© DEC 2023 | IRE Journals | Volume 7 Issue 6 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV7I6-1723104

Edge Computing and Artificial Intelligence in Internet of Things Systems: A Review

DR. USHA RANI
Associate Professor in Computer Science, Smt Sushma Swaraj Govt. College for Girls Ballabgarh

Abstract- Internet of Things (IoT) has empowered massive physical devices interconnection issuing on-the-fly data. But these cloud-centric IoT architectures are challenged by latency, network congestion and network dependence. These shortcomings are solved by edge computing which moves the computation close to data, while AI supports intelligent analysis of data and predictive decisions. In this paper, edge and AI integration in IoT systems have been summarized. It presents architecture evolution, key application domains, and underlying problems in the edge-based AI deployment. The survey shows how such a distributed view induces efficiency and responsivness but also sensitive open issues on security, scalability and resource constrants.

Keywords: Internet of Things (IoT), Artificial Intelligence, Edge Computing.

I. INTRODUCTION

Internet of Things (IoT) is a system of interrelated computing devices, mechanical and digital machines or objects provided with unique identifiers and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction [1]. IOT applications in smart home, healthcare, transportation, agriculture and industry automation are becoming popular. Such systems produce huge amounts of heterogeneous data that have to be processed and analyzed in real time [2].

In its early years, IoT applications leaned heavily on cloud for data storage and processing. While cloud computing can deliver scalability and abundant computational resources, it also creates latency and bandwidth issues making it unsuitable for real-time applications in IoT [3]. It is in these domains (health monitoring or industrial control, e.g.) that the bounded-delay model becomes all important.

Locally computational, close to the IoT devices is permitted with edge computing which is a complementary computing concept in addition to cloud computing [4]. And meanwhile, AI offers learning and reasoning functions to enable IoT systems to process data smartly. The combination of edge computing and AI make IoT systems capable to act autonomously, context-aware with low latency [7][10]. This paper provides a survey of current research on edge-AI-centric IoT systems and its applications and challenges.

II. BACKGROUND: IOT AND EDGE COMPUTING AND ARTIFICIAL INTELLIGENCE

In the IoT, three-tier architectures can be seen; which include (1) sensing, (2) network and (3) application layers: sensors collect data to transfer to centralized servers for computation. When the number of connected devices are significantly large, centralized processing is unfit because of communication overhead and delayed responses [8].

Edge computing adds intermediary process levels between IoT devices and the cloud servers. The edge nodes are responsible for processing and analytics tasks as well as data filtering from the devices, which in turn lessens the amount of data brought to the cloud [3]. This mechanism enhances system performance and alleviate network traffic.

IoT applications benefit from the inclusion of Artificial Intelligence as it enables data driven decision making. The system would be able to learn from the historical data and adapt in different environments using machine learning and pattern recognition [12]. The deployment of AI models at the edge lets IoT systems perform distributed intelligence and real-time analytics [10].

III. EDGE AI ARCHITECTURES EVOLUTION IN IOT SYSTEMS

Early IoT architectures are cloud-based, with most data being first sent to some central points and then analyzed. Though efficient on batch processing, they are not meant for latency aware applications [5]. The urge for real-time analytics resulted in the trend of edge and fog computing [15].

Edge-AI architectures distribute intelligence to multiple levels standing from inferencing devices through edge nodes up to cloud servers. Light AI models are deployed at device-level analytics, and complex processing is done on the edge nodes [4].The cloud still plays a role in archival and large-scale model training.

This hierarchical organization trades off computational efficiency against the analytical power. The researches show that edge-AI architectures can substantially enhance, system performance and reliability in dynamic IoT environment [3].

IV. ARTIFICIAL INTELLIGENCE FOR EDGE-BASED INTERNET OF THINGS

Machine learning is one of the most popular AI methodologies commonly used in edge-oriented IoT infrastructures. Prediction and classification problems such as fault detection, health monitoring, are solved using supervised learning methods [6]. The unsupervised learning is specialized for anomaly or unusual pattern detection in sensor data [5].

Deep learning models provide better feature extraction, more complexly for unstructured data (e.g., images and audio streams). But they are difficult to deploy at the edge due to their compute complexity. One approach to overcome this issue is through the development of lightweight deep learning models and model compression methods [10].

Reinforcement learning is designed to be able to adapt by learning the best action actions based on the interactions carried out by a system with its environment. Such a method also finds some application in intelligent traffic control and energy management systems [10].

V. USE CASES FOR EDGE-AI IN IOT SOLUTIONS

The combination of edge computing and artificial intelligence has greatly broadened the horizon as well as improved the performance for IoT applications. Edge-AI systems address this shortcoming of cloud-centric IoT designs by providing localized intelligence and real-time analytics. This section discusses key application spaces where edge-AI is known to have made significant contributions from current literature. But the one that is going to make a huge difference, with regards to storing and sharing health data in healthcare, will be:

5.1 Smart Healthcare and Remote Patient Monitoring

One of the most significant sectors which is applicable to edge-AI-based IoT system is healthcare. Wearable sensors and medical Internet of things (IoT) devices which have become widely available over the last years capture physiological data on an ongoing basis including heart rate, blood pressure, oxygen saturation and activity. With typical cloud-based solutions - this data needs to be sent over a network to remote servers for analysis which can generate latency and reliability issues.

Edge-AI can process medical data in real-time near the patient, and provide instant detection of abnormal patterns and health-related risks [6]. IoT-based machine learning deployed at the edge can become the early warning system for conditions such as cardiac irregularities or respiratory distress, alerting healthcare providers when to check in and potentially prevent adverse health events. This method lessens the reliance on constant internet connection and increases patient safety.

Further, edge-based processing mitigates the privacy risks of transmitting sensitive health data to centralized cloud platforms [11]. Research indicates that edge-AI systems are well-suited for elderly care, chronic disease monitoring and emergency response operations requiring rapid decision-making [4].

5.2 Industrial Internet of Things and Predictive Maintenance

In the Industrial IoT space, machine and process data are collected in massive quantities from sensors which are placed on machines and manufacturing lines. Edge-AI technologies are fundamental to the enablement of predictive maintenance through local sensor data analysis and pattern recognition for equipment degradation or failure.

Edge-based machine learning algorithms can track vibration, temperature and pressure data in real time to predict failures before they happen [2]. This prevents unplanned downtime, reduces maintenance costs and makes the workplace safer. In contrast to cloudbased monitoring, edge-AI makes it feasible and practical to achieve low latency that is mandatory for real-time industrial control systems.

What's more, edge-AI enables manufacturing plants to continue working optimally in low or erratic network conditions. Decentralised intelligence leads to an enhanced operational resilience and is beneficial for smart manufacturing and Industry 4.0[9].

5.3. Smart Cities and Urban Infrastructure

The smart city, largely built on the Internet of Things (IoT), involves applications such as traffic management, environmental monitoring, public safety and energy management. Edge-AI is an important embodiment that makes these applications become more and provides real-time process of data and intelligent decision at border.

For example, traffic sensors and cameras produce streams of data that can be immediately analyzed at the edge to manage traffic signals, decreasing congestion [10]. The on-board AI (Edge-AI) can also quickly detect accidents or traffic violations, so emergency services and law enforcement have the ability to act immediately.

In Headphones and Audio- Environmental monitoring systems that use edge-AI to analyze the quality of air, noise levels or weather conditions. Due to local processing, network utilization is minimised and timely warnings are provided in dangerous circumstances. It has been reported in a number of research studies that edge analytics enhance the scalability and performance [18] of IoT smart city deployments.

5.4 Smart agriculture and precision farming

IoT applications for enhancing productivity and sustainability in agriculture. The agricultural sector has embraced IoT technologies more and more to increase the productivity of crops is boosting. In fields, sensors capture information on the moisture of the soil, its temperature and humidity in addition to the crop's health. Edge-AI allows processing of these data in real time to support precision farming.

Through local data processing, edge-AI systems can make rapid recommendations on irrigation, fertilization, and pest control – even in rural areas with low connectivity [17]. Machine learning-based models can help diagnose crop stress and predict yield trends to inform management decisions.

Studies indicate that edge-AI reduces dependence on cloud resources and the cost of operations [3], which can lower barriers to sustainable smart agriculture for smallholder farmers. Such local knowledge leads to better ways of managing resources in a sustainable manner and adds to food security.

5.5 Intelligent Surveillance and Public Safety

Intelligent surveillance systems employ IoT sensors and cameras to supervise public areas. Edge-AI brings edge-side video and image analytics, meaning you can get immediate alerts about suspicious activities or safety threats.

When we bring AI to edge, the latencies are reduced and high volume video data does not need to be transmitted to cloud servers [4]. This method also reduces latency and increases privacy because it works with sensitive information locally. Researches indicate that edge AI based surveillance systems are more scalable and efficient, as compared to the centralized [3].

5.6 Energy Control and Smart Grids

Edge-AI is also significant for energy management and smart grid. IoT sensors track patterns of energy demand consumption, and AI models implemented at the edge further analyse usages and better distribute energies.

The use of local decision-making also allows for rapid load balancing and demand response which can minimize energy waste and grid destabilization [10]. Edge-AI systems can also facilitate the integration of renewable energy by forecasting any changes and modifying supply in response.

It is evident from the surveyed literature that edge-AI substantially improves IoT applications in various areas through enabling real-time intelligence, lowering latency and increasing system reliability. Turning IoT systems into proactive and self-adaptive ones, edge-AI realizes a range of applications from healthcare, industrial automation to smart city and agriculture. But to be effective, these applications need solutions to challenges involving resource constraints, security, and system management as discussed in later sections.

VI. ADVANTAGES OF THE INTEGRATION OF EDGE COMPUTING AND AI IN IOT

Edge computing with AI features also decreases the latency and enhances the real-time response ability in IoT environment [13]. On device processing also reduces bandwidth, and highly cost for operations. Moreover, edge-AI potentially strengthens data privacy as there are less sensitive data forwarded into the centralized servers [11].

The next benefit is the increased reliability of the system due to its ability to keep on working under some network failures, using distributed intelligence [4].

VII. CHALLENGES AND LIMITATIONS

Edge-AI empowered IoT systems have to address few challenges despite having an upper hand. Restricted resources available at edge devices restrict the types of complex AI models which can be deployed [16][10]. However, manually updating and managing distributed AI models over the large-scale networks is still resource-consuming.

Security and privacy issues are the most critical, because edge nodes are prone to cyber-attacks [11]. Interoperability and non-standardizedness hinder large scale deployment even more [8].

VIII. FUTURE RESEARCH DIRECTIONS

In the next step large and lightweight AL models to process on edge devices should be designed, since better transaction hypotheses can be provided by a consolidated model. Federated learning is a potential means to achieve collaborative model training while still protecting the privacy of raw data [10].

And research on explainable AI would be important to enhance the explainability and trust in edge-based decision-making systems [14]. Ethical and governance considerations should be included in the design of systems to support responsible AI deployment.

IX. CONCLUSION

In conclusion, this review has provided a broad perspective on the integration of edge computing and artificial intelligence in IoT systems. Edge-AI architectures help overcome considerable limitations of traditional cloud-based IoT solutions due to local and distributed intelligence. The paper illustrated leading applications and the major pros and cons of this integration. Solving the existing challenges in resource management, security, and system coordination is critical to the widespread adoption of intelligent IoT concepts. New studies should focus on developing scalable and energy-efficient edge-AI frameworks and enhancing the existing security models. This way, the relevant modifications will allow for real-time decision-making and create conditions for intelligent IoT in various applied areas.

REFERENCES

[1] Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of Things: A survey. Computer Networks, 54(15), 2787–2805. https://doi.org/10.1016/j.comnet.2010.05.010

[2] 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. https://doi.org/10.1016/j.future.2013.01.010

[3] Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. IEEE Internet of Things Journal, 3(5), 637–646. https://doi.org/10.1109/JIOT.2016.2579198

[4] Satyanarayanan, M. (2017). The emergence of edge computing. IEEE Computer, 50(1), 30–39. https://doi.org/10.1109/MC.2017.9

[5] Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. Mobile Networks and Applications, 19(2), 171–209. https://doi.org/10.1007/s11036-013-0489-0

[6] Islam, S. M. R., Kwak, D., Kabir, M. H., Hossain, M., & Kwak, K. S. (2015). The Internet of Things for health care: A comprehensive survey. IEEE Access, 3, 678–708. https://doi.org/10.1109/ACCESS.2015.2437951

[7] Li, S., Xu, L. D., & Zhao, S. (2015). The Internet of Things: A survey. Information Systems Frontiers, 17(2), 243–259. https://doi.org/10.1007/s10796-014-9492-7

[8] Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. (2015). Internet of Things: A survey on enabling technologies, protocols, and applications. IEEE Communications Surveys & Tutorials, 17(4), 2347–2376. https://doi.org/10.1109/COMST.2015.2444095

[9] Shi, W., & Dustdar, S. (2016). The promise of edge computing. IEEE Computer, 49(5), 78–81. https://doi.org/10.1109/MC.2016.145

[10] McElhannon, J., Zhang, Q., Chen, M., & Li, L. (2018). Future edge cloud and edge computing. IEEE Network, 32(1), 96–101. https://doi.org/10.1109/MNET.2018.1700207

[11] Khan, M. A., & Salah, K. (2018). IoT security: Review, blockchain solutions, and open challenges. Future Generation Computer Systems, 82, 395–411. https://doi.org/10.1016/j.future.2017.11.022

[12] Russell, S. J., & Norvig, P. (2021). Artificial intelligence: A modern approach (5th ed.). Pearson.

[13] Perera, C., Liu, C. H., & Jayawardena, S. (2015). The emerging Internet of Things marketplace from an industrial perspective. IEEE Transactions on Emerging Topics in Computing, 3(4), 585–598. https://doi.org/10.1109/TETC.2015.2390034

[14] Manyika, J., Chui, M., Bughin, J., Dobbs, R., Bisson, P., & Marrs, A. (2017). Artificial intelligence: The next digital frontier. McKinsey Global Institute.

[15] Bonomi, F., Milito, R., Zhu, J., & Addepalli, S. (2012). Fog computing and its role in the Internet of Things. In Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing (pp. 13–16). ACM. https://doi.org/10.1145/2342509.2342513

[16] Kshetri, N. (2017). Can blockchain strengthen the Internet of Things? IT Professional, 19(4), 68–72. https://doi.org/10.1109/MITP.2017.3051335

[17] Riggins, F. J., & Wamba, S. F. (2015). Research directions on the adoption, usage, and impact of the Internet of Things. Journal of the Association for Information Systems, 16(8), 525–541.