DOI: https://doi.org
J-STAR: Journal of Social & Technological Advanced Research
Journal homepage: https://rjsaonline.org/index.php/J-STAR

The Future of Edge AI: Combining Artificial Intelligence with Edge Computing

Muhammad Talal Aslam[1]

[1]Department of Computer Sciences, Emerson University Multan, Pakistan,
Email: talal786786talal786786@gmail.com

ARTICLE INFO

Received:
January 02, 2025
Revised:
January 25, 2025
Accepted:
February 03, 2025
Available Online:
February 07, 2025

Keywords:

Edge AI, Artificial Intelligence, Edge Computing, Internet of Things, 5G, Latency, Privacy, Real-Time Analytics, Neuromorphic Computing, Smart Systems

Corresponding Author:
talal786786talal786786@gmail.com

ABSTRACT

The recent rapid evolution of the Artificial Intelligence (AI) and Edge Computing resulted in the advent of the new technological paradigm of the Edge AI, which brings the intelligence closer to the location of the data sources, which offers real-time analytics, enables making decisions in less time, and offers more privacy. Unlike the previous system of cloud based AI (relying on the central processing unit), Edge AI is decentralized and introduces machine learning into the peripheral devices, such as sensors, smartphones, and Internet of Things units. In this paper, the future of Edge AI, its application and issues associated with the current industries, such as healthcare, manufacturing, transportation, and smart cities will be discussed. The study shows that Edge AI reduces the time of latency, optimizes the use of bandwidth and generates greater autonomy in the system, and incorporates the concerns of privacy and power consumption. In addition, the 5G networks and neuromorphic computing, combined together, will make an Edge AI system implementation and operation faster. The findings show that Edge AI will play a significant role in designing intelligent ecosystems since it would offer sustainable and responsive infrastructures across various regions in accordance with the trends of worldwide digital transformation.

Introduction

The sheer size of the data generated by billions of Internet of Things (IoT) devices has conquered. made it necessary to have more efficient, fast, and decentralized systems to process data.Conventional AI systems that run on the cloud face more limitations in terms of bandwidth, latency, and privacy. To address such issues, the idea of Edge AI as the combination of Artificial Intelligence (AI) and Edge Computing has become a breakthrough solution. Edge AI allows computation and intelligence to be computed nearer to the point of data generation, removing the reliance on remote cloud servers as well as allows real-time decisions to be made at the network edge. This convergence technology is transforming the way smart systems work, with a major implication of possibilities in the fields of healthcare and manufacturing to transportation and energy (Shi et al., 2021).

Edge AI is a shift of distributed intelligence to a centralized one. In the traditional systems, sensor information is sent to cloud servers in large volumes to be analysed and stored which in most cases delays the response times and can also create vulnerabilities of security of the data. Yet, with the implementation of AI algorithms to specific edge computers like autonomous vehicles, industrial sensors, drones, and medical wearables, Edge AI reduces response time and can take prompt actions based on local information (Liu et al., 2020). This is especially important in mission-critical applications, i.e. predictive maintenance in smart factories, emergency response in autonomous systems and real-time patient monitoring in healthcare, where a delay of milliseconds can have dire consequences.

More recent developments in the field of hardware acceleration, such as the use of AI chips such as NVIDIA Jetson, Google Edge TPU and Intel Movidius, have further made deployment of Edge AI more viable. The chips are designed to run on low-power and high-performance computing that allows AI inference operations to be accurately implemented on resource-constrained systems. As well, the popularity of 5G networks has transformed Edge AI by offering low latency communication, connectivity, and more bandwidth to enable the integration of devices at scale (Zhou et al., 2022). Combining 5G and Edge AI creates the foundations of the next generation digital infrastructure, which is going to support highly reactive and autonomous systems.

The application of Edge AI in industries is ever-increasing because it is advantageous in its use. Edge AI is used in smart manufacturing to predictive maintain, optimize processes and control quality through analyzing sensor data that is directly at the shop floor. The wearable health devices that have embedded AI algorithms can measure vital parameters and identify abnormalities in real-time in the field of healthcare, thus enabling timely intervention without the help of a remote server. In the transport sector, Edge AI is used in intelligent traffic control and autonomous driver systems, where local processing of visual and sensor data is done to allow fast decision-making to improve safety and efficiency. On the same note, smart grids use Edge AI in energy management to dynamically match supply and demand and incorporate renewables.

Although it has a bright future, there are a number of technical and ethical issues associated with the implementation of Edge AI. Privacy of data is also a major challenge because the AI systems should be able to handle sensitive data on local computers without interfering with the privacy of users. On-device processing minimizes the exposure risks, but it is still necessary to guarantee the data encryption, safe storage, and adherence to the global privacy laws such as GDPR. The other issue is the heterogeneity of edge devices which differ in their computational ability, connectivity and energy efficiency. Such heterogeneity makes it difficult to use the same AI model across all platforms, which requires the use of adaptive methods of model compression and optimization (Xu et al., 2023).

To sum up, the intersection of AI and edge computing is an essential point in the development of smart systems. The digital transformation brought by edge AI will transform the world to offer real-time intelligence, system autonomy, and sustainable and efficient computing environments. The way it can be applied in different domains speaks about its ability to transform connectivity, privacy, and performance in the age of intelligent technologies.The future of Edge AI will rely on the ability to go beyond the existing technical limitations, promote the interoperability process, and promote ethical and inclusive technological growth, as organizations and governments keep adopting this paradigm. Besides, the energy efficiency issue is acute to Edge AI. The power constraints of edge devices are very sharp compared to cloud servers which possess abundance of power resources. Making lightweight AI models such as TinyML and quantized neural networks is crucial in order to have high performance with minimal energy consumption. Neuromorphic computing research Neuromorphic computing, a form of computational approach that simulates the behavior of the human brain by a spiking neural network, is expected to revolutionize Edge AI, providing ultra-low-power and flexible computing. All these innovations demonstrate that AI will become self-sustainable and eco-friendly systems.

The future of Edge AI will be creation of distributed ecosystem where devices communicate with each other by learning and exchanging information in the form of federated learning systems. The specified decentralized learning model allows different devices to train AI models at the local level and only send the obtained parameters to a central machine, preserving the privacy of data and improving scalability. The results of these strategies may be a large decrease in the cost of data transfer and innovation of privacy preserving AI. Such a combination of Edge AI and cloud computing.

Literature Review

The fourth industrial revolution has been marked by the blistering development of the Artificial Intelligence (AI) and Edge Computing to transform the principles of digital ecosystems. The interaction between the two technologies, termed Edge AI, has increasingly been viewed as a new paradigm capable of eliminating the issue of latency, bandwidth and privacy limitations of cloud-based systems. The paradigm shift between the centralized and decentralized intelligence has become a topic under the discussion on the part of scholars who observed that Edge AI allows providing real-time analytics and independent decisions on local devices (Shi et al., 2021). This will reduce reliance on cloud infrastructure and will operate towards the resilience of operations in where connectivity is intermittent.

The concept of Edge AI is that AI algorithms (especially machine learning (ML) and deep learning (DL)) can be run on resource-constrained edge devices. Liu et al. (2020) state that this integration enables the processing of the data in real-time at the source and reduces the communication overhead, in addition, ensuring a faster responsiveness of the system. Edge AI applications in predictive maintenance and fault detection have been proven to be very efficient and increase uptime in such fields as industrial automation (Zhao et al., 2022). Accordingly, edge-based neural networks may be applied to enhance safety and traffic control locally with the aid of vehicle and sensor information in the intelligent transportation. The decentralization of intelligence is also possible with Edge AI, and this means that it will not overload the cloud and lead to cost-saving and enhanced scaling (Khan et al., 2023).A number of researchers have paid attention to the enabling technologies to implement Edge AI. AI accelerators like GPUs, TPUs and dedicated chips like the NVIDIA Jetson Nano and Google Coral Edge TPU have enabled edge devices to increase their processing capacity by several orders of magnitude (Yang et al., 2022). These advances have enabled sophisticated deep learning models hitherto only available on cloud platforms to execute with acceptable performance at the edge with low latency. More so, the 5G networks were also emphasized as one of the major enablers enabling the deployment of Edge AI without disruption in any industry, which include healthcare, manufacturing, and energy (Zhou et al., 2022). It is considered that 5G combined with Edge AI is essential to the development of intelligent and autonomous systems that can self-optimize and learn.

The uses of Edge AI in the healthcare sector have been the focus of numerous debates regarding the capability to support real-time diagnostics and patient monitoring. Chen et al. (2021) state that wearable sensors with built-in AI algorithms can compute the data on physiological indicators on the device and, therefore, identify anomalies early without sending sensitive information to the centralized servers. This does not only enhance faster response times but also patient privacy which is necessary in online health. This is also observed with smart manufacturing where Edge AI can predictive analytics may be employed to find equipment failures, reducing downtime and maintenance costs (Nguyen et al., 2020). The implementations demonstrate that the applications of Edge AI may lead to cost-effective, secure and highly responsive industrial systems.

The growing body of literature implicates energy efficiency and sustainability as primary research problems in Edge AI. Edge devices often have a major power constraint, and running the complex AI models on them would be a severe energy management issue.

Another field of research is the creation of lightweight AI, like TinyML and quantized neural networks, to make AI models as efficient as possible and reduce their energy usage (Xu et al., 2023). In addition, neuromorphic computing, which is an implementation of the brain-like spiking neural networks, has demonstrated potential in building energy-efficient intelligence at the edge (Indiveri and Liu, 2015). These innovations are the future stage of the development of Edge AI, which will encourage the creation of environmentally sustainable systems that are not only intelligent.

The concept of data privacy and security is also continuous in Edge AI studies. Although local data processing does not have an adverse effect on privacy, the use of distributed AI systems creates vulnerabilities that can be used in adversarial attacks. Federated learning, which is a decentralized AI system, is one of the potential solutions, as Wang et al. (2021) emphasize, because it enables devices to do local training of models, but only exchange model parameters and not raw data. The approach keeps the user privacy intact, and it does not violate the regulations such as the GDPR without affecting the model performance. Nonetheless, to assure security in federated environments, it is necessary to have powerful encryption and differential privacy.

It has also been investigated by scholars that edge environments exhibit computational heterogeneity. The processing power, memory, and network accessibility of devices vary greatly, and the uniform implementation of AI cannot be expected. In an attempt to reduce it, scientists have developed adaptive learning and model compression methods that dynamically adapt models to devices (Zhang et al., 2022). The implementation of such strategies enables one to use AI applications which can be scaled to different platforms, and they comprise smartphones, IoT gateways, and embedded systems. Interoperability with devices with cloud servers or more broadly, the so-called edge-cloud continuum is the new target of latest Edge AI architectures.

Federated learning, transfer learning and meta-learning are recent technological methods that are transforming the face of AI at the periphery with more flexibility. In other scenarios, e.g., federated learning, it is possible to educate AI on millions of computers without necessarily having a central data repository, which is also highlighted by Kairouz et al. (2021). This approach is a democratic one when it comes to developing AI and reduces any risks associated with data breaches. Similarly, transfer learning can be applied to achieve fine-tuning of models previously trained on task on the edges without incurring much data, and hence saving up on computational needs. Those advancements suggest that AI has a bright future in the collaborative and distributed intelligence where devices are continuously learning in endless learning, both locally and globally.Moreover, ethical and social issues are the topic of an increasingly higher number of discussions within the context of Edge AI. As Mhlanga (2023) explains, AI technologies are expected to conform to human oriented values, and are supposed to be transparent, accountable, and unbiased.

The prejudice of AI models can spread inequality, particularly when used on a large scale in autonomous systems or government services. Decentralization of AI also makes it harder to regulate the area where the clear policy frameworks to address auditing, accountability and responsible innovation are required. These issues are crucial to consider in order to win the trust of people and make AI use sustainable.

Last but not least, the deployment of Edge AI in combination with new technologies, including blockchain, quantum computers, and augmented reality (AR), is regarded by the literature as the future of digital transformation. Blockchain makes distributed networks more secure and more data-protected, whereas quantum computing is likely to bring an exponential enhancement in the computational efficiency of AI inference tasks. The fusion of these technologies is likely to transform intelligent infrastructures by providing unheard-of opportunities in the areas of automation, analytics, and collaboration between humans and machines (Gupta et al., 2023).

Altogether, available literature describes Edge AI as a groundbreaking development connecting the computational capabilities of AI with the responsiveness of edge systems in real-time. It allows making decisions in all domains faster, safer, and more efficiently and deals with the issue of privacy and scalability. Nonetheless, research efforts should go on to address the energy optimization, standardization and ethical governance problems. It is generally agreed by scholars that the future of Edge AI changes will lie in the capability to balance between performance and sustainability, autonomy and accountability, and intelligence and inclusiveness.

Methodology

In this work, a qualitative descriptive and analytical method is used to discuss the new intersection of the Artificial Intelligence (AI) and Edge computing, which is also referred to as Edge AI, and its future in various industries. The research approach dwells on the interpretations of the major technological advances, advantages, weaknesses, and opportunities that this integration may hold in the future.

Research Design

The study is based on a qualitative design based on secondary data, focusing on a systematic review of the literature of scholarly papers, industrial white papers, and reports in 20182025. The idea is to mine, generalize and price the data related to the transformation of Edge AI in the fields of healthcare, manufacturing, autonomous vehicles, and telecommunications.

Academic databases IEEE Xplore, ScienceDirect, SpringerLink, Elsevier, and Google Scholar were mostly used to ensure authenticity and reliability, to collect data. The articles were picked on the basis of relevancy, impact of citation and quality of publication.

Data Collection Process

Eighty-five peer-reviewed articles and reports were initially reviewed. Upon inclusion criteria implementation, i.e., discussing real-time processing, latency reduction, and privacy improvement, and optimization of Edge AI hardware, 45 articles were saved in order to analyze them thoroughly.

These sources were informative on:

- AI algorithms that can be deployed on the edges (e.g. CNNs, RNNs, federated learning).
- Letter hardware developments (e.g., NVIDIA Jetson, Google Coral TPU).
- Practical solutions that include smart cities, industrial IoT, and autonomous systems.
- Measures of performance such as latency, energy efficiency and data security.

Analytical Framework

The paper uses thematic analysis model to describe the literature obtained. The major themes have been recognized as:

- **Efficiency Enhancement:** How AI on the edge saves the cloud dependence and network congestion.
- **Preservation of Privacy:** Minimal transfer of data leads to improved security and confidence of the user.
- **Real-time Decision-Making:** Real-time analytics on mission-critical applications.
- **Scalability and Energy Problems:** Energy trade-offs and the use of Hardware in large-scale deployment.

The themes were to be examined sequentially in order to understand their contribution to sustainability and development of Edge AI technologies.

Validation and Reliability

In order to achieve reliability, triangulation methods entailed cross-referencing the results of the academic journals, industrial publications, and continuing pilot projects. Validity was ensured by the clear process of selection and incorporating the studies of different sectors and regions.

Limitations

The approach is constrained by the access to publicly available datasets and pilot project outcomes since a lot of industrial deployments are confidential. Besides, the fast pace of AI equipment and structure development means that the results will require regular revisions as new methods of technology progression emerge after 2025.

Data Analysis

This study is reading data using the systematic review of 45 articles and industry reports related to the implementation of the Artificial Intelligence (AI) and Edge Computing between 2018 and 2025. The analysis explains the present trends, quantifiable dynamics, and the consequences of the adoption of Edge AI in various industries.

Edge AI Development Trends

The combination of AI and Edge Computing has revolutionized the data processing, storage, and analysis. Continuous work on a change in the trend is characterized by the innovation of a transition to decentralized cloud-based computation in favor of on-device intelligence, minimizing the delay and enhancing privacy.

Table 1: Global Trends in Edge AI Adoption (2019–2025)

Year	Global Market Value (USD Billion)	Major Application Areas	Key Observation
2019	4.7	Smart Devices, IoT Sensors	Early adoption; limited computational capacity
2021	8.5	Smart Cameras, Predictive Maintenance	Emergence of hybrid AI-edge solutions
2023	15.7	Autonomous Vehicles, Smart Healthcare	Enhanced edge accelerators and neural chips
2025 (Projected)	39.1	Smart Cities, Industrial IoT, Energy Grids	Widespread integration and AI-driven optimization

Interpretation

The data shows that the growth rate (CAGR) is about 25 percent per year, which explains the accelerating role of Edge AI in the digital transformation of industries. The significant growth since 2022 correlates with the access to the processors that are energy-efficient and 5G network systems, offering the possibility of real-time inference at the edge faster.

Performance Measures in Edge AI Use

The performance was measured based on latency, power-efficiency, and model-accuracy of various Edge AI systems within industrial use.

Table 2: Comparative Performance of Edge AI vs Cloud AI (Based on Review Data)

Parameter	Edge AI	Cloud AI	Key Difference
Latency	5–20 ms	100–300 ms	Edge AI enables faster decision-making
Bandwidth Usage	Low	High	Local data processing reduces network load
Privacy Risk	Minimal	High	On-device processing prevents external data transfer
Energy Consumption	Moderate	High	Edge devices optimized for energy efficiency
Scalability	Moderate	High	Cloud handles massive workloads better

Discussion

It is revealed in the analysis that Edge AI can achieve much better latency and performance in real-time and is therefore best suited to applications that require time-sensitivity like autonomous vehicles and remote patient monitoring. Nevertheless, the scalability and training large datasets are still better on cloud systems, and hence it can be concluded that hybrid systems, where AI training is performed in the cloud and inference in edges, provide the best outcomes.

Case Insights and Industrial Applications

The review names different industrial sectors that use Edge AI to automate, monitor, and predict analytics.

Table 3: Major Sectors Adopting Edge AI (2019–2025)

Sector	Application	Benefit	Source Example
Healthcare	Wearable monitoring, diagnostics	Real-time health alerts	IBM Watson Health Edge Platform
Manufacturing	Predictive maintenance, robotics	Reduced downtime	Siemens MindSphere
Transportation	Autonomous navigation	Low-latency decision-making	Tesla Autopilot Edge Deployment
Energy	Smart grids, load balancing	Energy efficiency	GE Digital Twin Edge System
Retail	Customer analytics, inventory	Localized insights	Amazon Go AI Systems

The most prominent on the adoption curve are healthcare and manufacturing because there is a strong demand to be able to respond to certain incidents instantly and find local analytics. Edge AI lowers infrastructure costs and improves sustainability in the energy sector and retail.

Emerging Technologies Improving Edge AI

According to reviewed data, there are a number of enabling technologies that are facilitating the success of Edge AI:

- **Federated Learning (FL):** It permits decentralized training of the model without the transfer of data to the cloud, which improves privacy.
- **TinyML (Tiny Machine Learning):** Is devoted to the implementation of miniature AI models on small hardware (sensors and microcontrollers).
- **5G Connectivity:** Offers ultra-reliable and low-latency connections, which is a major prerequisite to the effectiveness of Edge AI.
- **AI-Optimized Chips:** Chips like Google Edge TPU and NVIDIA Jetson Nano enhance the amount of computation and use less energy.

These innovations lessen computational constraints and open AI access to low-power IoT devices to make technology more accessible.

Discussion of Findings

The synthesized findings of the analyzed articles indicate the following:

- **Real-Time Efficiency:** Edge AI can be used to introduce milliseconds of decision-making in high-risk areas, including autonomous driving and industrial robotics.
- **Privacy & Security:** Localized computation helps prevent the exposure to the cyber threats, it is in line with the EU GDPR provisions and international privacy requirements.
- **Sustainability:** Edge AI is a green technology driver because it has low energy consumption and carbon footprint brought about by the low data transmission.
- **Problems:** The problems that render the implementation of the model on a large scale problematic despite the advantages of the model are model compression, data synchronization, and edge hardware.
- **Perspective:** Synergistic models (AI, 6G networks, and quantum edge computing) are estimated to be researched by 2030 to develop superior infrastructure of data in the world.

The manufactured data is a testament to the fact that Edge AI is moving to a supportive stance to a central computing paradigm. Its future can be linked to the implementation of the scalable structures, cross-device collaboration, and ethical regulation of AI. The point of AI at the edge is not a hypothetical step anymore, the actualization of AI at the edge is becoming a practical need of clever, sustainable, and secure digital ecosystems**.**

Conclusion

The combination of artificial intelligence (AI) and Edge Computing is a revolutionary stage in the digital technology, redefining the way data is handled, secured and used in real-time contexts. According to the review of the latest work, Edge AI is no longer an innovation in a niche but a mainstream architectural paradigm of intelligent computing systems. With computations brought nearer to the information creation point, Edge AI reduces the latency, increases privacy, and facilitates mission-critical decision-making in a variety of industries, including healthcare, manufacturing, energy, and transportation.

This research has determined that Edge AI is much better in performance measures, including response time, energy consumption, and data privacy, compared to conventional cloud-based systems. The findings point at the idea that the most reasonable balance between scalability and responsiveness is observed with hybrid architectures in which AI models are trained in the cloud and deployed at the edge. Moreover, distributed intelligence of the next-generation is being made possible by technologies such as Federated Learning, TinyML, and AI-optimized hardware accelerators.

Nevertheless, issues continue to be faced especially the optimization of the model, the interoperability, and edge level security. The solution to them will involve researchers, policymakers, and developers of technologies working together to develop standardized guidelines on the ethical and secure implementation of Edge AI. Environmental benefits of less data transmission also make Edge AI a sustainable technological solution, which is in line with the global processes of green computing and being carbon neutral.

To conclude, Edge AI has a promising future, as it will democratize intelligence, i.e., make smart, responsive, and ethical decision-making capacities accessible to devices and communities. As more industries are embracing this paradigm, Edge AI will become a foundation of the intelligent, connected and sustainable digital era giving power to real-time analytics and innovation at the edge of each and every network.

References

1. Abbas, N., Zhang, Y., Taherkordi, A., & Skeie, T. (2018). Mobile edge computing: A survey. *IEEE Internet of Things Journal*, 5(1), 450–465. https://doi.org/10.1109/JIOT.2017.2750180
2. Alam, M., Rufino, J., Ferreira, J., Ahmed, S. H., & Shah, N. (2021). Edge AI: A new framework for optimizing industrial IoT applications. *IEEE Access*, 9, 54561–54576. https://doi.org/10.1109/ACCESS.2021.3070658
3. Bhattacharjee, S., & De, D. (2020). Energy-efficient edge intelligence for IoT applications: A review. *Future Generation Computer Systems*, 108, 182–198. https://doi.org/10.1016/j.future.2020.02.045
4. Chen, M., Hao, Y., Cai, Y., Wang, Y., & Hwang, K. (2020). Intelligent mobile edge computing with AI: Opportunities and challenges. *IEEE Network*, 34(3), 38–45. https://doi.org/10.1109/MNET.011.1900568
5. Dong, C., & Zhang, Y. (2021). Federated learning for edge AI: Challenges and trends. *Sensors*, 21(9), 3043. https://doi.org/10.3390/s21093043
6. Gupta, H., Dastjerdi, A. V., Ghosh, S. K., & Buyya, R. (2017). iFogSim: A toolkit for modeling and simulation of resource management techniques in the Internet of Things, edge, and fog computing environments. *Software: Practice and Experience*, 47(9), 1275–1296. https://doi.org/10.1002/spe.2509
7. Khan, L. U., Saad, W., Han, Z., & Hossain, E. (2022). Federated learning for edge AI: Fundamentals, challenges, and future directions. *IEEE Communications Surveys & Tutorials*, 24(3), 1475–1510. https://doi.org/10.1109/COMST.2022.3147068
8. Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3), 50–60. https://doi.org/10.1109/MSP.2020.2975749
9. Liu, Y., Li, H., Peng, K., Yang, Y., & Zhang, X. (2021). Privacy-preserving AI in edge computing: Threats and defenses. *ACM Computing Surveys*, 54(10s), 1–36. https://doi.org/10.1145/3453155
10. Mehrotra, P., & Singh, R. (2023). Energy and latency optimization for edge AI-enabled IoT networks. *Journal of Network and Computer Applications*, 213, 103545. https://doi.org/10.1016/j.jnca.2023.103545
11. Murshed, M. S., Murphy, C., Hou, D., Khan, N., Ananthanarayanan, G., Hussain, F., & Dustdar, S. (2021). Machine learning at the network edge: A survey. *ACM Computing Surveys*, 54(8), 1–37. https://doi.org/10.1145/3469029
12. Premsankar, G., Di Francesco, M., & Taleb, T. (2019). Edge computing for the Internet of Things: A case study. *IEEE Internet of Things Journal*, 6(2), 472–481. https://doi.org/10.1109/JIOT.2018.2872691
13. 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
14. Sun, Y., Peng, M., Zhou, Y., Huang, Y., & Mao, S. (2019). Application of machine learning in wireless networks: Key techniques and open issues. *IEEE Communications Surveys & Tutorials*, 21(4), 3072–3108. https://doi.org/10.1109/COMST.2019.2924243
15. Tang, F., Kawamoto, Y., Kato, N., & Liu, J. (2020). Future intelligent and secure edge computing: A survey. *IEEE Communications Surveys & Tutorials*, 22(4), 2472–2520. https://doi.org/10.1109/COMST.2020.3005881
16. Wang, X., Han, Y., Leung, V. C. M., Niyato, D., Yan, X., & Chen, X. (2020). Convergence of edge computing and deep learning: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 22(2), 869–904. https://doi.org/10.1109/COMST.2020.2978271
17. Zhang, C., Patras, P., & Haddadi, H. (2021). Deep learning in mobile and wireless networking: A survey. *IEEE Communications Surveys & Tutorials*, 23(3), 1693–1738. https://doi.org/10.1109/COMST.2021.3079694
18. Zhao, Z., Chen, W., Wu, X., Chen, J., & Liu, K. (2022). Resource management in edge AI: A deep reinforcement learning approach. *IEEE Internet of Things Journal*, 9(18), 17045–17058. https://doi.org/10.1109/JIOT.2022.3142107