MANTECH Publications

Journal of Research in Computer Science and Engineering
Volume 10, Issue 3, September-December, 2025
ISSN: 2457-0818 (Online)

AI Enabled Edge and Fog Computing for Intelligent, Scalable, and Real-Time Data Processing in Next-Generation Networks

Dr. Ritesh Kumar[1], Prithvi Singh[2], Kanika Rastogi[3]

Assistant Professor[1], Students[2, 3]

Department of Computer Science and Engineering

JIS College of Engineering, Kalyani

***Email ID**: riteshkumar1989@gmail.com[1]*

ABSTRACT

The rapid proliferation of Internet of Things (IoT) devices and the exponential growth of data traffic have driven the need for decentralized computing paradigms. Traditional cloud computing architectures face significant challenges related to latency, bandwidth consumption, and real-time responsiveness. To address these limitations, ***Edge and Fog Computing*** *have emerged as complementary paradigms, extending computational intelligence closer to data sources. With the integration of* ***Artificial Intelligence (AI)****, these distributed frameworks gain the ability to autonomously process, analyze, and optimize data flows at the network edge. This paper provides an in-depth exploration of* ***AI-enabled Edge and Fog Computing****, discussing their architectures, applications, and synergy in achieving intelligent, low-latency, and context-aware computing. The paper also highlights major challenges, existing solutions, and potential research opportunities to strengthen the future of AI-driven distributed computing infrastructures.*

KEYWORDS*: AI-enabled computing, Edge computing, Fog computing, Distributed intelligence, IoT, Cloud continuum, Low-latency processing, Autonomous systems*

INTRODUCTION

The unprecedented growth of connected devices in the **Internet of Things (IoT)** ecosystem has led to an enormous amount of data generation. These devices, ranging from smart sensors

Page 155-166 © MANTECH PUBLICATIONS 2025. All Rights Reserved

and autonomous vehicles to wearable devices and industrial machines, continuously produce heterogeneous data that demand real-time processing and intelligent decision-making. Traditional **cloud-centric architectures** rely on remote servers for computation and data analytics, leading to network congestion, higher latency, and limited scalability.

Edge and Fog Computing address these issues by decentralizing data processing and analytics, bringing computation closer to data sources. The **Edge layer** processes data directly on or near devices, while the **Fog layer** acts as an intermediate network infrastructure, providing additional computational resources and coordination between edge devices and the cloud. Integrating **Artificial Intelligence (AI)** into these layers transforms them from passive data collectors into **autonomous, learning-enabled computing systems** capable of performing predictive analytics, resource optimization, and decision-making locally.

The convergence of AI with Edge and Fog Computing not only enhances system intelligence and responsiveness but also empowers new applications in **smart cities, autonomous transportation, healthcare monitoring, and industrial automation**.

Table 1: Comparison between Cloud, Fog, and Edge Computing

Feature	Cloud Computing	Fog Computing	Edge Computing
Processing Location	Centralized data centers	Intermediate nodes (gateways, routers)	Near or on IoT devices
Latency	High	Moderate	Very Low
Bandwidth Usage	High	Medium	Low
Scalability	Very High	High	Moderate
Data Security	Centralized (less secure)	Distributed (more secure)	Localized (high security)
Example Use Cases	Big data analytics, storage	Smart grid, connected vehicles	Real-time monitoring, robotics

Page 155-166 © MANTECH PUBLICATIONS 2025. All Rights Reserved

LITERATURE REVIEW

Evolution of Edge and Fog Computing

Edge computing evolved as an extension of cloud computing, designed to overcome latency and bandwidth limitations by relocating computation closer to data sources. Fog computing, introduced by Cisco, acts as a horizontal layer extending the cloud to the network edge, enabling distributed services such as computation, storage, and control within intermediate nodes like gateways, routers, or base stations.

AI Integration into Edge-Fog Architecture

Recent studies demonstrate the integration of **machine learning (ML)** and **deep learning (DL)** models into edge and fog nodes for intelligent data analytics. Researchers have proposed federated learning models where training occurs locally on edge devices, reducing the dependency on centralized data storage. Fog nodes, with their moderate computational resources, support model aggregation and collaborative learning.

For example, AI-driven **predictive maintenance** in smart industries uses local edge intelligence to identify anomalies in real time, while fog nodes coordinate insights across distributed sensors. Similarly, **healthcare monitoring systems** leverage on-device neural networks for early health anomaly detection, minimizing data transfer delay to cloud servers.

Architectural Advancements

Recent architectures, such as **Hierarchical Edge-Fog-Cloud Continuum**, enable seamless distribution of computational tasks based on resource capacity and latency constraints. AI algorithms dynamically allocate workloads, ensuring energy efficiency and task reliability. Emerging frameworks like **Edge AI inference engines** (e.g., TensorRT, OpenVINO, and TinyML) facilitate lightweight AI model deployment on resource-constrained edge devices.

AI-ENABLED EDGE AND FOG COMPUTING ARCHITECTURE

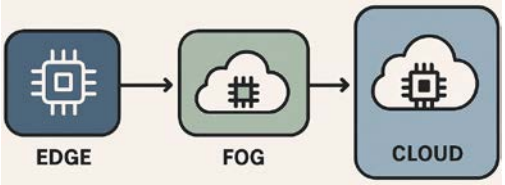


Figure 1: AI-Enabled Edge–Fog–Cloud Architecture

Page 155-166 © MANTECH PUBLICATIONS 2025. All Rights Reserved

Table 2: AI Techniques Used in Edge and Fog Computing

AI Technique	Edge Layer Applications	Fog Layer Applications	Advantages
Machine Learning (ML)	Sensor data classification, anomaly detection	Aggregated model training	Improved adaptability
Deep Learning (DL)	Image and video analysis	Distributed inference	High accuracy in perception tasks
Reinforcement Learning (RL)	Autonomous control in robots, vehicles	Task scheduling, resource management	Dynamic decision optimization
Federated Learning (FL)	On-device model updates	Model aggregation across nodes	Privacy preservation
Knowledge Distillation	Lightweight model deployment	Performance tuning	Energy-efficient AI inference

Edge Layer

The edge layer consists of IoT devices and local gateways with limited computing capacity. By embedding AI models, devices can perform lightweight tasks such as object detection, anomaly recognition, or signal classification. Edge AI enables low-latency decision-making essential for applications like **autonomous driving**, **industrial robotics**, and **remote patient monitoring**.

Fog Layer

The fog layer bridges the edge and cloud environments, comprising intermediate nodes with higher computing capabilities. AI at the fog layer supports resource orchestration, workload balancing, and distributed learning. It aggregates data insights from multiple edge nodes and provides context-aware processing.

Cloud Layer

While AI inference predominantly occurs at the edge and fog levels, the cloud layer remains vital for large-scale model training, historical data analysis, and long-term storage. The synergy

Page 155-166 © MANTECH PUBLICATIONS 2025. All Rights Reserved

among these layers creates a **distributed intelligence hierarchy** that enhances system reliability and responsiveness.

AI-Driven Orchestration

AI algorithms manage **task allocation, network scheduling, and energy optimization** across layers. Reinforcement learning and meta-learning techniques are employed to dynamically adjust computing tasks in response to changing network conditions.

APPLICATIONS OF AI-ENABLED EDGE AND FOG COMPUTING

Smart Cities

Edge AI assists in traffic prediction, pollution monitoring, and intelligent surveillance by analyzing data locally from cameras and sensors. Fog nodes coordinate city-wide optimization such as adaptive traffic light control and emergency response routing.

Healthcare and Remote Monitoring

AI models embedded in wearable devices or hospital equipment can detect vital sign anomalies and alert healthcare providers in real time. Fog nodes aggregate medical data securely for analysis while ensuring data privacy and regulatory compliance.

Industrial Automation (Industry 4.0)

AI-enabled edge devices in manufacturing monitor equipment health and process efficiency. Predictive maintenance reduces downtime, while fog-based coordination improves supply chain and production planning.

Autonomous Vehicles and Transportation Systems

Vehicles equipped with edge AI models process sensor data for obstacle detection and navigation. Fog infrastructure supports vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication, improving traffic efficiency and safety.

Smart Agriculture

Edge AI devices in farms analyze soil moisture, temperature, and crop health using local machine learning models. Fog nodes integrate this information for regional yield prediction

Page 155-166 © MANTECH PUBLICATIONS 2025. All Rights Reserved

and water management optimization.

CHALLENGES IN AI-ENABLED EDGE AND FOG COMPUTING

Table 3: Major Challenges and Corresponding AI Solutions

Challenge	Description	AI-Driven Solution
Resource Limitation	Edge devices have restricted compute and storage capacity	Model compression, pruning, and quantization
Security & Privacy	Sensitive local data may be exposed	Federated learning and differential privacy
Network Latency	Communication delay between layers	Intelligent task scheduling and caching
Scalability	Large-scale device management is complex	Decentralized orchestration using RL
Dynamic Environments	Models require frequent updates	Continual and meta-learning mechanisms

Resource Constraints

Edge devices have limited processing power, memory, and energy capacity. Running AI models requires optimization through **model compression**, **quantization**, or **knowledge distillation** to ensure efficient operation.

Security and Privacy

Data generated at the edge often contains sensitive information. Ensuring privacy-preserving AI techniques such as **federated learning** and **differential privacy** is essential for secure distributed intelligence.

Scalability and Interoperability

Heterogeneous devices, protocols, and communication standards complicate large-scale deployment. Unified frameworks are needed to ensure seamless interoperability across diverse hardware.

Page 155-166 © MANTECH PUBLICATIONS 2025. All Rights Reserved

Latency and Network Congestion

While edge processing reduces latency, coordination among fog and cloud layers can still introduce communication overheads. Intelligent task scheduling and bandwidth management are vital for performance optimization.

Model Updating and Adaptation

AI models deployed on the edge must adapt to changing environments without frequent cloud retraining. **Continual learning** and **online adaptation techniques** are being explored to support dynamic model updates.

ADVANTAGES OF AI IN EDGE AND FOG COMPUTING

The integration of Artificial Intelligence (AI) with Edge and Fog Computing introduces a wide range of advantages that enhance system performance, responsiveness, and scalability. By embedding intelligence into distributed nodes, the computing infrastructure becomes capable of autonomous decision-making, adaptive resource management, and real-time analytics. The following subsections elaborate on the key benefits provided by AI-enabled Edge and Fog systems.

Low Latency and Real-Time Processing

One of the most significant advantages of integrating AI into edge and fog architectures is the reduction in latency and the ability to perform real-time data processing. In traditional cloud-based systems, data must travel to distant servers for computation and return with results, causing delays that are unacceptable for time-sensitive applications. In contrast, AI at the edge allows for immediate inference and decision-making close to the data source.

For instance, in autonomous vehicles, edge AI models process camera and sensor inputs directly within milliseconds to detect obstacles, predict motion, and make driving decisions without waiting for cloud feedback. Similarly, in industrial automation, robots and smart machinery rely on local AI inference for rapid fault detection and process control. Fog nodes further enhance this process by coordinating multiple edge devices to maintain real-time synchronization across distributed environments.

Page 155-166 © MANTECH PUBLICATIONS 2025. All Rights Reserved

This distributed intelligence ensures instantaneous responses, minimizing downtime, and increasing system reliability, which is crucial for critical operations such as healthcare monitoring, emergency response, and predictive maintenance.

Bandwidth Efficiency

AI-enabled edge and fog systems substantially improve bandwidth utilization by performing data analytics locally and transmitting only essential or aggregated insights to higher network layers. In conventional architectures, massive amounts of raw IoT data are continuously uploaded to the cloud for processing, leading to congestion and high communication costs.

By deploying AI-driven data filtering and feature extraction techniques at the edge, unnecessary data transfer is avoided. For example, a surveillance camera equipped with edge AI can analyze footage in real-time and send only relevant clips or detected anomalies to fog or cloud servers. This selective transmission significantly reduces network load, conserves bandwidth, and allows cloud resources to focus on long-term analytics instead of routine processing.

Furthermore, fog nodes can perform intelligent data aggregation and compression, ensuring that only summarized or context-aware information is transmitted upstream. This hierarchical data management results in efficient bandwidth utilization, supporting scalability for large-scale IoT networks and 5G-based applications.

Enhanced Security and Privacy

AI integration enhances both data security and user privacy within edge and fog computing ecosystems. Since data processing and decision-making occur closer to the data source, sensitive information remains within local networks rather than being transmitted to remote cloud servers, thereby minimizing the risk of interception or unauthorized access.

Techniques such as federated learning, homomorphic encryption, and differential privacy empower AI models to train collaboratively across distributed nodes without exposing raw data. For example, in smart healthcare systems, patient data remains encrypted and processed locally on hospital or wearable devices, with only model updates being shared for global

Page 155-166 © MANTECH PUBLICATIONS 2025. All Rights Reserved

learning through fog nodes.

Additionally, AI-driven intrusion detection systems (IDS) deployed at the edge can identify abnormal network activities, cyber threats, or data breaches in real-time, enabling proactive security responses. The decentralized nature of fog architectures further provides resilience against single points of failure, ensuring data protection even under attack scenarios. Hence, AI not only strengthens security but also reinforces trust in distributed computing environments.

Context Awareness and Adaptive Intelligence

AI enables edge and fog devices to achieve context-aware computing, allowing systems to interpret and adapt to environmental, spatial, and temporal factors dynamically. By utilizing sensor fusion, pattern recognition, and reinforcement learning, edge AI can understand user behavior, operational context, and environmental conditions to make more informed and situation-specific decisions.

For instance, smart city traffic systems use context-aware AI models at the edge to analyze vehicle density, weather, and pedestrian movement, adjusting signal timings in real time. In smart homes, devices can adapt lighting, temperature, and energy usage based on occupants' habits and preferences learned over time. Fog nodes further contribute by integrating multi-context data from various edge sources to coordinate large-scale adaptive responses across systems.

This level of contextual understanding transforms passive IoT devices into proactive, intelligent entities capable of learning, predicting, and optimizing performance autonomously. As a result, systems become more responsive, personalized, and efficient, catering to user-specific and environmental variations seamlessly.

Energy Efficiency and Sustainable Operation

Energy efficiency is a critical concern in distributed computing environments, particularly for battery-powered or resource-limited IoT devices. AI enhances energy management through intelligent workload distribution, dynamic resource allocation, and power-aware task

Page 155-166 © MANTECH PUBLICATIONS 2025. All Rights Reserved

scheduling across the edge-fog-cloud hierarchy.

Machine learning models can predict workload variations and adaptively offload computational tasks between layers to maintain optimal energy usage. For example, non-urgent or heavy computations can be shifted from edge devices to nearby fog nodes when energy levels are low, ensuring system longevity without performance degradation.

Moreover, AI-driven cooling and power control mechanisms in fog data centers can optimize operational energy consumption based on real-time demand patterns. In smart agriculture, edge AI reduces power consumption by selectively activating sensors or communication modules only when environmental thresholds are exceeded.

These strategies collectively lead to sustainable and energy-efficient computing, supporting the deployment of large-scale AI-enabled infrastructures in remote or power-constrained environments. The synergy between AI and distributed computing thus ensures not only operational performance but also environmental sustainability.

Improved Scalability and Reliability

As a complementary advantage, AI also improves the scalability and reliability of edge-fog systems. Through intelligent orchestration, AI can monitor network performance, predict failures, and automatically rebalance workloads across devices and nodes. This capability prevents system overloads and enables fault tolerance even in large and dynamically changing networks.

For instance, reinforcement learning agents can determine optimal task placement strategies based on network latency, energy consumption, and node availability. This adaptability ensures continuous operation and service quality, even under varying workloads or hardware failures. Consequently, AI transforms the edge-fog ecosystem into a self-optimizing and resilient computing environment, capable of scaling **seamlessly as network demands evolve.**

SCOPE AND FUTURE DIRECTIONS

The future of **AI-enabled Edge and Fog Computing** lies in developing self-managing,

Page 155-166 © MANTECH PUBLICATIONS 2025. All Rights Reserved

adaptive, and explainable computing systems capable of handling dynamic and unpredictable environments. **Federated edge learning**, **multi-agent reinforcement learning**, and **neuromorphic computing** are expected to redefine decentralized AI by providing learning capabilities with minimal energy requirements.

Integration with **5G and 6G networks** will enhance data transfer speeds, enabling ultra-reliable, low-latency communication for time-sensitive applications like remote surgery and autonomous drone networks. The evolution of **TinyML** and **quantum-inspired AI models** will make it feasible to deploy high-performance AI on low-power edge devices.

Moreover, incorporating **blockchain technology** can secure data exchange among distributed nodes, ensuring transparency and trust. Emerging paradigms such as **fog federation** and **AI-driven resource orchestration** will further enhance scalability and reliability in large-scale deployments.

CONCLUSION

AI-enabled Edge and Fog Computing represent a transformative step toward decentralized intelligence and real-time analytics. By integrating AI capabilities at the network edge, these systems overcome the inherent limitations of cloud-based infrastructures, achieving lower latency, greater efficiency, and enhanced privacy. The synergy of AI with distributed computing layers not only supports scalable IoT ecosystems but also unlocks new opportunities for innovation in industries, healthcare, and smart cities.

As research continues, the focus will shift toward optimizing resource allocation, ensuring security, and enabling continuous learning at the edge. The convergence of AI, Edge, and Fog technologies will ultimately shape the foundation of next-generation computing architectures, driving the evolution of intelligent, autonomous, and sustainable digital ecosystems.

REFERENCES

1. Bonomi, F., Milito, R., Zhu, J., & Addepalli, S. (2012). *Fog computing and its role in the Internet of Things.* Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing, 13–16. ACM.

Page 155-166 © MANTECH PUBLICATIONS 2025. All Rights Reserved

2. Satyanarayanan, M. (2017). *The emergence of edge computing.* Computer, 50(1), 30–39. IEEE.
3. Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). *Edge computing: Vision and challenges.* IEEE Internet of Things Journal, 3(5), 637–646.
4. Chiang, M., & Zhang, T. (2016). *Fog and IoT: An overview of research opportunities.* IEEE Internet of Things Journal, 3(6), 854–864.
5. Liu, J., Kato, N., Ma, J., Ucar, S., & Zhang, J. (2020). *AI-enabled edge computing: An intelligent computing architecture for future wireless networks.* IEEE Wireless Communications, 27(2), 14–21.
6. Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., & Zhang, J. (2019). *Edge intelligence: Paving the last mile of artificial intelligence with edge computing.* Proceedings of the IEEE, 107(8), 1738–1762.
7. Li, E., Zhou, Z., & Chen, X. (2018). *Edge intelligence: On-demand deep learning model co-inference with device-edge synergy.* Proceedings of the 2018 Workshop on Mobile Edge Communications, 31–36. ACM.
8. Park, J., Samarakoon, S., Bennis, M., & Debbah, M. (2019). *Wireless network intelligence at the edge.* Proceedings of the IEEE, 107(11), 2204–2239.
9. Kaur, K., Garg, S., Kaddoum, G., & Guizani, M. (2020). *Security and privacy in fog computing: Challenges and solutions.* IEEE Communications Surveys & Tutorials, 22(3), 1827–1857.
10. McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). *Communication-efficient learning of deep networks from decentralized data.* Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), 1273–1282.

Page 155-166 © MANTECH PUBLICATIONS 2025. All Rights Reserved