ai accelerator design

ai accelerator design is a critical area of development in modern computing, focusing on specialized hardware that enhances the performance of artificial intelligence workloads. As AI applications grow increasingly complex and data-intensive, traditional processing units like CPUs and GPUs often fall short in terms of efficiency and speed. AI accelerator design addresses these challenges by optimizing hardware architectures specifically for machine learning algorithms, neural networks, and deep learning tasks. This article explores the fundamentals of AI accelerator design, including its key components, architectural considerations, and the latest trends shaping the industry. Additionally, it discusses common challenges in designing AI accelerators and the future outlook for this rapidly evolving technology. The following sections provide a structured overview of these topics to offer a comprehensive understanding of AI accelerator design.

Fundamentals of AI Accelerator Design

AI accelerator design involves creating hardware solutions tailored specifically to the computational demands of artificial intelligence algorithms. These accelerators are engineered to handle matrix multiplications, convolutions, and other operations common in neural network computations with greater efficiency than general-purpose processors. The primary goal of AI accelerator design is to maximize throughput while minimizing power consumption and latency.

Purpose and Importance

The purpose of AI accelerator design is to enable faster and more energy-efficient AI processing. Traditional CPUs are not optimized for the parallelism required by AI workloads, and while GPUs offer improvements, AI accelerators push performance even further by specializing in AI-specific operations. This specialization is critical for applications such as autonomous vehicles, natural language processing, and real-time image recognition.

Types of AI Accelerators

There are several types of AI accelerators, each designed for different use cases and performance requirements. Common types include:

- **Application-Specific Integrated Circuits (ASICs):** Custom-built chips optimized for specific AI tasks, offering high performance and low power consumption.
- **Field-Programmable Gate Arrays (FPGAs):** Flexible hardware that can be reconfigured for various AI workloads with moderate performance and efficiency.
- **Graphics Processing Units (GPUs):** Originally designed for graphics, GPUs are widely used for AI due to their parallel processing capabilities.
- **Tensor Processing Units (TPUs):** Specialized ASICs developed primarily for deep learning tensor operations.

Key Architectural Components

The architecture of an AI accelerator significantly impacts its efficiency and performance. Understanding the key components involved in AI accelerator design is essential for optimizing hardware for artificial intelligence applications.

Processing Elements

Processing elements (PEs) are the fundamental computational units in AI accelerators. These elements execute arithmetic operations such as multiply-accumulate (MAC), which are critical to neural network computations. The number, structure, and interconnection of PEs determine the parallelism and throughput of the accelerator.

Memory Hierarchy

Efficient memory design is crucial in AI accelerator design because data movement often consumes more power than computation. Memory systems typically include multiple levels, such as on-chip registers, scratchpad memory, and off-chip DRAM, to balance speed and capacity. Optimizing data locality reduces latency and power consumption.

Interconnects and Dataflow

Interconnects enable communication between processing elements and memory modules. AI accelerator architectures often employ specialized dataflow strategies to optimize how data moves through the system. Examples include weight-stationary, output-stationary, and row-stationary dataflows, each designed to reduce data movement and enhance efficiency.

Design Considerations and Optimization Techniques

Several considerations must be addressed to achieve optimal AI accelerator design. These factors influence performance, power consumption, scalability, and compatibility with various AI models.

Power Efficiency

Power consumption is a critical design constraint, especially for edge devices and mobile applications. Designers employ techniques such as voltage scaling, clock gating, and approximate computing to reduce energy usage without significantly impacting accuracy or performance.

Scalability and Flexibility

AI models evolve rapidly, requiring accelerators to support various network architectures and sizes. Flexibility can be achieved through reconfigurable hardware or modular designs that allow scaling processing elements and memory to meet different workload demands.

Precision and Numerical Formats

Choosing the appropriate numerical precision is vital in AI accelerator design. Lower precision formats like INT8 or mixed precision can accelerate computation and reduce memory bandwidth while maintaining acceptable accuracy for many AI tasks. Designers must balance precision with performance requirements.

Optimization Techniques

Designers utilize various optimization techniques to maximize accelerator efficiency:

- Loop unrolling and pipelining to increase parallelism.

- Data quantization to reduce memory footprint.
- Memory access pattern optimization to minimize latency.
- Hardware-software co-design to align architecture with AI frameworks.

Challenges in AI Accelerator Development

Developing AI accelerators involves overcoming several technical and practical challenges that impact design complexity and deployment.

Balancing Performance and Energy Consumption

Achieving high computational throughput while maintaining low power consumption is a persistent challenge. Excessive energy use limits the applicability of accelerators in battery-powered or heat-sensitive environments.

Compatibility with Diverse AI Models

AI models vary widely in structure and computational patterns. Designing accelerators that efficiently support a broad range of models, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformers, requires adaptable architectures and software support.

Manufacturing and Cost Constraints

Custom AI accelerator chips must balance performance improvements against manufacturing costs and time-to-market. High design complexity and fabrication expenses can hinder widespread adoption, especially in cost-sensitive markets.

Software Ecosystem and Support

Effective AI accelerator design includes developing a robust software stack, including compilers, libraries, and frameworks that enable developers to optimize AI workloads for the hardware. Lack of mature software tools can limit the usability of accelerators.

Emerging Trends and Future Directions

The field of AI accelerator design is rapidly evolving, driven by advances in AI algorithms and hardware technologies. Several emerging trends are shaping the future landscape of AI accelerators.

Integration of AI Accelerators with System-on-Chip (SoC)

Modern designs increasingly integrate AI accelerators directly into SoCs alongside CPUs, GPUs, and other specialized units. This integration improves data exchange efficiency and reduces overall system latency.

Neuromorphic and Analog Computing

Research into neuromorphic computing and analog AI accelerators aims to mimic the human brain’s neural architecture for ultra-efficient processing. These designs promise significant energy savings and new computational paradigms.

Edge AI and TinyML

With the rise of edge computing, AI accelerators are being designed for low-power, compact devices capable of running machine learning models locally. TinyML focuses on enabling AI in microcontrollers and constrained environments.

Advanced Process Technologies

Adoption of cutting-edge semiconductor process nodes and 3D integration techniques allows AI accelerators to achieve higher density, better performance, and lower power consumption.

Frequently Asked Questions

What is an AI accelerator and why is it important in modern computing?

An AI accelerator is a specialized hardware designed to speed up artificial intelligence workloads, such as machine learning and deep learning tasks. It is important because it improves computational efficiency, reduces latency, and lowers power consumption compared to general-purpose processors, enabling faster and more efficient AI applications.

What are the key design considerations when developing an AI accelerator?

Key design considerations include optimizing for parallelism, memory bandwidth, power efficiency, scalability, and support for various AI models and precision formats. Designers also focus on minimizing data movement, integrating with existing systems, and ensuring programmability and flexibility.

How do AI accelerators handle different types of neural network architectures?

AI accelerators are designed with flexible compute units and memory hierarchies that can efficiently execute various neural network operations like convolutions, matrix multiplications, and activation functions. Some accelerators include configurable data paths or support multiple precision formats to accommodate different architectures such as CNNs, RNNs, and transformers.

What role does memory hierarchy play in AI accelerator design?

Memory hierarchy is critical in AI accelerator design as it affects data access speed and energy consumption. Efficient use of on-chip caches, scratchpads, and high-bandwidth memory reduces latency and power usage by minimizing expensive off-chip memory accesses, which is essential for high-performance AI computation.

How is power efficiency achieved in AI accelerator designs?

Power efficiency is achieved through specialized low-power hardware components, optimized dataflow architectures, reduced precision arithmetic, minimizing data movement, and dynamic voltage and frequency scaling. These techniques collectively reduce energy consumption while maintaining high computational throughput.

What are emerging trends in AI accelerator design for edge devices?

Emerging trends include designing compact, low-power accelerators that support on-device AI processing, integration of AI accelerators with sensors, use of neuromorphic computing principles, and development of heterogeneous architectures combining CPUs, GPUs, and AI-specific cores to enable real-time, efficient AI inference on edge devices.

Additional Resources

1. *AI Accelerator Architectures: Principles and Practice*
This book provides a comprehensive overview of the fundamental principles behind AI accelerator design. It covers hardware architectures, design trade-offs, and optimization techniques for deep learning workloads. Readers will gain insights into how accelerators improve performance and energy efficiency in AI applications.

2. *Designing Efficient AI Accelerators for Deep Learning*
Focused on the practical aspects of building AI accelerators, this book delves into the hardware-software co-design process. It explores dataflow models, memory hierarchies, and computation units tailored for neural networks. Case studies highlight real-world implementations and performance benchmarks.

3. *Hardware for AI: Accelerators and Beyond*
This title explores the evolving landscape of AI hardware, including GPUs, TPUs, and specialized accelerators. It discusses emerging technologies such as neuromorphic computing and in-memory processing. The book aims to equip readers with knowledge on selecting and designing hardware for various AI workloads.

4. *Deep Learning Accelerator Design: Architectures and Algorithms*
Merging algorithmic insights with hardware design, this book examines how deep learning models influence accelerator architecture. It covers convolutional neural networks, recurrent networks, and transformer models from a hardware perspective. Optimization strategies for latency and throughput are also discussed.

5. *Energy-Efficient AI Accelerator Design*
Addressing the critical aspect of power consumption, this book focuses on techniques to reduce energy usage in AI accelerators. Topics include voltage scaling, approximate computing, and thermal management. It is ideal for engineers seeking sustainable and high-performance AI hardware solutions.

6. *FPGA-Based AI Accelerators: Design and Implementation*
This book provides an in-depth guide to designing AI accelerators using field-programmable gate arrays (FPGAs). It covers hardware description languages, design flow, and performance optimization specific to FPGA platforms. Practical examples demonstrate how to implement various neural network architectures.

7. *AI Accelerator Design for Edge Computing*
Targeting the challenges of deploying AI at the edge, this book discusses compact and low-power accelerator designs. It explores trade-offs between computational capability and resource constraints in edge devices. Readers will learn about hardware considerations for real-time and privacy-sensitive applications.

8. *Neuromorphic Computing and AI Accelerators*

```
This book introduces neuromorphic architectures inspired by the human brain
for AI processing. It covers spiking neural networks, event-driven
computation, and emerging hardware designs. The text highlights how
neuromorphic accelerators offer new paradigms for energy-efficient AI.

9. AI Accelerator Design: From Algorithms to Silicon
Providing an end-to-end perspective, this book guides readers through the
entire process of AI accelerator development. Topics range from algorithmic
requirements and hardware architecture to physical design and fabrication. It
is a valuable resource for understanding the integration of AI models with
silicon technology.
```

Ai Accelerator Design

ai accelerator design: Artificial Intelligence and Hardware Accelerators Ashutosh Mishra, Jaekwang Cha, Hyunbin Park, Shiho Kim, 2023-03-15 This book explores new methods, architectures, tools, and algorithms for Artificial Intelligence Hardware Accelerators. The authors have structured the material to simplify readers' journey toward understanding the aspects of designing hardware accelerators, complex AI algorithms, and their computational requirements, along with the multifaceted applications. Coverage focuses broadly on the hardware aspects of training, inference, mobile devices, and autonomous vehicles (AVs) based AI accelerators

ai accelerator design: *AI-Enabled Electronic Circuit and System Design* Ali Iranmanesh, Hossein Sayadi, 2025-01-27 As our world becomes increasingly digital, electronics underpin nearly every industry. Understanding how AI enhances this foundational technology can unlock innovations, from smarter homes to more powerful gadgets, offering vast opportunities for businesses and consumers alike. This book demystifies how AI streamlines the creation of electronic systems, making them smarter and more efficient. With AI's transformative impact on various engineering fields, this resource provides an up-to-date exploration of these advancements, authored by experts actively engaged in this dynamic field. Stay ahead in the rapidly evolving landscape of AI in engineering with "AI-Enabled Electronic Circuit and System Design: From Ideation to Utilization," your essential guide to the future of electronic systems. !--[endif]--A transformative guide describing how revolutionizes electronic design through AI integration. Highlighting trends, challenges and opportunities; Demystifies complex AI applications in electronic design for practical use; Leading insights, authored by top experts actively engaged in the field; Offers a current, relevant exploration of significant topics in AI's role in electronic circuit and system design. Editor's bios. Dr. Ali A. Iranmanesh is the founder and CEO of Silicon Valley Polytechnic Institute. He has received his Bachelor of Science in Electrical Engineering from Sharif University of Technology (SUT), Tehran, Iran, and both his master's and Ph.D. degrees in Electrical Engineering and Physics from Stanford University in Stanford, CA. He additionally holds a master's degree in business administration (MBA) from San Jose State University in San Jose, CA. Dr. Iranmanesh is the founder and chairman of the International Society for Quality Electronic Design (ISQED). Currently, he serves as the CEO of Innovotek. Dr. Iranmanesh has been instrumental in advancing semiconductor technologies, innovative design methodologies, and engineering education. He holds nearly 100 US and international patents, reflecting his signifi cant contributions to the field. Dr. Iranmanesh is the Senior life members of EEE, senior member of the American Society for Quality, co-founder and Chair Emeritus of the IEEE Education Society of Silicon Valley, Vice Chair Emeritus of the IEEE PV chapter, and recipient of IEEE Outstanding Educator Award. Dr. Hossein Sayadi is a Tenure-Track Assistant Professor and Associate Chair in the Department of Computer Engineering and Computer Science at California State University, Long Beach (CSULB). He earned his Ph.D. in Electrical and Computer Engineering from George Mason University in Fairfax, Virginia, and an M.Sc. in Computer Engineering from Sharif University of Technology in Tehran, Iran. As a recognized researcher with over 14 years of research experience, Dr. Sayadi is the founder and director of the Intelligent, Secure, and Energy-Efficient Computing (iSEC) Lab at CSULB. His research focuses on advancing hardware security and trust, AI and machine learning, cybersecurity, and energy-efficient computing, addressing critical challenges in modern computing and cyber-physical systems. He has authored over 75 peer-reviewed publications in leading conferences and journals. Dr. Sayadi is the CSU STEM-NET Faculty Fellow, with his research supported by multiple National Science Foundation (NSF) grants and awards from CSULB and the CSU Chancellor’s Office. He has contributed to various international conferences as an organizer and program committee member, including as the TPC Chair for the 2024 and 2025 IEEE ISQED.

ai accelerator design: *Hardware Accelerator Systems for Artificial Intelligence and Machine Learning* , 2021-03-28 Hardware Accelerator Systems for Artificial Intelligence and Machine Learning, Volume 122 delves into artificial Intelligence and the growth it has seen with the advent of Deep Neural Networks (DNNs) and Machine Learning. Updates in this release include chapters on Hardware accelerator systems for artificial intelligence and machine learning, Introduction to Hardware Accelerator Systems for Artificial Intelligence and Machine Learning, Deep Learning with GPUs, Edge Computing Optimization of Deep Learning Models for Specialized Tensor Processing Architectures, Architecture of NPU for DNN, Hardware Architecture for Convolutional Neural Network for Image Processing, FPGA based Neural Network Accelerators, and much more. - Updates on new information on the architecture of GPU, NPU and DNN - Discusses In-memory computing, Machine intelligence and Quantum computing - Includes sections on Hardware Accelerator Systems to improve processing efficiency and performance

ai accelerator design: Artificial Intelligence Chips and Data: Engineering the Semiconductor Revolution for the Next Technological Era Botlagunta Preethish Nandan, 2025-05-07 The 21st century is witnessing a profound technological transformation, with artificial intelligence (AI) at its epicenter. As AI algorithms become increasingly sophisticated, their insatiable demand for processing power and data throughput is pushing the boundaries of what traditional computing infrastructures can offer. At the heart of this evolution lies the semiconductor industry—reimagining its core principles to engineer chips that are not only faster and more efficient but also intelligent and adaptable. This book is born out of the urgent need to explore the critical intersection between AI and semiconductor innovation. It provides a comprehensive view of how custom-designed AI chips—such as GPUs, TPUs, FPGAs, and neuromorphic processors—are redefining performance benchmarks and unlocking capabilities that were once the realm of science fiction. We delve into the fundamental principles behind AI-centric chip design, the data pipelines that feed them, and the architectural innovations enabling real-time learning, inference, and massive parallelism. From edge computing to hyperscale data centers, the book investigates how data movement, storage, and processing are being reengineered to support the next wave of AI applications, including autonomous systems, natural language understanding, predictive analytics, and more. Equally important, this work sheds light on the global semiconductor ecosystem, including the geopolitical, economic, and environmental factors shaping chip manufacturing and supply chains. As AI continues to permeate every sector—healthcare, finance, defense, education, and beyond—the role of AI chips becomes increasingly strategic. Whether you're a researcher, engineer, policymaker, or tech enthusiast, this book aims to equip you with a deep understanding of the technological forces propelling us into a new era of intelligent machines. It is both a chronicle of current breakthroughs and a roadmap for future innovation. Welcome to the frontier of AI and semiconductors, where data meets silicon to redefine what's possible.

ai accelerator design: Advanced Parallel Processing Technologies Chao Li, Zhenhua Li, Li Shen, Fan Wu, Xiaoli Gong, 2023-11-07 This book constitutes the refereed proceedings of the 15th International Symposium on Advanced Parallel Processing Technologies, APPT 2023, held in Nanchang, China, during August 4–6, 2023. The 23 full papers and 1 short papers included in this book were carefully reviewed and selected from 49 submissions. They were organized in topical sections as follows: High Performance Computing and Parallelized Computing, Storage Systems and File Management, Networking and Cloud Computing, Computer Architecture and Hardware Acceleration, Machine Learning and Data Analysis, Distinguished Work from Student Competition.

ai accelerator design: Artificial Intelligence in Wireless Sensors and Instruments Halit Eren, 2024-11-26 This book heralds a new era in instrumentation and measurements. It combines artificial intelligence (AI) and wireless communications technologies with instrumentation and measurement systems to function as a single unit. AI has advanced considerably due to deep learning utilizing artificial neural networks, availability of large and curated datasets, implementation of a new generation of fast processors having millions of transistors in chips, advanced algorithms, competitive commercial interests, and interests of governments to gain advantages. At the same time, new and highly advanced wireless technologies open new frontiers in communication systems, both technologically and in terms of applications aspects. Advanced technologies such as 5G and 6G networks enable easy use of communication systems by billions of people as well as by billions of machine-to-machine systems. In this book, the communication principles are explained and the implementation of AI on wireless networks is discussed. Many examples are provided. The author discusses instruments and instrumentation networks, modern sensors, and transducers in detail. AI is the technology humans have created where the machines do not only assist us but also think for us creatively in some cases, excelling humans thinking and reasoning. This book includes a chapter explaining how this is done, backed up with more than 50 figures. The security issues, fairness, efficiency, and social impact and acceptance of AI are highlighted. As explained in this book, AI and wireless communications are changing our lives in many ways, including entertainment, games, social interactions, medicine and healthcare, R&D, automated living, intelligent transport systems, finance and economy, and the Internet of Things.

ai accelerator design: Computational Science and Computational Intelligence Hamid R. Arabnia, Leonidas Deligiannidis, Farzan Shenavarmasouleh, Soheyla Amirian, Farid Ghareh Mohammadi, 2025-08-09 The CCIS book constitutes selected papers accepted in the Research Track on Education of the 11th International Conference on Computational Science and Computational Intelligence, CSCI 2024, which took place in Las Vegas, NV, USA, during December 11–13, 2024. The 26 full papers included in this book were carefully reviewed and selected from a total of 155 submissions. They were organized in topical sections on subject-specific education and curriculum design; education and artificial intelligence; teaching and learning strategies and related reserach studies.

ai accelerator design: *Embedded Machine Learning for Cyber-Physical, IoT, and Edge Computing* Sudeep Pasricha, Muhammad Shafique, 2023-10-09 This book presents recent advances towards the goal of enabling efficient implementation of machine learning models on resource-constrained systems, covering different application domains. The focus is on presenting interesting and new use cases of applying machine learning to innovative application domains, exploring the efficient hardware design of efficient machine learning accelerators, memory optimization techniques, illustrating model compression and neural architecture search techniques for energy-efficient and fast execution on resource-constrained hardware platforms, and understanding hardware-software codesign techniques for achieving even greater energy, reliability, and performance benefits. Discusses efficient implementation of machine learning in embedded, CPS, IoT, and edge computing; Offers comprehensive coverage of hardware design, software design, and hardware/software co-design and co-optimization; Describes real applications to demonstrate how embedded, CPS, IoT, and edge applications benefit from machine learning.

ai accelerator design: Complex, Intelligent and Software Intensive Systems Leonard Barolli, Tomoya Enokido, Isaac Woungang, 2025-07-18 This book aims to deliver a platform of scientific interaction between the three interwoven challenging areas of research and development of future ICT-enabled applications: Software intensive systems, complex systems, and intelligent systems. Software intensive systems are systems which heavily interact with other systems, sensors, actuators, devices, other software systems and users. More and more domains are involved with software intensive systems, e.g., automotive systems, telecommunication systems, embedded systems in general, industrial automation systems, and business applications. Moreover, the outcome of web services delivers a new platform for enabling software intensive systems. Complex systems research is focused on the overall understanding of systems rather than their components. Complex systems are very much characterized by the changing environments in which they act by their multiple internal and external interactions. They evolve and adapt through internal and external dynamic interactions. The development of intelligent systems and agents which can be characterized by ontologies and their logical foundations builds a fruitful impulse for both software intensive systems and complex systems. Recent research in the field of intelligent systems, robotics, neuroscience, artificial intelligence, and cognitive sciences are very important factors for the future development and innovation of software intensive and complex systems.

ai accelerator design: T-Bytes Hybrid Cloud Infrastructure IT-Shades, 2020-01-01 This document brings together a set of latest data points and publicly available information relevant for Hybrid Cloud Infrastructure. We are very excited to share this content and believe that readers will benefit immensely from this periodic publication immensely.

ai accelerator design: Proceedings of the 14th International Conference on Computer Engineering and Networks Guangqiang Yin, Xiaodong Liu, Jian Su, Yangzhao Yang, 2025-05-07 This conference proceedings is a collection of papers accepted for CENet2024 - the 14th International Conference on Computer Engineering and Networks, held in Kashi, China, 18-21 October 2024. The topics covered include Internet of Things and Smart Systems, Artificial Intelligence and Applications, Detection, Analysis and Application of Communication Systems, Cloud Computing and Security, and Medical Engineering and Information Systems. Each section of this book can serve as an excellent reference for industry practitioners, university faculty, research fellows, undergraduate and graduate students who wish to build a knowledge base of the latest advances and state-of-the-art practices in the topics covered. Using this knowledge, they will be able to design, implement and manage systems that are both complex and trustworthy. We would like to thank the authors for their hard work and dedication, and the reviewers for their efforts in ensuring that only the highest quality papers were selected. Without their contributions, the proceedings would not have been possible.

ai accelerator design: *Nanotechnology in Electronics for Miniaturized Devices and Circuits* Ms. Priyanka, Mr. Rajeev Ranjan, 2024-12-17 Nanotechnology in Electronics for Miniaturized Devices and Circuits the transformative impact of nanotechnology on electronic components, enabling ultra-compact, high-performance devices. This nanoscale materials, fabrication techniques, and emerging applications in transistors, memory storage, sensors, and quantum computing. It examines advancements in nanoelectronics, including carbon nanotubes, graphene-based circuits, and molecular electronics, highlighting their role in enhancing speed, efficiency, and energy consumption. With a focus on current research and future prospects, this serves as a comprehensive resource for students, researchers, and industry professionals navigating the rapidly evolving landscape of miniaturized electronic systems.

ai accelerator design: Energy Research Abstracts , 1989

ai accelerator design: Applied Cryptography and Network Security Workshops Jianying Zhou, Sridhar Adepu, Cristina Alcaraz, Lejla Batina, Emiliano Casalicchio, Sudipta Chattopadhyay, Chenglu Jin, Jingqiang Lin, Eleonora Losiouk, Suryadipta Majumdar, Weizhi Meng, Stjepan Picek, Jun Shao, Chunhua Su, Cong Wang, Yury Zhauniarovich, Saman Zonouz, 2022-09-23 This book constitutes the proceedings of the satellite workshops held around the 20th International Conference on Applied Cryptography and Network Security, ACNS 2022, held in Rome, Italy, in June 2022. Due to the Corona pandemic the workshop was held as a virtual event. The 31 papers presented in this volume were carefully reviewed and selected from 52 submissions. They stem from the following workshops: – AIBlock: 4th ACNS Workshop on Application Intelligence and Blockchain Security – AIHWS: 3rd ACNS Workshop on Artificial Intelligence in Hardware Security – AIoTS: 4th ACNS Workshop on Artificial Intelligence and Industrial IoT Security – CIMSS: 2nd ACNS Workshop on Critical Infrastructure and Manufacturing System Security – Cloud S&P: 4th ACNS Workshop on Cloud Security and Privacy – SCI: 3rd ACNS Workshop on Secure Cryptographic Implementation – SecMT: 3rd ACNS Workshop on Security in Mobile Technologies – SiMLA: 4th ACNS Workshop on Security in Machine Learning and its Applications

ai accelerator design: Big Data and Artificial Intelligence Anirban Dasgupta, Rage Uday Kiran, Radwa El Shawi, Satish Srirama, Mainak Adhikari, 2025-03-03 This book constitutes the proceedings of the 12th International Conference on Big Data and Artificial Intelligence, BDA 2024, held in Hyderabad, India, during December 17–20, 2024. The 16 full papers and 12 short papers presented here were carefully reviewed and selected from 106 submissions. These papers have been categorized under the following topical sections: Image Classification; Graph Analytics; Big Data Analytics; Applications; Data Science; Health-Care Analytics; eLearning; Prediction and Forecasting.

ai accelerator design: *Applied Reconfigurable Computing. Architectures, Tools, and Applications* Roberto Giorgi, Mirjana Stojilović, Dirk Stroobandt, Piedad Brox Jiménez, Ángel Barriga Barros, 2025-04-03 This book constitutes the proceedings of the 21st International Symposium on Applied Reconfigurable Computing. Architectures, Tools, and Applications, ARC 2025, held in Seville, Spain, during April 9–11, 2025. The 12 full papers presented in this book together with 1 short paper from the technical program were carefully reviewed and selected from 40 submissions. ARC 2025 covers a wide range of topics, including hardware acceleration, security and fault tolerance, energy-efficient architectures, and emerging applications in artificial intelligence and high-performance computing. The symposium fostered collaboration and pushed the boundaries of state-of-the-art research.

ai accelerator design: Computer Science and Education Wenxing Hong, Yang Weng, 2023-06-16 This three-volume set constitues selected papers presented during the 17th International Conference on Computer Science and Education, ICCSE 2022, held in Ningbo, China, in August 2022. The 168 full papers and 43 short papers presented were thoroughly reviewed and selected from the 510 submissions. They focus on a wide range of computer science topics, especially AI, data science, and engineering, and technology-based education, by addressing frontier technical and business issues essential to the applications of data science in both higher education and advancing e-Society.

ai accelerator design: Trends and Innovations in Information Systems and Technologies Álvaro Rocha, Hojjat Adeli, Luís Paulo Reis, Sandra Costanzo, Irena Orovic, Fernando Moreira, 2020-06-07 This book gathers selected papers presented at the 2020 World Conference on Information Systems and Technologies (WorldCIST'20), held in Budva, Montenegro, from April 7 to 10, 2020. WorldCIST provides a global forum for researchers and practitioners to present and discuss recent results and innovations, current trends, professional experiences with and challenges regarding various aspects of modern information systems and technologies. The main topics covered are A) Information and Knowledge Management; B) Organizational Models and Information Systems; C) Software and Systems Modeling; D) Software Systems, Architectures, Applications and Tools; E) Multimedia Systems and Applications; F) Computer Networks, Mobility and Pervasive Systems; G) Intelligent and Decision Support Systems; H) Big Data Analytics and Applications; I) Human–Computer Interaction; J) Ethics, Computers & Security; K) Health Informatics; L) Information Technologies in Education; M) Information Technologies in Radiocommunications; and N) Technologies for Biomedical Applications.

ai accelerator design: Bioinformatics in Agriculture Pradeep Sharma, Dinesh Yadav, R.K. Gaur, 2022-04-28 Bioinformatics in Agriculture: Next Generation Sequencing Era is a comprehensive volume presenting an integrated research and development approach to the practical application of genomics to improve agricultural crops. Exploring both the theoretical and applied aspects of computational biology, and focusing on the innovation processes, the book highlights the increased productivity of a translational approach. Presented in four sections and including insights from experts from around the world, the book includes: Section I: Bioinformatics and Next Generation Sequencing Technologies; Section II: Omics Application; Section III: Data mining and Markers Discovery; Section IV: Artificial Intelligence and Agribots. Bioinformatics in Agriculture: Next Generation Sequencing Era explores deep sequencing, NGS, genomic, transcriptome analysis and multiplexing, highlighting practices forreducing time, cost, and effort for the analysis of gene as they are pooled, and sequenced. Readers will gain real-world information on computational biology, genomics, applied data mining, machine learning, and artificial intelligence. This book serves as a complete package for advanced undergraduate students, researchers, and scientists with an interest in bioinformatics. - Discusses integral aspects of molecular biology and pivotal tool sfor molecular breeding - Enables breeders to design cost-effective and efficient breeding strategies - Provides examples ofinnovative genome-wide marker (SSR, SNP) discovery - Explores both the theoretical and practical aspects of computational biology with focus on innovation processes - Covers recent trends of bioinformatics and different tools and techniques

ai accelerator design: Number Systems for Deep Neural Network Architectures Ghada Alsuhli, Vasilis Sakellariou, Hani Saleh, Mahmoud Al-Qutayri, Baker Mohammad, Thanos Stouraitis, 2023-09-01 This book provides readers a comprehensive introduction to alternative number systems for more efficient representations of Deep Neural Network (DNN) data. Various number systems (conventional/unconventional) exploited for DNNs are discussed, including Floating Point (FP), Fixed Point (FXP), Logarithmic Number System (LNS), Residue Number System (RNS), Block Floating Point Number System (BFP), Dynamic Fixed-Point Number System (DFXP) and Posit Number System (PNS). The authors explore the impact of these number systems on the performance and hardware design of DNNs, highlighting the challenges associated with each number system and various solutions that are proposed for addressing them.

Related to ai accelerator design

Artificial intelligence | MIT News | Massachusetts Institute of 4 days ago AI system learns from many types of scientific information and runs experiments to discover new materials The new “CRESt” platform could help find solutions to real-world

Explained: Generative AI’s environmental impact - MIT News MIT News explores the environmental and sustainability implications of generative AI technologies and applications

Using generative AI, researchers design compounds that can kill Using generative AI algorithms, the research team designed more than 36 million possible compounds and computationally screened them for antimicrobial properties. The top

MIT researchers introduce generative AI for databases Researchers from MIT and elsewhere developed an easy-to-use tool that enables someone to perform complicated statistical analyses on tabular data using just a few

What does the future hold for generative AI? - MIT News Hundreds of scientists, business leaders, faculty, and students shared the latest research and discussed the potential future course of generative AI advancements during the

“Periodic table of machine learning” could fuel AI discovery After uncovering a unifying algorithm that links more than 20 common machine-learning approaches, MIT researchers organized them into a “periodic table of machine

Explained: Generative AI - MIT News What do people mean when they say “generative AI,” and why are these systems finding their way into practically every application imaginable? MIT AI experts help break down

A new generative AI approach to predicting chemical reactions The new FlowER generative AI system may improve the prediction of chemical reactions. The approach, developed at MIT, could provide realistic predictions for a wide

Photonic processor could enable ultrafast AI computations with Researchers developed a fully integrated photonic processor that can perform all the key computations of a deep neural network on a photonic chip, using light. This advance

AI simulation gives people a glimpse of their potential future self The AI system uses this information to create what the researchers call “future self memories” which provide a backstory the model pulls from when interacting with the user. For

Artificial intelligence | MIT News | Massachusetts Institute of 4 days ago AI system learns from many types of scientific information and runs experiments to discover new materials The new “CRESt” platform could help find solutions to real-world

Explained: Generative AI’s environmental impact - MIT News MIT News explores the environmental and sustainability implications of generative AI technologies and applications

Using generative AI, researchers design compounds that can kill Using generative AI algorithms, the research team designed more than 36 million possible compounds and computationally screened them for antimicrobial properties. The top

MIT researchers introduce generative AI for databases Researchers from MIT and elsewhere developed an easy-to-use tool that enables someone to perform complicated statistical analyses on tabular data using just a few

What does the future hold for generative AI? - MIT News Hundreds of scientists, business leaders, faculty, and students shared the latest research and discussed the potential future course of generative AI advancements during the

“Periodic table of machine learning” could fuel AI discovery After uncovering a unifying algorithm that links more than 20 common machine-learning approaches, MIT researchers organized them into a “periodic table of machine

Explained: Generative AI - MIT News What do people mean when they say “generative AI,” and why are these systems finding their way into practically every application imaginable? MIT AI experts help break down

A new generative AI approach to predicting chemical reactions The new FlowER generative AI system may improve the prediction of chemical reactions. The approach, developed at MIT, could provide realistic predictions for a wide

Photonic processor could enable ultrafast AI computations with Researchers developed a fully integrated photonic processor that can perform all the key computations of a deep neural network on a photonic chip, using light. This advance

AI simulation gives people a glimpse of their potential future self The AI system uses this information to create what the researchers call “future self memories” which provide a backstory the model pulls from when interacting with the user. For

Back to Home: https://explore.gcts.edu