

Hardware Challenges in AI Sensors and Innovative Approaches to Overcome Them[†]

Filip Tsvetanov 

Department of Communication and Computer Engineering and Technologies,
South-West University “Neofit Rilski”, 2700 Blagoevgrad, Bulgaria; ftsvetanov@swu.bg

[†] Presented at the International Conference on Electronics, Engineering Physics and Earth Science (EEPES 2025), Alexandroupolis, Greece, 18–20 June 2025.

Abstract

Intelligent sensors with embedded AI are key to modern cyber-physical systems. They find applications in industrial automation, medical diagnostics and healthcare, smart cities, and autonomous systems. Despite their significant potential, they face several hardware challenges related to computing power, energy consumption, communication capabilities, and security, which limit their effectiveness. This article analyzes factors influencing the production and deployment of AI sensors. The key limitations are energy efficiency, computing power, scalability, and integration of AI sensors in real-time conditions. Among the main problems are the high requirements for data processing, the limitations of traditional microprocessors, and the balance between performance and energy consumption. To meet these challenges, the article presents several practical and innovative approaches, including the development of specialized microprocessors and optimized architectures for “edge computing,” which promise radical reductions in latency and power consumption. Through a synthesis of current research and practical examples, the article emphasizes the need for intermediate hardware–software solutions and standardization for mass deployment of AI sensors.

Keywords: smart sensors; AI sensors; hardware; energy efficiency; edge AI; challenges



Academic Editor: Grigor Mihaylov

Published: 25 August 2025

Citation: Tsvetanov, F. Hardware Challenges in AI Sensors and Innovative Approaches to Overcome Them. *Eng. Proc.* **2025**, *104*, 19. <https://doi.org/10.3390/engproc2025104019>

Copyright: © 2025 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The development of technology is giving impetus to the production and application of smart sensors that collect and process multi-sensor signals and perform interactions required for monitoring and control in many fields such as medicine, healthcare, agriculture, military industry, autonomous aircraft and vehicles, security, marketing profiling of customers, etc., [1]. AI provides powerful tools and algorithms for processing and analysis that enable the huge amount of data collected from intelligent sensor monitoring of the physical world to be analyzed and used to generate insights, make highly accurate predictions, and take effective actions. Combining AI and sensors is a rapidly developing field and a powerful way to understand and influence the world around us. Research shows a strong, growing demand for energy-efficient, high-performance hardware tailored to the unique requirements of AI-based applications implemented on IoT devices [2].

The rapid development of AI and its penetration into peripheral devices pose significant hardware challenges for modern sensor systems. The integration of AI with AIoT devices has experienced rapid growth, largely due to the effective implementation of deep

neural networks (DNNs) and machine learning (ML) models across various domains. However, deploying ML and DNN on such devices comes with challenges, primarily related to computational capacity, storage, and energy efficiency [3]. The implementation of AI in sensors is associated with challenges such as temperature constraints caused by the huge computations and the heat generated during these intensive computations. At the same time, AI sensors require more memory. Integrating memory modules into small sensors requires creative solutions for data compression and optimization. The transmission of AI-based information requires high-speed and secure communication protocols. Efficient edge computing protocols are needed. The challenge is to minimize latency and maximize throughput and security [4,5].

This paper analyzes the key constraints related to energy efficiency, computing power, and the integration of AI sensors in real-time environments. The analysis highlights the main issues related to high data processing requirements, the limitations of traditional microprocessors when working with ML and neural networks, and the balance between performance and power consumption using high security in the IoT or cloud. Approaches are promising and will improve performance. By outlining and addressing the relevant issues, we hope to demonstrate the real benefits of these technologies and the need for a multidisciplinary expert approach and collaboration between researchers, industry participants, policymakers, and end users to find practical and actionable solutions for innovative sensor hardware.

To address these challenges, various innovative approaches can be used, leading to a reduction in data transfer and processing delays and, consequently, energy consumption. Therefore, the latest innovative technologies in the field of memories used in sensor devices, optimization of architectures for Edge AI, and implementation of specialized AI microprocessors can be used.

2. Related Works

2.1. Applicability of Sensors with AI Factors

According to research works [6–8], there is an increased consumer interest in the development of applications based on AI, which enable decisions to be made in various industrial areas, driven by sensor data utilizing the most adaptive AI technologies. Technological innovations in sensor technology, such as minimizing sensor size, increasing accuracy, and reducing power consumption, enable the integration of these sensors into smaller devices and collect more data in real-time. There is a trend towards the increased adoption of intelligent systems in various fields such as agriculture, medicine, home automation, and industrial applications of the Internet of Things [9,10]. Government initiatives, funding programs, and partnerships also play an important role in determining how AI sensors are used and adapted. For example, the European Union is allocating a lot of money to AI-related projects and AI sensors. The development of smart cities is also driving the growth of the AI sensor market [11].

The increasing use of cloud platforms for data storage and processing also contributes to the rise of the AI sensor market. Cloud infrastructures provide access to vast amounts of data, processed and analyzed in real-time by AI sensors, allowing them to make informed decisions. This increases the demand for AI sensors, especially in healthcare for patient diagnosis and treatment, as well as in drones and vehicles for obstacle detection and collision avoidance [12].

2.2. Problems Integrating AI Sensors with Industrial Systems

The authors in [13] consider sensor hardware from the perspective of architectural constraints caused by the requirements for limited energy resources, small size, and real-time

processing. In [14], the integration of artificial intelligence techniques into electrochemical biosensors for disease prediction and diagnosis is analyzed. In [15], an energy management system in smart buildings using artificial intelligence sensors is proposed, which more accurately predicts energy consumption, facilitates real-time monitoring, and detects anomalies. The authors in [16] discuss the issue of faulty AI sensors in the agricultural sector and the likelihood that they provide misleading data, which can reflect incorrect plant development and profitability. Defect detection in AI sensors for agriculture using neural networks is proposed in [17]. A study on hardware risks in AI sensors is proposed in [18]. The authors examine FPGA-based security approaches, utilizing AI for security in conjunction with traditional processors, and developing AI-based security solutions on FPGA platforms.

2.3. State of the AI-Powered Smart Sensor Market

Research conducted by consulting and marketing companies on the trends of the global AI sensor market [2,19,20], etc., focuses on the manufacturing technology of AI sensors, the market segments in which they are used, the key manufacturers of these sensors, and what is the global and regional market status of AI sensors. The results of the research by various marketing companies show an increase in their production from 31.4% to 45.3% by 2030. The global AI sensor market reached USD 4.06 billion in 2023 and is expected to reach USD 55.62 billion by 2030. The European AI sensor market is also expected to grow at a compound annual growth rate (CAGR) of 38.4% from 2024 to 2030. The dominant consumer of AI sensors is Germany, followed by the UK and the French market [21]. There is also a trend of increased interest in the production of chips for sensor-based intelligent devices, and according to [22], their market is expected to reach USD 400 billion by 2027, up from USD 45 billion in 2023. Nvidia is the leader in the production of all chips used for AI, producing more than 80% of the chips and holding a near monopoly in the network equipment used to combine the chips in data center servers. Table 1 summarizes the interest and applicability of sensors with AI.

Table 1. Interest in AI sensors.

Sensor Types	By Technology, Managing the Functionality	List of Companies, Leading AI Sensors	Purpose of the Application
Light Position Pressure Navigation Temperature Ultrasonic Optical Motions	Machine learning, Natural language processing, Computer vision, Context-aware Computing data	Teledyne Tech. Incor., Goertek Inc., Yokogawa Electric Corp., Robert Bosch GmbH, Sensata Tech., Excelitas Tech. Corp., Sony Corp., Baidu, Inc. Hokuriku El. Industry Co., BAE Systems PLC, Oracle Corp., RELX PLC, Inc. Sensirion AG, Sensortek Tech. Corp., Silicon Sensing Systems Limited data.	Industrial support and process management in various industries such as power supply, agriculture, mechanical engineering, etc., smart home, innovative city management; monitoring, diagnostics of patients in healthcare; Autonomous driving of cars, etc.

Figure 1 shows the interest in a specific type of AI sensor. Users show the greatest interest in pressure, temperature, optical, position, motion, and navigation sensors. AI motion sensors hold a leading position [23]. This is related to the tasks they perform in various systems, including industrial automation, fitness wearables, healthcare, and monitoring and security systems, which increases their demand.

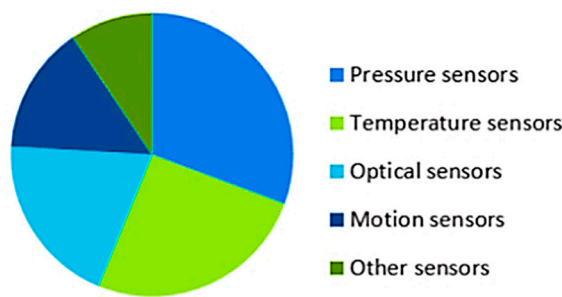


Figure 1. AI sensor market by type in 2023.

Leaders in the production of AI sensors include companies such as Teledyne Technologies Incorporated, Goertek Inc., Yokogawa Electric Corporation, Robert Bosch GmbH, and Sensata Technologies, among others. When familiarizing themselves with the activities of the companies, it is immediately noticeable that almost all of them develop research and innovation activities in the field of sensor technologies and are leaders in this field. For example, Yokogawa develops molecular spectroscopic sensors to visualize and detect invisible information. This technology is expected to boost various industries, including minimally invasive medical treatments for cancer diagnosis, blood sampling, and genetic testing, as well as logistics management. The authors in paper [23] consider the application of flexible, thin sensors that can be flexibly installed over large areas and adapted for various purposes and measurements, including industrial infrastructure, traceability, robots and cars, healthcare and other applications, as well as for optimal plant monitoring, quality control in logistics and vital data management.

3. Characteristics of Smart Sensors with AI

Sensors with integrated AI are designed to process raw data, identify patterns, and make sense of the information they collect. They have the ability to collect, understand, interpret, learn, predict, adapt, make decisions, respond to various interfaces, and function autonomously without human intervention. Data processing is performed by AI models that are trained on large datasets collected from various sensors. AI sensors with embedded intelligence can communicate collaboratively or over the internet.

3.1. Functional Difference Between an Intelligent Sensor and a Sensor with Built-In AI

The difference between a “smart sensor” and an “AI sensor” lies in their functionality and data processing capabilities, which are systematized in Table 2.

Table 2. Comparative analysis between an AI and a smart sensor.

Characteristics	Smart Sensors	AI Sensors
Data processing	Local processing of essential data with predefined algorithms. Basic filtering	Working with AI algorithms, machine learning, neural networks, and pattern recognition
Adaptivity	Adaptability is limited	Self-learning models
Noise filtering	Filtering through specific mathematical algorithms	Self-learning noise and detecting anomalies
Calibration	Manual or semi-automatic	Automatic with AI-based corrections
Real-time solutions	Fast responses, but with limited complexity	Autonomous decision-making based on more complex AI analyses
Network integration	Connectivity with IoT platforms and industrial protocols	Integration with AI-based networks and self-learning
Energy consumption	Lower	Higher

Table 2. Cont.

Characteristics	Smart Sensors	AI Sensors
Predictive analysis	Limited	Autonomous
Energy efficiency	Lower energy consumption	Algorithm-dependent
Communication	Local or cloud	Mostly local
Latency	Minimal	Depends on the algorithm
Complexity of implementation	Easy	Complex implementation and training of AI models
Applications	Industrial automation, medicine, IoT, and more	Autonomous vehicles, Industry 4.0, customized solutions
Benefits	Lower energy consumption Easier integration and reliability	Self-learning, adaptation, prediction, trends, autonomy, independence from external computing resources
Disadvantages	Limited adaptation to new conditions and inability to process complex unstructured data. Making complex decisions through external systems	Higher consumption, Requires training and setup, Complex implementation and maintenance

3.2. Hardware Components of an AI Sensor

Figure 2 presents a block diagram of an AI sensor, and Tables 3 and 4 describe the main modules.

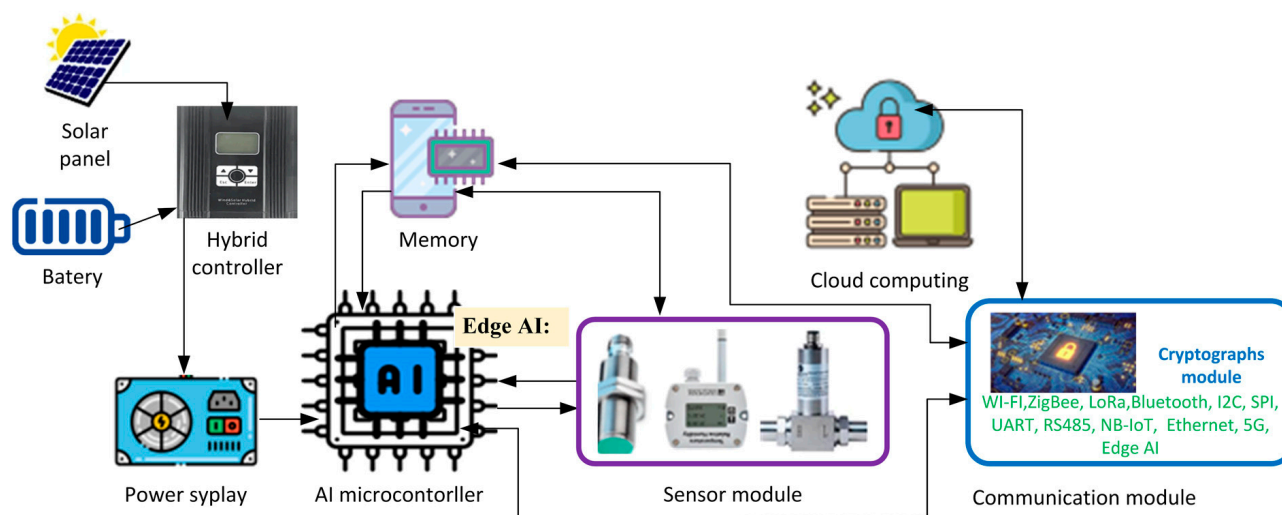


Figure 2. Main components of AI sensors.

Table 3. Hardware analysis and smart and AI sensors.

Characteristic	Smart Sensor	AI Sensor
Processor (CPU/MCU/AI chip)	Low-power microcontroller (MCU)	High-performance CPU/GPU/NPU for AI computing
Memory (RAM/Flash/EEPROM)	Small amount of RAM and Flash memory for basic operations	More RAM and NAND/SSD memory for processing complex AI algorithms
Analog-to-digital converter (ADC)	Built-in low-resolution ADC (10–12 bits)	High-precision ADC (16–24 bits) for processing complex data
Sensor module	Basic sensor element (temperature, pressure, motion)	Advanced sensors (lidar, infrared, multispectral)
Embedded algorithms	Fixed filtering and calibration algorithms	Dynamic algorithms for machine learning and neural networks

Table 3. *Cont.*

Characteristic	Smart Sensor	AI Sensor
Neural processing unit (NPU)	Lacks	Dedicated NPU for AI computing
Communication interfaces	2C, SPI, UART, RS-485	I2C, SPI, UART, RS485 + Ethernet, Wi-Fi, 5G, Edge AI
Power consumption	Low (few milliwatts)	Higher (watts to tens of watts) for complex processing
Power supply	Batteries, energy-efficient solutions	High-performance power supplies (12 V, 24 V, PoE)
Form factor	Compact, integrated into IoT devices	Larger size due to additional AI hardware
Thermal management	Minimal heat dissipation, no active cooling required	Higher due to powerful hardware and AI capabilities
Price	Lower, affordable mass production	May require heatsinks or active cooling (fans)

Table 4. Comparative analysis of smart and AI humidity sensors.

Characteristic	Smart Sensor	AI Sensor
Processor	Microcontroller (MCU) with basic processing	Powerful processor (MCU + NPU) for AI analysis
Memory	A little RAM (2–32 KB) and Flash (up to 256 KB)	Expanded RAM (512 MB–4 GB) and NAND/SSD
Sensor module	Capacitive or resistive sensor for moisture measurement	Advanced multi-channel capacitive sensor with ML processing
Analog-to-digital converter (ADC)	10–12 bits for basic measurements	16–24 bits for precise humidity analysis
Embedded algorithms	Fixed calibration and noise reduction algorithms	Dynamic machine learning for improved accuracy
Neural processing unit (NPU)	Missing	Specialized NPU for real-time data analysis
Communication interfaces	I2C, SPI, UART, RS-485	I2C, SPI, UART, RS485 + Ethernet, Wi-Fi, 5G, Edge AI
Power consumption	Low (1–5 mW)	Higher (200–500 mW) due to AI calculations
Power supply	3 V–5 V DC	5 V–12 V DC or PoE
Form factor	Compact, suitable for IoT devices	Bigger due to additional AI processing hardware
Thermal management	Minimal heat dissipation	Possible need for a heat sink for high computing power
Price	Lower (USD 5–USD 50)	Higher (USD 50–USD 300)

The data analysis in Tables 3 and 4 clearly outlines the hardware challenges of AI sensors.

As shown in Tables 2–4, smart sensors are more compact, energy-efficient, and utilize basic microcontrollers. They are suitable for simpler measurements and monitoring in standard measurement and control systems. Sensors with built-in AI feature powerful processors, larger memory, and advanced sensor modules. They usually have more complex data processing algorithms that allow them to recognize objects, predict trends, analyze behavior, and take automated actions in real-time, making them suitable for autonomous systems, predictive maintenance, and advanced data analysis.

3.3. Ways to Implement AI in Sensors

AI is most often integrated into the software part of the sensor device; it runs on the sensor's computing platform, and can be implemented in several ways:

3.3.1. Embedded Software

Directly integrated into the sensor software stack and running on the microcontroller or processor, they are often called Edge AI sensors. These sensors perform intelligent local feedback, such as real-time data analysis and interpretation without external computing resources. Embedded AI software in sensors may include the following components [24]:

- Machine learning algorithms train the sensor to recognize and analyze the objects, scenes, patterns, or behaviors in the collected data;
- Prediction and optimization models, based on past data to predict future events, optimize resources and processes, forecast trends, warn of anomalies, or optimize resources such as energy or time;
- Image and video processing algorithms that enable the recognition of objects, faces, motion, and other features in visual data;
- Interfaces for interaction and communication with other devices or systems allow the sensor to transmit data, receive commands, and integrate with various control and monitoring systems.

These embedded AI software components work together to enable the sensor to analyze and interpret data in real-time, delivering features and application capabilities with intelligent characteristics.

3.3.2. Cloud Data Synthesis

AI can be integrated into the sensor by connecting to cloud services for data processing and analysis. The sensor sends data to a cloud infrastructure for processing and analysis. This model is used when the sensor does not have sufficient computing resources or algorithms to process data on the device, but can connect to the Internet. Cloud-based data synthesis allows devices with limited resources to use complex analytical algorithms and data processing models, while providing the ability to centrally manage and monitor data from different devices. At the same time, it should be noted that cloud solutions have disadvantages, especially in terms of energy consumption, data security, and response speed.

3.3.3. Hybrid Approach

Sensors powered by both on-premises and cloud-based AI use a hybrid approach. Such sensors can process data in real-time, while more complex analytical operations are performed on the cloud platform. This approach increases the sensor's flexibility and ability to handle various requirements and scenarios. The choice of a specific approach depends on the application requirements, including real-time data processing, level of intelligence, internet accessibility, and other factors.

3.3.4. Comparative Analysis and Evaluation of Possible Applications and Real-World Implementations Between Edge AI and Cloud AI

Research shows that Edge AI sensors significantly reduce latency by processing data closer to the source, making them ideal for real-time applications such as autonomous vehicles and IoT devices that require immediate responses. Such sensors reduce the need for constant data transmission, saving bandwidth and lowering communication costs, especially when network quality is variable or data volumes are high [25]. Edge AI sensors are also believed to enhance privacy by keeping sensitive data locally, reducing exposure to external threats during transmission [26]. At the same time, criticism has also been made

regarding the suboptimal accuracy resulting from smaller training datasets and limited resources [27]. These sensors offer an advantage in reducing operational costs and fees for data transfer and cloud usage, but at the same time, may require investments in more powerful edge hardware.

Cloud AI sensors have been found to have higher latency due to data transmission to remote servers, which is considered unacceptable for time-sensitive tasks [28–31]. Also, processing sensor data by Cloud AI requires significantly more complex and large-scale models. Cloud AI relies on continuous data transfer, which can lead to congestion, higher costs, and delays, especially with high-frequency data streams [25,31]. Research in [28,31] on data security and privacy when processing sensor data with Cloud AI confirms the potential for an increased risk of large-scale breaches when the cloud is compromised.

Many studies recommend a hybrid edge-cloud model to dynamically balance latency, accuracy, resource utilization, and cost, thereby leveraging the strengths of both paradigms to meet different application needs [32–34].

The choice between Edge AI and Cloud AI involves latency, computing resources, privacy, accuracy, and cost trade-offs. Hybrid models that combine the two approaches are increasingly preferred to meet diverse application requirements.

3.3.5. Types of AI Integrated into Sensors

AI can be implemented differently depending on the specific applications and sensor requirements. Here are some of the main types of AI that can be embedded in sensor devices, shown in Figure 3.

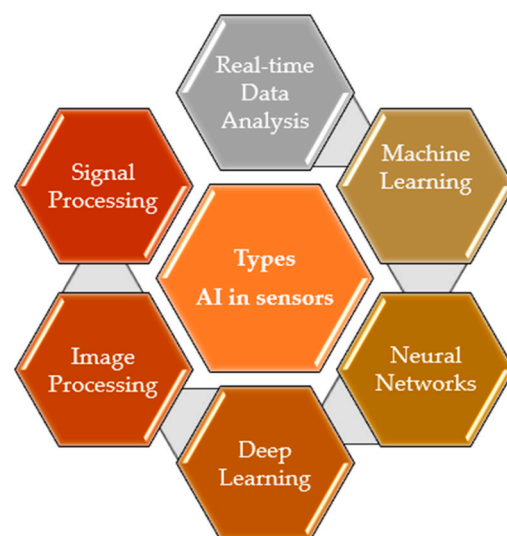


Figure 3. Types of AI embedded in sensors.

Machine learning is one of the primary methods for embedding AI in sensors. ML allows sensors to learn by analyzing data and identifying patterns without explicit programming. This includes supervised, unsupervised, and reinforcement learning [35,36].

Neural networks can recognize complex patterns and regularities in data, making them suitable for sensor data analysis applications [37].

Deep learning is a specific type of ML that uses neural networks with many layers to extract high-level abstractions from data. This method is used in sensors for object recognition, image classification, speech processing, and other tasks [38].

Image processing is applied to cameras that use image processing methods to analyze and recognize objects, faces, gestures, and other visual data. These methods include various techniques such as segmentation, object detection, feature extraction, and classification [39].

Signal processing is typical for sensors with microphones or accelerometers. It plays an important role in analyzing and interpreting data [40].

Real-time data analysis allows quick and effective response to events and data anomalies without delay [41].

4. Hardware Challenges in AI Sensors

As the comparative analysis in Tables 2–4 demonstrates, AI sensors are highly functional because they perform processing and communication with other devices in sensor networks, as well as measure and store data. This increases the complexity of the design and requires high-performance hardware. At the same time, they face several challenges, some of which are discussed below, as shown in Figure 4.

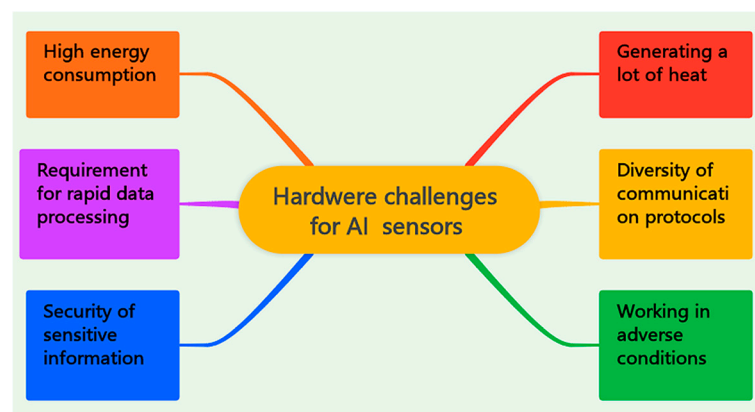


Figure 4. Hardware challenges for AI sensors.

Implementing AI sensors in extreme conditions, such as temperatures, humidity, vibrations, or pollution, requires specialized materials and housing designs. Embedded data processing, on the other hand, requires the protection of sensitive information. This includes adding cryptographic modules or specialized chips, increasing complexity and energy costs. AI processors generate a lot of heat. The small size of the sensors makes it difficult to effectively dissipate this heat, which can lead to overheating and malfunction. Overcoming this challenge requires mounting passive coolers on the sensor, which increases their size. Some AI algorithms require significant power, especially when continuously processing real-time data. Artificial intelligence sensors integrate with various IoT networks, including heterogeneous devices from multiple manufacturers that use different communication protocols and standards. Standardized interfaces, protocols, and interoperability frameworks are needed to seamlessly integrate these heterogeneous devices into a single energy-optimized system and enable coordination and communication [38,40]. Ensuring low latency in data processing is also essential for applications such as autonomous systems and industrial robots. Hardware must be optimized for maximum performance with minimal latency. Addressing diverse challenges has led to several hardware and architectural approaches and specific tactics that reduce power consumption and increase data security.

Therefore, we can summarize that balancing sensor performance and power consumption is a significant challenge. Less computation, fewer data, and less data movement mean less power consumption. It is these challenges that highlight the need to innovate in the hardware design of AI sensors. It should be noted that the development of AI sensor hardware requires high-quality and more expensive components, which leads to higher manufacturing costs.

5. Possible Approaches to Solving Sensor Hardware Challenges

5.1. Implementing AI Accelerators in Microprocessors of AI Sensors

The requirement for low power consumption is due to the fact that AI sensors are mostly battery-powered and must maintain a long lifespan. This limits the possibility of using powerful processors, resulting in limited computing resources. At the same time, the trend towards minimizing dimensions leads to their compactness and reduced hardware space for sensors, which in turn reduces the space available for more powerful chips and cooling systems in sensor designs. Integrating complex algorithms, especially DL algorithms, by embedding ML models in hardware requires significant storage capacity for large databases and powerful computational processing. Their implementation requires expensive, high-performance computing resources. They are developed and tested on development platforms to optimize the performance of AI algorithms for sensors [36]. Suitable platforms might be Azure OpenAI, Google Vertex AI, Microsoft Azure Machine Learning Studio, Amazon SageMaker, Hugging Face, TensorFlow, Keras, Caffe, or freely available ones such as Azure OpenAI. Some platforms are popular standard neural networks (NNs), such as AlexNet and VGG from the Visual Geometry Group. User research shows that as of January 2025, the Google Vertex AI platform has the highest approval rating at 18.8% [41]. NNs are developed for specific applications and then implemented in hardware called hardware accelerators. Hardware accelerators use parallelism to increase performance and provide much higher performance than traditional processors, and provide the computing power needed for AI and ML. An AI accelerator is a specialized hardware accelerator or automatic data processing system designed to accelerate computer applications, especially artificial neural networks, machine learning, and machine visualization. Typical applications include algorithms for AI, IoT, and various data-intensive or sensor-driven tasks [42]. Hardware implementations of accelerators are implemented using FPGAs, GPUs, and ASICs to accelerate computations and maintain the required accuracy [42,43]. Some companies have begun manufacturing specialized AI microcontrollers with built-in hardware accelerators for battery-powered applications, addressing these challenges. Hardware accelerators with AI minimize data movement and use parallelism to optimize power usage and performance. For example, the MAX78000, MAX78002 AI microcontrollers have a hardware-based convolutional neural network (CNN) accelerator and enable battery-powered applications to perform manufacturing tasks while consuming ultra-low power, only microjoules of energy [44]. Applications for the MAX78002 AI microcontroller include factory robots and drone navigation, industrial sensors and process control, embedded vision systems for quality assurance, smart security cameras, and portable medical diagnostic equipment. The Apollo5 and Apollo510 AI microcontroller families are claimed to deliver exceptional power efficiency with $300\times$ performance [45]. These ultra-low power microcontrollers are based on the industry-standard Arm Cortex-M and SPOT platforms, which include a range of cores to suit different levels of complexity. They also enable advanced processing for power-sensitive applications, maximizing battery life without compromising performance. These intelligent sensor microcontrollers also utilize lower voltages and a design with new deep sleep modes, providing unparalleled power efficiency and high connectivity. Their high security is realized through hardware isolation to protect sensitive data with Trust Zone (M23/M33), cryptographic accelerators such as AES, SHA, RNG, or STM32L5, and Secure Boot/Flash to prevent unauthorized access.

5.2. Reducing Energy Consumption by Reducing the Volume of Data Processed

AI sensors integrated in various sensor networks extract important and meaningful information about the observed processes and objects by collecting zettabytes of data daily, which is associated with the consumption of a large amount of energy by the sensors. The

approach of collecting and storing all possible data in the cloud with the expectation that it will later be structured, analyzed, and evaluated is widespread. In most cases, a significant amount of collected data is not processed, the collected and stored data is not used, or its subsequent use after a long period makes its storage expensive. Then it is more appropriate to make conceptual considerations earlier, and to determine which information is suitable for the application and where in the data flow the information can be extracted [35]. It is recommended to filter the data so that only the essential and helpful information for the specific task is stored and processed. Reducing the volume of data for processing, analysis, and decision-making can be carried out in several different ways, such as the following:

5.2.1. Transforming Sensor Data into Intelligent Data

If only data is available, but not essential information that can be described in mathematical equations, then data-driven approaches should be chosen. These algorithms directly extract the desired information, known as intelligent data, from sensor data, also referred to as big data. Two approaches to building a model are always part of an AI algorithm: intelligent data-driven and model-based [35]. Smart data is structured, contextualized, analyzed, and ready for making informed decisions, typical of IoT devices. It differs from ordinary raw data in that it has been processed, made meaningful, and often enriched with additional information through technologies such as AI, machine learning, and big data analytics, as shown in Figure 5. Using AI algorithms, sensor data can be transformed into intelligent data, resulting in reduced bandwidth and reduced power consumption compared to directly transmitting raw sensor data to the edge or cloud [46]. Thus, this approach sends only the relevant data to the application for storage and processing.

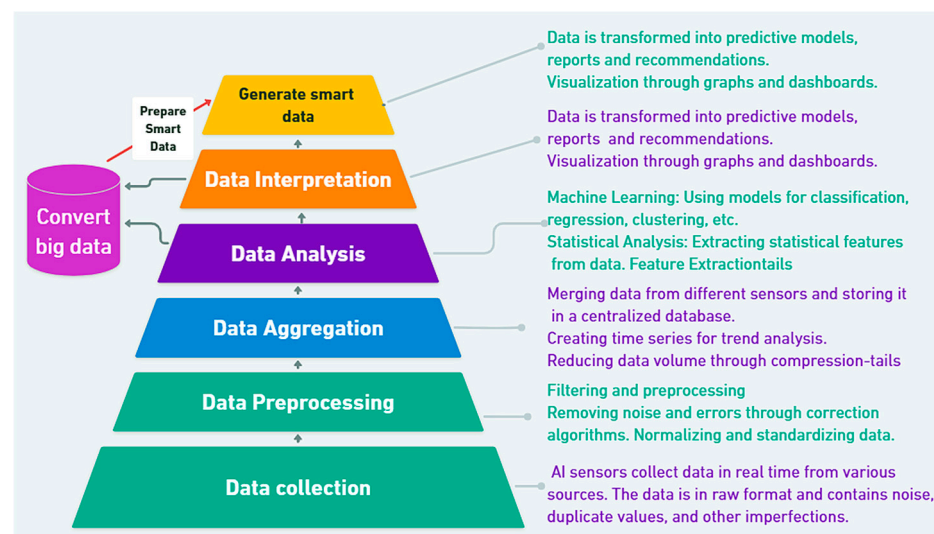


Figure 5. Transformation of raw data into intelligent data.

5.2.2. Local Embedded Processing near the Sensor, Edge AI

Another approach to reducing the energy consumption of sensor data processing is to perform local embedded processing near the sensor before processing the data in the cloud [47]. This approach is also useful when privacy, latency, or bandwidth constraints exist. The microcontroller performs local embedded processing and data storage in the sensor without sending the data to the cloud or a central server. As a result, the transmitted data volume is reduced, as only the results processed by the sensor processor are included.

In essence, Edge AI combines Edge Computing and AI, where the machine learning AI processes the data generated by the hardware device locally, with or without an internet connection. Therefore, Edge AI performs more advanced processing, such as running

neural networks on the device. This way, instead of sending data to the cloud for analysis, the device can make decisions or inferences on-site at the sensor. This would also reduce the data transmitted, because only the results are sent, not the raw data. This way, the data can be processed within a few milliseconds, providing real-time feedback [48]. Edge AI can run on a wide range of hardware platforms, from regular MCUs to AI microprocessors. Edge AI software has many ML algorithms that run on physical devices such as sensors or gateways. These algorithms process data locally, eliminating the need to send it to remote servers in real time. This also increases the security and privacy of sensitive information, as the data is not exported outside the device. This eliminates the exposure of sensitive data to cybercriminals while cryptography is performed on-site. Decision-making is low-latency, occurring in milliseconds, whereas in Cloud AI, latency can reach seconds or minutes [4,5,49]. Data filtering also reduces network load, as only relevant or aggregated results are transmitted to the cloud. Bandwidth requirements and costs are reduced, as Edge AI processes data locally on the device itself, reducing the cost of internet bandwidth and cloud storage [50]. Devices work even without the internet (e.g., sensors in remote areas). Moving data over fewer and shorter distances also reduces energy consumption. Since Edge AI processes data at the device level, it saves energy costs when moving it to the cloud for processing, which is recommended according to the study in [51]. Therefore, in most cases, implementing the algorithms in the sensor or as close to the sensor as possible is more profitable. This compresses and refines the data early, reducing communication and storage costs. In these cases, the essential information from the data is extracted earlier, and the algorithm development is not complicated. Therefore, both approaches involve local processing of the data, reducing the amount of data transmitted. Edge AI uses AI algorithms, while locally embedded processing can be any processing, even simple filtering or basic calculations. So, Edge AI is a subset of local processing with AI included. The hardware difference is in the type of microprocessor. Edge AI requires more powerful processors or specialized chips to run the models efficiently, while traditional embedded systems can use simpler microcontrollers.

5.2.3. Reducing the Accuracy of Collected Data

Such an approach to using more efficient ML models with reduced precision is presented in [40]. According to the authors, the default data size for programmable platforms, such as CPUs and GPUs, is often 64 or 32 bits with floating-point representation. In ML training, data can be represented with fixed points and significantly reduced bit widths to save power and storage area and increase performance. In hand-crafted approaches, the bit width can be drastically reduced to below 16 bits without affecting accuracy. The authors [52] published research on accelerators that support up to 16-bit fixed point, resulting in reduced data collection.

5.2.4. Data Compression and Optimization

Data compression is often recommended to reduce the cost of data movement and storage. It reduces the amount of information being transmitted and the data rate, an essential factor in reducing data volume. Various forms of lightweight compression have been investigated to reduce the amount of data being moved. Lossless compression can facilitate data transfer on and off the chip [53].

5.3. Using Energy-Efficient Sensors and Minimizing Their Number in Sensor Networks

Choosing low-power sensors explicitly designed for IoT and AI applications is advisable when building sensor systems. Integrating sensors with communication technologies such as Bluetooth Low Energy (BLE), Zigbee, LoRaWAN, and Smart Mesh, and operating in low-power modes are popular options for switching sensors to power-saving mode

when inactive. Another approach involves using smart timers to activate and deactivate sensors as needed. Using alternative methods to power sensors with solar panels or other renewable energy sources is an approach that has been gaining popularity recently. Minimizing the number of sensors and optimizing their number and placement in the sensor field is often applied to avoid unnecessary power consumption.

5.4. Using Modern Memory Technologies

Smart sensors with embedded AI require high-performance memory technologies that provide fast data processing, low power consumption, and high reliability. The choice of memory depends on the specific requirements of the smart sensor—speed, energy efficiency, endurance, and storage capacity. Some of the most advanced memory technologies used in these devices include Magnetoresistive RAM (MRAM), which combines DRAM's speed with flash memory's robustness. It is suitable for applications requiring energy efficiency and radiation resistance. According to the work in [54], this memory provides low power consumption and high speed of operation. It is suitable for neuromorphic calculations in AI sensors that mimic the human brain's work. It gives the ability to store the weights of neural networks directly in memory. Ferroelectric RAM (FeRAM) memory provides high write speed and low power consumption. The advantage is that it retains its data even when power is lost. It is used in IoT devices and sensors operating in energy-constrained environments [55]. This memory is based on carbon nanotubes and provides extremely high density and speed. Suitable for AI-based sensors that process large data sets. The 3D XPoint (Intel Optane) is faster than NAND flash memory but has a similar architecture. It is used for high-performance AI sensors that require high data bandwidth. Embedded Flash (eFlash) is often used in microcontrollers with embedded AI. It allows firmware and AI models to be stored directly in the sensor. PCM (phase-change memory) has high read and write speeds and good energy efficiency. It is used in AI sensors that require frequent updating of model parameters. It is used in sensors that store data when the power is turned off. RRAM and PCM are particularly promising for future AI applications in sensor technologies, as they balance performance and efficiency.

6. Conclusions

Using AI, smart sensor technologies that collect and process environmental data are changing the technology altogether. Smart sensors have hardware and functional differences and AI-based solutions that reveal the technological potential of this evolution in systems can be integrated into almost any device. Despite the existing hardware challenges, the proposed approaches to overcoming them show that the industry is actively optimizing and improving these systems. Implementing AI accelerators in microprocessors, innovative methods for reducing power consumption, and optimization of data processing outline a clear course for the future development of AI sensors. Local data processing solutions and Edge AI increase efficiency and solve significant problems with data latency and security.

It is essential to continue research to address the current technological limitations and adapt new hardware and software innovations to the specific needs of application areas. Solving these challenges is inconceivable without the joint efforts of specialists from academia and industry.

Funding: The research and all payments for conducting the experiments and subsequent fees were borne by me, the author of the publication.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are obtained in the article.

Acknowledgments: The study was conducted in the laboratories of South-West University “Neofit Rilski”, Blagoevgrad, Bulgaria.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Zhang, D.; Wei, B. Smart Sensors and Devices in Artificial Intelligence. *Sensors* **2020**, *20*, 5945. [CrossRef] [PubMed]
2. Stellar Market Research Pvt Ltd. Global Artificial Intelligence Sensor Market Analysis and Forecast (2024–2030). Available online: <https://www.linkedin.com/pulse/global-artificial-intelligence-sensor-market-analysis-wvovxf/> (accessed on 21 May 2024).
3. Mohanty, S.N.; Lydia, E.L.; Elhoseny, M.; Al Otaibi, M.M.G.; Shankar, K. Deep learning with LSTM based distributed data mining model for energy efficient wireless sensor networks. *Phys. Commun.* **2020**, *40*, 101097. [CrossRef]
4. Tsvetanov, F.A.; Pandurski, M. Selection of Protocols for Integration of Sensory Data Networks in Cloud Structures. *Int. J. Online Biomed. Eng.* **2022**, *18*, 29–40. [CrossRef]
5. Tsvetanov, F.; Pandurski, M. Delay of Transmitted Data in the Remote Patient Monitoring System through AMQP and CoAP. *Int. J. Online Biomed. Eng.* **2024**, *20*, 130–144. [CrossRef]
6. Thapa, R.; Poudel, S.; Krukiewicz, K.; Kunwar, A. A topical review on AI-interlinked biodomain sensors for multi-purpose applications. *Measurement* **2024**, *227*, 114123. [CrossRef]
7. Tsvetanov, F. Integrating AI Technologies into Remote Monitoring Patient Systems. *Eng. Proc.* **2024**, *70*, 54. [CrossRef]
8. Dimcheva, G. Opportunities for Application of Artificial Intelligence in Telecommunication Projects. *Eng. Proc.* **2024**, *70*, 18. [CrossRef]
9. Akosah, S.; Gratchev, I.; Kim, D.-H.; Ohn, S.-Y. Application of Artificial Intelligence and Remote Sensing for Landslide Detection and Prediction: Systematic Review. *Remote Sens.* **2024**, *16*, 2947. [CrossRef]
10. Tsvetanov, F.; Kuzmanov, K. Communication Challenges in Smart Irrigation Systems. In Proceedings of the 9th International Conference on Energy Efficiency and Agricultural Engineering, Ruse, Bulgaria, 27–29 June 2024; pp. 1–6. [CrossRef]
11. Nikolinakos, N.T. Launching a European Initiative on Artificial Intelligence. In *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies—The AI Act*; Springer International Publishing: Cham, Switzerland, 2023; pp. 23–98.
12. Akter, S.; Michael, K.; Uddin, M.R.; McCarthy, G.; Rahman, M. Transforming business using digital innovations: The application of AI, blockchain, cloud and data analytics. *Ann. Oper. Res.* **2022**, *308*, 7–39. [CrossRef]
13. Kumar, P.V.; Kulkarni, A.; Mendhe, D.; Keshar, D.K.; Tilak Babu, S.B.G.; Rajesh, N. AI-Optimized Hardware Design for Internet of Things (IoT) Devices. In Proceedings of the 5th International Conference on Recent Trends in Computer Science and Technology (ICRTCST), Jamshedpur, India, 9–10 April 2024; pp. 21–26. [CrossRef]
14. Nashruddin, S.N.A.B.M.; Salleh, F.H.M.; Yunus, R.M.; Zaman, H.B. Artificial intelligence-powered electrochemical sensor: Recent advances, challenges, and prospects. *Heliyon* **2024**, *10*, e37964. [CrossRef] [PubMed]
15. Alijoyo, F.A. AI-powered deep learning for sustainable industry 4.0 and internet of things: Enhancing energy management in smart buildings. *Alex. Eng. J.* **2024**, *104*, 409–422. [CrossRef]
16. Qazi, S.; Khawaja, B.A.; Farooq, Q.U. IoT-Equipped and AI-Enabled Next Generation Smart Agriculture: A Critical Review, Current Challenges and Future Trends. *IEEE Access* **2022**, *10*, 21219–21235. [CrossRef]
17. Shekarian, S.M.; Aminian, M.; Fallah, A.M.; Moghaddam, V.A. AI-powered sensor fault detection for cost-effective smart greenhouses. *Comput. Electron. Agric.* **2024**, *224*, 109198. [CrossRef]
18. Pham-Quoc, C. FPGA/AI-Powered Data Security for IoT Edge Computing Platforms: A Survey and Open Issues. In *Intelligence of Things: Technologies and Applications (ICIT 2023)*; Dao, N.N., Thinh, T.N., Nguyen, N.T., Eds.; Lecture Notes on Data Engineering and Communications Technologies; Springer: Cham, Switzerland, 2023; Volume 187. [CrossRef]
19. AI Sensor Market Size to Reach \$41.17 Billion By 2030. Available online: <https://www.grandviewresearch.com/press-release/global-artificial-intelligence-ai-sensor-market> (accessed on 21 May 2024).
20. Artificial Intelligence Sensors Market: Global Industry Analysis, Trends, Size, Share and Forecasts to 2030. 2023. Available online: <https://www.infiniumglobalresearch.com/market-reports/global-artificial-intelligence-sensors-market> (accessed on 25 January 2025).
21. Europe Artificial Intelligence (AI) Sensor Market Size, Share & Industry Trends Analysis Report By Type 2023–2030. Available online: <https://www.kbvresearch.com/europe-artificial-intelligence-sensor-market/> (accessed on 25 January 2025).
22. Empowering Intelligence: Exploring the Artificial Intelligence Sensor Market In The Latest Research. Available online: <https://www.whatech.com/og/markets-research/industrial/788318-empowering-intelligence-exploring-the-artificial-intelligence-sensor-market-2024-2030.html> (accessed on 25 April 2024).
23. Innovation Center: Yokogawa’s core R&D organization. Available online: https://www.yokogawa.com/eu/about/research-development/inv_center/laboratory-innovation-center/#Innovation-Center (accessed on 5 February 2025).

24. Dzianis, L. Embedded Sensor Platform with AI Algorithms—Locally from Big Data to Smart Data, Analog Devices, 1 February 2020. Available online: <https://www.analog.com/en/resources/technical-articles/embedded-sensor-platform-with-ai-algorithms-locally-from-big-data-to-smart-data.html> (accessed on 5 February 2025).
25. Shi, Y.; Yang, K.; Jiang, T.; Zhang, J.; Letaief, K. Communication-Efficient Edge AI: Algorithms and Systems. *IEEE Commun. Surv. Tutor.* **2020**, *22*, 2167–2191. [CrossRef]
26. Singh, R.; Gill, S. Edge AI: A survey. *Internet Things Cyber-Phys. Syst.* **2023**, *3*, 71–92. [CrossRef]
27. Yao, J.; Zhang, S.; Yao, Y.; Wang, F.; Zhang, J.; Chu, Y.; Ji, L.; Jia, K.; Shen, T.; Wu, A.; et al. Edge-Cloud Polarization and Collaboration: A Comprehensive Survey for AI. *IEEE Trans. Knowl. Data Eng.* **2021**, *35*, 6866–6886. [CrossRef]
28. Cherukuri, B. Edge Computing vs. Cloud Computing: A Comparative Analysis for Real-Time AI Applications. *Int. J. Multidiscip. Res.* **2024**, *6*, 1–17. [CrossRef]
29. Loseto, G.; Scioscia, F.; Ruta, M.; Gramegna, F.; Ieva, S.; Fasciano, C.; Bilenchi, I.; Loconte, D.; Sciascio, E. A Cloud-Edge Artificial Intelligence Framework for Sensor Networks. In Proceedings of the 9th International Workshop on Advances in Sensors and Interfaces (IWASI), Monopoli (Bari), Italy, 8–9 June 2023; pp. 149–154. [CrossRef]
30. Wang, Y.; Xue, J.; Wei, C.; Kuo, C. An Overview on Generative AI at Scale with Edge–Cloud Computing. *IEEE Open J. Commun. Soc.* **2023**, *4*, 2952–2971. [CrossRef]
31. Wu, Y. Cloud-Edge Orchestration for the Internet of Things: Architecture and AI-Powered Data Processing. *IEEE Internet Things J.* **2020**, *8*, 12792–12805. [CrossRef]
32. Yan, C.; Liu, S.; Liu, H.; Peng, X.; Wang, X.; Chen, F.; Fu, L.; Mei, X. Hybrid SD: Edge-Cloud Collaborative Inference for Stable Diffusion Models. *arXiv* **2024**, arXiv:2408.06646. [CrossRef]
33. Wang, W.; Zhang, Y.; Huang, R.; Ren, J.; Lyu, F.; Zhang, Y. Efficient Resource Management and Expansion Scheme for Collaborative Edge-Cloud Computing. *IEEE Trans. Mob. Comput.* **2024**, *23*, 2731–2747. [CrossRef]
34. Hosseinzadeh, M.; Tho, Q.; Ali, S.; Rahmani, A.; Sour, A.; Norouzi, M.; Huynh, B. A Hybrid Service Selection and Composition Model for Cloud-Edge Computing in the Internet of Things. *IEEE Access* **2020**, *8*, 85939–85949. [CrossRef]
35. Blasch, E.; Pham, T.; Chong, C.Y.; Koch, W.; Leung, H.; Braines, D.; Abdelzaher, T. Machine learning/artificial intelligence for sensor data fusion—opportunities and challenges. *IEEE Aerosp. Electron. Syst. Mag.* **2021**, *36*, 80–93. [CrossRef]
36. Gupta, N. Introduction to hardware accelerator systems for artificial intelligence and machine learning. *Adv. Comput.* **2021**, *122*, 1–21. [CrossRef]
37. Talaśka, T. Components of artificial neural networks realized in CMOS technology to be used in intelligent sensors in wireless sensor networks. *Sensors* **2018**, *18*, 4499. [CrossRef]
38. Yuan, S.M.; Hong, Z.W.; Cheng, W.K. Artificial Intelligence and Deep Learning in Sensors and Applications. *Sensors* **2024**, *24*, 3258. [CrossRef]
39. Zhang, Z.; Wang, L.; Lee, C. Recent advances in artificial intelligence sensors. *Adv. Sens. Res.* **2023**, *2*, 2200072. [CrossRef]
40. Hu, F.; Hao, Q. *Intelligent Sensor Networks: The Integration of Sensor Networks, Signal Processing and Machine Learning*; Taylor & Francis: New York, NY, USA, 2012.
41. Sharma, S.; Sharma, K.; Grover, S. Real-Time Data Analysis with Smart Sensors. In *Application of Artificial Intelligence in Wastewater Treatment*; Springer Nature Switzerland: Cham, Switzerland, 2024; pp. 127–153.
42. Best AI Development Platforms. Available online: <https://www.peerspot.com/categories/ai-development-platforms> (accessed on 26 January 2025).
43. Talib, M.A.; Majzoub, S.; Nasir, Q.; Jamal, D. A systematic literature review on hardware implementation of artificial intelligence algorithms. *J. Supercomput.* **2021**, *77*, 1897–1938. [CrossRef]
44. Karras, K.; Pallis, E.; Mastorakis, G.; Nikoloudakis, Y.; Batalla, J.M.; Mavromoustakis, C.X.; Markakis, E. A Hardware Acceleration Platform for AI-Based Inference at the Edge. *Circuits Syst. Signal Process* **2020**, *39*, 1059–1070. [CrossRef]
45. Ultra-Low Power Artificial Intelligence (AI) MCUs. Available online: <https://www.analog.com/en/product-category/ultralow-power-artificial-intelligence-ai-mcus.html> (accessed on 20 February 2025).
46. Apollo510, Advanced AI on a Chip. Available online: <https://ambiq.com/apollo510/> (accessed on 2 February 2025).
47. Betty Jane, J.; Ganesh, E.N. Big data and internet of things for smart data analytics using machine learning techniques. In *Proceeding of the International Conference on Computer Networks, Big Data and IoT*; Springer International Publishing: Berlin/Heidelberg, Germany, 2020; pp. 213–223.
48. Edge AI—What Is It and How Does It Work? Available online: <https://micro.ai/blog/edge-ai-what-is-it-and-how-does-it-work?> (accessed on 20 February 2025).
49. Tsvetanov, F.; Pandurski, M. Efficiency of integration between sensor networks and clouds. *Int. J. Electr. Comput. Eng. Syst. IJECES* **2022**, *13*, 427–433. [CrossRef]
50. Raj, N.; Sinha, A.; Bharddwaj, S.; Yadav, A.L. AI-Powered Energy Consumption Optimization for Smart Homes Using IoT. In Proceedings of the International Conference on Computational Intelligence and Computing Applications (ICCICA), Samalkha, India, 23–24 May 2024; pp. 231–236. [CrossRef]

51. Sze, V.; Chen, Y.H.; Emer, J.; Suleiman, A.; Zhang, Z. Hardware for machine learning: Challenges and opportunities. In Proceedings of the IEEE Custom Integrated Circuits Conference (CICC), Austin, TX, USA, 30 April–3 May 2017; pp. 1–8. [\[CrossRef\]](#)
52. Chen, T.; Du, Z.; Sun, N.; Wang, J.; Wu, C.; Chen, Y.; Temam, O. Diannao: A small-footprint high-throughput accelerator for ubiquitous machine-learning. *ACM SIGARCH Comput. Archit. News* **2014**, *42*, 269–284. [\[CrossRef\]](#)
53. Han, S.; Liu, X.; Mao, H.; Pu, J.; Pedram, A.; Horowitz, M.A.; Dally, W.J. EIE: Efficient inference engine on compressed deep neural network. *ACM SIGARCH Comput. Archit. News* **2016**, *44*, 243–254. [\[CrossRef\]](#)
54. Molas, G.; Nowak, E. Advances in emerging memory technologies: From data storage to artificial intelligence. *Appl. Sci.* **2021**, *11*, 11254. [\[CrossRef\]](#)
55. Reddy, B.V.; Chaudhary, T.; Singh, M.; Raj, B. Ferroelectric Random Access Memory (FeRAM). In *Integrated Devices for Artificial Intelligence and VLSI*; John Wiley & Sons, Inc.: New York, NY, USA, 2024; pp. 125–156.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.