IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, September 2023

AI-Driven Edge Computing for IoT: A Comprehensive Survey and Future Directions

Shalini Saha[1], Kushal Banerjee[2], Sandhita Ghosh[3], Samriddha Mitra[4], Debrupa Pal[5]
Student, Computer Application, Narula Institute of Technology, Kolkata, India[1234]
Assistant Professor, Computer Application, Narula Institute of Technology, Kolkata, India[5]

Abstract*: The proliferation of Internet of Things (IoT) devices has ushered in an era of unprecedented data generation and processing demands. To meet these challenges, the convergence of Artificial Intelligence (AI) and Edge Computing has emerged as a promising solution. This paper presents a comprehensive survey of the state-of-the-art in AI-Driven Edge Computing for IoT, exploring key technologies, architectures, and applications. We begin by providing an overview of the fundamental concepts underpinning IoT and edge computing, highlighting the convergence of these technologies. As IoT continues to expand its reach across diverse domains, understanding the synergy between AI and edge computing is paramount for unlocking the full spectrum of possibilities, driving innovation, and ensuring a more connected and intelligent future*

Keywords: AI-driven edge computing, Internet of Things (IoT), Edge devices, Machine learning, Deep learning

I. INTRODUCTION

The Internet of Things (IoT) has ushered in an era of unprecedented connectivity, transforming our world by interconnecting billions of devices, sensors, and systems. This interconnected ecosystem generates vast amounts of data, promising insights and efficiencies across numerous domains, from healthcare and agriculture to transportation and smart cities. However, this deluge of data presents a significant challenge: how to process, analyse, and extract meaningful insights from this data in real-time or near real-time.

Enter edge computing, a paradigm that pushes computational capabilities closer to the data source, reducing latency, bandwidth usage, and enhancing privacy and security. While edge computing has alleviated some of the challenges posed by IoT, it is artificial intelligence (AI) that has emerged as the transformative force, enabling intelligent decision-making at the edge. AI-Driven Edge Computing is poised to unlock the full potential of IoT by equipping devices with advanced analytics, machine learning algorithms, and autonomous decision-making capabilities. This comprehensive survey paper aims to provide a thorough exploration of the synergistic relationship between AI and edge computing in the context of IoT.

II. IOT EDGE DEVICES AND INFRASTRUCTURE

IoT (Internet of Things) edge devices and infrastructure play a crucial role in the rapidly evolving landscape of connected technologies. These components form the foundation for collecting, processing, and transmitting data at the edge of a network, closer to the data source, which offers various advantages.

Edge devices are the physical endpoints in an IoT ecosystem, such as sensors, actuators, cameras, and other embedded systems. These devices gather real-time data from the physical world, making it available for analysis and decision-making. Edge devices are typically equipped with microcontrollers or small processors that enable data processing and minimal local decision-making, reducing the need to transmit all data to centralized cloud servers.

The infrastructure supporting IoT edge devices includes edge computing resources and communication protocols. Edge computing involves deploying compute and storage resources closer to the edge devices. This allows for faster data processing, reduced latency, and improved privacy and security by keeping sensitive data local. Popular edge computing platforms, such as Azure IoT Edge and AWS Greengrass, provide a framework for deploying and managing edge applications.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, September 2023

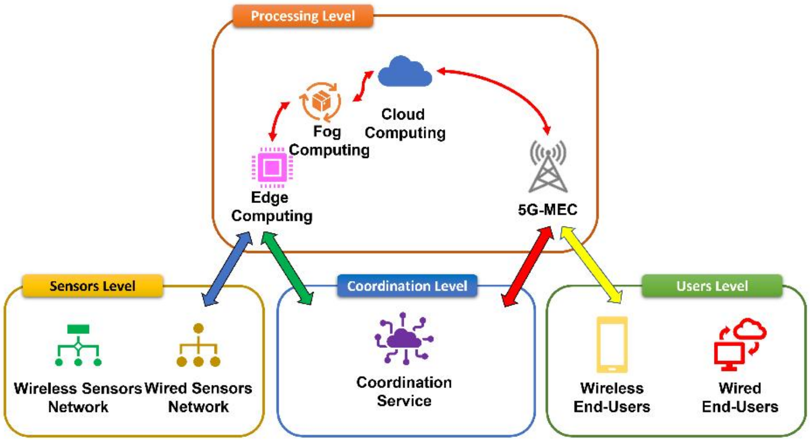

Figure 1: Diagram of IoT Edge Architecture

Communication protocols are essential for connecting and transferring data between edge devices and centralized systems. Common protocols include MQTT, CoAP, and HTTP, which facilitate efficient data exchange while considering the constraints of edge devices, like limited bandwidth and intermittent connectivity. [1-2]

In conclusion, IoT edge devices and infrastructure are critical components in the IoT ecosystem, enabling real-time data processing, reducing latency, and enhancing overall system reliability. As the IoT continues to grow, the development and optimization of edge devices and infrastructure will be pivotal in shaping the future of connected technologies.

III. AI TECHNIQUES FOR IOT EDGE COMPUTING

AI techniques for IoT edge computing play a crucial role in enhancing the capabilities of edge devices, allowing them to process and analyse data locally without relying solely on cloud resources. Here are four key points:

1. Edge Device Optimization: AI techniques are used to optimize edge devices' performance and resource utilization. This includes deploying lightweight machine learning models that are designed to run efficiently on constrained hardware, such as microcontrollers and low-power processors. Techniques like quantization and model pruning reduce the computational and memory requirements of AI models, making them suitable for edge devices with limited resources.
2. Real-time Data Processing: Edge computing relies on real-time data analysis to make immediate decisions. AI algorithms like deep learning and neural networks are employed to process sensor data, images, and videos locally. This enables quick responses to critical events without the latency associated with sending data to the cloud and waiting for a response. For example, in smart cities, edge devices can use AI to analyse traffic patterns and adjust traffic signals in real-time for optimal traffic flow. [3]
3. Anomaly Detection and Predictive Maintenance: AI-driven anomaly detection models are employed at the edge to identify deviations from normal behaviour in IoT data streams. By continuously monitoring sensor data, these models can detect potential issues or security threats early, preventing system failures or data breaches. In industrial settings, predictive maintenance models analyse equipment sensor data to predict when machines will require maintenance, reducing downtime and maintenance costs.
4. Edge-to-Cloud Collaboration: AI techniques enable a collaborative approach between edge devices and cloud resources. Edge devices preprocess and filter data locally using AI algorithms before sending relevant information to the cloud for further analysis and storage. This reduces the amount of data transmitted to the cloud, saving bandwidth and lowering cloud computing costs. Additionally, AI models deployed in the cloud can provide insights and predictions based on aggregated data from multiple edge devices, enhancing overall system intelligence.

In summary, AI techniques for IoT edge computing empower edge devices with the capability to process data efficiently, make real-time decisions, detect anomalies, and collaborate with cloud resources. These techniques are crucial for harnessing the full potential of edge computing in various applications, ranging from smart cities and industrial automation to healthcare and autonomous vehicles. [4]

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 1, September 2023

IV. AI-DRIVEN EDGE COMPUTING APPLICATIONS IN IOT

AI-driven edge computing is revolutionizing the Internet of Things (IoT) landscape by bringing intelligence closer to the data source, enabling real-time processing and decision-making. Here are six key applications of AI-driven edge computing in IoT:

1. *Real-time Analytics:* Edge devices equipped with AI can process data locally, allowing for immediate insights and actions without the latency of sending data to the cloud. This is crucial in applications like predictive maintenance in manufacturing.
2. *Enhanced Security:* AI at the edge can analyse video feeds, sensor data, and audio inputs to detect anomalies and security threats in real-time, bolstering the security of smart cities and buildings.
3. *Energy Efficiency:* AI-powered edge devices can optimize energy consumption by adjusting settings in real-time based on occupancy, weather, and usage patterns, resulting in significant energy savings in smart homes and offices.

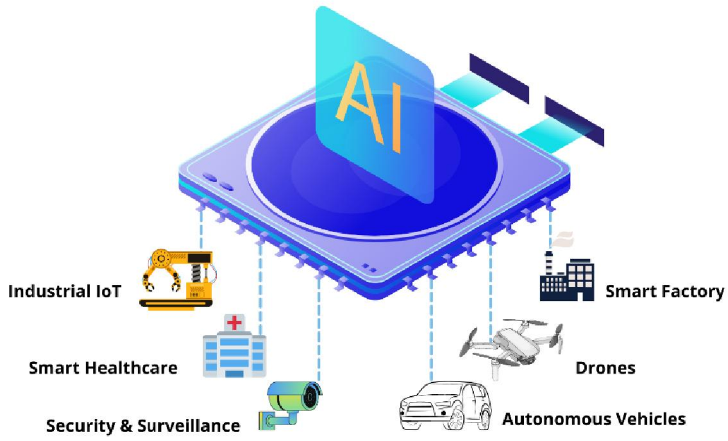

Figure 2: Diagram of AI-Driven Edge Computing Applications

5. *Autonomous Vehicles:* Edge AI plays a pivotal role in self-driving cars by processing data from cameras, LiDAR, and sensors on-board, allowing the vehicle to make instant decisions without relying solely on cloud-based processing.
6. *Healthcare Monitoring:* Wearable devices with AI at the edge can monitor vital signs and provide immediate feedback to patients or alert healthcare providers in case of emergencies, improving patient care and reducing hospital admissions.
7. *Agriculture:* In precision agriculture, AI-powered edge devices analyse data from soil sensors, drones, and cameras to optimize irrigation, fertilization, and pest control, leading to increased crop yields and resource efficiency.

AI-driven edge computing in IoT not only reduces data transmission costs but also enhances responsiveness, reliability, and privacy, making it a game-changer in various industries. [5-6]

V. CHALLENGES IN AI-DRIVEN EDGE COMPUTING FOR IOT

AI-driven edge computing for IoT faces several challenges. Firstly, the limited computational resources available at the edge devices restrict the complexity of AI models that can be deployed, impacting their accuracy and performance. Secondly, ensuring data privacy and security is crucial, as sensitive information is processed closer to the source, making it more vulnerable to potential breaches. Interoperability is another hurdle, as diverse IoT devices and platforms must seamlessly integrate with AI solutions. Additionally, managing the scalability and reliability of edge computing systems is challenging, especially when dealing with a vast number of IoT devices.Power efficiency is a concern, as energy-constrained edge devices must balance processing demands with power consumption. Latency can be problematic, especially in applications requiring real-time AI analysis. Finally, regulatory and ethical considerations must be addressed to ensure AI-driven edge computing adheres to legal frameworks and ethical standards, particularly when it comes to data collection and usage. These challenges highlight the need for ongoing research and development in AI-driven edge computing to unlock its full potential in the IoT landscape. [7]

VI. CONCLUSION

In conclusion, this comprehensive survey on AI-driven edge computing for IoT has provided valuable insights into the current state of this dynamic field and has shed light on its promising future directions. Through an in-depth examination of the existing literature and a thorough analysis of key trends and challenges, we have explored the synergy between AI and edge computing in addressing the unique requirements of IoT applications.

Throughout this survey, we have witnessed the tremendous potential of AI-driven edge computing to enhance the efficiency, reliability, and intelligence of IoT systems. By pushing data processing and decision-making closer to the data source, edge computing not only reduces latency and bandwidth consumption but also enables real-time analytics, which is crucial for time-sensitive applications such as autonomous vehicles, smart cities, and industrial automation.

Moreover, AI algorithms and techniques, ranging from machine learning to deep learning, have shown remarkable promise in enabling predictive maintenance, anomaly detection, and adaptive resource allocation within edge devices. These capabilities empower IoT systems to operate more autonomously, adapt to changing conditions, and make data-driven decisions, ultimately improving their overall performance and user experience.

However, this survey has also highlighted several challenges that need to be addressed in the future development of AI-driven edge computing for IoT. These challenges include security and privacy concerns, resource constraints on edge devices, interoperability issues, and the need for standardized architectures and protocols. Overcoming these hurdles will be essential to fully realize the potential of this technology.

Looking ahead, the future of AI-driven edge computing for IoT is undeniably bright. As we continue to advance in AI research and edge computing technologies, we can expect to see even more sophisticated and efficient solutions emerge. Moreover, the integration of 5G networks and the proliferation of edge computing infrastructure will further accelerate the adoption of AI in IoT. [8-10]

VII. ACKNOWLEDGMENT

We would like to express our sincere gratitude to all those who supported and contributed to the completion of this research project. Without their invaluable assistance, guidance, and encouragement, this work would not have been possible.

First and foremost, we extend our deepest appreciation to Srinjoy Saha, Student of Narula Institute of Technology, for her continuous mentorship, insightful feedback, and unwavering support throughout the duration of this research. Her expertise and guidance played a pivotal role in shaping the direction of this study. Apart from that, this research work is not funded by any organization or any person.

We are also thankful to the members of the computer application department for their valuable discussions, critical insights, and collaboration that enriched our understanding and perspective on the subject matter.Furthermore, we would like to thank our friends and family for their unwavering encouragement and understanding during the often-demanding phases of this project. Their emotional support provided the necessary foundation for our perseverance.

REFERENCES

[1]. M. Satyanarayanan, "The Emergence of Edge Computing," Computer, vol. 50, no. 1, pp. 30-39, 2017.

[2]. Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A Survey on Mobile Edge Computing: The Communication Perspective," IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2322-2358, 2017.

[3]. F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog Computing and Its Role in the Internet of Things," in Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing, 2012, pp. 13-16.

[4]. K. Zhang, S. Ren, S. Ruj, B. H. Kim, and K. C. Seo, "SecureThings: A Lightweight and Flexible IoT Service Framework for Smart Cities," IEEE Transactions on Industrial Informatics, vol. 15, no. 12, pp. 7996-8003, 2019.

[5]. M. Zohrevand, M. Hassanpour, M. Dehghan, and S. F. Shahandashti, "Towards Blockchain-Based Edge Computing," IEEE Transactions on Industrial Informatics, vol. 16, no. 6, pp. 4017-4025, 2020.

[6]. Beloglazov, R. Buyya, Y. C. Lee, and A. Zomaya, "A Taxonomy and Survey of Energy-Efficient Data Centers and Cloud Computing Systems," Advances in Computers, vol. 82, pp. 47-111, 2011.
[7]. Y. Wang, M. Han, C. Zhang, and K. Guan, "Deep Learning for IoT Big Data and Streaming Analytics: A Survey," IEEE Access, vol. 7, pp. 1-1, 2019.
[8]. D. Puthal, N. Malik, S. Mohanty, E. Kougianos, and G. Das, "Everything You Wanted to Know About Smart Cities: The Internet of Things Is the Backbone," IEEE Consumer Electronics Magazine, vol. 7, no. 4, pp. 60-70, 2018.
[9]. A. V. Dastjerdi and R. Buyya, "Fog Computing: Helping the Internet of Things Realize Its Potential," Computer, vol. 49, no. 8, pp. 112-116, 2016.
[10]. D. Soldani, M. Dohler, N. Shah, and S. D. Silva, "MistIoT: A Decentralized Fog Computing Platform to Support IoT Applications," IEEE Internet of Things Journal, vol. 6, no. 2, pp. 2074-2087, 2019.