AI at the Edge: Cloud-Edge Synergy

Kishan Raj Bellala[1]

[1]Independent Researcher, Austin, Texas, U.S.A.

ORCID ID:
https://orcid.org/0009-0007-2327-0993

Publication Date: 2025/05/31

Abstract: The new paradigm of AI at the Edge and its synergy with Cloud Computing to combine the advantages of local data processing and analysis with the scalability and resources offered by cloud systems is explained in this paper. AI capabilities deployed at the Edge enable real-time decision-making, reduced latency, and improved efficiency across various applications including healthcare, smart cities, industrial automation, and autonomous vehicles. Organizations can achieve maximum computational power and network bandwidth optimization and system performance enhancement through the combined strengths of Edge Computing and Cloud Computing. The advancement comes with security challenges and data privacy risks and requirements for effortless Edge-Cloud system integration. This paper conducts a thorough analysis of AI at the Edge with Cloud-Edge synergy use cases and advantages and limitations and future trends to explain the transformative potential of this relationship in artificial intelligence and computing.

Keywords: *Edge Computing, Cloud Computing, AI Applications, Real-Time Processing, Data Privacy, and System Performance, Shedding Light on the Transformative Potential of this Symbiotic Relationship.*

How to Cite: Kishan Raj Bellala. (2025). AI at the Edge: Cloud-Edge Synergy. *International Journal of Innovative Science and Research Technology*, 10(5), 2561-2567. https://doi.org/10.38124/ijisrt/25may967.

I. INTRODUCTION

The development of Artificial Intelligence (AI) has accelerated during recent years to enable machines to replicate human cognitive abilities while performing complex tasks with high precision (Manduva, 2021). AI algorithms needed centralized cloud servers for processing large data volumes until Edge Computing emerged as a solution which reduced latency and network bandwidth requirements. AI at the Edge represents a fundamental change in data processing practices because it moves computations to local positions near data origins which results in faster responses and decreased cloud dependency (Manduva, 2021). AI algorithms deployed on edge devices such as sensors and cameras and IoT devices enable real-time data analysis at the source through the implementation of AI at the Edge. The implementation of this method results in multiple benefits which include faster data processing times and better data privacy measures and improved system reliability and reduced network costs (Manduva, 2021). The main advantage of Cloud Computing includes its ability to scale resources and store data while providing powerful computation for handling demanding AI tasks (Manduva, 2021).

The combination of AI at the Edge with Cloud Computing enables organizations to harness both system strengths for maximum computational efficiency and system performance (Manduva, 2021). Organizations use Edge devices for fast data processing alongside cloud servers for extensive analysis of computationally intensive tasks to reach real-time responsiveness and cloud scalability (Manduva, 2021). This paper examines AI at the Edge alongside its cooperative potential with Cloud Computing to develop new solutions for different applications. The paper examines the applications, benefits, obstacles and future outlook of this combined system while demonstrating how AI at the Edge gains transformative value through Cloud-Edge synergy (Manduva, 2021).

II. OVERVIEW OF EDGE COMPUTING AND CLOUD COMPUTING

The modern computing landscape includes Edge Computing and Cloud Computing as separate paradigms which provide different benefits for distinct use cases (Sun, 2023). Edge Computing handles data processing near its source through local devices or edge devices positioned at network peripheries. The distribution system reduces latency because it allows decision-making and live data processing directly at the data collection points (Sun, 2023) (Manduva, 2021). Technology finds its best application in systems that need fast responses such as industrial automated systems and automated vehicles IoT systems and live monitoring systems. Cloud Computing provides on-demand computing resources such as servers and networking and storage and software and databases through internet access with pay-as-you-go pricing (Manduva, 2021). The cloud services platform provides scalability and high availability and cost efficiency which makes them suitable for handling large data volumes and running complex applications and enabling team collaboration across different locations (Sun, 2023). Organizations use a hybrid architecture that combines Edge Computing with Cloud Computing to access benefits from both paradigms. Businesses can optimize system performance and resource utilization by using cloud resources to extend the capabilities of edge devices which enables localized data processing and scalable cloud infrastructure (Sun, 2023). The hybrid approach gains special importance in AI at the Edge because it allows real-time inference on edge devices while using cloud servers for training and data analytics and resource-intensive tasks (Sun, 2023).

A. Edge AI

The distribution of artificial intelligence algorithms through Edge AI enables end-edge devices to make instant decisions while depending less on cloud resources. Live execution of different AI tasks like path planning, NLP (natural language processing), data anomaly detection, and image recognition can be achieved via this approach (Sathupadi, 2024). Because of this there will be an enhancement of operational efficiency and a decrease in latency. The advancement of Edge AI technology proves helpful because of the massive data or information processing requirements from IoT systems and connected devices (Sathupadi, 2024). AI applications used cloud infrastructure for both training and inference model operations before this time. The reason is that large amounts of data are collected and sent to robust cloud servers to train and deploy machine learning models require huge amounts of data that needs to be sent to robust cloud servers for training and deployment (Sathupadi, 2024). The process of sending large raw data objects to the cloud for real-time inference results in latency issues which make it inappropriate for time-sensitive applications. Self-driving vehicles need immediate decisions which makes cloud processing dangerous because it would lead to delayed responses (Sathupadi, 2024). To analyze data in real world applications of Edge AI are used. The need for edge-based decision making is the main factor in accepting edge AI. For example, in a factory setting, edge devices built in with AI technology can identify system failures and automatically take corrective actions to replace the machine before failure occurs, preventing huge financial losses (Sathupadi, 2024). For instance, in manufacturing space pre-integrated edge devices with AI technology can identify operational machinery faults and automatically take corrective actions by replacing machines before breakdowns occur and prevent massive financial losses (Sathupadi, 2024). Wearables equipped with AI technology provide healthcare monitoring functions that detect abnormal patterns in patient vital signs which get transmitted to a central server without needing data return or RN and medical team notification for urgent situations (Sathupadi, 2024). The device operates without requiring sails but uses them to transfer data back to a main server while AI tracks patient vital signs. These devices function autonomously in most situations where network coverage is limited or unreliable thus making them valuable for such regions (Chennupati, 2025).

The deployment of AI algorithms at edge nodes relies on their cautiously tailored work on small, less resourceful systems. The majority of available edge devices operate with limited processing power and memory capacity and energy resources which are inferior to typical cloud servers. Researchers and engineers resolved these problems by using model pruning and quantization and knowledge distillation methods to enhance AI model performance with minimal impact on speed and other dimensions (Chennupati, 2025). Optimization enables edge devices to execute intelligent computational operations with reduced power usage that is essential for specific battery-operated applications including drones and digital sensors (Chennupati, 2025) (Sathupadi, 2024).

The implementation of complex artificial intelligence models in edge devices remains challenging due to several issues (Sun, 2023). The application of high artificial neural network intelligence to device environments remains a challenge for edge AI systems. The basic processing capacity of edge devices remains lower than cloud devices thus restricting the size of deployable models. The deployment of various AI models across numerous edge devices requires efficient and consistent management systems (Sathupadi, 2024). The deployment becomes difficult when numerous edge devices operate independently in large numbers. Edge AI is the major advancement in the implementation of AI capability that becomes possible through Edge AI because it enables real-time operations with minimal delay and cloud-independent functionality (Chennupati, 2025). Technology enables completely new possibilities including self-driving cars and smart cities as well as industrial uses and emerging new technologies. The hardware development will continue alongside increasing model efficiency which together will boost Edge AI development into an advanced decentralized processing system (Chennupati, 2025).

B. Cloud Computing's Function in AI and Edge Systems

Cloud computing functions as a vital component for AI and edge systems because it delivers elastic infrastructure alongside high-performance computation to support complex AI algorithms (Zou et al., 2019). The system enables the deployment of extensive AI models through its available resources which serve applications needing large data processing and storage functionality (Banerjee, 2024; Rane et al., 2024). The combination of cloud computing technology with edge systems brings forth both potential advantages and technical difficulties. The capabilities of cloud computing do not align well with applications that need immediate responses or fast data processing (Rane et al., 2024). Edge computing solves these problems through data source proximity which enhances both energy efficiency and security as well as reducing latency (Zou et al., 2019). The development of hybrid cloud-edge architectures together with fog computing paradigms emerged to support edge processing that maintains cloud connectivity (Rane et al., 2024; Zou et al., 2019). The function of cloud computing in edge systems and AI transforms into a collaborative model. Intelligent Cooperative Edge (ICE) computing shows how AI core functions can be distributed from cloud to edge nodes to support privacy-preserving transfer learning and incremental model updates (Gong et al., 2020). AI technologies together with cloud and edge computing integration create new possibilities for efficient IoT applications and autonomous systems and human-machine interactions (Rong et al., 2021; Singh & Gill, 2023; Zou eet al. 2019).

C. Cloud and Edge Computing Working Together in AI Applications

Real-time data processing across industries undergoes revolutionary transformation through the combined power of Artificial Intelligence (AI) and Edge Computing and Cloud Computing. AI provides intelligent decision-making capabilities while Edge Computing enables low-latency localized processing and Cloud Computing delivers scalable infrastructure for storage and training and orchestration (Reddy, 2023). The combined system creates an efficient dynamic system that handles large data volumes and immediate insights to enhance real-time application effectiveness and efficiency. The central function of Artificial Intelligence within this ecosystem enables systems to learn from data while detecting patterns and making autonomous decisions (Reddy, 2023). Deep learning algorithms require large datasets for training which are usually stored in cloud infrastructure because of their ability to handle computations and store information (Reddy, 2023). The deployment of trained models to edge devices enables local real-time inference operations (Chennupati, 2025). The proximity of AI-driven decision-making to data sources through this method reduces information processing time which is essential for applications such as autonomous driving and industrial robotics and real-time health monitoring (Reddy, 2023).

The "edge" of the network receives computational tasks from AI through Edge Computing which mitigate latency by handling data closer to its source (Reddy, 2023). The data processing distance reduction through this shift enables substantial latency improvement since data no longer needs to travel to distant servers (Reddy, 2023). The network bandwidth usage decreases while data privacy improves because the system restricts the amount of sensitive information that needs external transmission. Smart cameras and industrial sensors along with wearables can analyze data directly on-site through AI models which were trained in cloud infrastructure. The system provides real-time responses through its ability to stop machinery during detected malfunctions and to notify emergency services about health anomalies (Reddy, 2023). The architecture depends on Cloud Computing to manage data collection and model training and system-wide analysis and long-term storage operations (Reddy, 2023). AI model training in the cloud benefits from diverse large-scale datasets through centralized processing. The cloud system functions as the central coordination point which distributes updates to edge devices and handles workload management and advanced analytics that edge nodes lack capability to perform (Reddy, 2023). The cloud enables continuous learning because it collects performance data and feedback from edge devices to retrain models and enhance their accuracy. The continuous data exchange between cloud and edge infrastructure allows AI systems to improve their performance and responsiveness through time (Reddy, 2023).

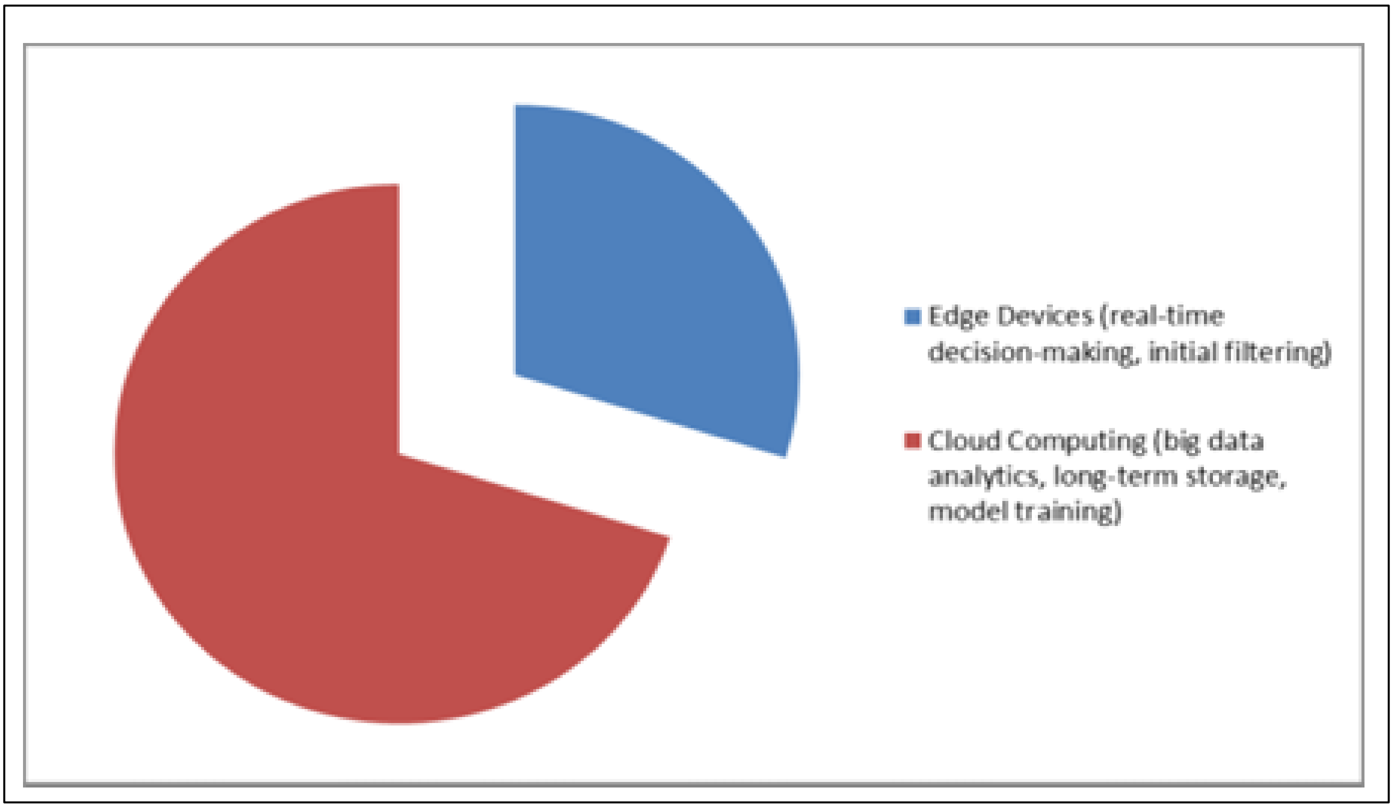

Fig 1: Workload Distribution in a Hybrid System between Edge and Cloud (Reddy, 2023).

These three technologies work together to enable various practical applications. The combination of edge AI technology in smart manufacturing enables real-time defect detection but cloud systems perform inventory management and production analytics functions (Reddy, 2023). Wearable devices perform immediate patient data analysis at the edge for alerts while the cloud maintains historical data storage to improve diagnostic accuracy. Autonomous vehicles use edge AI for navigation and real-time decision-making while cloud platforms analyze aggregated driving data to enhance system-wide improvements in the automotive industry (Reddy, 2023). The retail sector benefits from edge systems that provide individualized customer experiences while cloud analytics works to enhance supply chain operations and understand customer behavior patterns. AI together with edge computing and cloud infrastructure creates an adaptive and scalable system that generates ultra-low-latency responses (Chennupati, 2025) (Reddy, 2023). The convergence of these technologies brings multiple benefits to operations because it moves cloud-based work to edge computing reduces costs and improves system dependability through offline capabilities and protects privacy by reducing data exchange (Reddy, 2023). The growing presence of 5G networks together with Internet of Things expansion will deepen the integration of AI systems with edge computing and cloud infrastructure. The upcoming era will introduce highly decentralized systems that combine learning capabilities with real-time adaptation and action to extend current automation and connectivity and intelligence capabilities (Reddy, 2023).

III. APPLICATIONS IN REAL WORLD AND CASE STUDIES

Real-time data handling through AI combined with edge computing and cloud technologies has transformed industries which need instant decisions and performance optimization (Reddy, 2023). AI systems deployed at the edge with cloud connectivity enable new possibilities for healthcare systems and smart cities and IIoT applications and automated self-driving cars (Reddy, 2023). AI systems together with edge computing enable smart cities to manage their traffic systems and energy distribution and security needs. Smart traffic control systems employ cameras and sensors to process real-time traffic data through edge devices where AI algorithms make decisions to manage traffic lights and avoid congestion (Manduva, 2021). The cloud infrastructure maintains storage of permanent analysis data which helps plan future infrastructure development (Sun, 2023). The combination of edge AI traffic lights in Barcelona optimizes road efficiency and decreases pollution while cloud systems evaluate present-day traffic conditions to plan future urban progress. Edge AI technology has brought significant changes to health vital signs monitoring practices in healthcare settings (Manduva, 2021). Smart medical sensors and AI-powered smartwatches monitor patient health conditions instantly while cloud systems operate simultaneously to provide healthcare professionals with immediate notification access. Medtronic's insulin pump employs AI to measure glucose spike levels and perform dose administration directly from the edge (Reddy, 2023).

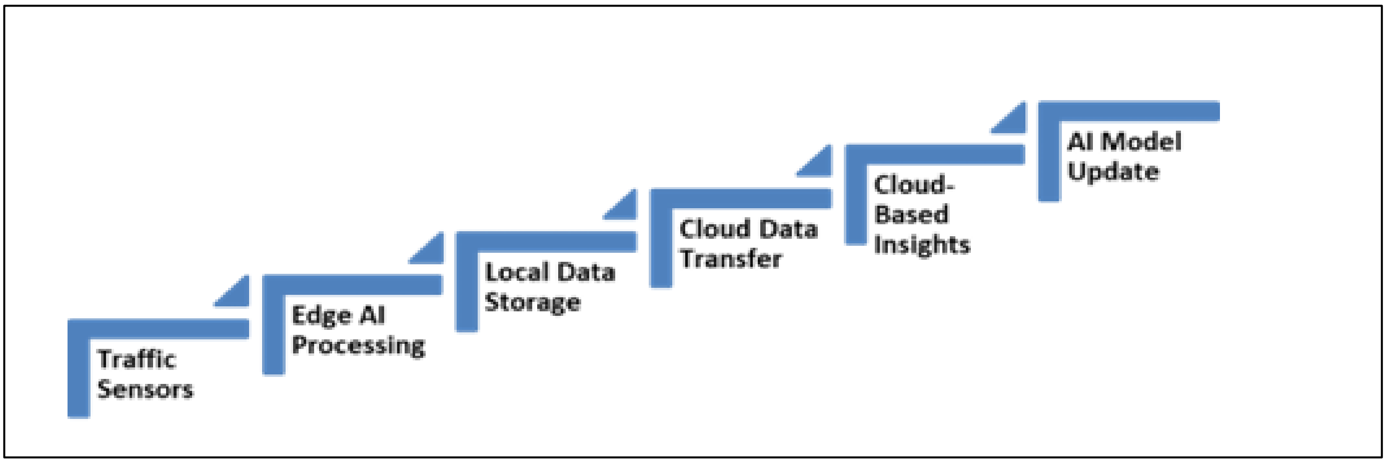

Fig 2: Smart City Data Processing Workflow Using Edge AI (Reddy, 2023).

The combination of AI with edge and cloud technologies creates major effects on predictive maintenance operations in IIoT (Reddy, 2023). Industrial settings need machine and sensor data interactions to track equipment status and manufacturing issues and performance improvements. Real-time data from edge AI systems enables systems to detect equipment failures through vibration and temperature monitoring thus minimizing operational time and repair duration (Reddy, 2023) (Manduva, 2021). Siemens Mind Sphere uses edge computing to analyze industrial machine performance in real-time. The cloud storage system maintains information for trend analysis which enables predictive maintenance to help companies schedule production and reduce maintenance costs. The model enhances machine operation through cloud data analysis for improvement purposes (Manduva, 2021). Self-driving cars demonstrate the importance of AI, edge computing, and cloud integration. The system requires sensor and camera and radar data processing to generate instant decisions regarding route changes and danger responses and speed management (Reddy, 2023). Edge AI functions as a critical element because delayed decisions can result in fatal outcomes. The Autopilot system of Tesla depends on edge AI to analyze camera and radar information for driving functions and obstacle prevention (Reddy, 2023). The cloud system collects data from all Tesla vehicles to update AI models through collective road experiences which enhances safety and efficiency for the entire fleet over time (Manduva, 2021).

IV. EVALUATION OF EDGE-CLOUD TECHNICAL ISSUES AND THE EFFICIENCY OF THEIR MITIGATION

Figure.3 demonstrates how different mitigation strategies perform to address technical challenges in edge-cloud computing (Chennupati, 2025). The results demonstrate that edge feature extraction with data reduction (97%) and differential privacy protection (95%) achieve high effectiveness. The results for 5G network latency (80%) and hardware-based security (74%) demonstrate strong performance. Predictive latency management (44%) and task partitioning latency reduction (47%) demonstrate limited success in their respective areas. The current systems face major challenges in federated learning accuracy (5%) and authentication vulnerabilities (0%) which need substantial development to address these issues (Chennupati, 2025).

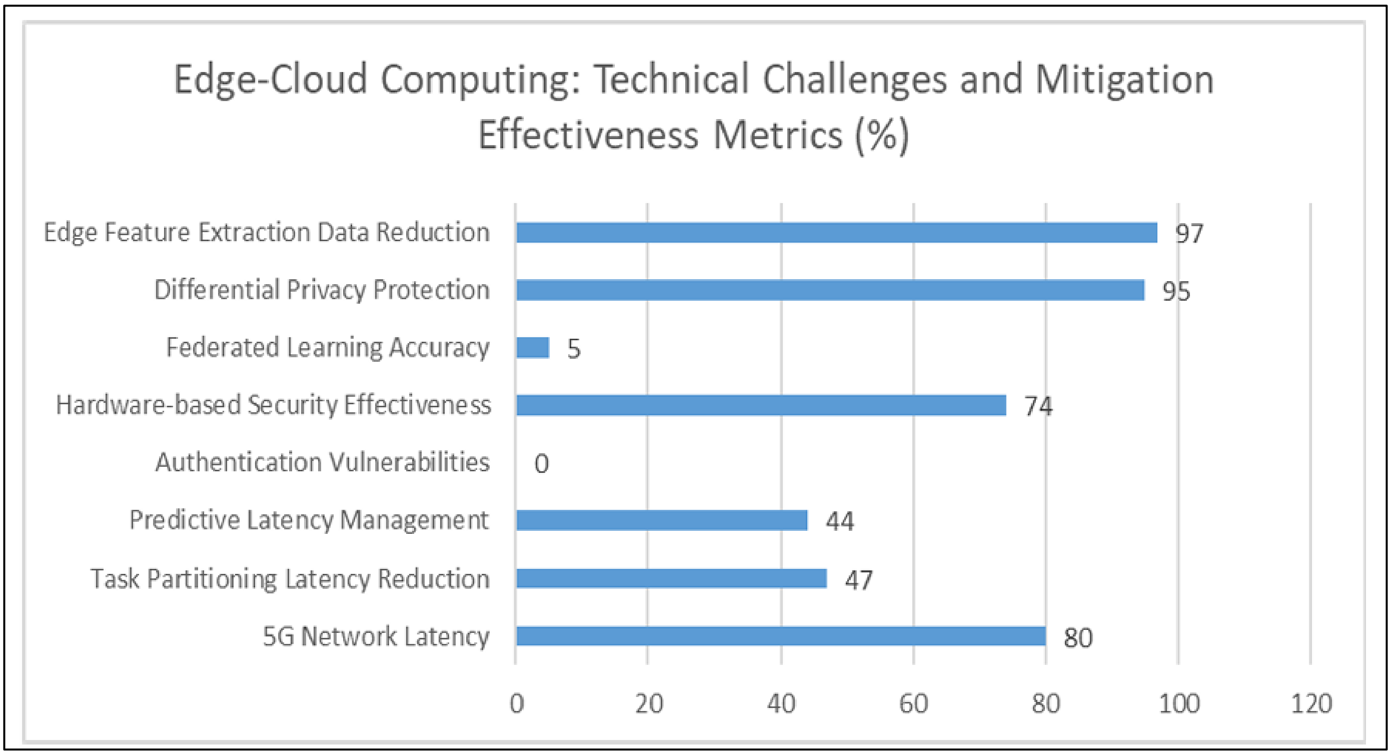

Fig 3: Technical Issues with Edge-Cloud Computing and the Efficiency of Mitigation (Chennupati, 2025).

V. CHALLENGES AND CONSIDERATIONS IN INTEGRATING AI AT THE EDGE WITH CLOUD SYSTEMS

The combination of AI at the edge with cloud systems creates multiple technical issues which must be solved to achieve successful deployment. The main difficulty stems from restricted edge computational power which faces challenges when running modern Deep Neural Networks (DNN) (Torres et al., 2021). The restricted capabilities require developers to create efficient AI systems and hardware solutions which function on edge devices. Researchers have developed Edge-Cloud architecture with Branchy Net capabilities to create fault-tolerant and low-latency AI predictions (Torres et al., 2021). The advancement of specialized hardware for edge machine learning systems from embedded systems to sub-mW "always-on" IoT nodes represents a crucial element for achieving energy-efficient processing at the edge (Zou et al., 2019). Security and privacy issues stand as major obstacles during the integration of edge-cloud AI systems. The distributed structure of edge computing creates security vulnerabilities because of its limited resources and distributed nature (Rupanetti & Kaabouch, 2024; Xu et al., 2020). Researchers have developed AI-based security and privacy preservation methods through decentralized trust measurement systems and IoT-enabled system security frameworks (Rupanetti & Kaabouch, 2024). Researchers are investigating blockchain technology and privacy-preserving methods for AI model training and deployment to boost security in edge-enabled IoT services (Xu et al., 2020).

The successful integration of AI at the edge with cloud systems demands solutions for computational resource limitations and energy efficiency problems and security and privacy issues. The development of hybrid cloud-edge architectures and dedicated edge hardware and AI-based security solutions shows promise for addressing these challenges. Future studies need to create adaptable and robust solutions which distribute AI workloads efficiently between edge devices and cloud infrastructure while protecting data privacy and system security (Rupanetti & Kaabouch, 2024; Xu et al., 2020; Zou et al., 2019).

VI. FUTURE TRENDS

The combination of AI technology operating at the edge with cloud-edge synergy creates promising future trends and computing prospects. The combination of edge AI technology with cloud computing enables real-time analytics improvements and reduces latency while supporting intelligent applications across multiple industries (Rane et al., 2024). The development of hybrid cloud-edge architectures together with distributed AI through federated learning represents an emerging trend according to Rane et al. (2024). The "In-Edge AI" framework uses device and edge node collaboration to exchange learning parameters which results in near-optimal performance with minimal learning overhead (Wang et al., 2019). The system achieves both application-level enhancement and system-level optimization through this method which minimizes pointless system communication traffic.

➢ *The Combination of Edge and Cloud Computing will Push Forward Multiple Essential Technological Developments:*

- Energy Efficiency: AIoT (Artificial Intelligence of Things) systems can reduce their energy consumption through optimized task scheduling between edge devices and cloud services (Zhu et al., 2022).
- Privacy and security: Edge AI enables local data processing which resolves privacy issues that arise from centralized cloud computing operations (Badidi, 2023; Singh & Gill, 2023).
- Healthcare applications: The predictive capabilities of early health prediction through Edge AI demonstrate potential by analyzing distributed large datasets while protecting privacy through federated learning (Badidi, 2023).
- Green edge intelligence: Green edge intelligence represents a new paradigm that has emerged because AI continues to expand its influence on mobile edge computing (Chen et al., 2019).
- Performance optimization: The performance of distributed cloud computing platforms will receive improvements through AI-powered applications which will benefit security and IoT integration domains (Zangana & Zeebaree, 2024).
- Communication efficiency: Future research will concentrate on developing communication-efficient methods which will operate during both training and inference procedures at network edges (Shi et al., 2020).
- Advanced edge AI topics: The current research investigates how pretraining models together with graph neural networks and reinforcement learning operate within edge computing systems (Yao et al., 2022).
- Hardware acceleration: The successful operation of AI inference models at the edge depends heavily on hardware acceleration through specialized accelerators and software platforms (Banjanović-Mehmedović & Husaković, 2023).

The future of AI at the edge combined with cloud-edge synergy demonstrates positive potential because it will develop efficient distributed computing systems with enhanced security and intelligence (Wang et al., 2019). The ongoing evolution of this field shows potential to transform different industries while advancing edge computing technology development (Wang et al., 2019).

VII. CONCLUSION

The future of AI at the edge with Cloud-Edge synergy demonstrates an exciting environment for advancement and innovation (Manduva, 2021). The combination of advanced edge-deployable AI algorithms with evolving edge computing capabilities for complex AI workloads will enable smooth data exchange and adaptive resource management. The AI edge landscape will transform through security improvements and AI model lifecycle management and edge-native AI service development (Manduva, 2021). The adoption and responsible deployment of AI solutions across various industries will be driven by ethical AI frameworks and 5G network integration. The market expansion potential of AI at the edge with Cloud-Edge synergy creates a path for transformative developments which will reshape businesses and communities and advance technologies during the upcoming years (Manduva, 2021).

REFERENCES

[1]. Manduva, V. C. (2021). Optimizing AI Workflows: The Synergy of Cloud Computing and Edge Devices. International Journal of Modern Computing, 4(1), 50-68.

[2]. Sun, P. (2023). Cloud-Edge-Network-Device Synergy, and Convergence of Communication, Sensing, and Computing. In A Guidebook for 5GtoB and 6G Vision for Deep Convergence (pp. 331-336). Singapore: Springer Nature Singapore.

[3]. REDDY, P. (2023). AI and Edge Computing: Synergistic Approaches for Real-time Data Processing in Cloud Environments.

[4]. Chennupati, N. S. (2025). Edge-Cloud Synergy in Real-Time AI Applications: Opportunities, Implementations, and challenges. International Journal of Scientific Research in Computer Science Engineering and Information Technology, 11(2), 2524–2539. https://doi.org/10.32628/cseit25112740

[5]. Sathupadi, K., Achar, S., Bhaskaran, S. V., Faruqui, N., Abdullah-Al-Wadud, M., & Uddin, J. (2024). Edge-cloud synergy for AI-enhanced sensor network data: A real-time predictive maintenance framework. Sensors, 24(24), 7918.

[6]. Zou, Z., Jin, Y., Huan, Y., Nevalainen, P., Heikkonen, J., & Westerlund, T. (2019). Edge and Fog Computing Enabled AI for IoT-An Overview. 51–56. https://doi.org/10.1109/aicas.2019.8771621

[7]. Rong, G., Fan, H., Xu, Y., & Tong, X. (2021). An edge-cloud collaborative computing platform for building AIoT applications efficiently. Journal of Cloud Computing, 10(1). https://doi.org/10.1186/s13677-021-00250-w

[8]. Rane, J., Mallick, S. K., Kaya, Ö., & Rane, N. L. (2024). Artificial intelligence, machine learning, and deep learning in cloud, edge, and quantum computing: A review of trends, challenges, and future directions. deep science. https://doi.org/10.70593/978-81-981271-0-5_1

[9]. Gong, C., Gong, X., Lu, Y., & Lin, F. (2020). Intelligent Cooperative Edge Computing in Internet of

Things. IEEE Internet of Things Journal, 7(10), 9372–9382. https://doi.org/10.1109/jiot.2020.2986015

[10]. Rane, J., Mallick, S. K., Kaya, Ö., & Rane, N. L. (2024). Artificial intelligence, machine learning, and deep learning in cloud, edge, and quantum computing: A review of trends, challenges, and future directions. deep science. https://doi.org/10.70593/978-81-981271-0-5_1

[11]. Wang, X., Chen, M., Wang, C., Zhao, Q., Han, Y., & Chen, X. (2019). In-Edge AI: Intelligentizing Mobile Edge Computing, Caching and Communication by Federated Learning. IEEE Network, 33(5), 156–165. https://doi.org/10.1109/mnet.2019.1800286

[12]. Zhu, S., Ota, K., & Dong, M. (2022). Energy-Efficient Artificial Intelligence of Things with Intelligent Edge. IEEE Internet of Things Journal, 9(10), 7525–7532. https://doi.org/10.1109/jiot.2022.3143722

[13]. Singh, R., & Gill, S. S. (2023). Edge AI: A survey. Internet of Things and Cyber-Physical Systems, 3, 71–92. https://doi.org/10.1016/j.iotcps.2023.02.004

[14]. Badidi, E. (2023). Edge AI for Early Detection of Chronic Diseases and the Spread of Infectious Diseases: Opportunities, Challenges, and Future Directions. Future Internet, 15(11), 370. https://doi.org/10.3390/fi15110370

[15]. Chen, Z., Lan, D., Mao, Z., He, Q., Chung, H.-M., & Liu, L. (2019). An Artificial Intelligence Perspective on Mobile Edge Computing. 100–106. https://doi.org/10.1109/smartiot.2019.00024

[16]. Zangana, H. M., & Zeebaree, S. R. M. (2024). Distributed Systems for Artificial Intelligence in Cloud Computing: A Review of AI-Powered Applications and Services. International Journal of Informatics, Information System and Computer Engineering (INJIISCOM), 5(1), 11–30. https://doi.org/10.34010/injiiscom.v5i1.11883

[17]. Shi, Y., Yang, K., Zhang, J., Letaief, K. B., & Jiang, T. (2020). Communication-Efficient Edge AI: Algorithms and Systems. IEEE Communications Surveys & Tutorials, 22(4), 2167–2191. https://doi.org/10.1109/comst.2020.3007787

[18]. Yao, J., Wang, F., Jia, K., Zhang, F., Yao, Y., Zhang, S., Wu, A., Shen, T., Chu, Y., Ma, J., Zhang, J., Tan, Z., Yang, H., Ji, L., Wu, F., Kuang, K., Zhou, J., & Wu, C. (2022). Edge-Cloud Polarization and Collaboration: A Comprehensive Survey for AI. IEEE Transactions on Knowledge and Data Engineering, 1. https://doi.org/10.1109/tkde.2022.3178211

[19]. Banjanović-Mehmedović, L., & Husaković, A. (2023, October 1). Edge AI: Reshaping the Future of Edge Computing with Artificial Intelligence. https://doi.org/10.5644/pi2023.209.07

[20]. Torres, D. R., Martín, C., Rubio, B., & Díaz, M. (2021). An open-source framework based on Kafka-ML for Distributed DNN inference over the Cloud-to-Things continuum. Journal of Systems Architecture, 118, 102214. https://doi.org/10.1016/j.sysarc.2021.102214

[21]. Rupanetti, D., & Kaabouch, N. (2024). Combining Edge Computing-Assisted Internet of Things Security with Artificial Intelligence: Applications, Challenges, and Opportunities. Applied Sciences, 14(16), 7104. https://doi.org/10.3390/app14167104

[22]. Zou, Z., Jin, Y., Huan, Y., Nevalainen, P., Heikkonen, J., & Westerlund, T. (2019). Edge and Fog Computing Enabled AI for IoT-An Overview. 51–56. https://doi.org/10.1109/aicas.2019.8771621

[23]. Xu, Z., Liu, W., Tan, H., Lu, J., Huang, J., & Yang, C. (2020). Artificial Intelligence for Securing IoT Services in Edge Computing: A Survey. Security and Communication Networks, 2020, 1–13. https://doi.org/10.1155/2020/8872586