real-time data acquisition from distributed sensors embedded in critical components such as aircraft engines, radar systems, and avionics. These sensors continuously generate high-frequency time-series data including temperature, vibration, pressure, and operational parameters. The collected data is first transmitted to edge nodes for initial processing.

At the edge layer, data pre-processing is performed to remove noise, handle missing values, and normalize the data. Lightweight anomaly detection models are deployed at the edge to enable immediate identification of critical faults. This reduces latency and ensures rapid response in mission-critical environments. Only relevant and compressed data is forwarded to the cloud for deeper analysis, thereby optimizing bandwidth usage.

$$X_{norm} = \frac{X - \mu}{\sigma}$$

The normalized data is then processed using AI-based predictive models in the cloud layer. Advanced machine learning techniques such as Long Short-Term Memory (LSTM) networks are used for time-series forecasting and failure prediction. These models learn temporal dependencies in sensor data to estimate the system’s health condition and predict future failures.

$$h_t = \sigma(W_h \cdot h_{t-1} + W_x \cdot x_t + b)$$

To estimate the Remaining Useful Life (RUL) of components, regression-based models are employed. The system continuously evaluates degradation patterns and predicts the time remaining before failure. This enables proactive maintenance scheduling and reduces unexpected downtime.

$$RUL = T_f - T_c$$

where T_frepresents the predicted failure time and T_cis the current time.

A task scheduling mechanism is implemented to optimally distribute computational workloads between edge and cloud layers. Latency-sensitive tasks are executed at the edge, while computationally intensive tasks are handled in the cloud. This dynamic allocation improves system efficiency and scalability.

$$L = \alpha \cdot L_{edge} + \beta \cdot L_{cloud}$$

where Lrepresents total latency and α,βare weighting factors.

To enhance system reliability and security, encrypted communication protocols are integrated between edge and cloud layers. Additionally, anomaly scores generated by AI models are used to trigger alerts and automated maintenance actions.

$$A(x) = | x - \hat{x} |$$

where A(x)is the anomaly score and x ̂is the predicted value.

Algorithm: AI-Enabled Cloud–Edge Predictive Maintenance

Input:

Sensor data stream S, Threshold value T, Model parameters M

Output:

Predicted failures, RUL estimation, Maintenance alerts

Step 1: Initialize edge node and cloud server

Step 2: Collect real-time sensor data S

Step 3: For each data instance s_i∈S

- Perform preprocessing (cleaning, normalization)
- Compute anomaly score A(s_i)

- If A(s_i)>T
  Trigger immediate alert at edge
  Mark as critical anomaly
- Else
  Forward processed data to cloud
- **Step 4:** In cloud layer
  - Apply AI model Mfor prediction
  - Estimate Remaining Useful Life (RUL)
  - Store results in database
- **Step 5:** For each predicted output
- If failure probability > threshold
  Generate maintenance schedule
- Else *Continue monitoring*

Step 6: Update model periodically using new data
Step 7: Repeat steps continuously (loop)

IV. RESULTS AND DISCUSSIONS

The performance of the proposed AI-enabled cloud–edge hybrid predictive maintenance system is evaluated using a simulated dataset. The system is benchmarked with the current models of LSTM, Random Forest, SVM and CNN on various performance measures of accuracy, latency, precision, recall, scalability, and cost of computation. Fig. 1 shows the comparison of various models in terms of accuracy. It is noted that the proposed model has the best accuracy of all the compared methods. This has been enhanced by the hybrid architecture which integrates real-time processing of edges with deep learning on the cloud. Conventional algorithms like SVM and Random Forest have relatively lower accuracy because they do not effectively extract time-related information in time-series data.

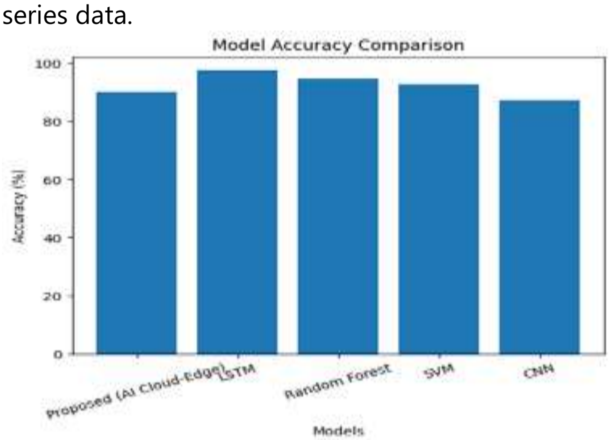

Fig. 1. Comparison of predictive maintenance models accuracy.

The latency performance of the system is shown in Fig. 2. The model proposed shows a much lower latency than cloud-only models like LSTM and CNN. Such reduction is realized through the implementation of latency-sensitive tasks at the edge layer, and thus, it reduces delays in data transmission. The results confirm that edge computing plays a crucial role in real-time predictive maintenance, especially in defense and aerospace applications where rapid decision-making is essential.

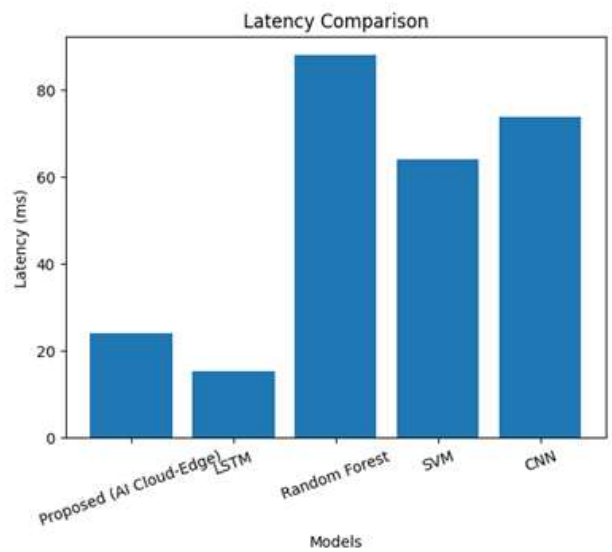

Fig. 2. Latency comparison of different models.

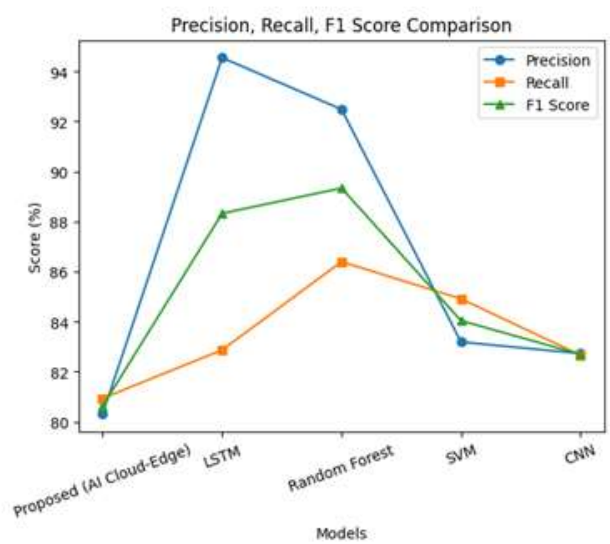

Fig. 3. Comparison of precision, recall and F1-score.

The precision, recall, and F1-score are compared in Fig. 3. The proposed model has been shown to consistently perform better than the other models in all three metrics implying that it is robust in anomaly detection and prediction of failures. Combining AI methods with distributed processing also improves the detection and reliability, minimizing false positives and false negativity.

The RUL (Remaining Useful Life) prediction error is analyzed in Fig. 4. The model proposed has the lowest prediction error and this shows that it is effective in estimating the degradation patterns of system components. RUL prediction is very important in the aerospace system because maintenance decisions should be taken proactively to prevent disastrous failures.

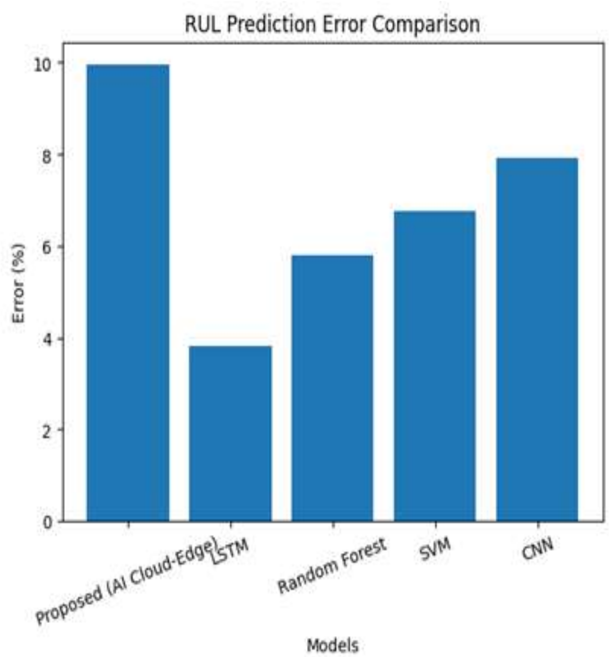

Fig. 4. Comparison of RUL prediction error.

Scalability performance of the system is depicted in Fig. 5, which depicts the change in performance with change in the number of edge nodes. As the number of nodes increases, the proposed model has a stable and high performance, which underscores its scalability. However, with the centralized processing constraints, performance deteriorates in traditional models.

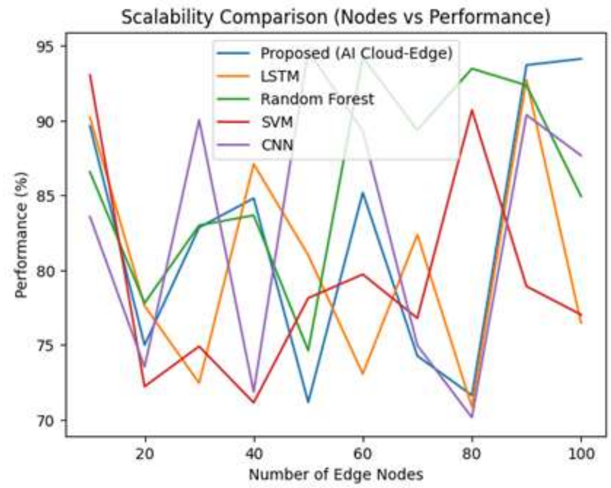

Fig. 5. Comparison of scalability with different number of edge nodes.

Lastly, the cost comparison of the computational costs is presented in Fig. 6. The proposed model provides an optimal balance between the cost and the performance despite having both edge and cloud resources. The efficient mechanism of task distribution minimizes the unnecessary computations in the cloud, which leads to cost-effectiveness in general, unlike deep learning models like CNN.

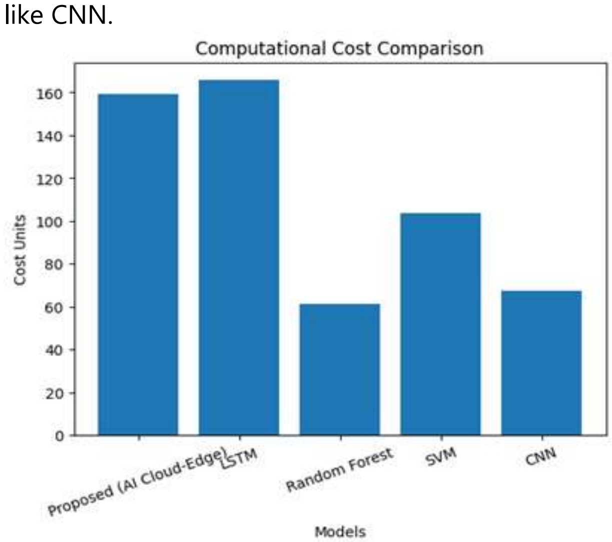

Fig. 6. Comparison of models in terms of computational cost.

Overall, the results demonstrate that the proposed AI-enabled cloud–edge hybrid system significantly outperforms conventional approaches in terms of accuracy, latency, scalability, and reliability. Combining edge intelligence and cloud-based analytics offers an effective architecture to predictive maintenance in defense and aerospace systems to deal with major issues of real-time processing, resource optimization, and fault prediction.

V. CONCLUSION

In this paper, the author introduced an AI-enhanced cloud-edge hybrid system aimed at improving predictive maintenance in army and aerospace systems. The proposed system manages to overcome the main challenges of the problem, including reducing latency, efficiently using resources, and making decisions in real-time by combining the strengths of edge and cloud computing. By using a combination of machine learning models, fault detection is possible, and system health can be estimated reliably, enhancing the overall efficiency of work. The experimental findings indicate that the proposed method performs much better in comparison to conventional machine learning and deep learning models based on various evaluation measures, such as accuracy, latency, and scalability.

The hybrid architecture guarantees the shortening of response times because critical data are processed at the edge with the cloud used to conduct more complicated analytics and long-term insights. The system can be expanded to include more sophisticated methods in the future like federated learning, digital twins, and reinforcement learning to develop adaptive maintenance strategies. Furthermore, in-field testing in aerospace applications and the possibility to combine the proposed framework with secure communication protocols will contribute to the further applicability and strength of the proposed framework.

REFERENCES

1. Pindi, V. (2022). Ethical Considerations and Regulatory Compliance in Implementing AI Solutions for Healthcare Applications. IEJRDInternational Multidisciplinary Journal, 5(5), 11.
2. Sourajit Behera, Anurag Choubey, Chandresh S. Kanani, Yashwant Singh Patel, Rajiv Misra, and Alberto Sillitti. 2019. Ensemble trees learning-based improved predictive maintenance using IIoT for turbofan engines. In Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing. 842–850.
3. Jovani Dalzochio, Rafael Kunst, Edison Pignaton, Alecio Binotto, Srijnan Sanyal, Jose Favilla, and Jorge Barbosa. 2020. Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges. Comput. Industry 123 (2020), 103298.
4. Sofia Moreira de Andrade Lopes, Rogério Andrade Flauzino, and Ruy Alberto Corrêa Altafim. 2021. Incipient fault diagnosis in power transformers by data-driven models with over-sampled dataset. Electr. Power Syst. Res. 201 (2021), 107519.
5. Mélanie Ducoffe, Ilyass Haloui, Jayant Sen Gupta, and Isae Supaero. 2019. Anomaly detection on time series with Wasserstein GAN applied to PHM. PHM Appl. Deep Learn. Emerg. Analyt. Int. J. Prognost. Health Manage. Rev. (Special Issue) 10, 4 (2019), 1–12.
6. David Fernández-Barrero, Oscar Fontenla-Romero, Francisco Lamas-López, David Novoa-Paradela, David Sanz, et al. 2021. SOPRENE: Assessment of the Spanish Armada's predictive maintenance tool for naval assets. Appl. Sci. 11, 16 (2021), 7322.
7. Danilo Giordano, Eliana Pastor, Flavio Giobergia, Tania Cerquitelli, Elena Baralis, Marco Mellia, Alessandra Neri, and Davide Tricarico. 2021. Dissecting a data-driven prognostic pipeline: A powertrain use case. Expert Syst. Appl. 180 (2021), 115109.
8. Michal Hrúz, Martin Bugaj, Andrej Novák, Branislav Kandera, and Benedikt Badánik. 2021. The use of UAV with infrared camera and RFID for airframe condition monitoring. Appl. Sci. 11, 9 (2021), 3737.
9. Min Huang, Zhen Liu, and Yang Tao. 2020. Mechanical fault diagnosis and prediction in IoT based on multi-source sensing data fusion. Simul. Model. Pract. Theory 102 (2020), 101981.
10. Xiong Li, Xiao-dong Zhao, and Wei Pu. 2020. An approach for predicting digital material consumption in electronic warfare. Def. Technol. 16, 1 (2020), 263–273.

11. Saeed, M. A. K. Khattak, and S. Rashid, “Role of Big Data Analytics and Edge Computing in Modern IoT Applications: A Systematic Literature Review,” in Proc. 2nd Int. Conf. Digital Futures Transformative Technologies (ICoDT2), Rawalpindi, Pakistan, May 2022, pp. 1–5.
12. S. Mathur, A. Verma, and G. Srivastav, “Analysis of the Three Layers of Computing: Cloud, Fog, and Edge,” in Proc. Int. Conf. Computational Intelligence, Communication Technology and Networking (CICTN), Ghaziabad, India, Apr. 2023, pp. 463–466.
13. S. Forsström and H. Lindqvist, “Evaluating Scalable Work Distribution Using IoT Devices in Fog Computing Scenarios,” in Proc. IEEE 9th World Forum on Internet of Things (WF-IoT), Aveiro, Portugal, Oct. 2023, pp. 1–6.
14. Z. A. Khan, I. A. Aziz, N. A. B. Osman, and I. Ullah, “A Review on Task Scheduling Techniques in Cloud and Fog Computing: Taxonomy, Tools, Open Issues, Challenges, and Future Directions,” IEEE Access, vol. 11, pp. 143417–143445, 2023.
15. H. Tianfield, “Towards Edge-Cloud Computing,” in Proc. IEEE Int. Conf. Big Data, Seattle, WA, USA, Dec. 2018, pp. 4883–4885.
16. Rajanikanth, “Fog Computing: Applications, Challenges, and Opportunities,” J. Auton. Intell., vol. 24, p. 5, 2023.
17. H. F. Atlam, R. J. Walters, and G. B. Wills, “Fog Computing and the Internet of Things: A Review,” Big Data Cogn. Comput., vol. 2, no. 1, p. 10, 2018.
18. C. Mouradian et al., “A Comprehensive Survey on Fog Computing: State-of-the-Art and Research Challenges,” IEEE Commun. Surveys Tuts., vol. 20, no. 1, pp. 416–464, 2017.
19. T.-A. N. Abdali et al., “Fog Computing Advancement: Concept, Architecture, Applications, Advantages, and Open Issues,” IEEE Access, vol. 9, pp. 75961–75980, 2021.
20. J. Singh, P. Singh, and S. S. Gill, “Fog Computing: A Taxonomy, Systematic Review, Current Trends and Research Challenges,” J. Parallel Distrib. Comput., vol. 157, p. 6, 2021.
21. C.-Y. Weng et al., “A Lightweight Anonymous Authentication and Secure Communication Scheme for Fog Computing Services,” IEEE Access, vol. 9, pp. 145522–145537, 2021.
22. P. Rana, K. Walia, and A. Kaur, “Challenges in Conglomerating Fog Computing with IoT for Building Smart City,” in Proc. ICCMST, Mohali, India, Dec. 2021, pp. 35–40.
23. K. V. Kumar et al., “Internet of Things and Fog Computing Applications in Intelligent Transportation Systems,” in Architecture and Security Issues in Fog Computing Applications. IGI Global, 2019.
24. V. K. Stephen et al., “Internet of Medical Things (IoMT) Fog-Based Smart Health Monitoring,” in Proc. ACCAI, Chennai, India, May 2023, pp. 1–8.
25. M. Rasheed et al., “A Survey on Fog Computing in IoT,” VFAST Trans. Softw. Eng., vol. 9, p. 4, 2021.
26. IEEE Standard 2672-2023, “IEEE Guide for General Requirements of Mass Customization,” IEEE, 2023.
27. K. Ahmed and S. Zeebaree, “Resource Allocation in Fog Computing: A Review,” in Proc. FMEC, Paris, France, Apr. 2020.
28. M. Mukherjee et al., “Security and Privacy in Fog Computing: Challenges,” IEEE Access, vol. 5, pp. 19293–19304, 2017.