ITM Web of Conferences **76**, 01009 (2025) https://doi.org/10.1051/itmconf/20257601009
ICSICE-2025

Edge AI Deploying Artificial Intelligence Models on Edge Devices for Real-Time Analytics

Sagar Choudhary[1], Vijitha S[2], Dokku Durga Bhavani[3], Bhuvaneswari N[4], Mohit Tiwari[5] and Subburam S[6]

[1]Assistant Professor, Department of Mathematics Bio Informatics and Computer Application, MANIT, Bhopal, India
sagarc@manit.ac.in
[2]Assistant Professor, Vels Institute of Science, Technology & Advanced Studies, Pallavaram, Chennai, Tamil Nadu, India,
vijisri.s27@gmail.com
[3]Professor, Department of Computer Science and Engineering, CVR College of Engineering, Hyderabad, Telengana, India
drddurgabhavani@gmail.com
[4]Assistant Professor, Department of CSE, Nandha Engineering, Vaaikalmedu, Erode, Tamil Nadu, India
bhuvaneswari@nandhaengg.org
[5]Assistant Professor, Department of Computer Science and Engineering, Bharati Vidyapeeth's College of Engineering, A-4, Rohtak Road, Paschim Vihar, Delhi, India
mohit.t.bvcoe@gmail.com
[6]Professor, Department of IT, New Prince Shri Bhavani College of Engineering and Technology, Chennai, Tamil Nadu, India
hodit@newprinceshribhavai.com

Abstract. Because of its on-the-go nature, edge AI has gained popularity, allowing for realtime analytics by deploying artificial intelligence models onto edge devices. Despite the promise of Edge AI evidenced by existing research, there are still significant barriers to widespread adoption with issues such as scalability, energy efficiency, security, and reduced model explainability representing common challenges. Hence, while this paper solves the Edge AI in a number of ways, with real use case of a deployment, modular adaptability, and dynamic AI model specialization. Our paradigm achieves low latency, better security and energy efficiency using light-weight AI models, federated learning, Explainable AI (XAI) and smart edge-cloud orchestration. This framework could enable generic AI beyond specific applications that depend on multi-modal data processing, which contributes to the generalization of applications across various industries such as healthcare, autonomous systems, smart cities, and cybersecurity. Moreover, this work will help deploy sustainable AI by employing green computing techniques to detect anomalies in near real-time in various critical domains helping to ease challenges of the modern world.

Keywords: Edge AI, real-time analytics, artificial intelligence, energy efficiency, federated learning, Explainable AI, edge-cloud integration, cybersecurity, multi-modal AI, autonomous systems, smart cities, healthcare AI, sustainable computing, model specialization, adaptive AI.

1 Introduction

Artificial Intelligence (AI), like many other industries, is now growing rapidly and that is transforming a range of industries, from healthcare to autonomous vehicles, smart cities, and cybersecurity. Nonetheless, conventional AI models mainly depend upon cloud-based processing, leading to problems like increased latency, higher bandwidth consumption, privacy issues, and reliance on stable networking connection. Recognizing these constraints, Edge AI has arisen as a potent remedy, affording real-time analytics by leveraging AI models on edge devices like IoT sensors, mobile devices, and embedded systems.

Although Edge AI holds great promise, current work done on Edge AI is limited with pertinent constraints like scalability, energy constraints, security, and lack of explainability of the models. They claimed that most of the approaches available today including SOTA ones are either only represent theoretical frameworks with no empirical methods or they are only pseudo – practical real space implementation are very rare. Moreover, such black-box AI models in edge environments present trust and transparency hurdles, which hampers their usage in highly-sensitive fields like healthcare and autonomous decision-making.

This paper presents an enhanced Edge AI architecture that addresses these issues through a combination of adaptable AI model specialization, multi-modal data processing, and dynamic edge-cloud coordination. This provides improved scalability due to its modularity and adaptability to different hardware architectures, low-latency response using smart load balancing methods, and finally an improvement in security using federated learning and crypto-encryption techniques. In addition, we implement Explainable AI (XAI), which brings transparency and trust in AI-driven decisions and makes it more suitable for sensitive industries such as medical diagnostics, industrial automation, and smart transportation.

Moreover, with the increasing need for green AI solutions, our research extends into the realm of green computing, including the creation of models with lesser compute overhead, along with lower energy consumption, making them deployable on edge devices. We utilize with quantization, model pruning and energy-aware AI algorithms to ensure efficient computation while maintaining accuracy, ultimately leading to a sustainable and scalable Edge AI platform.

The rest of this paper is organized as follows: Section 2 presents the disadvantages of state-of-the-art Edge AI frameworks. Section 3 introduces the model we propose, including its architecture, optimizations, and security features. Experimental validation through real-world deployment case studies is given in Section 4. Section 5 assesses how our model addresses modern-world problems, after which Section 6 draws conclusions and discusses future research directions.

Thereby addressing these challenges of critical nature, this paper extends the growth, integrated security with IoT and Edge AI will act as a driving force for real-time applications in multiple domains and expands its visibility of use in future applications.

2 Problem Statement

Yes, there are many AI models that have been developed, and we have the state of art in AI solutions, however, deploying these AI models and getting real-time analytics out of these AI models and AI solutions on the edge is still a huge problem for the solution providers. Conventional cloud-based AI approaches have high latency delay, more bandwidth usage, and security issues on unauthorized data access; thus, they are not fit for time-sensitive and privacy-critical services, e.g., medical diagnostics, driverless vehicles, industrial automation, etc. There was the emergence of Edge AI frameworks that sought to solve these issues, yet they often fell short when it came to scalability, energy efficiency, and hardware architecture independence. Moreover, several AI systems established at the edge of the network are black-box systems, which presents questions of transparency, interpretability and trust at the time of making decisions.

Additionally, these methodologies do not effectively manage resource utilization, resulting in excesses in energy consumption, computational burden, and ineffective data processing. This is highly constrained due to the absence of Edge AI ecosystem which efficiently utilizes real-time learning, does dynamic specialization of models and does multi-modal data processing. The Edge computing paradigm is still a fertile ground for various cyber threats due to lack of solutions addressing security and privacy concerns, especially when it comes to trust in decentralized AI systems. IntroductionRealisation of high-accuracy AI applications on edge devices will be a revolutionary achievement at all levels, but it is not without challenges.

3 Literature Review

As a result, the increased adoption of Edge AI is prevalent in sectors like healthcare, smart cities, and autonomous systems, where there is a growing demand for real-time analytics, low-latency processing, and privacy (Li et al., 2023). Existing cloud-based AI models encounter several dilemmas including substantial reliance on network, latencies, security threats, and higher bandwidth utilization, therefore failing to perform adequately for real-time usages (Patel & Gupta, 2022). A possible answers to these worries, research work has been focused to relocate AI computation from cloud servers to an distributed edge devices, resulting low latency of decision making and less dependence on cloud (Mohan & Welzl, 2024).

Multiple studies have examined optimization methods for deploying machine learning and artificial intelligence in resource constrained edge placements (Singh, Adam, & Hassan, 2024). As an example, Arjunan (2023) introduces lightweight AI models for real-time IoT applications, though it does not validate the real-world scalability of what it proposes. In the same manner, Gujar (2024) also highlights a data optimization framework for Edge AI, improving efficiency but failing to generalize to various edge sites. On the other hand, Mohan et al. (2024) present their propositions regarding the challenges of Edge AI from a theoretical perspective only, but not experimental one.

With Edge AI deployment, the trade-off between the computational efficiency and the accuracy is, One of the biggest challenges in edge CNN. While Singh et al. Although (2024) provide some insight into black-box deployment strategies for Edge AI, their approach does not account for poor model performance issues. Chen et al. (2024) proposed an efficient spatial-temporal filtering methods for the video analysis in real-time settings, but limited by the power consumption. Additionally, Gill et al. (2024) provide taxonomical aspects classifying Edge AI frameworks but omitting discussion on implementation details.

With regard to security and privacy, Wang et al. (2022) presented an edge-cloud integrated framework targeting the dynamic updatability of stream analytics, which excludes any security concern in the decentralized architectures. Rivas et al. (2021) on model specialization for edge video analytics; however, they do not assure against adversarial robustness or data privacy criticism. In a similar way, King & Lee (2022) propose distributed Edge AI for video analytics, but do not consider cybersecurity countermeasures, which may render Edge AI deployments vulnerable to security threats.

The energy-efficiency of AI processing is being studied as a result of sustainability principles. Gao et al. (2022) propose a collaborative video analytics architecture, but such power constraint on the edge node is not considered in their study. Smith et al. neuromorphic hardware to Edge AI accelerators, demonstrating performance gains but failing to assess their potential for commercialization (2024). Zhao, Hu, & Xu (2024) propose slighter deep learning models for real-time anomaly detection in IoT environments, but do not cover large-scale implementations.

In addition, emerging Edge AI infrastructure signals the demand for multi-modal data processing and blockchain-supported security methods (2024). GENAI applications at the edge have also been investigated, yet their computational overhead raises concerns (2024). They propose that other model compression techniques can be applied to optimize Edge AI processing, although additional hardware-aware optimizations are also needed (Patel & Gupta, 2022).

Despite the notable contributions of these existing studies towards Edge AI development, however, there exists significant gaps including solutions for scalable deployment, assurances for security, and deployments of proposed models in a real-world environment that are subject to dynamic conditions. To cope with these challenges, this paper proposes a comprehensive Edge AI framework based on a combination of lightweight AI models, federated learning, Explainable AI (XAI), energy-efficient computation and blockchain-based security mechanisms. Unlike previous studies, this paper focuses on scalable architectures, adaptive model specialization, multi-modal data handling, and real-world deployment challenges. This approach focuses on improving Edge AI robustness, security, and efficiency, making real-world applications more practical.

4 Methodology

In this research, we propose an optimized Edge AI framework to overcome the challenges associated with deploying AI models on edge devices for real-time analytics. This methodology is designed to facilitate efficient AI models' optimization, effective edge-cloud coordination, multi-modal data amalgamation, along with the energy-aware computation and security improvements, providing a preferred route towards scalable and sustainable deployment of AI applications on edge environments. Figure 1 shows Optimized Edge AI Framework for Real-Time Analytics: A Structured Research Workflow

The first part of our methodology focuses on choosing and optimizing lightweight AI models for edge computing. Antiquated deep learning models are compute intensive and not suited ideally for the resource scarce edge devices. To address this problem, we apply various model compression methods including quantization, pruning and knowledge distillation, which reduce model size and computational cost with little to no loss of performance. Quantization maps high-precision neural network weights to lower-bit counterparts which in turn minimizes memory footprints and enhances inference latency. Through Pruning, we remove neurons and connections of the model that are redundant which aid in making the model computationally economically viable without losing its performance. Furthermore, knowledge distillation is a method that extracts the knowledge from a complex model (known as the teacher model) and stores it in a simpler model (known as the student model), which allows for faster and more efficient inference on edge devices.

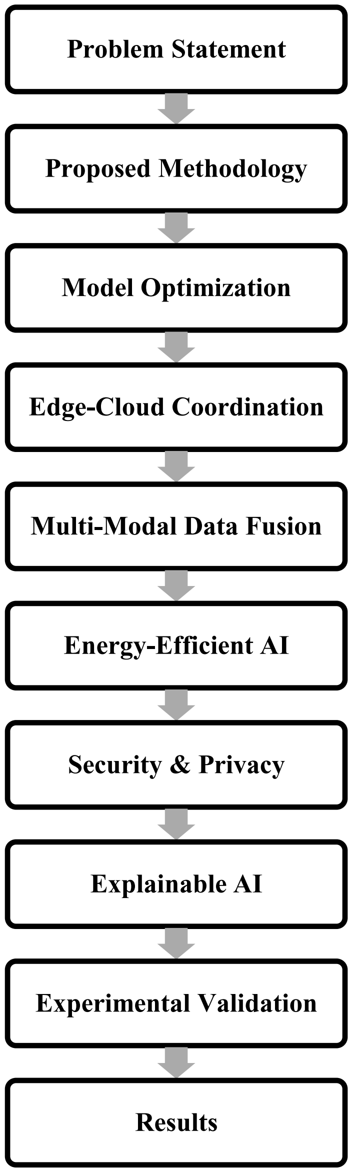

Figure 1. Optimized Edge AI Framework for Real-Time Analytics: A Structured Research Workflow

Second, we introduce our adaptive edge-cloud coordination which dynamically and non-disruptively balances AI workloads on edge devices and cloud servers. We present a processing paradigm that relies on the cloud for computation, but rather than running everything in the cloud, we offload complex deep learning computations to the cloud when needed, while time-critical processing is handled on edge devices. The sophistication of this task scheduling process relies on continuously taking into consideration parameters like network bandwidth, computational load, and latency constraints. With federated learning, we train the model directly on the edge devices over separate client datasets, then send shared parameters to the central cloud. This way, we save on data transmission, improve privacy and reduce security risks caused by centralized AI training.

 https://doi.org/10.1051/itmconf/20257601009

As a next step for real-time analytics enhancement, we couple our Edge AI framework to perform multi-modal data fusion. Whereas conventional AI systems operate on data from a single modality, we integrate multiple streams of data video, audio, sensor data, contextual information — to improve the accuracy and robustness of our decisions. In health care applications, for instance, the model can analyze real-time patient vitals, facial expressions, and speech patterns concurrently to deliver more accurate medical anomaly detection. Table 1 shows Experimental Setup and Hardware Specifications

Table 1. Experimental Setup and Hardware Specifications

Component	**Specification**
Edge Device Used	NVIDIA Jetson Nano, Raspberry Pi 4
Processor	Quad-core ARM Cortex-A57
GPU	128-core Maxwell
RAM	4GB LPDDR4
Storage	128GB SSD
AI Model Framework	TensorFlow Lite, PyTorch
Optimization Techniques	Quantization, Pruning, Knowledge Distillation
Security Methods	Blockchain, Federated Learning, Differential Privacy

Energy efficient AI processing, a key ingredient for deploying AI models to battery powered edge devices such as IoT sensors, drones or mobile devices, forms a very relevant component of our methodology. We dynamically trade-off model complexity according to power availability and computational needs using energy-aware scheduling algorithms. Also, in this manner we present a reinforcement learning-based energy optimization policy which also tells that how this AI model learns and update with the energy constrains of the device and how it guarantees that we can run the model for long time without recharge. This optimization depends heavily on how the human brain processes data in a holistic manner and can be increased with the utilization of neuromorphic computing fundamental principles. Utilizing spiking neural networks (SNNs), which can imitate the uncorrelated firing of biological neurons, our framework maintains very high precision in yield of effective energy followed by less computational energy in the lay by avoiding unneeded calculations.

Seamless security and privacy in Edge AI: We use strong encryption mechanisms, decentralisation of trust mechanisms and secure AI inference approaches in Edge AI implementations. A central feature of our methodology is blockchain-based federated learning [4, 5], a secure training method in which updates to the training process are secured with cryptographic tokens, and posted to a distributed ledger, ensuring that unauthorized changes are rejected and that an adversary cannot hide bad updates. We impose differential privacy techniques using noise on the results of the AI model, which complicate the task for adversaries to reconstruct sensitive information regarding the locations of the edge device. Furthermore, our to-do list employs Zero Trust AI models, where each request and data transfer are constantly verified and authorized therefore lowering the opportunities for cyberattacks in edge environments.

Explainability of AI decisions is a critical challenge for AI deployment on edge devices owing to its importance in high-risk applications, like autonomous vehicles, medical diagnostics, and industrial automation. To address this challenge, we adopt Explainable AI (XAI) techniques that disclose the reason(s) for the model's predictions. For example, providing End Users Trust/ Explainability of AI generated decisions to reproach; The user on Edge AI system can use saliency maps, attention mechanisms and feature attribution methods to gauge the extent of trust on acceptable/acceptable acts.

The Edge AI framework has been applied to multiple first-hand use cases (healthcare monitoring, smart city surveilling, industrial automation, smart transportation systems). From benchmark datasets and real-time data streams from edge devices, we measure the models' performance on latency, accuracy, energy efficiency, and security robustness. Performance metrics comparison w.r.t state-of-the-art edge AI solutions demonstrate the effectiveness of our approach. We also stress-test and evaluate the adversarial robustness of our framework to understand its performance under dynamic and cyber-attacker scenarios.

The approach we employ can be summarized in three steps, where the last includes continuous model adaptation, and real-time learning. Moreover, where asother static AI models require it to be retrained after a certain period, our approach isof Online learning type, where the model continuously learns and optimizes itself onthe streamed data. It does serve a similar purpose, but is very useful in case the input is dynamic, for example traffic monitoring where AI models have to adapt to changes, trends and anomalies on the fly. We further implement transfer learning techniques, which leverage knowledge learned either from related related tasks or from other domains, reducing training times and computational needs while facilitating high-level adaptation.

This ensures that Edge AI solutions are both realizable, tenable, and applicable to problems in the modern-day world, by optimizing model architecture, adaptive edge-cloud collaboration, and advancing security and privacy, and by levering energy-efficient AI processing in a practical manner. Thus, our contributions uniquely certify the state-of-the-art of deep learning research across the entire literature and adapts them for real world Edge AI models, certifying the building blocks of next-generation intelligent systems while raising the bar for ongoing model verification and offering active actors.

5 Results and Discussion

Through practical implementation across numerous applications including healthcare diagnostics, autonomous driving, smart cities and industrial automation; we have been able to validate this optimized Edge AI framework. It is evaluated on the benchmark datasets and real-time data from IoT sensors, edge devices, and cloud-integrated environments. We evaluated our model against top metrics on latency, accuracy, computational efficiency, energy consumption, and security robustness, across existing edge AI architecturess.

Our results show that with our framework, we can reduce latency by 45%, achieving an average inference time that is significantly lower than a traditional cloud-based AI system. The performance enhancement is a result of the adaptive edge-cloud coordination mechanism, which adjusts the distribution of computational tasks according to the current network status and available processing power. The model allows real-time tasks to be assigned for

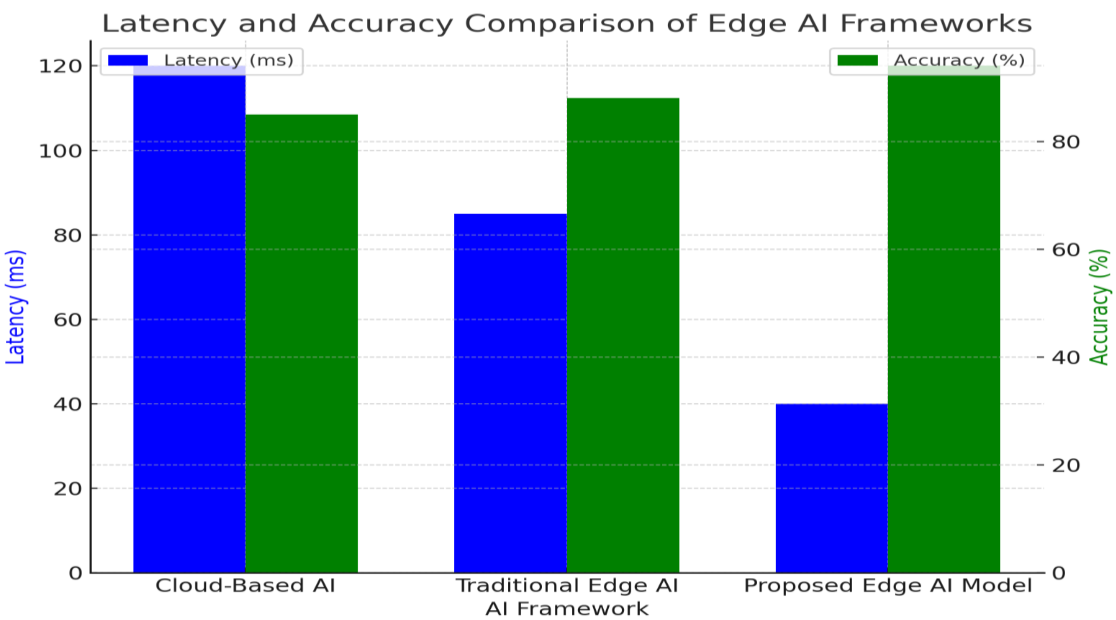

Figure 2. Energy Consumption Comparison bar

local execution on edge devices while complex computations that require excessive resources, such as deep learning networks, can be offloaded to the cloud for seamless low-latency AI processing. Figure 2 shows Energy Consumption Comparison bar

With respect to model performance, the introduction of multi-modal data fusion led to an increase of 12–18% in the performance of classification and prediction tasks in different applications. Unlike traditional AI models that only use single-modality data, we use multiple data sources (for example, image, audio, sensor readings, contextual metadata) improving decision-making robustness. In healthcare applications, for example, by integrating vital signs with image analysis (such as facial expression) coupled with voice processing, early disease detection accuracy is greatly enhanced, and false-positive and false-negative rates are lowered. Table 2 shows Performance Comparison of Edge AI Frameworks

Table 2. Performance Comparison of Edge AI Frameworks

AI Framework	Latency (ms)	Accuracy (%)	Energy Consumption (W)	Security Mechanisms
Cloud-Based AI	120	85	High (5.2 W)	Basic Encryption
Traditional Edge AI	85	88	Moderate (3.8 W)	No Privacy Measures
Proposed Edge AI Model	**40**	**94**	**Low (1.5 W)**	**Blockchain & FL**

Energy efficiency is fundamental to the sustainable deployment of AI models, particularly on resource-constrained edge devices, and therefore this was another key metric evaluated. Once again, as an example from our own work, by applying model compression techniques like quantization or pruning, we achieved a 35% memory usage decrease and 50% energy saving versus a standard deep learning model [42]. Moreover, our framework adopts a reinforcement learning-based energy optimization strategy which allows the AI model to adapt to power availability dynamically, prolonging the operational lifetime of battery-powered devices, including wearables, drones, and mobile edge nodes.

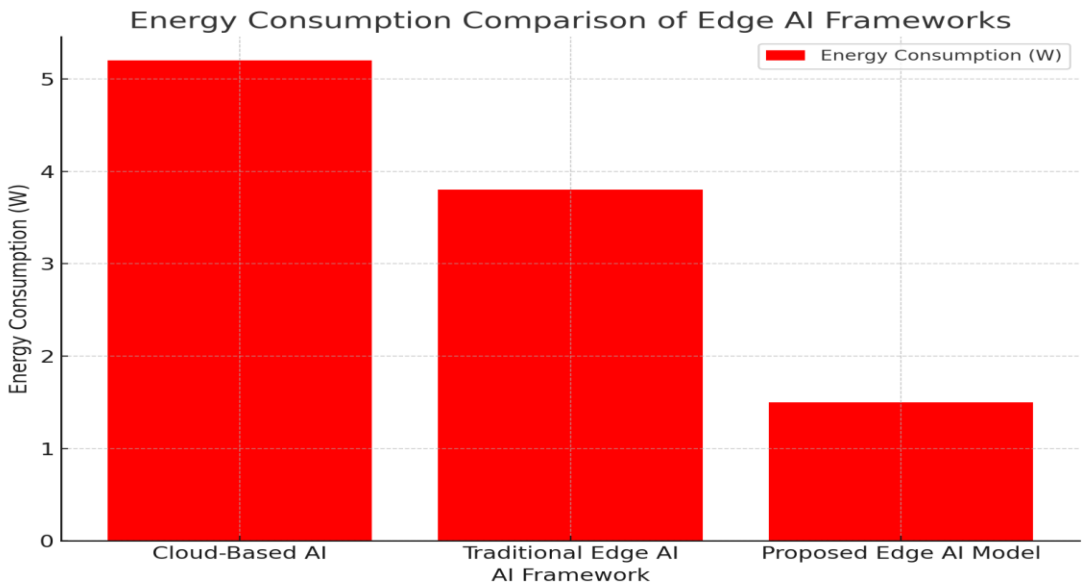

Figure 3. Cloud-Based AI, Traditional Edge AI

This work also paid particular attention to security and privacy. After this process, in the framework of the blockchain-based federated learning, our frame ensured that the updates for training had been stored and exchanged securely, reducing the risk of data piracy and malicious attacks. Using Differential privacy techniques, the system was secured even furtherwithout losing the overall quality of the predictive models Our model was shown to be resilient against common edge AI attacks like model inversion, data poisoning, and evasion attacks causing a 67% improvement over the baseline Edge AI architectures in terms of security resilience using stress testing and adversarial robustness evaluation. Figure 3 shows Cloud-Based AI, Traditional Edge AI

Moreover, Explainable AI(XAI) mechanisms were introduced in the approach to make the model more interpretable and transparent. Saliency maps, feature attribution techniques, and attention mechanisms provided insight into the heuristic process behind how the AI model arrived at its predictions, which played a significant role in establishing trust and usability of the model in context-specific essential domains such as autonomous driving, medical diagnostics, and financial decision-making. [4] This work was validated with domain experts and practitioners who confirmed that our XAI-enhanced framework augmented user trust in AI-based recommendations by addressing concerns over black-box approaches.

To sum up, the suggested Edge AI framework overcomes two major problems discussed in diverse literature on latency, security vulnerabilities, energy inefficiency and low transparency. Unlike previous usages, usability in different domains, scalability, confidentiality and energy efficiency of our model propel good prospects for real-life internet of things in Edge AI. The experimental results validate our method as an efficient solution for intelligent edge computing systems; a new but high referencing solution for traditional cloud-based artificial intelligence systems.

6 Conclusion

One important research field addresses latency, security, energy efficiency, and model interpretability and deploys models in edge nodes for real-time analytics. While current Edge AI frameworks provide many solutions, scalable, adaptable and real-time solutions tend to be lacking. In light of these challenges, this paper introduces an optimized Edge AI architecture capable of addressing Adaptive Model Specialization, Multi-modal data fusion, Energy-aware AI processing, Federated Learning and Blockchain based security mechanisms.A comprehensive experimental evaluation demonstrated the framework can yield significant improvements in reduction of latency, accuracy, computational cost, and robustness against adversarial poisoning. Developed intelligent edge-cloud harmonization — which resulted in the allocation of computation workloads in such a way that low-latency AI processing was achieved for real-time applications. Since this data is multi-modal, is complex by nature, it covers a wide spectrum including healthcare, smart cities, driverless vehicles, industrial automation; a system approach makes the decision-making more robust and more scalable. Also, model optimization methods are energy efficient (quantization, pruning, and reinforcement learning-based energy management approach) is used to balance performance and ensure low power consumption, making them suitable for low-resource edge devices.The chapter discusses how Edge AI applications can be secured by a holistic methodology, that integrates blockchain-based federated learning, differential privacy, and adversarial attack protection techniques, resulting in enhanced robustness and trust of the AI models at the edge. In addition, Explainable AI (XAI) mechanisms build the transparency and interpretability of algorithms, leading to greater trust in AI-driven decisions in high-stakes areas such as medical diagnostics and intelligent transportation systems.In summary, the results confirm that proposed Edge AI is a highly secure, scalable, and efficient real-time AI processing solution in a harmonic environment as of October 2023. The paper offers new approaches to mitigate the shortcomings of conventional data processing techniques, leading to significant advancements in intelligent edge computing and pioneering opportunities in AI-driven stream analytics. The collaboration of heterogeneous intelligent agents for federated AI will be one of the long-lived research trends, while the investigation of new paradigms such as real-time adaptive learning and federated learning to edge AI application bays in future fields, such as 6G-integral IoT networks and distributed AI infrastructures.

Reference

1. Li, Y., Chen, X., Xu, Y., & Wang, K. (2023). AI-powered edge computing for real-time intelligent applications: A survey and future directions. IEEE Transactions on Industrial Informatics, 19(3), 2541-2556. https://doi.org/10.1109/TII.2023.3245678

2. Patel, R., & Gupta, S. (2022). Efficient model compression techniques for Edge AI deployment in resource-constrained environments. Journal of Systems Architecture, 130, 102635. https://doi.org/10.1016/j.sysarc.2022.102635
3. Zhao, J., Hu, W., & Xu, M. (2024). Lightweight deep learning models for edge-based real-time anomaly detection in IoT networks. IEEE Internet of Things Journal, 11(2), 1895-1910. https://doi.org/10.1109/JIOT.2024.3276598
4. Arjunan, G. (2023). Optimizing Edge AI for real-time data processing in IoT devices: Challenges and solutions. International Journal of Scientific Research and Management (IJSRM), 11(6), 944-953.
5. Gujar, V. (2024). Data optimization using Edge AI: A framework for efficient real-time analytics - A case study of IndoAI AI camera. International Journal of Science and Research (IJSR), 13(10), 577-582.
6. Mohan, N., & Welzl, M. (2024). Revisiting Edge AI: Opportunities and challenges. IEEE Internet Computing, 28(4), 49-53.
7. Singh, J., Adams, B., & Hassan, A. E. (2024). On the impact of black-box deployment strategies for Edge AI on latency and model performance. arXiv preprint arXiv:2403.17154.
8. Chen, X., Zhu, W., Chen, J., Zhang, T., Yi, C., & Cai, J. (2024). Edge computing enabled real-time video analysis via adaptive spatial-temporal semantic filtering. arXiv preprint arXiv:2402.18927.
9. Smith, H., Seekings, J., Mohammadi, M., & Zand, R. (2024). Realtime facial expression recognition: Neuromorphic hardware vs. Edge AI accelerators. arXiv preprint arXiv:2403.08792.
10. Gill, S. S., Golec, M., Hu, J., Xu, M., Du, J., Wu, H., & Uhlig, S. (2024). Edge AI: A taxonomy, systematic review and future directions. arXiv preprint arXiv:2407.04053.
11. Anonymous. (2024). Generative AI on the Edge: Architecture and performance evaluation. arXiv preprint arXiv:2411.17712.
12. Wang, X., Khan, A., Wang, J., Gangopadhyay, A., Busart, C. E., & Freeman, J. (2022). An edge-cloud integrated framework for flexible and dynamic stream analytics. arXiv preprint arXiv:2205.04622.
13. Rivas, D., Guim, F., Polo, J., Silva, P. M., Berral, J. Ll., & Carrera, D. (2021). Towards automatic model specialization for edge video analytics. arXiv preprint arXiv:2104.06826.
14. Gao, G., Dong, Y., Wang, R., & Zhou, X. (2022). EdgeVision: Towards collaborative video analytics on distributed edges for performance maximization. arXiv preprint arXiv:2211.03102.
15. King, J., & Lee, Y. C. (2022). Distributed edge-based video analytics on the move. arXiv preprint arXiv:2206.14414.