Wisdom Journal of Science and Technology 1(2), 98-105, September 2025
https://doi.org/10.64848/WJST.1.2.2025.98-105

WISDOM JOURNAL of Science & Technology

REVIEW ARTICLE

Exploring the potential of adaptive computing architectures in building efficient infrastructure systems for artificial intelligence

Arunkumar Narayanan[1*] and J. Meenakumari[2]

[1]Research Scholar, International School of Management Excellence (ISME), Research Centre, Bangalore
Affiliated with the University of Mysore. Email: arunkumar.phd20@isme.in
[2]Research Supervisor, International School of Management Excellence (ISME), Research Centre, Bangalore
Affiliated with the University of Mysore. Email: j_meenakumari@yahoo.com

Abstract: The accelerating demands of artificial intelligence workloads necessitates a paradigm shift in computing infrastructure, moving beyond traditional static and monolithic architectures towards more dynamic and adaptive solutions. This research delves into the potential of adaptive computing architectures to address the increasing computational requirements of AI, focusing on how these architectures can optimize resource utilization, enhance performance, and reduce operational costs by dynamically reconfiguring hardware resources to match the specific needs of AI algorithms, adaptive computing offers a compelling alternative to fixed hardware platforms, promising significant improvements in efficiency and scalability. This paper explores various adaptive computing technologies, including Edge AI, field-programmable gate arrays (FPGA), configurable system-on-chips, and software-defined hardware, analyzing their strengths and weaknesses in the context of AI applications. Further, explored the impact of adaptive computing on key AI domains such as deep learning, machine vision, and natural language processing, highlighting the significance of potential for customized hardware acceleration, and auto scaling resources to meet data processing, availability, provisioning of resources to unlock new levels of performance and energy efficiency and provides a comprehensive overview of the challenges and opportunities associated with deploying adaptive computing in AI infrastructure, offering insights into future research directions and practical considerations for adoption.

Keywords: *Artificial Intelligence, Enterprise infrastructure systems, Computing architecture, Adaptive computing*

Introduction

The rapid increase of artificial intelligence and its subset solutions such as machine learning, deep learning, adoption across diverse industry sectors has provoked an unmatched demand for computational resources, requiring the development of efficient and scalable AI infrastructure [1].

The traditional computing architectures, more often depicted by their static and homogeneous nature, and are increasingly strained by the dynamic and heterogeneous demands of modern AI workloads [2].

Adaptive computing architectures, distinguished by their ability to dynamically reconfigure their hardware and software resources in response to changing application requirements, offer a promising solution to address these challenges [3].

These architectures are capable of optimizing resource allocation, enhancing performance, and improving energy efficiency across a wide spectrum of AI applications, ranging from deep learning and machine learning to computer vision and natural language processing [4].

[*]Corresponding author: International School of Management Excellence (ISME), Research Centre, Bangalore
Affiliated with the University of Mysore. Email: arunkumar.phd20@isme.in
Received: 26/07/25, Accepted: 29/09/25, Published Online: 07/10/25

The inherent flexibility of adaptive computing allows for the creation of systems that can operate effectively in resource-constrained environments, which is particularly relevant in edge computing scenarios and military applications where timely and precise decision-making are paramount [5]. By dynamically adjusting the hardware and software configurations, these architectures can cater to the specific needs of AI algorithms, leading to substantial improvements in computational speed and energy consumption [6].

Further, the adaptability of these systems facilitates the integration of novel AI algorithms and techniques, ensuring that the infrastructure remains relevant and efficient over time. The development of adaptive computing architecture necessitates a holistic approach, encompassing advancements in hardware design, software orchestration, and AI algorithm optimization. This exploration delves into the potential of adaptive computing architectures in constructing efficient AI infrastructure and systems, examining their underlying principles, key technologies, and potential applications.

Adaptive Computing for Enhanced AI Efficiency

The adaptive computing architectures hold immense potential for transforming AI infrastructure by providing the flexibility and efficiency required to address the evolving demands of AI workloads [7]. These architectures dynamically adjust their hardware and software resources to optimize performance, power consumption, and resource utilization, adapting to the specific needs of different AI tasks [8]. This adaptability is crucial in handling the diversity of AI algorithms and datasets, allowing for tailored solutions that maximize efficiency.

For example, in edge computing scenarios, where resources are limited and energy efficiency is paramount, adaptive computing can enable the deployment of complex AI models by dynamically optimizing the hardware configuration to match the computational requirements of the specific task [9], [10].

Adaptive learning systems, which tailor educational content and delivery methods to individual learner needs to exemplify the potential of AI to personalize and optimize experiences [11]. Moreover, adaptive computing can facilitate the integration of emerging AI techniques, such as neuromorphic computing and quantum computing, into existing AI infrastructure, enabling the development of more powerful and efficient AI systems [12].

The development of these architectures requires a multifaceted approach, considering hardware design, software orchestration, and AI algorithm optimization, and by dynamically adjusting hardware and software configurations, adaptive computing architectures can cater to the specific needs of AI algorithms, resulting in significant improvements in computational speed and energy consumption [13].

Edge Computing for AI

The emergence of edge AI, driven by advancements in AI efficiency and the proliferation of various Internet of Things devices, highlights the critical importance of adaptive computing in enabling real-time data management, processing, analysis and distribution of the edge network [14], [15].

The computational gap between resource-intensive deep learning algorithms and resource-constrained edge systems poses a significant challenge in edge intelligence [6]. The use of AI at the edge facilitates a multitude of applications, spanning from autonomous systems and human-machine interactions to IoT and beyond, capitalizing on the advantages of both edge and cloud computing [8].

Edge computing addresses the challenges of latency, bandwidth, and autonomy by processing data closer to the source of generation [16]. This paradigm reduces the need to transfer vast amounts of data to centralized cloud servers, minimizing latency and enabling real-time decision-making [15].

The edge computing augments data privacy and security by retaining the sensitive data within the chosen local network, and reduces the risk of unauthorized access, that include permission & privileges elevation and the data breaches. The design of lightweight deep neural networks to reduce the number of floating-point operations and parameters for execution on edge devices has become increasingly important [17], [18].

The edge computing for AI involves works based on certain permutations and computations, therefore the nearby users data gets cached at the network's edge location, versus instead of centralized location like colocation datacenter or the cloud service provider data center [19], [20]. Edge AI systems should also include mechanisms for continuous learning and adaptation, enabling them to improve their performance and accuracy over time [21]. The edge network typically comprises of many numbers of highly distributed and connected systems that are purposed for data collection & processing, analysis and caching to the nearest locations which are close to where the data is originally stored and managed [4].

Research Gaps

Although adaptive and edge computing presents numerous opportunities yet several research gaps exists and it must be addressed to fully realize its potential, such as developing energy-efficient AI algorithms and hardware architectures that can operate within the enhanced power budgets of edge devices is crucial [22], current edge AI systems often lack the ability to handle dynamic and unpredictable changes in the environment, which can degrade their performance and reliability [23], more research is needed to develop robust and adaptive edge AI systems that can handle unexpected events and maintain their accuracy and performance in dynamic environments [24], [25].

The need for specialized hardware to accelerate AI workloads on edge devices is also a major area of focus [26]. It is also essential to address the challenges related to data privacy and security in edge AI deployments, as edge devices are often deployed in unsecured environments and can be vulnerable to cyberattacks [27]. There is need to investigate how to effectively partition and distribute AI models across multiple edge devices, while minimizing communication overhead and ensuring data consistency [28]. The development of specialized hardware accelerators, such as GPUs, FPGAs, and ASICs, has played a crucial role in accelerating AI workloads [29], [30].

Challenges

Despite the significant potential of adaptive computing architectures in AI, several challenges remain that need to be addressed to fully realize their benefits and to enable their widespread adoption in AI infrastructure and systems. One of the major challenge is the complexity of designing and implementing these architectures, which requires expertise in hardware design, software engineering, and AI algorithm optimization [23].

The development of efficient and scalable software tools and frameworks for managing and orchestrating adaptive computing resources is crucial for simplifying the deployment and management of AI applications on these platforms. Another challenge is the need for standardized interfaces and protocols to facilitate interoperability between different adaptive computing components and platforms [31].

Further, ensuring the security and reliability of adaptive computing architectures is vital importance, as the systems are often deployed in critical applications where failures can bring significant and uneventful consequences [32]. The development of unique hardware technologies, such as neuromorphic computing and quantum computing, could further enhance the capabilities of adaptive computing architectures for AI.

Developing efficient and automated design tools that can translate high-level AI algorithms into optimized hardware configurations for adaptive computing architectures is crucial [33]. Furthermore, ensuring the security and reliability of adaptive computing systems is essential, as they are often deployed in critical applications where failures can have significant consequences. There is a huge problem of noisy and unlabeled data in heterogeneous platforms [3].

Furthermore, challenges from a company perspective can include: security, budget, lack of talent to implement and run AI, big data and data analytics, integration with existing systems, and procurement limitations [34].

One of the significant challenges is the lack of standardized programming models and software tools for adaptive computing architectures, which can hinder the development and deployment of AI applications on these platforms [35]. A standard programming model can significantly reduce the complexity of developing AI applications on adaptive computing architectures and facilitate code portability across different platforms.

Potential Applications

AI has a vast potential to be extremely helpful in manufacturing, especially in applications like predictive maintenance, quality assurance, and process optimization [36]. The integration of AI into manufacturing operations can lead to significant cost and efficiency benefits, especially when combined with edge computing and fog computing paradigms [1].

Adaptive computing architectures hold immense potential for revolutionizing AI infrastructure and systems across various domains. The convergence of technologies like the Internet of Things, AI, edge-fog-cloud computing, and block chain is driving digital transformation, opening doors for innovative applications in diverse sectors such as healthcare, finance, and industry 4.0 and many more [32].

Integrating AI technologies into existing systems is difficult due to a lack of understanding about what a particular type of AI technology can or cannot do [37]. Thus, the adaptive solutions must be integral to developing resilient, and scalable systems, aiding to the wider adoption of sustainability measures in the physical infrastructures such as data center and as well as digital eco system.

Findings

Adaptive computing allows direct stacking and provisioning of compute, storage, and acceleration resources, largely fostering the infrastructure utilization and minimizing idle hardware. This adaptable approach leads to the resource optimization by avoiding the allocation of static capacity in servers, enabling tailored and dynamic configuration for each application [38].

Data processing and availability requires an improved performance and latency, such as navigating the computation resources to make available the data sources, using adaptive platform solutions such as edge networks, FPGA, reduces latency distinctively and the network bottlenecks [39], especially on data intensive and sensitive operation's to real-time capacity and performance efficiency demands like streaming.

Applications and systems that are built on adaptive architectures could be reconfigured or upgraded in relevance to varying workload requirements without the need for onboarding immersive hardware, promoting scalability and future ready [40].

The architectures & application solutions that are built with artificial intelligence and provisioning of resources could be converted to imitation of human-centric technology systems [39], therefore, by assisting with the dependability, adaptability and robustness of critical infrastructure in domains like banking, finance, energy, healthcare, and FMCG.

Opportunities and Future Directions

As AI continues to permeate various aspects of modern life, the computational demands of AI models are growing exponentially. Adaptive computing architectures offer a promising solution to address the evolving demands of AI, as they can dynamically adjust their hardware resources and configurations to optimize performance for different AI tasks [30]. By dynamically mapping computational tasks to the most suitable hardware resources, adaptive computing architectures can achieve significant performance gains while reducing energy consumption and improving resource utilization. In advanced manufacturing, AI techniques are used to address unique manufacturing problems in order to significantly improve productivity, quality, flexibility, safety, and cost [41]. Also, the convergence of AI and Fog Continuum presents a massive opportunity for research and enterprise [25]. The new intelligent applications are feasible due to fog computing, which provides computing capabilities closer to edge devices, enabling new applications and services [42].

The convergence of adaptive computing and AI holds immense promise for creating intelligent systems that can learn, adapt, and evolve in response to changing environments and demands, and the integration of AI into edge devices is expected to grow significantly in the near future [43]. Overcoming these challenges will require collaborative efforts from researchers, engineers, and industry stakeholders to develop innovative solutions and standards that pave the way for the widespread adoption of adaptive computing architectures in AI infrastructure. Also, the development of a unified data model is needed at the edge nodes, as the data belongs to different subsystems [44]. The goal is to leverage the strengths of both cloud and edge computing paradigms to enable a wide range of AI applications with stringent latency, bandwidth, and privacy requirements [5]. By strategically allocating computational tasks between edge and cloud resources, the hybrid approach optimizes overall system performance while adhering to application-specific constraints. Also, edge AI systems should also include mechanisms for continuous learning and adaptation, enabling them to improve their performance and accuracy over time.

Further research is essential and should focus on developing robust encapsulation layer, micro-services (i.e., kubernetes), AI-driven, event-driven sustainable architecture [45], automated provisioning and optimization methods, to make adaptive computing accessible to architects, developers, and across the enterprises.

Discussion

Despite major advancement on the hardware platforms, software applications and tools, the encapsulation layer to fully utilize adaptive infrastructure remain a bottleneck [46], this the intelligence aware automated deployments, and effective compiler solutions are critical for mainstream.

The adaptive computing's principles should be applied across the industry and affiliated domains, to effectively address the uncertainties, resilience, and to the advancement of sustainable practices and future solutions such as AI/ML for DevOps must be engaged [47].

The adaptive architectures are dynamic in nature, which highlights the significance of blended hardware & software designs [48], cultivating the communication efficiently between the varying hardware and software components, which reinforces innovation in design methodologies and system architecture paradigms.

Implications

The adaptive computing architectures offers several theoretical and practical implications for the advancement of efficient infrastructure and some of these are listed in this section.

Theoretical Implications

Redefining architecture boundaries - The emergence of modern adaptive architecture, challenges the traditional disassociation between hardware and software, prompting for a re-assessment of their interrelation [49].

Observability metrics and paradigms – Instead prioritizing decisions based on performance metrics, the adaptive architectures spotlights for a comprehensive viewpoints, such as introducing new metrics – performance, energy efficiency, availability, reliability, and security first approach as core design principles [25].

This prompts for a new and enhanced research on contemporary areas, such as reconsidering the von Neumann model [33] and devising flexible interfaces that allow for dynamic, application-aware configuration.

Practical Implications

Dynamic resource allocation: The adaptive computing empowers real-time provisioning, with that the infrastructure can scale the resources elastically based on demand without involving any intervention. This greatly improvises the efficiency, cost effectiveness, and agreed SLA (service-level) for cloud platforms, edge devices, and deployments [39].

Energy and performance efficiency – The adaptive computing reduces both latency and power consumption, when the systems are configured to auto-scale the resources capacity such as compute and memory, which is an essential for the data intensive operations [44].

Security and Reliability - Infusing the security fist approach and resilience at the core of architecture level will be more relevant to design infrastructure which would dynamically respond to threats, failures, thus enhancing the security and resilience of critical systems [51].

Conclusion

As AI models become more complex and datasets continue to grow, the very need for specialized hardware accelerators has become increasingly apparent, with that the adaptive computing architectures offer a promising approach to building efficient AI infrastructure and systems, with the potential to significantly improve performance, reduce latency, and lower energy consumption for a wide range of AI applications by dynamically auto scaling the hardware resources through configurations to match the specific needs of each AI requirements, the adaptive architectures could very well overcome the limitations of standard CPU and GPU based systems [30].

The design and implementation of adaptive computing systems can be complex and require specialized expertise in hardware and software development.

The development of efficient and scalable programming models and tools for adaptive computing architectures is also essential to enable developers to easily create and deploy AI applications. AI technologies have witnessed exponential growth in recent years, driven by advancements in machine learning, particularly deep learning. However, implementing AI models on edge devices may cause considerable trials and encounters due to the inadequacy in available resources.

The trade-offs between adaptive computing architectures and traditional systems, showing that adaptive architectures are particularly well-suited for applications that require high performance, low latency, and energy efficiency yet time and again look out for opportunities to improvise [51] and thus, the arising key challenges must be addressed to fully realize the potential of adaptive computing architectures in building efficient AI infrastructure and systems.

References

[1] Alexander T et al., AutoML in the Age of Large Language Models: Current Challenges, Future Opportunities and Risks. https://doi.org/10.48550/arXiv.2306.08107

[2] Tadimalla et al., Tadimalla, S. Y., & Maher, M. L. (2024, May). Implications of identity in ai: Creators, creations, and consequences. In Proceedings of the AAAI Symposium Series (Vol. 3, No. 1, pp. 528-535) (2024) https://doi.org/10.1609/aaaiss.v3i1.31268

[3] Joloudari, J. H et al., The state-of-the-art review on resource allocation problem using artificial intelligence methods on various computing paradigms. (2022). https://doi.org/10.48550/arXiv.2203.12315

[4] Xiao, Y et al., Toward Self-Learning Edge Intelligence in 6G. IEEE Communications Magazine, 58(12), 34. Dec 2020, https://doi.org/10.1109/mcom.001.2000389

[5] Sander, J. W et al., On Accelerating Edge AI: Optimizing Resource-Constrained Environments. arXiv (Cornell University). (2025). https://doi.org/10.48550/arxiv.2501.15014

[6] Liu, D et al., (2020). Bringing AI To Edge: From Deep Learning's Perspective. Neurocomputing, 485, 297-320. May 2022. https://doi.org/10.1016/j.neucom.2021.04.141

[7] Briggs, F. C. Field Programmable Gate Array (FPGA) Design Strategies and Applications. AIAA Infotech@Aerospace 2007 Conference and Exhibit. https://doi.org/10.2514/6.2007-2958

[8] Fanariotis, A et al., Power Efficient Machine Learning Models Deployment on Edge IoT Devices. Sensors, 23(3), 1595. 2023. https://doi.org/10.3390/s23031595

[9] Bhattacharya, G. From DNNs to GANs: Review of efficient hardware architectures for deep learning. (2021). https://doi.org/10.48550/arxiv.2107.00092

[10] Isik, M et al., Design optimization for high-performance computing using FPGA. Research Square, 2023. https://doi.org/10.21203/rs.3.rs-2679691/v1

[11] Gligorea, I et al., Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review [Review of Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review]. Education Sciences, 13(12), 1216. Multidisciplinary Digital Publishing Institute, 2023. https://doi.org/10.3390/educsci13121216

[12] Wang, X., & Jia, W. Optimizing Edge AI: A Comprehensive Survey on Data, Model, and System Strategies. 2025. https://doi.org/10.48550/arxiv.2501.03265

[13] Biswas, P et al., . AI-Driven Approaches for Optimizing Power Consumption: A Comprehensive Survey. https://doi.org/10.48550/arxiv.2406.15732

[14] Gill, S. S et al., Edge AI: A Taxonomy, Systematic Review and Future Directions. https://doi.org/10.48550/arxiv.2407.04053

[15] Hayyolalam, V et al., Edge Intelligence for Empowering IoT-based Healthcare Systems. https://doi.org/10.48550/arxiv.2103.12144

[16] Garcia-Perez, A et al., Analysing Edge Computing Devices for the Deployment of Embedded AI. Sensors, 23(23), 9495. 2023. https://doi.org/10.3390/s23239495

[17] Choi, S., & Choi, J. (2023). Arithmetic Intensity Balancing Convolution for Hardware-aware Efficient Block Design. https://doi.org/10.48550/arxiv.2304.04016

[18] Zhou, Z et al., Edge Intelligence: Paving the Last Mile of Artificial Intelligence with Edge Computing. https://doi.org/10.48550/arxiv.1905.10083

[19] Richins, D et al., AI Tax: The Hidden Cost of AI Data Center Applications. https://arxiv.org/pdf/2007.10571

[20] Shi, Y et al., Communication-Efficient Edge AI: Algorithms and Systems. IEEE Communications Surveys & Tutorials, 22(4), 2167. 2020. https://doi.org/10.1109/comst.2020.3007787

[21] Deng, S et al., Edge Intelligence: The Confluence of Edge Computing and Artificial Intelligence. IEEE Internet of Things Journal, 7(8), 7457. 2020. https://doi.org/10.1109/jiot.2020.2984887

[22] Himeur, Y et al., Edge AI for Internet of Energy: Challenges and perspectives. Internet of Things, 25, 101035. 2023. https://doi.org/10.1016/j.iot.2023.101035

[23] Ding, A. Y et al., Roadmap for Edge AI: A Dagstuhl Perspective. https://doi.org/10.48550/arxiv.2112.00616

[24] Saha, S et al., AI-Driven Edge Computing for IoT: A Comprehensive Survey and Future Directions. International Journal of Advanced Research in Science Communication and Technology, 117. 2023. https://doi.org/10.48175/ijarsct-12921

[25] Talati, D. V. Scalable AI and data processing strategies for hybrid cloud environments. World Journal of Advanced Research and Reviews, 10(3), 482. 2021. https://doi.org/10.30574/wjarr.2021.10.3.0289

[26] Capra, M et al., An Updated Survey of Efficient Hardware Architectures for Accelerating Deep Convolutional Neural Networks. Future Internet, 12(7), 113. 2020. https://doi.org/10.3390/fi12070113

[27] Vyhmeister, E et al., A responsible AI framework: pipeline contextualisation. AI and Ethics, 3(1), 175. 2022. https://doi.org/10.1007/s43681-022-00154-8

[28] Jin, Y., & Yang, Z. Scalability Optimization in Cloud-Based AI Inference Services: Strategies for Real-Time Load Balancing and Automated Scaling. 2025. https://doi.org/10.48550/ARXIV.2504.15296

[29] Baczmanski, M et al., Implementation of a Perception System for Autonomous Vehicles Using a Detection-Segmentation Network in SoC FPGA. In Lecture notes in computer science (p. 200). Springer Science+Business Media. 2023. https://doi.org/10.1007/978-3-031-42921-7_14

[30] Zhao, T. FPGA-Based Machine Learning: Platforms, Applications, Design Considerations, Challenges, and Future Directions. Highlights in Science Engineering and Technology, 62, 96. 2023. http

[31] Biran, Y., & Kissos, I. Adaptive Orchestration for Large-Scale Inference on Heterogeneous Accelerator Systems Balancing Cost, Performance, and Resilience. 2025. https://doi.org/10.48550/ARXIV.2503.20074

[32] Gao, X et al., Artificial Intelligence Applications in Smart Healthcare: A Survey. Future Internet, 16(9), 308. 2024. https://doi.org/10.3390/fi16090308

[33] Peter S. Pacheco, Matthew Malensek, An Introduction to Parallel Programming (Second Edition). 2022. https://doi.org/10.1016/B978-0-12-804605-0.00009-9

[34] Véstias, M et al., Moving Deep Learning to the Edge. Algorithms, 13(5), 125. 2020. https://doi.org/10.3390/a13050125

[35] Tuli, S et al., AI augmented Edge and Fog computing: Trends and challenges. Journal of Network and Computer Applications, 216, 103648. 2023. https://doi.org/10.1016/j.jnca.2023.103648

[36] Plathottam, S. J et al., A review of artificial intelligence applications in manufacturing operations [Review of A review of artificial intelligence applications in manufacturing operations]. Journal of Advanced Manufacturing and Processing, 5(3). Wiley. 2023. https://doi.org/10.1002/amp2.10159

[37] Chen, M., & Décary, M. Artificial intelligence in healthcare: An essential guide for health leaders [Review of Artificial intelligence in healthcare: An essential guide for health leaders]. Healthcare Management Forum, 33(1), 10. SAGE Publishing. 2019. https://doi.org/10.1177/0840470419873123

[38] Zhang, M et al., Deep Learning in the Era of Edge Computing: Challenges and Opportunities. https://doi.org/10.48550/arxiv.2010.08861

[39] Bourechak, A et al., At the Confluence of Artificial Intelligence and Edge Computing in IoT-Based Applications: A Review and New Perspectives. Sensors, 23(3), 1639. Multidisciplinary Digital Publishing Institute. 2023. https://doi.org/10.3390/s23031639

[40] Joshi, P et al., Enabling All In-Edge Deep Learning: A Literature Review [Review of Enabling All In-Edge Deep Learning: A Literature Review] https://doi.org/10.48550/arxiv.2204.03326

[41] Arinez, J et al., Artificial Intelligence in Advanced Manufacturing: Current Status and Future Outlook. Journal of Manufacturing Science and Engineering, 142(11). 2020. https://doi.org/10.1115/1.4047855

[42] Rodríguez, I et al., Horizontally Distributed Inference of Deep Neural Networks for AI-Enabled IoT [Review of Horizontally Distributed Inference of Deep Neural Networks for AI-Enabled IoT]. Sensors, 23(4), 1911. Multidisciplinary Digital Publishing Institute. 2023. https://doi.org/10.3390/s23041911

[43] Yang, H., & Han, X. (2021). National Sports AI Health Management Service System Based on Edge Computing. Wireless Communications and Mobile Computing, 2021(1). https://doi.org/10.1155/2021/5536329

[44] Xu, D et al., Edge Intelligence: Architectures, Challenges, and Applications. https://doi.org/10.48550/arxiv.2003.12172

[45] Gao, W et al., Artificial Intelligence Accelerators Based on Graphene Optoelectronic Devices. Advanced Photonics Research, 2(6). 2021. https://doi.org/10.1002/adpr.202100048

[46] D'Hollander, E. Empowering Parallel Computing with Field Programmable Gate Arrays. In Advances in parallel computing. https://doi.org/10.3233/apc200020

[47] Liang, Q et al., AI on the Edge: Rethinking AI-based IoT Applications Using Specialized Edge Architectures. https://doi.org/10.48550/arxiv.2003.12488

[48] Mungoli, N. Scalable, Distributed AI Frameworks: Leveraging Cloud Computing for Enhanced Deep Learning Performance and Efficiency. https://doi.org/10.48550/arxiv.2304.13738

[49] Reuther, A et al., Survey and Benchmarking of Machine Learning Accelerators. https://doi.org/10.1109/hpec.2019.8916327

[50] Singh, R., & Gill, S. S. Edge AI: A survey. Internet of Things and Cyber-Physical Systems, 3, 71. 2023. https://doi.org/10.1016/j.iotcps.2023.02.004

[51] Radanliev, P., & Roure, D. D. Advancing the cybersecurity of the healthcare system with self-optimising and self-adaptative artificial intelligence (part 2). Health and Technology, 12(5), 923. 2022. https://doi.org/10.1007/s12553-022-00691-6