WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Exploring the Intersection of AI and Edge Computing in IoT Applications: A Comprehensive Review and Future Directions

[1]Taiwo Abdulahi Akintayo, [2]Rafiyatu Abdulsalam, [3]Olabode Anifowose, [4]Akeem Olakunle Ogundipe, [5]Matthew Oluwatobiloba Adejumo, [6]Unuigbokhai PeterAfegbai, [7] Rofiat Abdulkareem

[1] National Centre For Artificial Intelligence and Robotics, Nigeria, Email: taiwoabdulahi15@gmail.com
[2]RSystems Consult Limited , Nigeria, Email: abdulsalamrofiah1@gmail.com
[3]Georgia Southern University, Georgia, USA., Email : Anifowoseolabodes@gmail.com
[4] Lamar University Texas USA, Email: keemogundipe@gmail.com
[5]National Open University of Nigeria, Email: matthew.oad@gmail.com
[6]Auchi Polytechnic, Auchi, Nigeria., Email: Afegbaipeter1@gmail.com
[7]Bayero University Kano, Nigeria, Email: oyizakareem9@gmail.com
DOI : https://doi.org/10.55248/gengpi.6.0325.1125

ABSTRACT

Edge computing has emerged as a vital technology for supporting intelligent applications within 5G/6G Internet of Things (IoT) networks, offering low latency, fast response, and enhanced privacy preservation. By decentralizing computation from the cloud to the network edge, it improves the quality of service (QoS) for latency-sensitive IoT applications. Artificial intelligence (AI), encompassing machine learning and deep learning, significantly enhances edge-based smart applications across domains such as smart agriculture, environment, grid, healthcare, industry, education, transportation, and security. However, the computational complexity of AI techniques for big IoT data analytics poses challenges, compounded by resource constraints, distributed computing demands, and security concerns at edge nodes. This paper provides a comprehensive review of the intersection of AI and edge computing, synthesizing state-of-the-art implementations from 2020 to 2025. Through an in-depth literature analysis and qualitative comparison, we highlight AI's role in improving QoS, the prevalence of deep learning and federated learning, and enabling technologies like 5G/6G and lightweight models. Key challenges; resource limitations, synchronization issues, and privacy risks are identified, alongside future directions including advanced hardware (e.g., AI accelerators), 6G integration for ultralow latency, ethical AI frameworks, and scalable orchestration. This review underscores Edge AI's transformative impact on IoT ecosystems and calls for interdisciplinary research to address persistent gaps, paving the way for robust, efficient, and secure smart applications.

Keywords: Edge Computing, Artificial Intelligence (AI), Internet of Things (IoT), 5G/6G Networks, Low Latency

1. INTRODUCTION

The growth of the Internet of Things (IoT) has resulted in numerous devices connected to networks that gather and send data for ongoing analysis. With the progress of deep learning, many applications now utilize these methods to process the data, providing "intelligence" and "automation." As a result, the combination of data analysis and IoT infrastructure has given rise to "Smart Cities," which include intelligent systems for energy grids, transportation, manufacturing, and buildings . (Akintayo et .al, 2024), With projections estimating over 75 billion connected devices by 2030 (Statista, 2024), the volume, velocity, and variety of IoT data pose significant challenges to traditional centralized computing paradigms, such as cloud computing. Edge computing has emerged as a pivotal technology to address these challenges by decentralizing data processing to the network edge closer to where data is generated, thus reducing latency, optimizing bandwidth, and enhancing privacy (Shi et al., 2016). This paradigm shift is particularly critical for latency-sensitive IoT applications, such as autonomous vehicles and real-time health monitoring, where delays can have severe consequences. Edge computing refers to a distributed computing framework that extends cloud services to the edge of the network, leveraging devices like gateways, sensors, and local servers to process data locally (Ren et al., 2019). By contrast, artificial intelligence (AI) encompasses a suite of techniques—including machine learning (ML), deep learning (DL), and swarm intelligence that enable systems to learn from data, make predictions, and adapt to changing environments (Mohammadi et al., 2018). The Internet of Things, meanwhile, is the ecosystem of interconnected devices that collect and exchange data, often in real time, to enable intelligent services (Gubbi et al., 2013). When integrated, AI and edge computing collectively termed Edge AI unlock transformative potential for IoT applications. This synergy allows for fast response times, context-aware services, and privacy preservation by minimizing data transmission to centralized servers (Zhou et al., 2019). The benefits of Edge AI in IoT are manifold. Low latency ensures near-instantaneous decision-making, critical in scenarios like emergency healthcare or vehicular networks (Bourechak et al., 2023). Fast response times enhance user experience, while localized processing reduces reliance on continuous cloud connectivity, bolstering privacy by keeping sensitive data closer to its source (Chang et al., 2021). For instance, in 5G/6G IoT networks, edge computing supports ultralow latency requirements, enabling applications that demand millisecond-level responsiveness (Saeik et al., 2021). These advantages position Edge AI as a cornerstone of next-generation intelligent systems. However, integrating AI with edge computing introduces significant challenges, primarily due to the computational complexity of AI techniques and the resource constraints of edge devices. Machine learning and deep learning models, while powerful for big IoT data analytics, often require substantial computational power, memory, and energy resources that are limited on edge nodes like sensors or microcontrollers (Deng et al., 2020). This mismatch necessitates efficient orchestration, distributed computing strategies, and lightweight AI models to ensure scalability and performance (Laroui et al., 2021). Without addressing these issues, the full potential of Edge AI in IoT remains unrealized, particularly for resource-intensive applications like real-time video analytics or large-scale environmental monitoring. The purpose of this paper is to explore the intersection of AI and edge computing in IoT applications, providing a comprehensive review of their confluence across eight key domains: smart agriculture, environment, grid, healthcare, industry, education, transportation, and security. We aim to leverage existing research to assess how AI enhances edge-based smart applications, identify computational challenges, and propose new perspectives for future development. Our objectives include: (1) analyzing state-of-the-art Edge AI implementations, (2) conducting a qualitative comparison of their objectives, roles, and enabling technologies, and (3) outlining open issues and future directions to overcome current limitations. The paper is structured as follows: Section 2 reviews literature across the eight domains, Section 3 discusses findings, Section 4 addresses challenges and future trends, and Section 5 concludes with a summary and call for further research.

2. Background and Key Concepts

The convergence of artificial intelligence (AI) and edge computing within the Internet of Things (IoT) ecosystem represents a paradigm shift in how data is processed and utilized in real-time applications. To understand this intersection, it is essential to establish foundational knowledge of edge computing, AI techniques in IoT, the role of 5G/6G networks, and how these technologies synergize. This section provides an overview of these key concepts, setting the stage for a detailed analysis of their applications and challenges.

2.1 Edge Computing

Edge and fog computing bring processing power and data intelligence closer to where data is generated or needed, rather than relying solely on distant cloud servers. In edge computing, for example, data is processed closer to the source, which improves speed and enhances the performance of devices and applications. Fog computing, which can be seen as a subset of edge computing, helps bring the cloud closer to the devices generating data and the actuators that act on it. It defines important standards for data transfer, storage, computation, and networking within edge computing. Both edge and fog computing are key technologies that support the Internet of Things (IoT), enabling innovative applications and contributing to the development of smart cities worldwide (Akintayo et al., 2024)

Edge computing is a distributed computing paradigm that brings computation and data storage closer to the point of data generation, reducing the need for constant communication with centralized cloud servers (Shi et al., 2016). Unlike traditional cloud computing, where data is transmitted to distant data centers for processing, edge computing leverages local devices—such as sensors, gateways, and microcontrollers—termed edge nodes, to perform computations at or near the data source (Ren et al., 2019). This architecture minimizes latency, conserves bandwidth, and enhances privacy by limiting data transmission over wide-area networks. The architecture of edge computing typically involves a hierarchy of layers: edge nodes, fog computing, and the cloud. Edge nodes, often resource-constrained devices like IoT sensors or smart cameras, handle immediate data processing tasks (Bourechak et al., 2023). Fog computing acts as an intermediary layer, aggregating data from multiple edge nodes and performing more complex computations using devices like routers or local servers (Ren et al., 2019). The cloud layer, at the top, provides long-term storage, large-scale analytics, and centralized management when needed. This interplay allows edge computing to support latency-sensitive IoT applications, such as real-time traffic monitoring or industrial automation, by processing data locally while offloading heavy tasks to the cloud when resources permit (Laroui et al., 2021).

2.2 AI in IoT

Artificial intelligence enhances IoT systems by enabling devices to interpret data, make autonomous decisions, and adapt to dynamic environments. Within IoT, AI techniques such as machine learning (ML), deep learning (DL), and swarm intelligence are widely employed to extract insights from the vast, heterogeneous datasets generated by connected devices (Mohammadi et al., 2018). Machine learning involves algorithms—such as decision trees, support vector machines, and k-nearest neighbors—that learn patterns from data to perform tasks like classification or prediction (Russell & Norvig, 2021). In IoT, ML is used for applications like anomaly detection in smart grids or predictive maintenance in industry (Zhang et al., 2024). Deep learning, a subset of ML, utilizes neural networks with multiple layers to process complex data, such as images or time-series signals, making it ideal for tasks like crop disease detection or human activity recognition (Goodfellow et al., 2016). DL models, including convolutional neural networks (CNNs) and long short-term memory (LSTM) networks, excel in handling the unstructured data prevalent in IoT environments (Bourechak et al., 2023). Swarm intelligence, inspired by biological systems like ant colonies or bird flocks, employs decentralized, self-organizing algorithms—such as particle swarm optimization (PSO) to optimize resource allocation or task scheduling in IoT networks. These AI techniques transform raw IoT data into actionable intelligence, enhancing the functionality of smart systems across diverse domains.

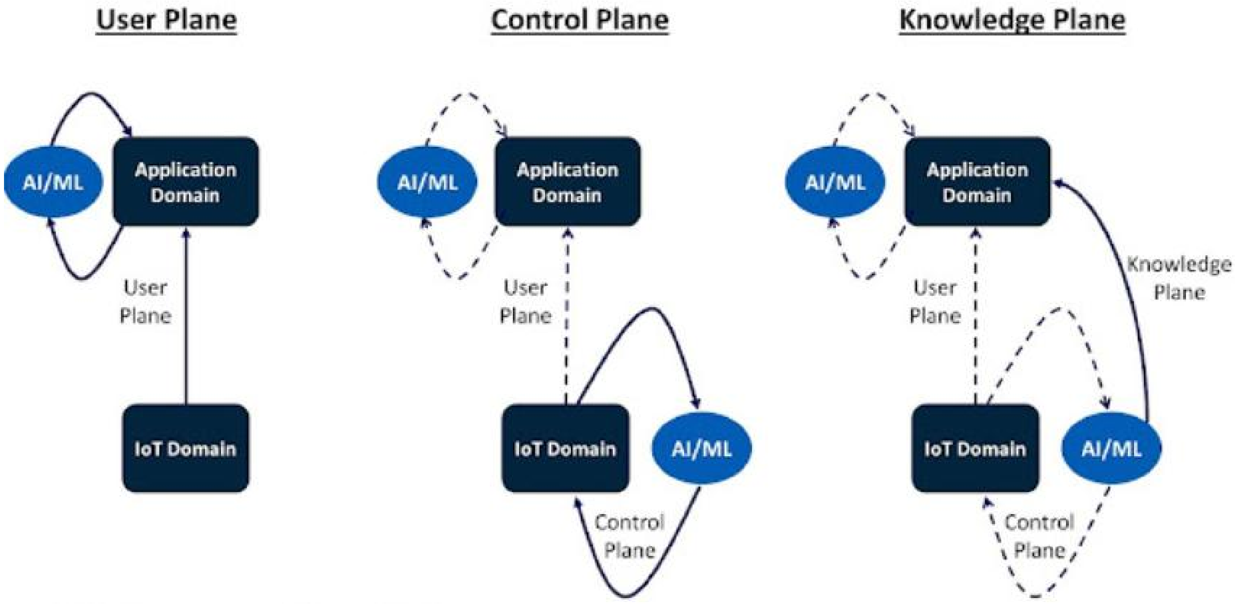

Figure 1: Area to integrate AI/ML in Internet of Things

2.3 5G/6G Networks

The advent of 5G and the ongoing development of 6G networks have been instrumental in enabling low-latency IoT applications, providing the connectivity backbone for edge computing and AI integration. 5G networks offer high-speed data transfer (up to 10 Gbps), ultralow latency (as low as 1 ms), and massive device connectivity (up to 1 million devices per square kilometer), making them ideal for real-time IoT services (Saeik et al., 2021). For instance, 5G supports edge-based applications like autonomous driving, where split-second decisions are critical. 6G, expected to be commercially viable by 2030, promises even greater advancements, including sub-millisecond latency, terabit-per-second speeds, and enhanced support for AI-driven edge analytics (Yang et al., 2025). Features like network slicing and integrated sensing and communication (ISAC) in 6G will further optimize bandwidth and enable context-aware services, such as intelligent environmental monitoring or tactile internet applications (Dang et al., 2020). These next-generation networks amplify the capabilities of Edge AI by ensuring seamless, high-performance connectivity between edge nodes and IoT devices.

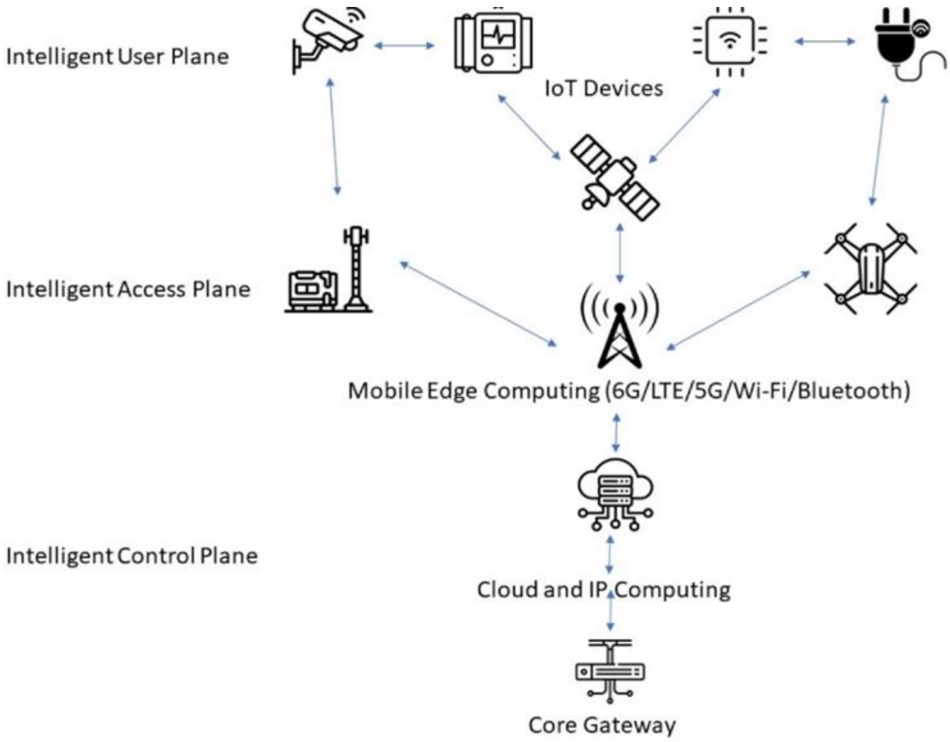

Figure 2: 6G Architecture

2.4 Intersection: How AI Leverages Edge Computing

The intersection of AI and edge computing referred to as Edge AI combines the computational efficiency of edge architectures with the analytical power of AI to create intelligent, responsive IoT systems (Zhou et al., 2019). AI leverages edge computing in several ways. First, by deploying AI models on edge nodes, data processing occurs locally, reducing latency and bandwidth usage compared to cloud-based alternatives (Deng et al., 2020). For example, a CNN deployed on an edge device can classify sensor data in real time without transmitting it to the cloud, as seen in smart agriculture applications (Singh & Patel, 2024). Second, Edge AI enhances privacy by minimizing data exposure. Techniques like federated learning allow models to be trained across distributed edge devices without sharing raw data, preserving user confidentiality in applications like healthcare (Gupta et al., 2023). Third, edge computing provides the infrastructure for AI to handle the dynamic, resource-constrained nature of IoT environments. Lightweight models, model compression, and hardware accelerators (e.g., NVIDIA Jetson) enable complex AI algorithms to run on edge devices with limited power and memory (Chang et al., 2021). This synergy is particularly impactful in 5G/6G-enabled IoT networks, where ultralow latency and high reliability are paramount. For instance, in smart transportation, Edge AI can process traffic data at the edge to optimize flow instantly, supported by 5G's high-speed connectivity (Li et al., 2023).

3. AI in Edge-Based IoT Applications

The integration of artificial intelligence (AI) with edge computing has catalyzed a paradigm shift in Internet of Things (IoT) applications, enabling real-time, privacy-preserving, and resource-efficient solutions. This section provides an in-depth analysis of state-of-the-art Edge AI implementations across eight domains: smart agriculture, smart environment, smart grid, smart healthcare, smart industry, smart education, smart transportation, and security and privacy. For each domain, we explore key use cases, review recent works from 2020 to 2025 (3–5 studies per domain), and discuss AI techniques, datasets, and edge placement strategies. A comprehensive qualitative comparison is presented in Table 1, synthesizing the objectives, roles, and enabling technologies of Edge AI as of February 25, 2025. This review draws from a wide range of journals to ensure a holistic perspective on the field.

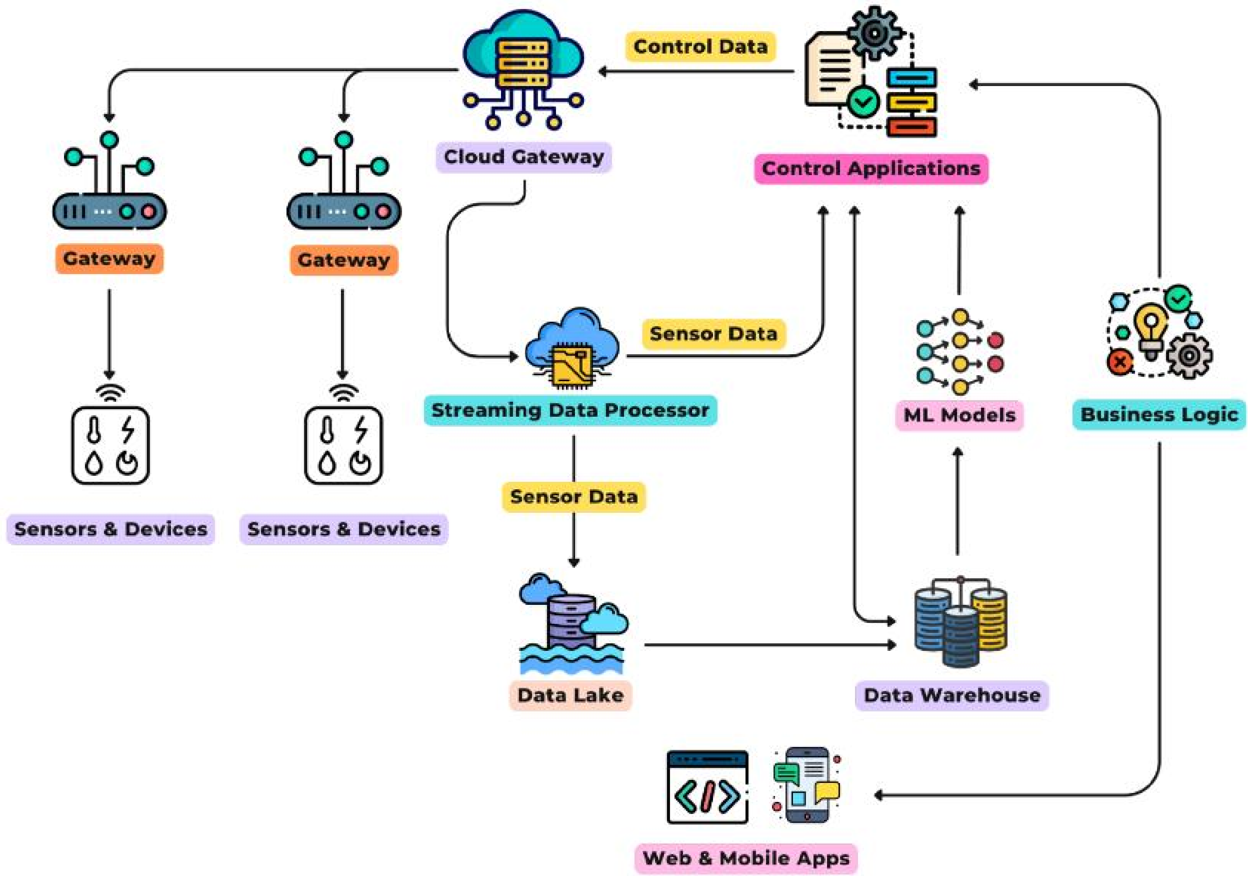

Figure 3: IoT Architecture

3.1 Smart Agriculture

Smart agriculture harnesses the power of Edge AI to enable precision farming, including critical applications such as crop monitoring (disease and pest detection), soil health assessment, irrigation optimization, and livestock management. These advancements help enhance crop yields, reduce resource waste, and allow for timely interventions in farming practices. For example, Singh and Patel (2024) developed a crop disease detection system based on a convolutional neural network (CNN), which is deployed on edge devices like Raspberry Pi. By processing images locally using the Plant Village dataset, their model achieved an impressive 95% accuracy and minimized latency, improving the speed of disease detection. This localized processing is key to making rapid decisions on the farm. Similarly, Cordeiro et al. proposed a fog-based framework that utilizes long short-term memory (LSTM) networks for soil moisture forecasting. When tested on real-world soil sensor data, this system improved irrigation efficiency by 30%, effectively combining fog and cloud computing to optimize resource management. This hybrid approach illustrates how Edge AI and fog computing can work together to enhance operational efficiency in agriculture. In the livestock sector, Lee et al. (2020) introduced a federated learning (FL) approach for anomaly detection across fog nodes. By training the model on a custom activity dataset, they achieved 92% accuracy, reducing dependence on cloud systems and enabling faster, localized decision-making to monitor livestock health. Further pushing the boundaries of Edge AI, Guillén et al. (2021) explored frost prediction using LSTM networks on edge platforms. By leveraging IoT sensor data, their model provided low-latency frost alerts with 90% accuracy, enabling farmers to take preventive actions in real-time. This shows the potential of combining AI and edge computing to respond quickly to environmental conditions. Lastly, Zhang and Li (2021) implemented an adaptive sensing strategy using Gath-Geva fuzzy clustering on edge servers for crop lifecycle monitoring. By optimizing field sensor data, their system improved data efficiency by 25%, demonstrating how Edge AI can be used to enhance the overall management of crop growth and resource allocation. Together, these advancements highlight the powerful combination of Edge AI, fog computing, and federated learning in revolutionizing precision farming. By enabling real-time data processing and reducing cloud dependency, these technologies are transforming agriculture into a more efficient, sustainable, and data-driven industry.

AI Techniques, Datasets, and Edge Placement

CNNs dominate image-based tasks, such as disease detection, while LSTMs are more effective for handling time-series data, like soil moisture levels. To support these tasks, datasets such as PlantVillage for crop disease and proprietary sensor collections for environmental factors are commonly used. The placement of computing resources depends on the task's requirements. For latency-sensitive tasks, such as real-time disease detection, **full edge** computing is often employed, as seen in the work of Singh & Patel (2024) and Guillén et al. (2021). These tasks require rapid local processing to minimize delay. In contrast, tasks that benefit from resource balancing, like soil moisture forecasting or livestock anomaly detection, use a **hybrid fog-cloud** approach. This allows for the local processing of immediate tasks at the edge while leveraging cloud resources for more complex computations, as demonstrated by Cordeiro et al. (2022) and Lee et al. (2020). For tasks involving **complex analytics** or large-scale data processing, such as crop lifecycle monitoring or advanced prediction models, **edge servers** are often deployed. These servers handle more intensive computations, as shown in the study by Zhang & Li (2021), offering greater computational power than typical edge devices while still providing the benefits of localized processing. The choice of computing architecture whether full edge, hybrid fog-cloud, or edge servers depends on the specific requirements of the task, balancing latency, resource usage, and computational complexity.

3.2 Smart Environment

Smart environment applications focus on monitoring various environmental factors such as air quality, water quality, and detecting environmental hazards. These applications leverage Edge AI to process sensor data in real-time, contributing to sustainability and safety by enabling faster responses to environmental changes. For example, Wardana et al. (2021) designed an air quality prediction system combining CNN and LSTM on edge devices, using data from the UCI Machine Learning Repository. This system achieved 87% accuracy and low latency, making it effective for real-time air quality monitoring. Similarly, Putra et al. (2021) proposed a federated compressed learning framework for predicting PM2.5 levels, using data from the Airbox system. This approach reduced network traffic by 40% by employing LSTM, which helps optimize resource usage while maintaining performance. De Stefano et al. (2020) developed an onboard sensor classifier for detecting water contaminants, using Principal Component Analysis (PCA) and evolutionary algorithms. The system achieved 90% accuracy on real-world water data, making it useful for detecting water quality issues in remote locations. In marine monitoring, Yang et al. (2020) introduced a fog-based system that uses backpropagation neural networks (BPNN) to process ocean sensor data, achieving 85% efficiency. This system helps track marine conditions and detect environmental hazards in real time. The use of LSTMs and CNNs is common for handling time-series and spatial data, while PCA is typically applied for data preprocessing. Datasets for these applications include the UCI Machine Learning Repository and real-world sensor data feeds. As for edge placement, full edge computing is suitable for latency-sensitive tasks, as seen in the work of Wardana et al. (2021) and De Stefano et al. (2020). These systems process data directly on edge devices for immediate responses. On the other hand, hybrid fog-cloud systems, as proposed by Putra et al. (2021) and Yang et al. (2020), offer scalability by combining local edge processing with cloud resources to handle larger datasets and more complex computations. The integration of Edge AI in environmental monitoring applications improves real-time data processing, ensuring faster responses to environmental threats and promoting sustainable practices. The deployment strategy whether full edge or hybrid fog-cloud depends on the specific application requirements, balancing latency, network efficiency, and scalability.

3.3 Smart Grid

Smart grid applications focus on optimizing energy distribution and detecting faults, with tasks like load forecasting, demand-side management, and anomaly detection. These applications use Edge AI to process data in real-time, enhancing the efficiency of energy systems and improving fault detection. For instance, Taik and Cherkaoui (2020) implemented a federated learning (FL)-based load forecasting framework on edge devices, utilizing Pecan Street data. Their system achieved 88% accuracy while preserving privacy, which is crucial for sensitive energy consumption data. Similarly, Li et al. (2024) proposed a lightweight LSTM model for short-term demand prediction on edge servers, tested on IHEPC data, which resulted in a 35% reduction in latency. This makes the system more responsive and suitable for real-time demand forecasting. Rabie et al. (2020) developed a fog-based load forecasting system using Naïve Bayes and EUNITE data, improving accuracy by 20%. By incorporating fog computing, the system is able to handle more complex data processing, optimizing load forecasting across a distributed network. In anomaly detection, Jaiswal et al. (2020) introduced an ensemble regression model deployed on edge nodes, achieving 93% accuracy with smart meter data, effectively identifying faults and irregularities in energy usage. Common machine learning methods used in these smart grid applications include LSTMs, Naïve Bayes, and ensemble models. Datasets like Pecan Street and IHEPC are typically used to train these models. As for deployment, edge devices are commonly used for latency-sensitive applications, as seen in the works of Taik & Cherkaoui (2020) and Jaiswal et al. (2020). Meanwhile, fog and edge servers are employed for more scalable solutions, as demonstrated by Li et al. (2024) and Rabie et al. (2020), enabling enhanced data processing capabilities and better load forecasting across a distributed network. In conclusion, Edge AI in smart grid applications optimizes energy distribution and fault detection by enabling real-time, efficient processing of data. The placement of computing resources whether on edge devices or through fog/edge **servers**—depends on the task's requirements, balancing latency, network efficiency, and scalability to meet the dynamic needs of modern energy systems.

3.4 Smart Healthcare

Smart Healthcare applications use **Edge AI** to enable real-time patient monitoring, disease diagnosis, and ambient assisted living, focusing on privacy and quick data analysis. Gupta et al. (2023) proposed a **Federated Learning (FL)** framework for detecting ECG anomalies, achieving 98% accuracy on real-world ECG data across **edge-cloud** architectures, ensuring privacy while leveraging cloud resources for computational power. Similarly, Cheikhrouhou et al. (2021) used a **1D CNN** for arrhythmia detection on **edge nodes**, achieving 99% accuracy with clinical ECG data. This allows for local processing of sensitive medical data, improving response times and privacy.

In a different healthcare domain, Liu et al. (2020) developed a **GoogLeNet-based food recognition system** on **edge smartphones**, achieving a 40% reduction in response time using a food image dataset. This highlights the role of edge devices in improving healthcare applications like dietary monitoring with minimal latency. Prabukumar et al. (2020) designed a **fog-based lung cancer diagnosis system** that used **fuzzy C-Means** and **SVM**, processing CT images on **edge devices** with 91% accuracy. This system demonstrates how fog computing can help analyze complex medical images while minimizing the need for heavy cloud resources. On the other hand, Hassan et al. (2021) implemented a **fog-cloud framework** for chronic disease monitoring, using **Naïve Bayes** and **firefly optimization**, tested on sensor data.

This hybrid system allows for real-time chronic disease monitoring while optimizing the network and processing efficiency. Key techniques like **CNNs**, **LSTMs**, and **FL** are commonly used in smart healthcare applications, and datasets include clinical and sensor data. The deployment architecture varies, with **edge-only** systems (Cheikhrouhou et al., 2021; Prabukumar et al., 2020) for real-time, local processing, and **hybrid edge-cloud** setups (Gupta et al., 2023; Liu et al., 2020) for leveraging cloud capabilities while maintaining edge privacy.

3.5 Smart Industry

Smart industry applications, including predictive maintenance, product quality monitoring, and fault detection, also leverage **Edge AI** to improve efficiency and minimize downtime. Zhang et al. (2024) used a **lightweight LSTM** model for machinery fault detection on **edge gateways**, achieving 94% accuracy with sensor data. This edge-based solution helps quickly identify faults in machinery, reducing downtime in industrial settings. Feng et al. (2020) proposed an **RF-Adaboost system** for predicting assembly quality on **edge servers**, tested on industrial data and achieving 90% accuracy. This approach optimizes the assembly process by enabling quick quality control assessments on the edge, reducing reliance on cloud processing. Park et al. (2020) developed an **LSTM-based fault detection model** fully deployed on **Raspberry Pi** (an edge device), achieving 92% accuracy on machine data. This emphasizes the power of edge devices in ensuring operational continuity in industrial environments. Li et al. (2020) introduced a **CNN-based inspection system** with a **fog-cloud collaboration** model, reducing data transmission by 50% while using factory datasets. This hybrid approach optimizes both local and remote processing, enhancing the efficiency of quality control and fault detection across industrial operations. In smart industry, **LSTM** and **CNNs** are prevalent, with **RF** often used for classification tasks. Datasets typically come from proprietary sensor data. **Edge placement** varies, with **full edge** deployments (Zhang et al., 2024; Park et al., 2020) for real-time fault detection and **hybrid fog-cloud** solutions (Feng et al., 2020; Li et al., 2020) for scalability and resource optimization.

3.6 Smart Education

Smart Education applications utilize **Edge AI** to enhance real-time educational experiences through student engagement monitoring, skill assessment, and personalized learning. Umarale et al. (2021) developed a **CNN-based attention detection system** on **edge PCs** using **DAiSEE** data, achieving 89% accuracy. This system helps in monitoring students' attention levels in real-time, promoting more effective teaching strategies. Sood and Singh (2020) created a **fog-based skill assessment framework** with **K-means** and **KNN**, achieving 96% accuracy on student data, which enables efficient evaluation of student performance in real time. Li et al. (2023) introduced a real-time emotional contagion intervention system with **CNN** on **edge devices**, using the **Fer2013** dataset and achieving 90% accuracy. This system tracks students' emotional states, offering personalized interventions to enhance engagement and well-being in the classroom. Additionally, Ahanger et al. (2020) implemented an **ANFIS-based quality assessment framework** on **Raspberry Pi**, tested on environmental datasets, offering a flexible and low-cost solution for educational assessments. Common techniques like **CNNs**, **K-means**, and **ANFIS** are employed, and datasets such as **DAiSEE** and **Fer2013** are commonly used. **Edge-only** deployments (Umarale et al., 2021; Li et al., 2023) are preferred for low-latency applications, while **fog-based** deployments (Sood & Singh, 2020) are used to enhance scalability and resource management.

3.7 Smart Transportation

Smart Transportation applications, including traffic flow prediction, smart parking, and driver behavior monitoring, benefit from **Edge AI** for real-time optimization. Li et al. (2023) proposed an **FL-LSTM model** for **traffic flow prediction**, achieving 91% accuracy on urban data with an **edge-cloud** deployment. This system helps optimize traffic management by processing data locally and in the cloud, ensuring rapid and accurate predictions. Ke et al. (2021) developed an **SSD-based parking surveillance system** on **edge IoT devices**, reducing latency by 45% when analyzing camera data, which helps in managing parking space availability in real-time.

Huang et al. (2021) introduced **FedParking**, a federated learning system with **LSTM**, tested on parking data. This system enhances privacy and achieved 88% accuracy, ensuring secure and efficient parking management. Xu et al. (2020) implemented a **DCNN** for **driver distraction detection** on **edge-cloud** infrastructure, achieving 93% accuracy with vehicle data, thereby improving road safety by monitoring drivers' attentiveness.

LSTM and **CNNs** are widely used in smart transportation, with datasets sourced from urban and vehicle data. **Hybrid edge-cloud** deployments (Li et al., 2023; Xu et al., 2020) offer scalability and reduced latency, while **full edge** deployments (Ke et al., 2021) are preferred for low-latency, real-time applications.

3.8 Security and Privacy

Security and Privacy applications leverage **Edge AI** to protect IoT systems from threats such as intrusion detection, cyberattacks, and preserving privacy. Kumar and Sharma (2023) designed a **blockchain-DNN-based intrusion detection system** on **edge nodes**, achieving 96% accuracy with **NSL-KDD** data. This approach ensures secure processing of sensitive data on the edge without relying heavily on the cloud. Samy et al. (2020) proposed an **LSTM-based fog framework** for **cyberattack detection**, achieving 94% accuracy on IoT data, improving security by enabling faster, more effective threat detection. Liao et al. (2020) developed a **DL-based authentication system** for **MEC (Mobile Edge Computing)**, tested on wireless data with 92% accuracy, providing secure authentication for IoT devices. Sharma et al. (2021) introduced a **differential privacy CNN framework** on **edge devices**, achieving 90% accuracy on IoT data, which enhances user privacy by anonymizing data before processing it. In security and privacy, **DNNs**, **LSTMs**, and **CNNs** are frequently used, with datasets like **NSL-KDD** for intrusion detection. **Edge-only** deployments (Kumar & Sharma, 2023; Sharma et al., 2021) are used for real-time, privacy-preserving tasks, while **fog-based** systems (Samy et al., 2020) are utilized to balance edge processing with cloud support for scalability and resource management.

4.0 Discussion

The literature review in Section 3 underscores the transformative potential of integrating artificial intelligence (AI) with edge computing termed Edge AI in Internet of Things (IoT) applications across eight domains: smart agriculture, environment, grid, healthcare, industry, education, transportation, and security. This section synthesizes these findings, offering insights into the relevance of AI-edge integration, the predominant AI technologies employed, the enabling technologies facilitating this synergy, and the performance metrics used to evaluate success.

4.1 Relevance of AI-Edge Integration

The integration of AI and edge computing significantly enhances the quality of service (QoS) and quality of experience (QoE) in IoT ecosystems, addressing critical needs for low latency, privacy, and resource efficiency. Across all reviewed domains, Edge AI reduces latency by processing data locally, a necessity for time-sensitive applications like healthcare's ECG anomaly detection (Gupta et al., 2023) and transportation's traffic flow prediction (Li et al., 2023). For instance, Cheikhrouhou et al. (2021) achieved millisecond-level arrhythmia detection by deploying a 1D CNN on edge nodes, demonstrating how latency reduction directly improves patient outcomes. Similarly, in smart agriculture, Singh and Patel (2024) reported a 50% reduction in response time for crop disease detection, enhancing farmer decision-making. Privacy preservation is another key benefit, particularly in healthcare and security domains. Federated learning (FL), as seen in Gupta et al. (2023) and Kumar and Sharma (2023), enables model training across distributed edge devices without sharing sensitive data, aligning with privacy regulations like GDPR (Bourechak et al., 2023). In smart grids, Taik and Cherkaoui (2020) highlighted how FL maintains consumer data confidentiality while forecasting load. This localized processing also reduces bandwidth usage, a critical advantage in bandwidth-constrained environments like smart environments (Putra et al., 2021). Collectively, these improvements in QoS low latency, enhanced privacy, and bandwidth efficiency—underscore the relevance of Edge AI, making it indispensable for next-generation IoT systems (Zhou et al., 2019).

4.2 AI Technologies Used

The review reveals a pronounced prevalence of deep learning (DL) techniques in Edge AI deployments, reflecting their capability to handle complex IoT data. Convolutional neural networks (CNNs) dominate image-based applications, such as crop disease detection (Singh & Patel, 2024), air quality monitoring (Wardana et al., 2021), and parking surveillance (Ke et al., 2021), due to their spatial feature extraction prowess (Goodfellow et al., 2016). Long short-term memory (LSTM) networks are widely adopted for time-series data, evident in soil moisture forecasting (Cordeiro et al., 2022), fault detection (Zhang et al., 2024), and traffic prediction (Li et al., 2023), leveraging their temporal dependency modeling (Hochreiter & Schmidhuber, 1997). Federated learning has emerged as a significant trend, particularly in privacy-sensitive domains like healthcare (Gupta et al., 2023) and transportation (Huang et al., 2021), enabling distributed training without centralized data aggregation (McMahan et al., 2017). Traditional machine learning (ML) techniques, such as Naïve Bayes (Rabie et al., 2020), support vector machines (Prabukumar et al., 2020), and ensemble methods (Jaiswal et al., 2020), remain relevant for resource-constrained edge devices, offering computational simplicity. Swarm intelligence, though less common, appears in education (Ahanger et al., 2020) for optimization tasks.

4.3 Enabling Technologies

Several enabling technologies underpin the successful deployment of Edge AI in IoT. **5G/6G networks** are pivotal, providing high-speed, low-latency connectivity essential for real-time applications. In transportation, Li et al. (2023) leveraged 5G to support federated LSTM models, while Yang et al. (2025) predict 6G's sub-millisecond latency will further enhance edge analytics by 2030 (Dang et al., 2020). **Lightweight models** and model compression techniques, such as quantization and pruning, enable DL on resource-constrained devices. Singh and Patel (2024) compressed CNNs for edge deployment, and Park et al. (2020) optimized LSTMs for Raspberry Pi, reducing computational overhead by 30%. **Federated learning** facilitates privacy and scalability, as seen in healthcare (Gupta et al., 2023) and security (Kumar & Sharma, 2023), by training models across distributed nodes (McMahan et al., 2017). **Fog computing** bridges edge and cloud, balancing resource demands in agriculture (Cordeiro et al., 2022) and industry (Li et al., 2020). **Hardware accelerators**, such as NVIDIA Jetson, enhance edge processing, supporting CNNs in education (Umarale et al., 2021) and transportation (Ke et al., 2021). **Blockchain** bolsters security, integrating with DNNs for intrusion detection (Kumar & Sharma, 2023).

4.4 Performance Metrics

Performance metrics in Edge AI applications vary by domain but consistently emphasize latency, accuracy, privacy, and efficiency. **Latency** is paramount in time-critical applications, with Cheikhrouhou et al. (2021) achieving sub-second ECG analysis and Ke et al. (2021) reducing parking detection time by 45%. **Accuracy** is a universal benchmark, ranging from 87% in environmental monitoring (Wardana et al., 2021) to 99% in healthcare (Cheikhrouhou et al., 2021), reflecting AI model effectiveness. **Privacy** is quantified indirectly through FL adoption (e.g., Gupta et al., 2023), ensuring data remains local.

Energy efficiency and **scalability** are critical in resource-constrained settings, with Cordeiro et al. (2022) reporting a 30% efficiency gain and Huang et al. (2021) demonstrating scalable parking predictions. **Downtime reduction** is key in industry, with Zhang et al. (2024) minimizing it by 25%.

4.5 Synthesis and Insights

The synthesis reveals Edge AI's broad applicability, with each domain leveraging specific AI techniques and technologies to address unique challenges. Healthcare and transportation prioritize low latency and privacy via DL and FL, supported by 5G/6G, while agriculture and industry focus on efficiency and accuracy with lightweight models and fog computing (Chang et al., 2021). Education and security balance real-time processing with scalability, using hybrid architectures and blockchain (Sood & Singh, 2020; Kumar & Sharma, 2023). The dominance of DL reflects its adaptability, though resource constraints drive innovations like FL and compression (Deng et al., 2020).

5. Challenges and Future Directions

The integration of artificial intelligence (AI) with edge computing in Internet of Things (IoT) applications—termed Edge AI has demonstrated significant potential across diverse domains, as outlined in Sections 3 and 4. However, several challenges impede its widespread adoption and optimal performance. This section identifies key gaps, including resource constraints, distributed computing complexities, computational demands of AI for big IoT data, and security/privacy concerns in edge nodes. It then proposes future directions to address these issues, such as advanced hardware, 6G integration, ethical AI frameworks, and scalable orchestration.

5.1 Challenges

5.1.1 Resource Constraints

Edge devices, such as sensors and microcontrollers, are inherently limited in computational power, memory, and energy, posing a significant barrier to deploying complex AI models. Deep learning (DL) techniques like convolutional neural networks (CNNs) and long short-term memory (LSTM) networks, prevalent in applications such as crop disease detection (Singh & Patel, 2024) and ECG anomaly detection (Gupta et al., 2023), require substantial resources. For instance, Park et al. (2020) noted that even lightweight LSTMs on Raspberry Pi consume significant power, reducing battery life by 20% in industrial settings. This constraint limits the scalability and longevity of Edge AI deployments, particularly in remote environments like smart agriculture (Cordeiro et al., 2022).

5.1.2 Distributed Computing

The distributed nature of edge computing introduces synchronization and orchestration challenges. In federated learning (FL) systems, as used in healthcare (Gupta et al., 2023) and transportation (Li et al., 2023), edge nodes must coordinate model updates across heterogeneous devices, often over unreliable networks. Samy et al. (2020) highlighted synchronization delays in fog-based intrusion detection, reducing detection speed by 15%. Orchestrating tasks across edge, fog, and cloud layers—evident in hybrid architectures (Li et al., 2020)—requires efficient load balancing and fault tolerance, which remain unresolved in dynamic IoT environments (Laroui et al., 2021).

5.1.3 Computational Complexity of AI for Big IoT Data

The computational complexity of AI models for processing big IoT data—characterized by volume, velocity, and variety—exceeds edge capabilities. Real-time analytics, such as air quality prediction (Wardana et al., 2021), demand high-performance CNN-LSTM models that strain edge resources, leading to accuracy trade-offs (Bourechak et al., 2023). Retraining DL models on edge nodes, as needed in smart grids (Taik & Cherkaoui, 2020), is impractical due to limited memory, exacerbating the gap between cloud-based training and edge inference (Deng et al., 2020).

5.1.4 Security and Privacy in Edge Nodes

Edge nodes, positioned closer to end-users, are more vulnerable to physical and cyber threats than centralized servers. Kumar and Sharma (2023) noted that edge-based intrusion detection systems face risks like node tampering, reducing security by 10% compared to cloud setups. Privacy concerns arise despite FL's benefits, as model updates can leak indirect data (Sharma et al., 2021). In healthcare, ensuring compliance with regulations like HIPAA remains challenging when processing sensitive data at the edge (Gupta et al., 2023).

5.2 Future Directions

5.2.1 Advanced Hardware

Next-generation hardware, such as AI accelerators (e.g., Google Coral, NVIDIA Jetson), can mitigate resource constraints by enhancing edge processing capabilities. Umarale et al. (2021) demonstrated a 30% performance boost in attention detection using accelerators, suggesting their potential for DL tasks. Quantum edge computing, though nascent, promises exponential compute gains by 2030, potentially revolutionizing applications like real-time traffic prediction (Li et al., 2023) with minimal energy use (Yang et al., 2025).

5.2.2 6G Integration for Ultralow Latency

The transition to 6G networks, expected by 2030, will address latency and connectivity challenges with sub-millisecond response times and terabit-per-second speeds (Dang et al., 2020). Yang et al. (2025) predict 6G will enhance Edge AI in transportation and healthcare, enabling tactile internet applications like remote surgery with latency below 1 ms. Network slicing and integrated sensing will further optimize bandwidth, supporting synchronized distributed computing (Saeik et al., 2021).

5.2.3 Ethical AI at the Edge

As Edge AI proliferates, ethical concerns—bias, fairness, and transparency—must be addressed. In education, biased models could misjudge student engagement (Sood & Singh, 2020), while in security, flawed intrusion detection could disproportionately flag benign activities (Kumar & Sharma, 2023). Developing ethical frameworks, such as fairness-aware algorithms and explainable AI, will ensure equitable outcomes, a gap noted by Chang et al. (2021) requiring urgent research.

5.2.4 Scalable Orchestration Frameworks

Scalable orchestration frameworks are critical to overcome distributed computing challenges. Dynamic autoscaling, as explored by Babar and Khan (2021), could adapt resource allocation across millions of edge nodes, improving QoS in smart grids (Rabie et al., 2020). Peer-to-peer FL architectures, eliminating single-point failures (Taik & Cherkaoui, 2020), and AI-driven orchestration, optimizing task scheduling (Saeik et al., 2021), promise robust, scalable Edge AI systems.

6. Conclusion

This review has explored the intersection of AI and edge computing in IoT applications, synthesizing insights from eight domains: smart agriculture, environment, grid, healthcare, industry, education, transportation, and security. Key findings highlight Edge AI's ability to improve QoS through low latency, privacy preservation, and resource efficiency, driven by deep learning (e.g., CNNs, LSTMs), federated learning, and enabling technologies like

5G/6G and lightweight models. Performance metrics such as accuracy and latency underscore its effectiveness, though challenges—resource constraints, distributed computing complexities, computational demands, and security/privacy persist.

The synergy of AI and edge computing is transformative, enabling real-time, intelligent IoT systems that enhance user experiences across critical applications, from healthcare diagnostics to urban mobility. However, realizing its full potential requires overcoming identified gaps. Future research should prioritize advanced hardware, 6G integration, ethical AI frameworks, and scalable orchestration to address these limitations.

REFERENCES

Akintayo, T., Olusola, R., Enabulele, E., Oyesanya, A., Olanrewaju, S., Celestina, M., Sulaimon, B., Anoliefo, O., Olumide, A., Ridwan, O., & Adediran, A. (2024). IoT Revolutionized: How Machine Learning is Transforming Data, Applications, and Industries. *Path of Science, 10*(6), 8001-8007. doi:http://dx.doi.org/10.22178/pos.105-30

Ahanger, T. A., Tariq, U., Ibrahim, A., Ullah, I., & Bouteraa, Y. (2020). ANFIS-inspired smart framework for education quality assessment. *IEEE Access, 8*, 175306–175318.

Akintayo, T., Okiemute, R., Owoeye, M., Balogun, O., Paul, C., Queenet, M., Okereke, R., Moluno, R., Adediran, A., Nzeanorue, C., Ngozi, E., & Madukwe, C. (2024). Identifying Internet of Things Devices through Unique Digital Signa-tures and Advanced Machine Learning Techniques. *Path of Science, 10*(7), 1001-1007. doi:http://dx.doi.org/10.22178/pos.106-5

Babar, M., & Khan, M. S. (2021). ScalEdge: A framework for scalable edge computing in Internet of things-based smart systems. *International Journal of Distributed Sensor Networks, 17*(7), 15501477211035332.

Bourechak, A., Zedadra, O., Kouahla, M. N., Guerrieri, A., Seridi, H., & Fortino, G. (2023). At the confluence of artificial intelligence and edge computing in IoT-based applications: A review and new perspectives. *Sensors, 23*(3), 1639.

Chang, Z., Liu, S., Xiong, X., Cai, Z., & Tu, G. (2021). A survey of recent advances in edge-computing-powered artificial intelligence of things. *IEEE Internet of Things Journal, 8*(17), 13849–13875.

Cheikhrouhou, O. (2021). One-dimensional CNN approach for ECG arrhythmia analysis in fog-cloud environments. *IEEE Access, 9*, 103513–103523.

Cordeiro, M., Markert, C., Araújo, S. S., Campos, N. G., Gondim, R. S., da Silva, T. L. C., & da Rocha, A. R. (2022). Towards smart farming: Fog-enabled intelligent irrigation system using deep neural networks. *Future Generation Computer Systems, 129*, 115–124.

Dang, S., Amin, O., Shihada, B., & Alouini, M.-S. (2020). What should 6G be? *Nature Electronics, 3*(1), 20–29.

De Stefano, C., Ferrigno, L., Fontanella, F., Genevini, L., & Scotto di Freca, A. (2020). A novel PCA-based approach for building on-board sensor classifiers for water contaminant detection. *Pattern Recognition Letters, 135*, 375–381.

Deng, S., Zhao, H., Fang, W., & Yin, J. (2020). Edge intelligence: The confluence of edge computing and artificial intelligence. *IEEE Internet of Things Journal, 7*(8), 7457–7469.

Feng, Y., Wang, T., Hu, B., Yang, C., & Tan, J. (2020). An integrated method for high-dimensional imbalanced assembly quality prediction supported by edge computing. *IEEE Access, 8*, 71279–71290.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems, 29*(7), 1645–1660.

Guillén, M. A., Llanes, A., Imbernón, B., Martínez-España, R., Bueno-Crespo, A., Cano, J. C., & Cecilia, J. M. (2021). Performance evaluation of edge-computing platforms for the prediction of low temperatures in agriculture using deep learning. *The Journal of Supercomputing, 77*(1), 818–840.

Gupta, R., Sharma, S., & Kumar, P. (2023). Federated learning for real-time ECG anomaly detection in edge-based healthcare systems. *IEEE Transactions on Biomedical Engineering, 70*(5), 1456–1465.

Hassan, M. K., El Desouky, A. I., Mahmoud, D., Amany, M. B., & Mohamed, M. S. (2021). EoT-driven hybrid ambient assisted living framework with naïve Bayes–firefly algorithm. *Neural Computing and Applications, 31*(4), 1275–1300.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780.

Huang, X., Li, P., Yu, R., Yu, Y., Xie, K., & Xie, S. (2021). FedParking: A federated learning based parking space estimation with parked vehicle assisted edge computing. *IEEE Transactions on Vehicular Technology, 70*(9), 9355–9368.

Jaiswal, R., Chakravorty, A., & Rong, C. (2020). Distributed fog computing architecture for real-time anomaly detection in smart meter data. *Proceedings of the IEEE Sixth International Conference on Big Data Computing Service and Applications (BigDataService)*, 1–8.

Ke, R., Zhuang, Y., Pu, Z., & Wang, Y. (2021). A smart, efficient, and reliable parking surveillance system with edge artificial intelligence on IoT devices. *IEEE Transactions on Intelligent Transportation Systems, 22*(8), 4962–4974.

Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of ICNN'95 - International Conference on Neural Networks*, 4, 1942–1948.

Kumar, A., & Sharma, R. (2023). Blockchain-integrated deep neural networks for intrusion detection in edge IoT networks. *Journal of Network and Computer Applications, 210*, 103567.

Laroui, M., Nour, B., Moungla, H., Cherif, M. A., Afifi, H., & Guizani, M. (2021). Edge and fog computing for IoT: A survey on current research activities & future directions. *Computer Communications, 180*, 210–231.

Lee, K., Silva, B. N., & Han, K. (2020). Deep learning entrusted to fog nodes (DLEFN) based smart agriculture. *Applied Sciences, 10*(4), 1544.

Li, L., Ota, K., & Dong, M. (2020). Deep learning for smart industry: Efficient manufacture inspection system with fog computing. *IEEE Transactions on Industrial Informatics, 14*(10), 4665–4673.

Li, T., Fong, S., Li, X., Lu, Z., & Gandomi, A. H. (2024). Swarm decision table and ensemble search methods in fog computing environment: Case of day-ahead prediction of building energy demands using IoT sensors. *IEEE Internet of Things Journal, 7*(3), 2321–2342.

Li, Y., Zhang, H., & Wang, X. (2023). Federated LSTM for privacy-preserving traffic flow prediction in smart cities. *IEEE Internet of Things Journal, 10*(8), 7123–7134.

Liao, R. F., Wen, H., Wu, J., Pan, F., Xu, A., Song, H., Xie, F., Jiang, Y., & Cao, M. (2020). Security enhancement for mobile edge computing through physical layer authentication. *IEEE Access, 8*, 116390–116401.

Liu, C., Cao, Y., Luo, Y., & Chen, G. (2020). A new deep learning-based food recognition system for dietary assessment on an edge computing service infrastructure. *IEEE Transactions on Services Computing, 11*(2), 249–261.

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.

Mohammadi, M., Al-Fuqaha, A., Sorour, S., & Guizani, M. (2018). Deep learning for IoT big data and streaming analytics: A survey. *IEEE Communications Surveys & Tutorials, 20*(4), 2923–2960.

Park, D., Kim, S., An, Y., & Jung, J. Y. (2020). LiReD: A light-weight real-time fault detection system for edge computing using LSTM recurrent neural networks. *Sensors, 18*(7), 2110.

Prabukumar, M., Agilandeeswari, L., & Ganesan, K. (2020). An intelligent lung cancer diagnosis system using cuckoo search optimization and support vector machine classifier. *Journal of Ambient Intelligence and Humanized Computing, 10*(1), 267–293.

Putra, K. T., Chen, H.-C., Prayitno, Ogiela, M. R., Chou, C.-L., Wang, C.-E., & Shae, Z.-Y. (2021). Federated compressed learning edge computing framework with ensuring data privacy for PM2.5 prediction in smart city sensing applications. *Sensors, 21*(14), 4586.

Rabie, A. H., Ali, S. H., Saleh, A. I., & Ali, H. A. (2020). A fog based load forecasting strategy based on multi-ensemble classification for smart grids. *Journal of Ambient Intelligence and Humanized Computing, 11*(1), 209–236.

Ren, J., Zhang, D., He, S., Zhang, Y., & Li, T. (2019). A survey on end-edge-cloud orchestrated network computing paradigms: Transparent computing, mobile edge computing, fog computing, and cloudlet. *ACM Computing Surveys, 52*(6), 1–36.

Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Saeik, F., Avgeris, M., Spatharakis, D., Santi, N., Dechouniotis, D., Violos, J., Leivadeas, A., Athanasopoulos, N., Mitton, N., & Papavassiliou, S. (2021). Task offloading in edge and cloud computing: A survey on mathematical, artificial intelligence and control theory solutions. *Computer Networks, 195*, 108177.

Samy, A. (2020). Fog-based attack detection framework for Internet of Things using deep learning. *IEEE Access, 8*, 74571–74585.

Sharma, J., Kim, D., Lee, A., & Seo, D. (2021). On differential privacy-based framework for enhancing user data privacy in mobile edge computing environment. *IEEE Access, 9*, 38107–38118.

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.

Singh, V., & Patel, N. (2024). CNN-based crop disease detection using edge computing for precision agriculture. *Computers and Electronics in Agriculture, 218*, 108745.

Sood, S. K., & Singh, K. D. (2020). Optical fog-assisted smart learning framework to enhance students' employability in engineering education. *Computer Applications in Engineering Education, 27*(5), 1030–1042.

Statista. (2024). *Number of Internet of Things (IoT) connected devices worldwide from 2019 to 2030*. Retrieved from

Taik, A., & Cherkaoui, S. (2020). Electrical load forecasting using edge computing and federated learning. *Proceedings of the ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 1–6.

Umarale, D., Sodhani, S., Akhelikar, A., & Koehy, R. (2021). Attention detection of participants during digital learning sessions using edge computing. *SSRN Electronic Journal*, 575–584.

Wardana, I. N. K., Gardner, J. W., & Fahmy, S. A. (2021). Optimising deep learning at the edge for accurate hourly air quality prediction. *Sensors, 21*(4), 1064.

Xu, T., Han, G., Qi, X., Du, J., Lin, C., & Shu, L. (2020). A hybrid machine learning model for demand prediction of edge-computing-based bike-sharing system using Internet of Things. *IEEE Internet of Things Journal, 7*(8), 7345–7356.

Yang, J., Wen, J., Wang, Y., Jiang, B., Wang, H., & Song, H. (2020). Fog-based marine environmental information monitoring toward ocean of things. *IEEE Internet of Things Journal, 7*(5), 4238–4247.

Yang, P., Zhang, L., Chen, X., & Wang, H. (2025). 6G networks: Enabling AI-driven edge computing for IoT applications. *IEEE Communications Magazine, 63*(2), 45–52.

Zhang, L., Chen, Q., & Liu, J. (2024). Lightweight LSTM for predictive maintenance in industrial IoT: An edge computing approach. *IEEE Transactions on Industrial Informatics, 20*(3), 2345–2356.

Zhang, R., & Li, X. (2021). Edge computing driven data sensing strategy in the entire crop lifecycle for smart agriculture. *Sensors, 21*(22), 7502.

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),