

A Comprehensive Review on Edge Computing - Architecture, Algorithm, and Devices

A Madhumitha¹, Aishwarya G P², B P Swathi³, B Vanaja⁴, Govinda Raju M⁵

Department of Electronics and Communication Engineering RV College of Engineering

Abstract—Traditional cloud-based solutions face issues such as excessive latency, network congestion, and security flaws, making edge computing an essential option for latency-sensitive applications. This study investigates current improvements in edge computing architectures, algorithms, and hardware platforms designed specifically for embedded systems. We investigate significant architectural concepts such as Fog Computing, Mobile Edge Computing (MEC), and Cloudlet Computing, emphasizing their importance in boosting processing efficiency and network performance. In addition, we look at developing edge computing strategies that include dynamic service placement, resource allocation, and redundancy management, which improve computational efficiency and adaptability. The article also explores edge computing's various applications, including video surveillance, gearbox condition monitoring, and bedside healthcare systems, demonstrating its usefulness in real-time data processing.

Index Terms—Edge computing, Edge devices, Cloud computing, Fog computing, Cloud edge collaboration

I. INTRODUCTION

Edge computing is rapidly transforming traditional computing paradigms by relocating data processing and storage closer to where data is generated. This shift addresses the latency, bandwidth, and reliability challenges inherent in centralized cloud models, enabling faster response times and more efficient network resource utilization. By decentralizing computational tasks, edge computing minimizes the need for constant data transmission to remote data centers. It leverages a layered architecture that integrates various devices—from sensors and smartphones to industrial machines—alongside specialized algorithms and robust platforms, ensuring seamless data flow and optimized resource allocation in real-time environments. Driven by the explosive growth of connected devices and the surge in data generation, edge computing has become essential. These use cases demand immediate analytics and decision-making capabilities that centralized systems

often cannot provide. Additionally, by distributing workloads and easing the burden on central infrastructures, edge computing not only enhances scalability and performance but also establishes a robust foundation for future advancements in distributed computing and intelligent systems. This paper examines the architectural designs, algorithmic innovations, and diverse array of devices and platforms that underpin edge computing, highlighting its critical role in meeting modern technological demands.

II. LITERATURE SURVEY

In ever changing Digital world, need for computing technologies to evolve is greater than ever today. Cloud Computing and edge computing are two types of computing methodologies that are greatly used in today's world. This paper conducts a survey to understand characteristics of both the computing methodologies. This research is more of an attempt to compare the methodologies on their contrasting Architectures yet addressing the user requirements. While cloud computing conducts the computing activities at the server end, Edge computing brings these computational activities closer to the user, and hence improves the response speed and efficiency. Both of these computing systems ensure they utilize the hardware resources to their maximum extent. Effective utilisation of resources like network bandwidth and hardware capabilities is much needed for better performance. Both these methodologies excel in this aspect. Cloud computing, developed by companies like AWS, Google, and Microsoft, has evolved alongside edge computing. Both technologies share data storage and processing methods, but differ in physical location, speed, and volume of data processed. Outstanding computational power, high data storage capacity, and highly scalability are the advantages of cloud computing while edge computing serves users where latency, real-time analysis are needed and less dependency needed on network connectivity.

[1].

Limited bandwidth load, slower response time, and concerns around privacy and security are day-to-day issues that users are facing in the world where internet dependency is increasing every moment. Edge computing addresses majority of these problems and focuses on being closer to user and data. It ensures it provides immediate response to applications by processing data quickly and making it usable in various applications where real time data handling is important to take quick decisions. Ex : surveillance, traffic monitoring, health care and virtual reality. With models like cloudlets, mobile edge computing and Fog computing it addressed issues like mobility and latency. Edge computing also allows computing tasks to be performed at the network's edge, reducing latency and providing a better user experience.

Collaborations with other technologies, such as AR/VR, IoT, bio-sensing neurotechnology, data science, and mobile computing, have led to new improvements and challenges in edge computing. Another study presents a broad classification of edge computing architectures, emphasizing its advantages in latency reduction and efficiency. However, it does not investigate the security concerns in depth, which are crucial for deployment in critical applications [2].

Research on the architectural implications of edge computing explores emerging technologies such as blockchain and 5G, but lacks an in-depth discussion on the practical complexities of deployment and resource allocation strategies [3].

In smart cities, edge computing supports healthcare, transportation, and energy management, yet regulatory and interoperability challenges hinder large-scale deployment [4]. Similarly, a study on edge computing's opportunities and challenges provides a comprehensive analysis but does not critically assess workload management and scalability concerns [5]. Edge computing enhances localized processing and reduces latency, addressing cloud computing limitations. Research highlights its applications, security risks, and future trends. Security concerns include threats such as denial-of-service attacks, malware injection, and authentication vulnerabilities. While potential solutions like Zero Trust Mechanisms and

homomorphic encryption are proposed, empirical validation of their effectiveness remains limited, highlighting the need for real-world implementation and performance analysis [6].

A case study on Multi-Access Edge Computing (MEC) implementation in Indonesia's telecommunication towers highlights collaborative efforts in deploying a proof-of-concept infrastructure. A comparison with American Tower's MEC deployment reveals challenges such as limited scalability, reliance on repurposed devices, and inadequate security and power management. The lack of standardized troubleshooting frameworks and technical literature further complicates deployment [7]. Fault tolerance in edge computing ensures reliability despite failures in devices, networks, or resources. Challenges include limited scalability, error propagation, and performance failures such as network congestion and resource bottlenecks. Key research gaps involve Byzantine fault tolerance, adaptive consistency models, and failure prediction frameworks for dynamic workload management. Existing fault management systems also lack solutions for multi-tenancy and cross-layer dependencies. Addressing these issues is crucial to enhancing scalability, security, and efficiency, especially in latency-sensitive applications like autonomous systems, healthcare, and industrial IoT [8].

III. EDGE COMPUTING ARCHITECTURE

Edge computing is typically structured into a three-layered architecture, allowing seamless data processing and management between end devices and cloud infrastructure. The End Layer consists of devices that generate and collect data, such as IoT sensors (temperature, motion, and environmental sensors), smartphones, drones, smart cameras, autonomous vehicles, and industrial robots. These devices transmit data to higher layers for processing. Some high-performance devices, like smartphones and industrial robots, can perform limited local processing before sending data forward. The Edge Layer includes edge nodes responsible for processing data closer to the source. Common edge nodes include 5G base stations, which process mobile network data before sending it to cloud servers, gateways and routers that filter and analyze data, and local edge servers that store and process

data for real-time applications. This layer reduces latency, minimizes cloud dependence, and enhances security by keeping sensitive data closer to its source. In this Layer, data remains at the origin hence decreasing the probability of security risks. This also allows the architecture to increase the speed and response time. While edge computing minimizes cloud reliance, certain tasks, such as deep learning model training and big data analytics, still require high-performance computing (HPC) provided by cloud infrastructure [1].

A. Architectures in Edge computing

Fog Computing extends cloud capabilities closer to the edge by using fog nodes such as industrial controllers, smart routers, and edge servers to process data before reaching the cloud. It enables distributed processing between IoT devices and the cloud, low-latency data processing for time-sensitive applications, and support for multiple communication protocols beyond IP-based networks. For example, in smart cities, fog nodes at traffic intersections analyze real-time traffic data from sensors and dynamically adjust signal timings to reduce congestion without cloud involvement. Mobile Edge Computing (MEC) enhances mobile network performance by deploying edge computing resources within 5G base stations and telecom infrastructure. It ensures ultra-low latency by processing data at the network edge, improves Quality of Service (QoS) for mobile applications, and is optimized for 5G and IoT networks. A practical example is cloud gaming platforms such as Google Stadia and NVIDIA GeForce Now, which use MEC to reduce lag and ensure smooth gameplay by processing game data at 5G base stations instead of distant cloud servers.

Cloudlet Computing introduces mini-clouds (cloudlets) that are deployed closer to users for fast and localized computing. These cloudlets act as small data centers and provide computing power without relying on remote cloud servers. With a single-hop connection for ultra-fast processing, localized cloud infrastructure reduces network congestion and is ideal for real-time AI applications such as AR/VR. Hospitals, for instance, use cloudlets to process MRI scan results quickly, allowing doctors to access diagnostic insights without waiting for cloud-based processing.

Collaborative Edge Computing dynamically distributes workloads between the edge layer and the cloud, ensuring optimal performance for both

real-time and large-scale data processing. It balances workloads dynamically between the Edge and Cloud layers, adapts to computing demands for different applications, and enhances scalability and fault tolerance. In self-driving cars, for example, local AI systems process immediate navigation decisions, while long-term AI learning models are updated via cloud servers [3].

B. Comparison of Edge Computing Architectures

Each architecture has unique advantages depending on its application. Fog computing is most effective for smart cities, IoT, and manufacturing, while MEC is optimized for 5G networks and cloud gaming. Cloudlet computing excels in healthcare and AR/VR applications, whereas collaborative edge computing is best suited for AI-driven applications such as self-driving cars [3]. Ungurean et al. (2024) propose an Advanced Metering Infrastructure (AMI) that integrates edge and fog computing with cloud technology to improve real-time processing in smart grids. The study highlights how edge computing enables local data processing, reducing reliance on centralized cloud servers while ensuring timely decision-making. Meanwhile, edge and fog computing introduce low latency and localized processing, allowing data analysis and decision-making at the source [9].

IV. ALGORITHMS OF EDGE COMPUTING

Edge computing has emerged as a pivotal paradigm to reduce latency, enhance efficiency, and optimize resource utilization by processing data closer to its source. Recent research has introduced various algorithms aimed at improving edge computing performance across key areas such as dynamic service placement, resource allocation, microservice deployment, redundancy management, and data caching.

A. Dynamic Service Placement Algorithms

A recent study [10] evaluates the effectiveness of nature-inspired algorithms, including genetic algorithms and particle swarm optimization, for dynamic service placement in edge computing environments. These algorithms consider workload variations, latency constraints, and resource heterogeneity to optimize service placement, effectively balancing computational costs and response times.

B. Resource Allocation in Satellite Edge Computing

Efficient resource allocation is a critical challenge in satellite edge computing. The study [11] introduces an approach that incorporates spatiotemporal load factors to model satellite load variations. This method aims to optimize user-perceived delay and energy consumption, ensuring high availability and energy efficiency in edge environments.

C. Deep Reinforcement Learning for Microservice Deployment

Microservice-based architectures in edge computing demand intelligent deployment strategies. The paper [12] explores the deployment and optimization of deep learning algorithms on edge computing platforms. It focuses on model pruning and optimization to align with the computational capabilities of edge devices, leading to improved efficiency and reduced operational costs.

D. Redundancy

Redundancy management is vital for minimizing service delays in edge computing systems. The research [13] proposes redundancy-aware scheduling algorithms that utilize multiple edge nodes to minimize service delay. The study discusses the trade-off between redundancy and resource consumption, proposing strategies that maximize response speed while maintaining efficient resource usage.

E. Data Caching Strategies in Edge Computing

Data caching plays a pivotal role in reducing latency in edge networks. The article [14] examines how deep learning techniques can predict content popularity, enabling edge devices to proactively cache content. This approach reduces access latency and improves user experience by intelligently storing frequently accessed data at the network edge. Recent advancements in edge computing algorithms focus on optimizing dynamic service placement, resource allocation, microservice deployment, redundancy management, and data caching strategies. The integration of nature-inspired algorithms, deep reinforcement learning, and predictive modeling has significantly enhanced the efficiency and adaptability of edge computing systems, paving the way for more intelligent and autonomous operations.

V. APPLICATIONS OF EDGE COMPUTING

The applications of edge computing in a variety of sectors, such as health care, autonomous systems are reviewed in this section. Potential applications where edge computing is utilized to address critical challenges are condition monitoring, healthcare and remote patient monitoring, image processing, smart city, natural language processing, transfer learning applications, and fault detection. Some popular hardware platforms for edge computing include: NVIDIA Jetson Nano series, Google Coral edge TPU, Microsoft Azure IoT Edge, AWS IoT Greengrass, Siemens MindSphere (now Insights Hub), all offering powerful processing capabilities in a compact form factor, suitable for running AI and data analysis close to the data source at the network edge.

A. Design of Video Monitoring system

Edge computing can optimize the delay and efficiency of the video monitoring system, and perform well in bandwidth, storage, delay, etc., while edge cloud collaboration technology can improve the overall function and performance of the video monitoring system. The edge computing function on both the video capture terminal and the edge server makes full use of the characteristics of edge computing, such as short distance, flexibility, and high efficiency. It reduces the traffic between edge computing servers and cloud centres. The system proposed in paper [15] resolves the problems of video surveillance applications using edge computing. Various capture nodes capture the data of the monitoring target such as faces, license plates, parts of the vehicle. The four modules in the edge computing server analyses the data and preprocesses the video data nearby. The video receiving module receives the integrated video sent by the nodes. The target extraction module obtains the image of the target and extracts the feature information. Generation module generates recognition result when the recognition target matches with the target to be recognized. Later, the result sending modules transfers the data to the cloud centre. The original video data is localized and identified through edge nodes and edge servers to avoid uploading data of non-monitoring targets to the cloud. When the monitoring target involves people or objects closely associated with people, the privacy protection function can also be realized

B. Gearbox monitoring system

A new Jacobi-type data parallel processing

technique is used in a gearbox monitoring system [16] that computes the gearbox's life and efficiency using an edge service application. By implementing industrial Internet applications and analysing vast amounts of heterogeneous data in real-time, edge computing helps the industry meet its professional application requirements. Lightweight containers like Kubernetes and Docker are frequently used in edge computing to offer a microservices runtime environment that includes dependencies and applications. Sensors gather a variety of data about the gearbox of mining equipment. Data is sent to Docker on Jetson Nano edge devices via Modbus. The management of loud edge, the update of computing data, and data gearbox are among the effective devices with an integrated cloud edge end node that have been developed.

C. Bedside monitoring system

A thermopile array sensing element is used by the edge computing-based bedside monitoring system [17] to continually take thermal pictures. The AI chip's neural network model can identify the bounding box of the bed, the human's bounding box, and the first-stage posture of the image. The computational workload is moved to the edge devices to minimize the need of data transmission. The delays typically caused by network latency is removed by localized processing allowing for faster identification of fall events. To create the final stage, the AI chip's CPU then adjusts the first stage posture. An event involving a fall happens when the subject's body transitions from a non-falling to a falling state and remains in that state for a while. As soon as a fall event is recognized, the system instantly sends a Wi-Fi warning to the central control console.

An AI development board and a thermopile array sensing element make up the platform's hardware. Graphics processing units, or FPGAs, have been used in several related experiments to speed up the inference of AI models using an edge computing platform. The YOLO-LW model was used on the edge computing platform to infer neural networks in the int8 precision format.

VI. HARDWARE PLATFORMS

A. Edge hardware platform

Using the low power Jetson nano[16] as a computer device, applications and alert models are created at

the edge device. The development languages for the Gearbox application are Python and C++. Real-time communication is made possible via the gearbox matrix technique and a GPU processor with NVIDIA CUDA cores, which together provide 472 GFLOPS of AI computing power. Rich interfaces that allow data collection and gearboxing by sensors of different protocols include UART, SPI, and I2S.

B. Cloud Platform

Cloud edge collaboration can be facilitated via open-source edge computing platforms such as Alibaba Cloud, Baidu Cloud, and KubeEdge. In addition to MQTT, Modbus, and OPC UA drivers, the Alibaba cloud's connectivity IoT edge provides rapid monitoring, simpler deployment, and easy access to edge devices. Web applications, including making twin models of edge devices, are made with frameworks like Flask, Echart, and Python [16]. To generate digital twin models, cloud systems may compile edge data and transmit it to desktop CAE applications.

C. Deep learning for edge computing

Using TensorFlow Lite on the ODROID-XU4 platform, the article[18] offers a thorough method for deploying deep learning models on edge computing devices. It offers a useful manual for enabling AI on devices with limited resources by tackling issues like hardware compatibility and the difficulty of deploying machine learning models at the edge. The edge AI accelerator NVIDIA Jetson-TX1 supports four ARM cores and 256 CUDA Maxwell cores, while the Jetson-TX2 supports six ARM cores and 256 CUDA Pascal cores.

VII. CHALLENGES

A. Security and Privacy: Edge computing's decentralized nature exposes it to significant security risks, including denial-of-service attacks, malware injection, and unauthorized access. Since edge nodes operate in distributed and often unprotected environments, they are more vulnerable to cyber threats. The absence of standardized security frameworks across different edge platforms further complicates risk mitigation. To safeguard sensitive data and maintain data integrity, it is crucial to employ robust encryption, continuous authentication, and blockchain-based security solutions.

B. **Resource Limitations and Scalability:** Due to limited computational capacity, memory, and storage, edge computing devices encounter several challenges in managing large-scale applications efficiently. With the growing number of IoT devices, efficient resource allocation and workload balancing mechanisms are crucial to optimizing performance while avoiding overloading edge nodes.

C. **Data Consistency and Synchronization:** As data processing is distributed across multiple edge nodes, maintaining consistency and ensuring real-time updates become complex. Maintaining consistency is especially crucial in areas like industrial automation and autonomous vehicles, as even small discrepancies can cause significant operational issues.

D. **Energy Efficiency:** Many edge computing devices rely on battery power or have limited energy resources, making power consumption a crucial concern. High computational demands can quickly deplete energy reserves, necessitating the development of low-power hardware solutions and adaptive power management techniques to sustain edge operations efficiently.

E. **Interoperability and Standardization:** With edge computing incorporating various hardware and software platforms, achieving seamless integration becomes a challenge. The absence of uniform protocols and frameworks limits interoperability, complicating communication between devices from different manufacturers. To create a unified edge computing environment, industry-wide standardization initiatives are essential.

F. **Bandwidth Management:** With the increasing number of edge devices, efficiently managing network bandwidth becomes increasingly important. Variations in bandwidth can cause latency problems, especially in real-time scenarios like industrial IoT, video analytics, and smart cities. Optimizing bandwidth usage through edge caching and adaptive bandwidth allocation strategies is necessary to maintain seamless performance.

G. **Network Reliability and Latency:** Edge computing enhances performance by processing data near its source, but ensuring reliable network connectivity remains a concern. Sectors like remote

healthcare and autonomous vehicles are highly sensitive to even brief connectivity losses, which can cause major disruptions. Ensuring stable network connections and implementing failover mechanisms are essential to mitigate this risk.

H. **Management and Orchestration:** The large-scale deployment of edge devices requires efficient management and orchestration systems to monitor performance, update software, and allocate resources dynamically. In smart city applications, for example, managing thousands of interconnected IoT sensors for traffic control and environmental monitoring demands robust orchestration frameworks.

I. **Edge-to-Cloud Integration:** Achieving seamless integration between edge computing and cloud platforms is necessary for balancing computational loads and ensuring efficient data synchronization. However, challenges such as latency in data transfers, security vulnerabilities in hybrid environments, and the complexity of managing distributed workloads need to be addressed to enable effective edge-to-cloud collaboration.

VIII. OPPORTUNITIES AND FUTURE TRENDS

A. **Artificial Intelligence (AI) and Machine Learning at the Edge:** Combining AI with edge computing is transforming real-time decision processes. Lightweight AI models optimized for edge devices are enabling applications such as predictive maintenance, smart surveillance, and autonomous navigation. Current research aims to create AI models that are both energy-efficient and capable of functioning on edge devices with limited resources without compromising on performance.

B. **Adaptive Resource Management:** Dynamic resource allocation mechanisms are emerging to optimize bandwidth, processing power, and storage in response to workload variations. AI-driven predictive models are being explored to enhance workload distribution across edge, fog, and cloud layers, reducing processing delays and improving system efficiency.

C. **5G-Enabled Edge Computing:** Implementing 5G networks is anticipated to

greatly improve edge computing by delivering high-speed data transfer and ultra-low latency. These innovations will support advanced applications such as autonomous systems, interactive augmented reality, and telemedicine solutions. Ongoing research focuses on improving how 5G and edge computing work together for large-scale implementation.

D. Collaborative Edge-Cloud Architectures: Hybrid computing models that integrate edge and cloud environments are gaining traction. These architectures enable seamless workload distribution, allowing computationally intensive tasks to be offloaded to the cloud while keeping latency-sensitive processes at the edge. Future research aims to refine these hybrid models for improved efficiency and scalability.

E. Federated Learning for Privacy-Preserving AI: Federated learning is gaining attention as a privacy-centric AI training method, enabling models to be trained on edge devices directly without requirement that data needs to be sent to central servers. This approach strengthens data security in areas like financial fraud detection, healthcare analytics, and personalized recommendation platforms.

F. Energy-Efficient Edge Solutions: Sustainable edge computing architectures are being explored to reduce energy consumption. Innovations include renewable energy-powered edge devices and dynamic power management techniques that optimize energy use based on workload demands. These solutions are critical for extending the operational lifespan of edge networks.

G. MEC solutions: MEC improves real-time data processing by bringing computational resources closer to users in 5G networks. Research is focused on addressing scalability challenges and optimizing MEC for seamless integration across heterogeneous networks.

H. Edge AI Democratization: The growing availability of cost-effective edge AI hardware is making AI-powered edge computing accessible to smaller businesses and individual developers. Emerging devices, such as Nvidia's Jetson Orin Nano, provide high-performance computing capabilities at lower costs, expanding the adoption of AI at the edge.

I. Edge-to-Edge Networking: Future research is exploring direct communication between edge nodes without relying on centralized cloud servers. This peer-to-peer edge networking approach can improve efficiency in distributed applications such as smart grids, autonomous drone fleets, and decentralized IoT networks.

J. Quantum Computing at the Edge: Although still in its early stages, the integration of quantum computing with edge networks is being investigated to enhance complex computational tasks such as cryptography and large-scale data analytics. Quantum-enhanced edge computing has the potential to revolutionize fields requiring high-speed processing of massive datasets.

IX. CONCLUSION

Edge computing helps reduce dependency on centralized cloud infrastructure for data processing by shifting computation closer to the source. This improves latency, optimizes bandwidth, and strengthens security, making it valuable for applications in smart cities, healthcare, industrial automation, and autonomous systems. The paper explores key architectures including Fog computing, MEC, Cloudlet computing, and Collaborative Edge computing, which enhance efficiency by distributing workloads. It also examines algorithms for dynamic service placement, resource allocation, and failure prediction, which are crucial for optimizing edge performance. Various use cases, including real-time traffic monitoring, industrial automation, and edge-based healthcare diagnostics, highlight its practical advantages.

However, challenges such as computational constraints, network instability, interoperability issues, and energy limitations require further research into scalable and adaptive solutions. While significant progress has been made, several research gaps remain, particularly in seamless edge-to-cloud integration, robust security frameworks, and real-time workload distribution. Ongoing efforts focus on AI integration, energy-efficient hardware, and adaptive resource management. Additional research is needed in standardizing communication protocols, leveraging quantum computing for edge applications, and optimizing AI models for resource-constrained devices. Addressing these issues will enable more efficient, secure, and scalable edge computing solutions, driving

advancements in autonomous and data-driven technologies.

REFERENCES

- [1] H. Nouhas, A. Belangour and M. Nassar, "Cloud and Edge Computing Architectures: A Survey," 2023 IEEE 11th Conference on Systems, Process and Control (ICSPC), Malacca, Malaysia, 2023, pp. 210-215, doi: 10.1109/ICSPC59664.2023.10420123.
- [2] Deepak, M. K. Upadhyay and M. Alam, "Edge Computing: Architecture, Application, Opportunities, and Challenges," 2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS), Tashkent, Uzbekistan, 2023, pp. 695-702, doi: 10.1109/IC-TACS59847.2023.10390171.
- [3] K. Thakkar, A. Patel, A. Budhrani, K. Patel, S. Patel and A. Nayak, "Edge Computing: Architecture, Implications and Future with Latest Trends," 2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN), Salem, India, 2023, pp. 1065-1070, doi: 10.1109/ICPCSN58827.2023.00181.
- [4] L. U. Khan, I. Yaqoob, N. H. Tran, S. M. A. Kazmi, T. N. Dang and C. S. Hong, "Edge-Computing-Enabled Smart Cities: A Comprehensive Survey," in IEEE Internet of Things Journal, vol. 7, no. 10, pp. 10200- 10232, Oct. 2020, doi: 10.1109/IJOT.2020.2987070.
- [5] G. Kaur and R. S. Batth, "Edge Computing: Classification, Applications, and Challenges," 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2021, pp.254-259,doi: 10.1109/ICIEM51511.2021.9445331.
- [6] R. Anusuya, D. K. Renuka and L. A. Kumar, "Review on Challenges of Secure Data Analytics in Edge Computing," 2021 International Conference on Computer Communication and Informatics (ICCCI), Coim- batore, India, 2021, pp. 1-5, doi: 10.1109/ICCCI50826.2021.9402559.
- [7] P. A. Budiman, Marfani and D. M. Sari, "Multi-Access Edge Computing Implementation On Tower Ecosystem Indonesia: Challenges And Visibil- ity," 2022 13th International Conference on Information and Commu- nication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2022, pp. 158-162, doi: 10.1109/ICTC55196.2022.9952477.
- [8] M. Pourreza and P. Narasimhan, "Fault Tolerant Edge Computing: Challenges and Opportunities," 2023 IEEE 7th International Conference on Fog and Edge Computing (ICFEC), Bangalore, India, 2023, pp. 73- 80, doi: 10.1109/ICFEC57925.2023.00018.
- [9] I. Ungurean and N. C. Gaitan, "Integrating Cloud and Edge/Fog Computing in Advanced Metering Infrastructure," 2024 International Symposium on Sensing and Instrumentation in 5G and IoT Era (ISSI), Lagoa, Portugal, 2024, pp. 1-5, doi: 10.1109/ISSI63632.2024.10720496.
- [10] A. H. Kazmi, A. Staffolani, T. Zhang, C. Cabrera and S. Clarke, "Dynamic Service Placement in Edge Computing: A Comparative Evaluation of Nature-Inspired Algorithms," in IEEE Access, vol. 13, pp. 2653-2670, 2025, doi: 10.1109/ACCESS.2024.3520701.
- [11] J. Zhou, Y. Zhao, L. Zhao, H. Cai and F. Xiao, "Adaptive Task Offload- ing With Spatiotemporal Load Awareness in Satellite Edge Computing," in IEEE Transactions on Network Science and Engineering, vol. 11, no. 6, pp. 5311-5322, Nov.-Dec. 2024, doi: 10.1109/TNSE.2024.3368086.
- [12] X. Liu, Y. Zhang, Z. Wang and J. Yang, "Research on Deep Learning Model and Optimization Algorithm in Edge Computing," 2023 5th In- ternational Conference on Applied Machine Learning (ICAML), Dalian, China, 2023, pp. 242-246, doi: 10.1109/ICAML60083.2023.00054.
- [13] P. Peng and E. Soljanin, "Redundancy Management for Fast Service (Rates) in Edge Computing Systems," in IEEE Transactions on Net- working, doi: 10.1109/TNET.2024.3507372.
- [14] S. Rahman, M. G. R. Alam and M. M. Rahman, "Deep Learning- based Predictive Caching in the Edge of a Network," 2020 International Conference on Information Networking (ICOIN), Barcelona, Spain, 2020, pp. 797-801, doi: 10.1109/ICOIN48656.2020.9016437.
- [15] H. Yang, X. Guo, C. Zhang and J. Wang,

- "Design of Video Monitoring System Based on Edge Computing," 2024 IEEE 2nd International Conference on Image Processing and Computer Applications (ICIPCA), Shenyang, China, 2024, pp. 1411-1414, doi: 10.1109/ICIPCA61593.2024.10709132.
- [16] D. Lu, G. Gao, Y. Shen and Z. Tong, "Design of Gearbox Monitoring System Based on Edge Computing," 2022 4th International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), Hamburg, Germany, 2022, pp. 355-360, doi: 10.1109/AIAM57466.2022.00074.
- [17] Lin BS, Peng CW, Lee IJ, Hsu HK, Lin BS. System Based on Artificial Intelligence Edge Computing for Detecting Bedside Falls and Sleep Posture. *IEEE J Biomed Health Inform.* 2023 Jul;27(7):3549-3558. doi: 10.1109/JBHI.2023.3271463. Epub 2023 Jun 30. PMID: 37115834.
- [18] A. Alsalemi, A. Amira, H. Malekmohamadi and K. Diao, "Facilitating Deep Learning for Edge Computing: A Case Study on Data Classification," 2022 IEEE Conference on Dependable and Secure Computing (DSC), Edinburgh, United Kingdom, 2022, pp. 1-4, doi: 10.1109/DSC54232.2022.9888939.